UNIVERSITY OF HULL

IDENTIFICATION OF TUMOUR ANTIGENS THAT MAY FACILITATE EFFECTIVE CANCER DETECTION AND TREATMENT

By

Eithar Mohamed

Doctor of Philosophy

October 2023

Declaration

I, Eithar Mohamed, declare that the work submitted is my own and is not similar in content to, or based on the work of others, published or unpublished, nor does it include contribution of any AI technologies except with full and proper acknowledgement.

INSERT electronic signature here

Covid impact statement

Due to COVID, my doctoral plan has been significantly affected. The pandemic affected the scope of the study. The focus of Hull University Teaching Hospitals NHS Trust was directed towards treating NHS patients and not collecting samples. Sample collection has now started but due to the significant delay the project moved from a focus on incidental lung nodules and biomarkers of their propensity to develop into lung cancer onto a disease we already had samples for, which was adult B-cell acute lymphoblastic leukaemia.

Abstract

Tumour antigens (TAs) play a crucial role in terms of cancer diagnosis and targets for therapy. As TAs are normally expressed in restricted tissues such as testis, sensitivity and specificity may not be ideal, but they could still represent good biomarkers in combination with other diagnostic tools and could become targets for the treatment, with tolerable safety profiles. Thus, search of TAs continues to improve diagnostic performance when combined in panels rather than used as single antigens as carcinogenesis is a complex and heterogenous with single cancer types. This thesis aims to identify TAs acting as biomarkers and targets for cancer focusing on lung cancer that has developed from incidental lung nodules and B-acute lymphoblastic leukaemia (B-ALL). The malignancy of lung nodules is challenging to be identified with only 5% of these nodules developing into lung carcinoma within 2-years and significant overlap between the features of benign and malignant nodules and requires follow-up and invasive procedures in some cases of indeterminate nodules with high growth rate. A systematic review was performed to characterise the known antigens that could act as biomarkers for early lung cancer detection. Verification of biomarkers with high sensitivities (Ciz1, exoGCC2, ITGA2B), high specificities (CYFRA21-1, antiHE4, OPNV), or both (HSP90α, CEA) along with miR-15b and miR-27b were indicated as promising biomarkers for early lung cancer detection. COL11A1 was identified from RNAseq data using different algorithms such as linear regression and logistic regression with elastic net regularization, and presented an aggregated score calculated as the cumulative rank of variable importance.

Regarding B-ALL, we identified a number of antigenic targets for the treatment of adult B-ALL, based on serological analysis of recombinant cDNA expression libraries (SEREX), previous protoarray analysis, transcriptional (from GSE13204), epigenetic profiling (GSE38403) and cancer testis antigens (http://www.cta.lncc.br/). Pathways that were enriched included Wnt, Hippo, and TGFβ. Their expression in B-ALL versus healthy bone marrow were examined and associated with survival, using the BloodSpot database as well as literature searches. Prioritising the TAs using the pre-defined criteria described by Cheever *et al.* identified a panel of genes (SOX4, ROCK1, YAP1, TEAD4, SMAD3, and TCF4) that had a high cumulative score for 0.89, 0.41, 0.36, 0.34, 0.33, and 0.32

respectively. Upon examining the expression of the above genes in primary B-ALL samples by qPCR, TEAD4 and SOX4 were found to be significantly upregulated in adult B-ALL samples compared to healthy donors with $p<0.01$ and $p<0.05$ respectively. Immunocytochemistry identified high expression of TEAD4 in the cell nucleus of B-ALL samples and moderate to high expression of SMAD3 both in cell nuclei and cytoplasm. Future studies will examine how these antigens and/or their pathways can be targeted by immunotherapeutic strategies.

Acknowledgements

I would like to express my deepest gratitude and appreciation to God to complete the PhD work. I am deeply thankful for all those who have contributed to the successful completion of my PhD work.

First and foremost, I am immensely grateful to my supervisor, Dr Barbara Guinn for her guidance, support, and invaluable insights throughout this journey. Her expertise, patience, and encouragement have been instrumental in shaping the direction of my research and pushing me to achieve my best.

I am also indebted to my second supervisor Dr Simon Hart for his valuable feedback and constructive criticism. His expertise in his research field has greatly enriched the quality of my work.

I acknowledge Prof Ken Mills for microarray analysis (GSE38403), Dr Mark Wade for qPCR training, and Daniel Flechter for figure design in chapter 6.

Furthermore, I would like to acknowledge the contributions of all the participants who generously shared their time and knowledge for this study. Their willingness to participate has been crucial in gathering meaningful data and insights.

I am also grateful to my family and friends for their unwavering support throughout this challenging journey. Their love, understanding, and encouragement have been a constant source of motivation for me.

Lastly, I would like to express my gratitude towards all those who have directly or indirectly contributed to this work but may not be mentioned here. Your support has played a significant role in helping me achieve my goals.

Publications to date

PEER REVIEWED PUBLICATIONS:

Mohamed, E., García Martínez, D.J., Hosseini, M-S., Yoong, Y-S., Hart, S. & Guinn, B.A. (2023) Identification of blood biomarkers for the earlier detection of lung cancer: a systematic literature review. *Carcinogenesis*. doi: 10.1093/carcin/bgad091

Mohamed, E., Bature, F., Ali, I.H.M., Abdelaziz, F., Alsayed, S., Mukhtar, M.M., Osman, B.I. & Guinn, B.A. (2021) Ovarian Cancer in the Sudan - Identifying the Social and Clinical Factors that Prevent an Early Diagnosis. Austin Journal of Obstetrics and Gynecology, 8(5), 1181.

CONFERENCE PRESENTATIONS:

Mohamed, E., Goodman, S., Cooksey, L., Mills, K.I., Orchard, K.H. & Guinn, B.A. (2023) Adult B-cell acute lymphoblastic leukaemia (B-ALL): enriched pathways identify new targets for leukaemia stem cell immunotherapy. 23rd International Conference on Progress in Vaccination Against Cancer" (PIVAC-23). Oral presentation.

Mohamed, E., Goodman, S., Cooksey, L., Mills, K.I., Orchard, K.H. & Guinn, B.A. (2023) Genes involved in the Wnt and Hippo pathways predict survival for patients with adult B-cell acute lymphoblastic leukaemia (B-ALL). British Society for Haematology (BSH 2023) Annual Scientific Meeting, Birmingham, U.K.

Mohamed, E., Goodman, S., Cooksey, L., Mills, K.I., Orchard, K.H. & Guinn, B.A. (2022) Identification of tumour antigens in adult B-cell acute lymphoblastic leukaemia (B-ALL) at diagnosis: targets for treatment. British Association for Cancer Research Annual meeting, Nottingham, U.K.

Contents

Table of Figures

Table of Tables

Abbreviations

ADC	Adenocarcinoma
aB-ALL	adult B-Acute Lymphoblastic Leukaemia
AML	Acute myeloid leukaemia
AntiHE4	antibodies to HE4
AUC	area under curve
BMX	Bone marrow tyrosine kinase
BTS	British Thoracic Society
CAR-T	Chimeric antigen receptor T cell
cB-ALL	child B-Acute Lymphoblastic Leukaemia
CEA	Carcinoembryonic antigen
CI	confidence interval
CIZ1	CDKN1A Interacting Zinc Finger Protein 1
CLL	chronic lymphoblastic leukaemia
COL11A1	collagen type XI alpha 1
COPD	chronic obstructive pulmonary disease
CR	complete remission
CRLF2	Cytokine receptor-like factor 2
CRP	C-reactive protein
CT	computer tomography
CTAs	cancer-testis antigen
CTC	circulating tumour cells
CYFRA21-1	cytokeratin 19 fragments
DCTPP1	dCTP pyrophosphatase 1
DNA	Deoxyribonucleic acid
EDRN	Early Detection Research Network
EGFR	Epidermal Growth Factor Receptor
ELISA	Enzyme Linked Immunosorbent Assay
exoGCC2	exosomal GRIP And Coiled-Coil Domain Containing 2 protein
FPR	False Positive Result
GCC2	GRIP And Coiled-Coil Domain Containing 2 protein
HER2	human epidermal growth factor receptor 2
HSP90	Heat Shock Protein 90
ITGA2B	integrin alpha 2b (tumour-educated platelets)
KRAS	Kirsten rat sarcoma
LAA	leukaemia associated antigen
LC	lung cancer
LDCT	low dose computer tomography
LSC	leukaemia stem cell
MAGE	melanoma-associated antigen gene
MDR	multidrug resistance
MHC	Major Histocompatibility ty Complex
MILE	Microarray Innovations In LEukaemia

MIP	mRNA for major intrinsic protein
MLL	mixed lineage leukaemia
NLST	National Lung Screening Trial
NSCLC	non-small cell lung carcinoma
NSE	neuron-specific enolase
OPN	osteopontin
OPNV	osteopontin velocity
PET	positron emission tomography
pfu	the plaque forming units
pGGN	pure Ground Glass Nodules
PPV	positive predictive value
PRISMA	preferred reporting items for systematic review
PROSPERO	prospective register of systematic review
PTPNS	Protein tyrosine phosphatase non-receptors
QUADAS	quality assessment of diagnostic accuracy studies
R/R B-ALL	relapsed/refractory B-ALL
RNA	Ribonucleic acid
ROB	risk of bias
ROCK1	Rho associated coiled-coil containing protein kinase 2
SCC	squamous cell carcinoma
SEREX	Serological Analysis of Antigens by Recombinant Expression Cloning
SLR	Systematic literature review
SOX4	SRY-Box Transcription Factor 4
SSX	Synovial sarcoma X breakpoint
TA	tumour antigens
TAA	Tumour associated antigens
TAAb	tumour-associated antibodies
T-ALL	T-Acute Lymphoblastic Leukaemia
TCF4	T cell Factor 4
TCR	T cell receptor
TEAD4	TEA Domain Transcription Factor 4
TGFβ	Transforming Growth Factor-β
TKI	Tyrosine kinase inhibitor
TNF	tumour necrosis factor
TNM	Tumour Node Metastasis
TPR	True Positive Result
TSA	tumour specific antigen
TTF-1	Thyroid transcription factor 1
VDT	Volume-doubling time
VGLL4	vestigial like 4
YAP1	YES Associated Protein 1

Chapter 1: Introduction

1.1 Cancer

Cancer is a condition characterised by uncontrolled cellular activities. It is associated with aberrantly expressed genes or proteins that have an oncogenic activity or loss of tumour suppressor functions. Mutated and overexpressed proteins may elicit an immune response and are known as tumour antigens (TAs) (Anderson & LaBaer, 2005). TAs (Jiang et al., 2019) are molecular structures that can be recognised by antibodies or specific T cell receptors (TCRs) once TA-derived peptides are presented within major histocompatibility complexes (MHCs) molecules. TAs may be classified as tumour specific or tumour associated antigens (TSA, TAA) according to the parental gene expression. TSA includes neoantigens, oncoviral and endogenous retroviral elements. Unique TSAs are strictly found in the tumour that results from mutations such as single point mutation, DNA insertions or deletions and are patient-specific and not detected in healthy individuals (Jiang et al., 2019). Mutated p53 (Umano et al., 2001) is an example of this class of TSA, aberrantly expressed in many cancers due to the presence of mutated protein with reduced or loss of function in regulating tumour cell apoptosis. While TAAs comprise self-proteins found in healthy tissues in low amount, they are overexpressed in the cases of cancer. This group of tumour antigens consists of cancer testis antigens (CTAs), differentiation antigens and antigens that are derived from genes overexpressed in cancers. Examples of these antigens are melanoma-associated antigen gene (MAGE) overexpressed in renal cancer and melanoma (Simpson et al., 2005). Examples of overexpressed antigens are human epidermal growth factor receptor 2 (HER2) which has been detected in non-small cell lung cancer (NSCLC) (Lynch et al., 2004; Soma et al., 2014).

Cancer remains a challenge in terms of early detection, which is considered essential for effective treatment, improved patient quality of life and preventing relapse after treatment (Goebel et al., 2019). Better understanding of the molecular mechanisms involved in cancer, development pathways, tumour microenvironment, recurrence after treatment, risk factors and epidemiology, will collectively provide better disease diagnosis and management (Silva et al., 2020). According to (Coakley & Popat, 2020) some cancers are considered to be fatal, including lung, stomach and colorectal cancers

due to their often-late detection and poor prognosis. In the efforts of promoting early diagnosis of these fatal cancer types, TAs could be used as biomarkers although this depends on their sensitivity and specificity for cancer diagnosis (Ward et al., 2016). Importantly, these antigens play a crucial role in anti-tumour immunity as they may promote the recognition by tumour-specific T-cells. They may correlate with tumour burden and clinical outcomes. Clinically, they are representatives of a potential immunotherapeutic perspectives such as engineered T-cell based antigens or vaccine (Yarchoan et al., 2017).

In this study, two different types of cancer were examined. Both cancers are difficult to treat and need new treatment strategies to help improve survival rates. They are NSCLC and adult B-cell acute lymphocytic leukaemia (aB-ALL). We will look for TAs as biomarkers for early NSCLC and TAs as targets for aB-ALL.

1.2 Lung cancer (LC)

Lung cancer (LC) is the most common cause of cancer-related death in the UK and it ranks as the second overall cause of death with around 1.6 million deaths per year worldwide (Prabhakar et al., 2018). The number of LC cases accounts for 13% of all new cancer cases diagnosed annually (Jemal et al., 2010; Bray et al., 2012). The incidence rate has continued to rise in recent years; this is due to the peak in smoking in less developed countries together with an ageing population (Balata et al., 2019). Patients with LC have poor survival rates predominantly due to late detection (**Figure 1.1**). Before it's clinical diagnosis, this disease has often spread all over the organs of the body and it is more difficult to treat effectively at this stage. Advanced stages of LC cases are associated with serious side effects such as pulmonary fibrosis and cardiotoxicity, and at this stage, the disease managed symptomatically (Verma et al., 2017). About three-quarters of patients present with advanced LC stages and after ineffective treatment eventually die within three months of diagnosis resulting in high mortality rate (O'Dowd et al., 2015). Records have also shown that around 35% of LC patients are diagnosed immediately after emergency admission and not less than 90% of such patients are at stage III or IV LC.

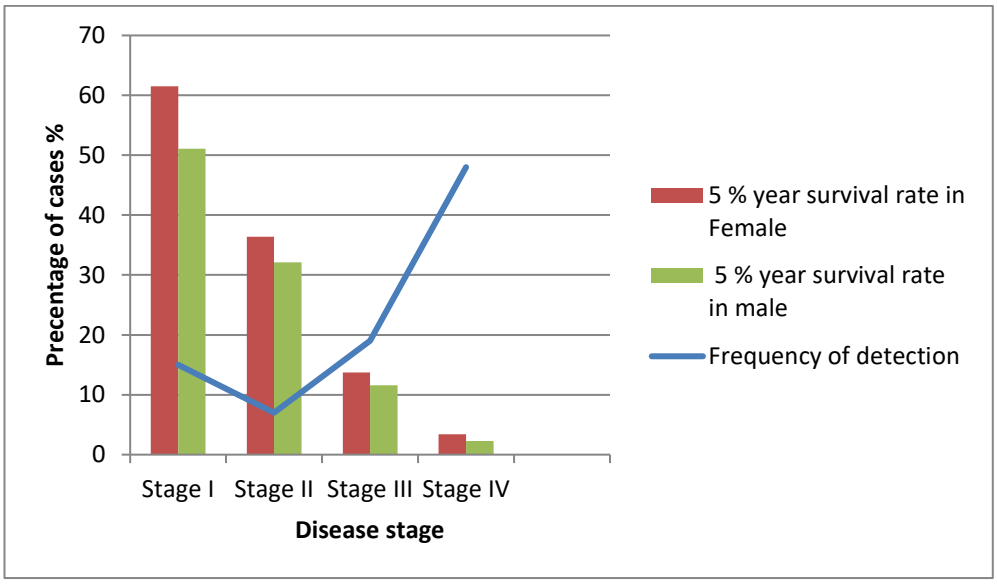


Figure 1.1 LC five-year survival and incidence by stage

Around 24-28% of LC patients are diagnosed at the early stages (stage I and II) with a high 5-year survival rate of 62% for females and 51% for males. While 49-53% of LC patients with are detected at advanced stage IV, 72-76% of patients with a known stage, are diagnosed at a late stage (stage III or IV). The 5-year survival for patients at stage IV is 3% and showing an absolute difference of 58 percentage points compared to early stages. Image generated using data from https://www.cancerresearchuk.org/about-cancer/lung-cancer/survival.

Comparatively, a high 5 – year survival rate of among the early stages (Stages I and II) LC can be as high up 75% following surgical resection (Balata et al., 2019). Although, LC has a relatively lower 5-year survival rate (56% on average) when compared to other common cancer types: 98.2% prostate, 89.6% breast and 64.5% colorectal cancers (Prabhakar et al., 2018). The metastatic LC has only less than 5% of 5-year survival rate, primarily because only 16% of LCs are detected at the early stage. These unfortunate outcomes could be averted by improving the tools and mechanisms for early detection of disease. Clinically, the detection of LC in the early stages is challenging because of the absence of specific symptoms which could enhance its presentation or detection. Patients with LC are characterised with non-specific symptoms such as shortness of breath, hoarseness, cough and blood in sputum in the advanced stages, which is due to the uncontrolled abnormal growth of lung cells resulting in progression of malignancy and metastasis to other organs via blood and lymph nodes (Raz et al., 2007).

Primarily LC is an epithelial carcinoma which develops from the major airways, but some can arise in the lung parenchyma. When considering treatment options and prognosis, it is further divided into two types: NSCLC and small cell lung cancer (SCLC) (Balata et al., 2019). The NSCLC accounts for 85% of LC while the rest (10-15%) are SCLC. The disease is predominant among the elderly individuals which are aged 75 years (Coakley & Popat, 2020).

1.2.1 LC staging system

The tumour/node/metastasis (TNM) staging system is used for LC staging as this helps clinicians to determine the extent of the primary tumour spread within the body. This system provides guidance for patient management and information that is associated with prognosis, eligibility for clinical trials and also facilitates international comparisons (Sobin et al., 2011). The TNM system is based on primary tumour characteristics, degree of lymph node involvement and the absence or presence of metastasis to overall stage (I-IV). It aims to cluster patients into stages of those with similar prognoses, but their treatment may vary from stage to stage. In addition, there are two common NSCLC staging, namely clinical and pathological TNM (Thomas & Gould, 2008). The clinical staging relies on taking patient history and physical examination, as well as laboratory, radiology and bronchoscopic findings before initiating any treatment. While the pathological staging depends on histological results after tissue sampling the confirmation of disease diagnosis. Some experts request that the pathological staging is performed after a complete surgical exploration of the hemithorax and mediastinum. However, some tissue sampling methods can allow for assessing tumour extent along airways and mediastinum (Tsim et al., 2010). On the other side, the SCLC differs from NSCLC in terms of TNM classification and it is limited to SCLC form. Advanced technology especially positron emission tomography (PET) images allow better principle of tumour site characterization and thus results in decrease of SCLC proportion (Bishnoi et al., 2011).

1.2.2 Risk factors of LC

There are several factors that are known to contribute to the initiation and progression of LC. The most dominant of these factors is smoking and continuous exposure to

tobacco which accounts for not less than 80% of LC (Wood et al., 2018). Smoking increases the risk of LC by 20-fold higher when compared to non-smoking patients. Although a number of studies have shown that inhalation of smoke can result in oncogene mutation and loss of tumour suppressor genes through the loss of heterozygosity, only 15% of smokers eventually develop LC (Kondo et al., 2006). For example, high frequencies of p53 mutations are found to increase with the number of cigarettes patients smoke (Gibbons et al., 2014). Also, evidence supports that LC can occur among second-hand smokers as well as non-smokers. It was reported that about 4000 non-smokers die of LC annually with 20% of this mortality linked with passive smoking (Taylor et al., 2007).

Other factors that are associated with LC include age. Like many other cancers, incidence of LC increases with age and it is most common among the elderly that are above the age of 70 years (McWilliams et al., 2013c; Loverdos et al., 2019), peaking between the age of 85 – 89 years. Furthermore, history of previous LC increases the risk of second primary disease, even after resection (Johnson, 1998). For example, stage I NSCLC patients treated with surgical resection have seven times higher risk than that of the initial LC diagnosis in the first year of disease after treatment (Surapaneni et al., 2012). Patient with other cancers such as head and neck cancer as well as other cancers associated with smoking including pancreatic and bladder cancers have a simultaneously increased LC risk (Morris et al., 2011; Walsh, 2016; Kwon et al., 2018).

Another risk factor contributing to LC is occupational exposure to toxic substances. Some studies have shown that metal, fumes and several dusts (e.g silica dust) correlate positively with the occurrence of lung nodules. Also, cigarette smoking enhances the risk of LC in synergy with other toxins. For instance, the risk of LC development with asbestos exposure is twice higher among nonsmoker than those without exposure, while it is nine times higher among smokers than those without exposure (Lubin et al., 2008; Pukkala et al., 2009; Leuraud et al., 2011; Ngamwong et al., 2015). Moreover, previous studies have also demonstrated a strong correlation between chronic obstructive pulmonary disease (COPD) and LC (Loverdos et al., 2019). This was reported in the National Lung Screening Trial (NLST) study, where patients with COPD recorded twice as high risk of developing cancer in comparison to individuals with normal lung function. Similarly, the

risk of LC increases in patients with emphysema when examined using CT scans and it remains high even after adjusting for the limitations in their ability to create airflow (Young et al., 2015). Similarly, comorbidity of idiopathic pulmonary fibrosis is associated with LC with around 10 % prevalence (Raghu et al., 2015).

1.2.3 Lung nodules as a risk factor for LC

Pulmonary nodules are small growth of cells inside the lung tissue and they are classified as benign or malignant. Tammemagi et al. reported incidental lung nodules at a rate of 25–51% among healthy volunteers and in patients undergoing LC screening (Tammemagi et al., 2013). These nodules have low malignancy potential and it is difficult to identify cancerous nodules at an early stage. Even, while using a differential diagnosis based on slight morphological changes, locations and clinical biomarkers, it is a very challenging task to measure the propensity of malignancy for those nodules. Early malignant lung nodules diagnosed so far were through different range procedures of clinical settings from computerised tomography (CT) scan analysis (morphological assessment), PET (metabolic assessments), to needle prick biopsy analysis. However, most of these procedures are invasive methods requiring biopsies or surgery and they increase patient anxiety as well as they come with a procedural risk (Gould et al., 2013).

1.2.4 Types of nodules

A nodule is defined as rounded or irregular opacity with measurement of up to 3 cm in diameter. Pulmonary nodules are divided according to their density (**Figure 1.2**), into solid and sub-solid nodules and later sub-grouped into part-solid nodule (PSN) and pure ground glass nodules (pGGN) according to British Thoracic Society (BTS) guideline (Baldwin & Callister, 2015).

1.2.4.1 Solid nodules

Solid nodules are well marginated, round, opaque and detached (discrete) with less than or equal to 3 cm in diameter (Hansell et al., 2008; Edey & Hansell, 2009; Baldwin & Callister, 2015). Lung parenchyma cells surround these nodules with no involvement of adenopathy, atelectasis, or pleural effusion. Nodules with diameter greater than 3 cm are known as masses with most of these considered malignant until further analysed to

be otherwise proven as benign nodules. The majority of small solid nodules are considered benign nodules, including most granulomas and intrapulmonary lymph nodes (80%), hamartomas (10%) and other benign lesions (10%). (Erasmus et al., 2000). Moreover, calcification related to non-malignancy are rarely observed in LC and appear to be amorphous or punctate as well as densely calcified in lung metastases from primary bone-forming malignancy (Grewal & Austin, 1994). Benign solid nodules are generally stable in size for 2 years and hence, the recommendation of 2 years for follow-up algorithms. Generally, smooth, regular margins and well-defined nodules are benign while the solid lung nodules without calcium, fat content or perifissural (benignancy characteristics) are considered in-determinate. However, 21% of these well-defined and regular margins nodules can be malignant (Erasmus et al., 2000).

1.2.4.2 Sub-solid nodules

Sub-solid nodules are frequent in the bronchial or vascular margins of the underlying lung parenchyma (underlying bronchial structures or pulmonary vessels) with higher malignancy tendency than solid nodules (Kakinuma et al., 2015). Sub-solid nodules are less than 3 mm in diameter and appear as focal regions with ground glass component in CT images. They are divided into two subtypes: pure ground glass nodules or “non-solid nodules” and part-solid nodules. Pure ground nodules have only ground glass attenuation while part-solid nodules contain both ground glass attenuation and solid component (Cho et al., 2013). Non-solid nodules manifest with ground-glass opacity with preservation of lung parenchyma and the bronchovascular structures. Part solid nodules also contain a solid component in ground-glass opacity. Ground-glass opacity is defined as a hazy increased opacity of the lung, that do not obscure the underlying bronchial and vascular margins of the lung (Hansell et al., 2008). Moreover, sub-solid nodules have been found in many diseases ranging from benign to malignant disorders mainly adenocarcinoma (ADC) (Kakinuma et al., 2016). These nodules are characterized by slow growth rate with low metastatic rate and good prognosis. The nodule size and solid component availability determine the aggressive nature of malignancy and are related to the invasive tumour component. Due to slow growth rate of these nodules, there is an over–assessment of nodules for malignancy, which could be avoidable by using the longitudinal CT examinations for temporal changes (Kauczor et al., 2015).

Nodules are assumed to be benign if they are transient while nodules that are larger in size and with solid component are suspicious for malignancy and require further screening (Silva et al., 2018).

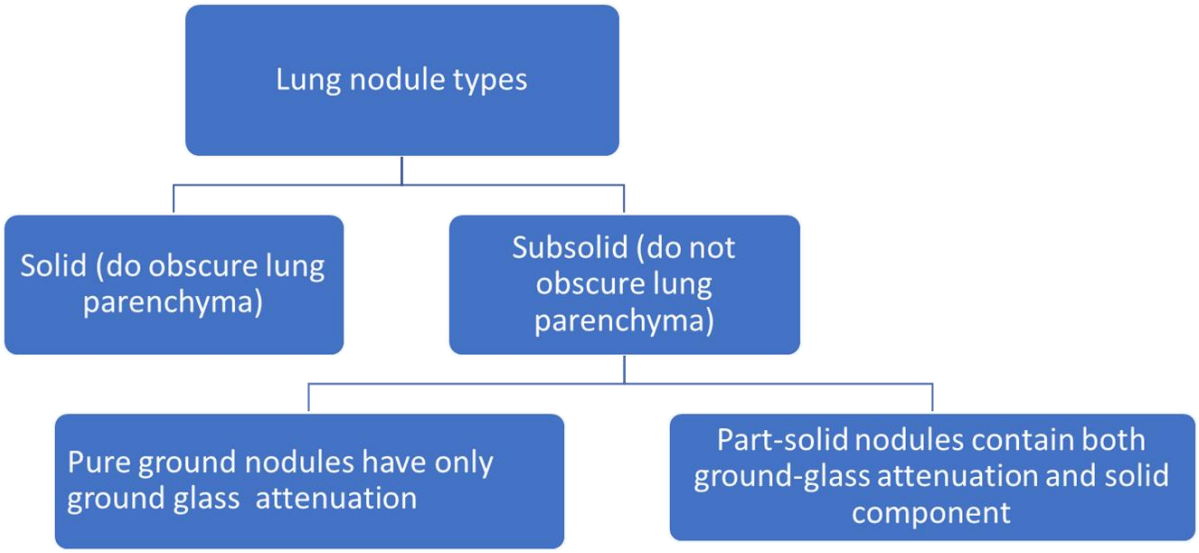


Figure 1.2 Schematic representation for types of pulmonary nodules

Solid nodules do not preserve the lung parenchyma but the sub-solid nodules do. Sub-solid nodules are sub-classified into two subgroups called pure ground and part-solid nodules.

1.2.5 Frequency of occurrence of pulmonary nodules

Detection of pulmonary nodules has been very common since the 1990s following the introduction of helical CT and multi-detector row CT (Al-Ameri et al., 2015). The majority of those detected have being benign nodules (Edey & Hansell, 2009). Thus, identification of incidental lung nodules increased and more information were obtained from several screening programs and this have continued to enhance the management of pulmonary nodules (Sánchez et al., 2018). Nodules (mainly multiple and frequently solids) are predominantly found on CT scan images of individuals 8-51% of nodule <10mm. 96% of these solid nodules <10mm are non-calcified nodules while 72% of them are <5mm. Recent imaging techniques such as CT cardiac image have identified many nodules without any association with smoking history (Henschke et al., 1999; Diederich et al., 2002; Burt et al., 2008).

Sub-solid nodules are less common than solid ones and incidentally detected in multi-detector CT scan. In the International Early LC Action Program (I-ELCAP), 30% (17,356 of 57,496 participants) of participants have solid nodules while 4.2% of cases are pGGN and 5% of patients have PSNs at baseline screening. Similarly, National Large randomized LC screening studies (NLST) reported that 9.4% of all enrolled individuals have sub-solid nodules of at least one pGGN, while 26% of sub-solid nodules detected in the I-ELCAP cohort either transient or decreasing in size at annual repeat screening and this number was higher than the prevalence of baseline screening (Felix et al., 2011; Walter et al., 2016; Yip et al., 2016). However, a high incidence was found in areas of where tuberculosis and histoplasmosis are endemic such as the southwestern USA. Although, this relation is not applicable to screening studies in North USA, Europe or Japan (Diederich et al., 2000; Nawa et al., 2002; Jett & Midthun, 2008). If large number of nodules are detected in these regions, it may reflect the incidence of chronic infective granulomas and exhibit distinctive calcification patterns. These nodules are considered benign and no further analysis would be required (Edey & Hansell, 2009).

1.2.6 LC screening

Chest radiography was used for LC diagnosis in the early 1960s. Unfortunately, the radiographic studies lack true control, and its mortality benefit was not clear (Marcus et al., 2006). In 2011, a study of Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) Cancer Screening Trial has identified more patients of stage I cancer using low dose computer tomography (LDCT). PLCO screening may have mortality benefits but over-diagnosis bias is the concern (Oken et al., 2011). NLST enrolled 53,454 individuals with a history of at least 30 packs-year of smoking either current smokers or quit within 15 years before enrolment in the study and their ages range was from 55-74 years. Individuals with a prior history of LC, haemoptysis, chest CT scan performed within 18 months of the study, and unexplained weight loss of more than 15 pounds in the last years were excluded. Three annual screenings with either chest radiography or LDCT were randomly performed for each participant with the endpoint of LC mortality. The study was performed over 3 years from the basal scan with two annual scans and followed up to 6 years without screening. NLST had 90% sensitivity for LC in the LDCT group and reduced the mortality rate by 20% compared to the radiography group.

Consequently, NLST reduced the absolute risk of LC death from 1.66 to 1.33% in high-risk smokers and previous smokers. The drawback of NLST is the high rate of 67,550 false positive results, 190 participants were over-diagnosed, and 910 benign lesions had invasive procedures, while 24 individuals died from the radiation (Aberle et al., 2011). Also, the Multicentre Italian Lung Detection (MILD) trial was a randomized trial comparing annual or biennial LDCT with a control arm. Eligible participants were age 49 years or above and smokers of at least 20 packs-year either current smokers or quit smokers within 10 years. The study also included participants with non-calcified nodules of >5mm which required follow-up. The two groups were composed of 1,190 and 1,186 for annual and biennial arms with a median of 5 and 3 scans respectively. About 65% (32) of patients were stage I detected from 49 scans and the death ranges from seven to 12 in the two groups following 5-years follow-up (Pastorino et al., 2019). Novel imaging technology (DANTE) included participants aged between 60-74 years with at least 20 packs-year of smoking. Participants had chest X-ray (CXR) at a baseline, 3-day sputum cytology with an LDCT scan. A total of 1,264 participants received LDCT, and 1,186 participants were the control arm. DANTE identified 45% of stage I LC in the LDCT group and 22% in the control arms. The mortality of LC has been evaluated at 50 cases per 100,000 (Infante et al., 2017). In another study, the Italian lung study (ITALUNG) was performed to assess LDCT efficacy in reducing LC mortality. The study had 3,206 participants randomized into two groups with regards to their annual LDCT, participants received at least four screens in one group and those that had no screening were assigned as a control group. This included individuals aged between 55-69 years with at least 20 years of smoking experience and who were current smokers or had quit within 10 years. The results identified 36% and 11% diagnosed with stage I in the screening and control arm respectively. The death cases were reported 43 and 60 in the screening and control arms respectively (Tanner & Silvestri, 2015).

The Dutch-Belgian LC Screening Trial (NELSON) is a randomized trial comparing LDCT screening to usual care over 10 years. The participants were aged between 50-70 years with smoking of 15 cigarettes or more per day in 25 year, or ≥10 in 30 years or quit smoking within past 10 years. The screening was based on volumetric measurements and doubling time detecting solid nodule ranges as negative, indeterminate, and positive malignancy. The participants were randomized 7,915 in the LDCT arm and 7,909

in the control group with no screening. 255 (69%) screens had detected stage I LC. Also, the UK Lung Cancer Screening Trial (UKLS) had 247,354 participants aged 50-75 years About 1.5% (8,729) participants were at high risk of lung LC within 5 years (≥5%) (Field et al., 2016), 2.1% (42) participants had LC, and the false diagnostic rate was 27%. The study suggests the possibility of using a risk prediction model to avoid low risk people and selecting individuals with indeterminate nodules. However, the number of patients with early LC was not high. UKLS was a pilot study and not driven for a long-term mortality which is the main drawback of the screening. Overall, the NELSON study was associated with a 26% reduction in 10-year cancer mortality in males who were screened by CT compared to those who received no screening (de Koning et al., 2020) while the MILD trial showed a 39% lower mortality from LC after 10 years (Pastorino et al., 2019). Other screening trials included the Danish LC screening trial (DLCST) (Wille et al., 2016), the German LC screening intervention trial (LUSI) (Becker et al., 2020), and the UK LC screening (UKLS) (Field et al., 2016) which had similar results. Low-dose CT screening of high-risk individuals results in decreasing LC mortality and associated with many drawbacks.

1.2.7 Assessing the probability of malignancy of pulmonary nodules

The diagnosis of lung nodules is extremely important for early LC detection. Malignancy probability estimation is crucial for follow-up and further evaluation (Snoeckx et al., 2018). The first step is the assessment of clinical parameters which should be performed as well as consideration of the risk factors such as smoking history, exposure to toxins, family cancer history as discussed previously (Section 1.2.1). Then, the radiological features including nodule size and growth rate are the main features for estimating LC probability and decision-making of management. In addition, to the evaluation of other predictors of benign or malignant aetiology (Callister et al., 2015). Common features of benign nodules include calcification (dense and uniform), fat content or location, usually peripheral and sub-pleural nodules; frequently benign solid nodules are representing about 61-100% sensitivity and specificity for nodule diagnosis. However, 70% of patients with LCs have nodules located in the upper lobes specifically in the right lung. Moreover, peripheral solitary pulmonary nodules constitute 50% of primary ADCs, while squamous cell carcinoma commonly manifest as a centralized lesion (Swensen et al.,

2000; Winer-Muram et al., 2002). In addition, benign nodules have well-defined, smooth and regular margins (Ost & Fein, 2000), yet well-defined and regular margins are observed in 21% of malignant nodules. Some characteristics overlap between malignant and benign nodules (Matsuoka et al., 2005). In contrast, common malignancy predictors include spiculated or lobulated margins frequently observed in 33-100% of malignant nodules. However, more than 50% malignant nodules have smooth margins and the presence/ absence of spiculation usually assists in nodule characterization (Wahidi et al., 2007). Well known features such as pleural retraction, spiculation, and thickening, or vessel leading directly to lesion, or partly solid components are all increasing malignancy risk in pulmonary nodules, in thin section CT scan (Seemann et al., 2000). However, the presence of one of these features represents high sensitivity for the risk of malignancy (91%) but with 57% low specificity as it may also be present in some inflammatory lesions. Sub-solid nodules identification is associated with pseudo-cavitation and may be an indicator of ADCs. However, cavitation alone may not help and CT scan showing thin or thick walls appear in equal proportions of benign and malignant nodules (Honda et al., 2007).

1.2.7.1 Contribution of nodule size to diagnosis

Nodule size is strongly related to increased malignancy risk and it represents the basis of nodule assessment in all algorithms. Generally, larger nodules have higher cancer probability and this risk varies in different studies (Network, 2013; Horeweg et al., 2014). For example, many articles reported that nodules larger than 20 mm in diameter are 64-82% linked with malignancy of lung. 6-28% of LC have reported nodules with size 5-10 mm and <1% of malignancy in nodules measure <5mm (Henschke et al., 2004). A cut diameter of less than 6mm is associated with a low cancer risk (<1%) according to the most recent guidelines (BTS and the Fleischner Society) in LC screening trials. This cut-off size is the same for solitary, multiple solid and sub-solid nodules. A second clinically important cut-off size is >8mm (McWilliams et al., 2013b). Around 80% of solitary nodules are benign lesions that are <5mm (Erasmus et al., 2000).

1.2.7.2 Nodule growth rate and cancer diagnosis

Lung nodule growth rate has been shown to be related to an increased malignancy risk. Any incidental nodules should be compared to present and previous imaging studies if

known, as well as time interval and nodules stability for further action. Rapid growth indicates LC on a theoretical basis for CT surveillance. Malignant nodules are generally growing at an exponential rate and they are mitotically active (Edey & Hansell, 2009). Volume-doubling time (VDT) is used for the estimation of growth rate and they are considered as the most sensitive marker in clinical practice (Edey & Hansell, 2009). VDT can be calculated using the equation: -

$$VDT=(t\times\log 2)/(3\times[\log(d2/d1])$$

Where:

VDT: is the doubling time in days

t: is the time in days between scans

d_2: is the diameter of the nodule at the time of the current scan

d_1: is the diameter at the time of the previous study

1.2.7.3 Risk assessment using different mathematical models

Overlapping of clinical and radiological features may conflict the diagnosis confirmation such as situation where patient is never-smoker, spiculated 15 mm upper lobe nodule. Due to given multiplicity, several risk prediction models have been used for malignancy assumption by performing multivariate logistic regression analysis (Xiao et al., 2013; Deppen et al., 2014; Al-Ameri et al., 2015). Different prediction models such as Mayo Clinic, Veterans Affairs, and Brock have been developed to assist with the calculation of malignancy risk (Maldonado et al., 2020). These models are based on factors such as age of patient, history of cancer, smoking status and nodule characteristics (size, morphology and location). In all models, the risk of malignancy increases with age and nodule size.

The BTS guidelines recommend that the Brock model is used for initial risk assessment, then patients undergo PET/CT scan and the Herder model is used when risk estimates are ≥10% in Brock model. The BTS integrated two models (Brock and Herder) in its guideline in 2015 (Callister et al., 2015). The American College of Chest Physicians (ACCP) guidelines do not recommend any predictive models and advise risk of malignancy

estimation for solid nodules >8 mm. However, guidelines have not shown any significant performance of prediction models over clinical judgment and their use is not solidly suggested (Balekian et al., 2013). For instance, prediction models are highly not recommended in most recent Fleischner society guidelines. Instead, a dichotomous risk stratification scheme is suggested and sub-classified into a low-risk (<5%) with association of less smoking, younger age, smaller, smooth and non-upper lobe nodules as well as intermediate (5–65%) and high (>65%) being associated with all of the opposite features (MacMahon et al., 2017).

1.2.8 Molecular Pathology of NSCLC

NSCLC develops due to variety of distinct somatic mutations occurring in a heterogeneous population of tumour progenitor cells. NSCLC is further sub-classified into: 35% ADC, 30% squamous carcinoma (SCC), 10% large cell carcinoma (LCC), bronchoalveolar carcinoma (<5%), adenosaquamous, and 1% carcinoid (Chikwe et al., 2013). NSCLCs (Chen et al., 2014b) is associated with a high frequency of p53 mutations (50%) and 30-60% of the epidermal growth factor receptor (EGFR) mutations. Other mutations such as 20% Kirsten rat sarcoma viral oncogene homolog gene (KRAS), 7% mesenchymal epithelial transition growth factor gene (MET) and 4-5% HER2 mutations have been detected in NSCLC (Lynch et al., 2004; Soma et al., 2014). The differences in the frequency of common mutations in ADC and SCC (**Figure 1.3**) have been reported. ADC arises from epithelial cells in the terminal respiratory tract frequently expresses thyroid transcription factor 1 (TTF1) and cytokeratin (Mitsudomi, 2014). While SCC develops in the central airways and susceptibility to form large solid carcinoma and associated with PTEN and PIK3CA mutations (Chen et al., 2014b). Interestingly, non-smokers with ADCs have higher frequency of EGFR, ALK and ROS mutations (Soda et al., 2007) while smokers show high frequency of mutations in KRAS, this however suggest that there are different pathogenetic pathways of tumour development between non-smokers and smoker LC patients (Sanders & Albitar, 2010). In the same vein, the mutation frequency is also affected by ethnic/geographic factors. For example, 60% of patients from Asian background have EGFR and ALK driver mutations while only 10-15% of white Europeans show these pattern of abnormalities (Mitsudomi, 2014; Gridelli et al., 2015). SCC develops in the central airways and with susceptibility to form large solid

carcinoma, which sometimes cavitate. Its prevalence among human correlates positively with smoking habit (Fong et al., 2003). SCC is classically a central lung tumour; however, some of SCC is found in the periphery (Tomashefski Jr et al., 1990).

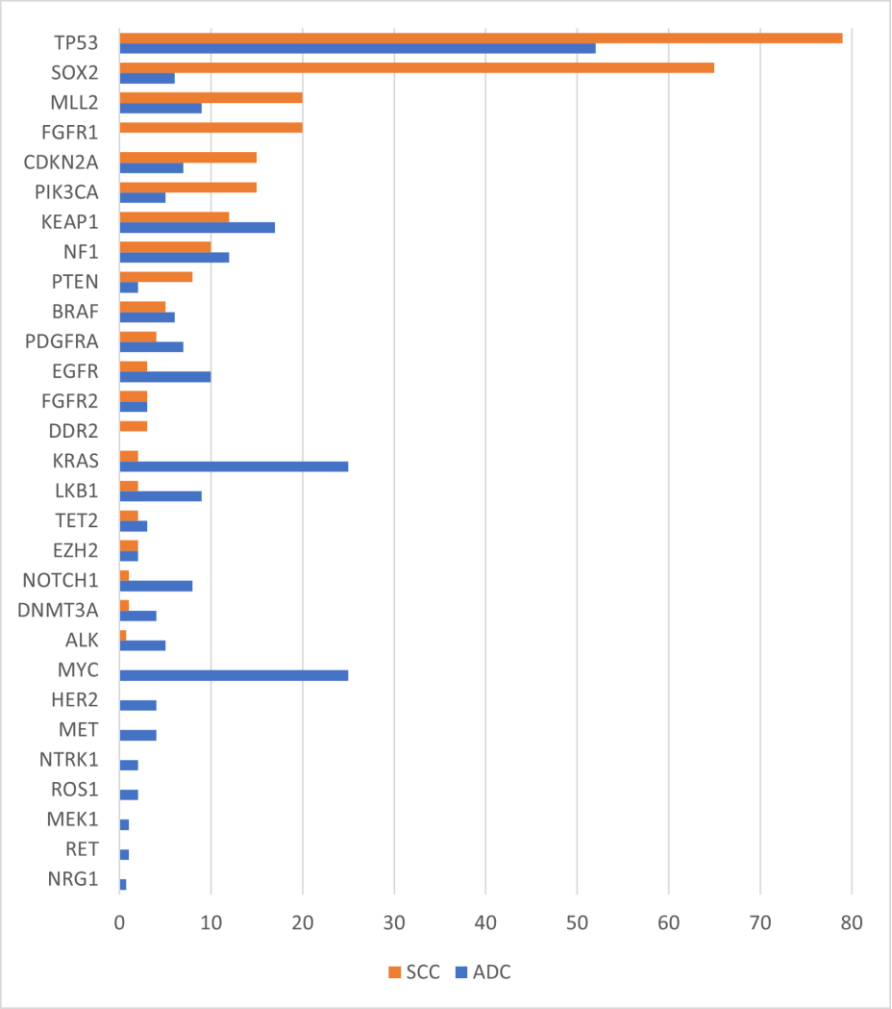


Figure 1.3 Common mutations in ADC and SCC.

KRAS, LKB1 and EGFR mutations are predominant in ADC while FGFR, PIK3CA, PTEN are common in SCC. P53 is common in both with 52% and 79% in ADC and SCC respectively. Data taken from (Chen et al., 2014b).

1.2.9 LC diagnosis

Conventional methods used in the diagnosis of LC include CT images, biopsy, cytology and bronchoscopy (**Figure 1.4**). Traditionally, LC diagnosis (Sharma et al., 2015) was

based on histological examination of resected tumours. The disease detection is associated with risks and mishandling that causes damage to patients as lung is a fragile organ (Patlak & Nass, 2013). CT scan shows high efficiency of LC when combined with biopsy and sputum cytology. The sensitivity of chest radiography (91.3%) was higher than LDCT sensitivity (73.5%), but specificity was higher in LDCT (93.8%) than chest radiography (72.4%) in NLST study (NLST, 2013). However, the chest radiograph increases the risk of exposure to high-intensity radiation (Rubin, 2014). PET-CT scan uses tracer elements such as 18FDG and Fluorine-18-methyl-tyrosine used to provide higher accuracy (Counts & Kim, 2017). CT-scan limitations are associated with poor patient compliance and less accurate diagnosis due to high false positive results (Prabhakar et al., 2018).

In addition, needle biopsy is a commonly used technique for detecting LC. Trans-thoracic needle biopsy (Kalanjeri & Gildea, 2016) is a useful tool for detecting malignancy in lung nodules but its use depends on nodule size (2cm or more). It is also used to determine the presence of tumour in lung pleura, mediastinum, or in the lung parenchyma (Prabhakar et al., 2018). Cutting needles have better specificity than aspiration needles, but are associated with higher incidence of complication (Guimaraes et al., 2014).

Navigational bronchoscopy is a recent technique and it has diagnostic efficiency of 84% when combined with PET (Lamprecht et al., 2012). However, the limitation of bronchoscopy is its lower specificity. It is prone to high of false positive results due the inability to differentiate between cancer and lung inflammation (McWilliams et al., 2013a).

Cytology investigation includes both sputum and pleural fluid cytology. Sputum cytology is usually performed for early detection of LC as it is non-invasive and it is also a quantitative technique. Characteristically, more malignant cells are found in the morning sputum sample compared to fresh sputum, with 25% higher yield and recovery of malignant cells from morning samples (Patriquin et al., 2015). Moreover, pleural fluid cytology is used as a LC indicator for the presence of malignant cells in the pleural fluid lining of the lungs in its late clinical stage. Pleural fluid examination has sensitivity of 60-70% (Antonangelo et al., 2015). Comparatively, high levels of tumour markers are found

in the pleural fluid of patients with malignant type of tumour in the lung. When radiology fails to detect the malignant cells, pleural fluid could be an alternative for LC detection. This technique is considered an invasive method and it is associated with poor patient compliance. It has better sensitivity when compared to sputum cytology. It also reduces misdiagnosis of lung tumour as it can differentiate between inflammatory, benign and malignant lesions (Kremer et al., 2010; Kaur et al., 2017).

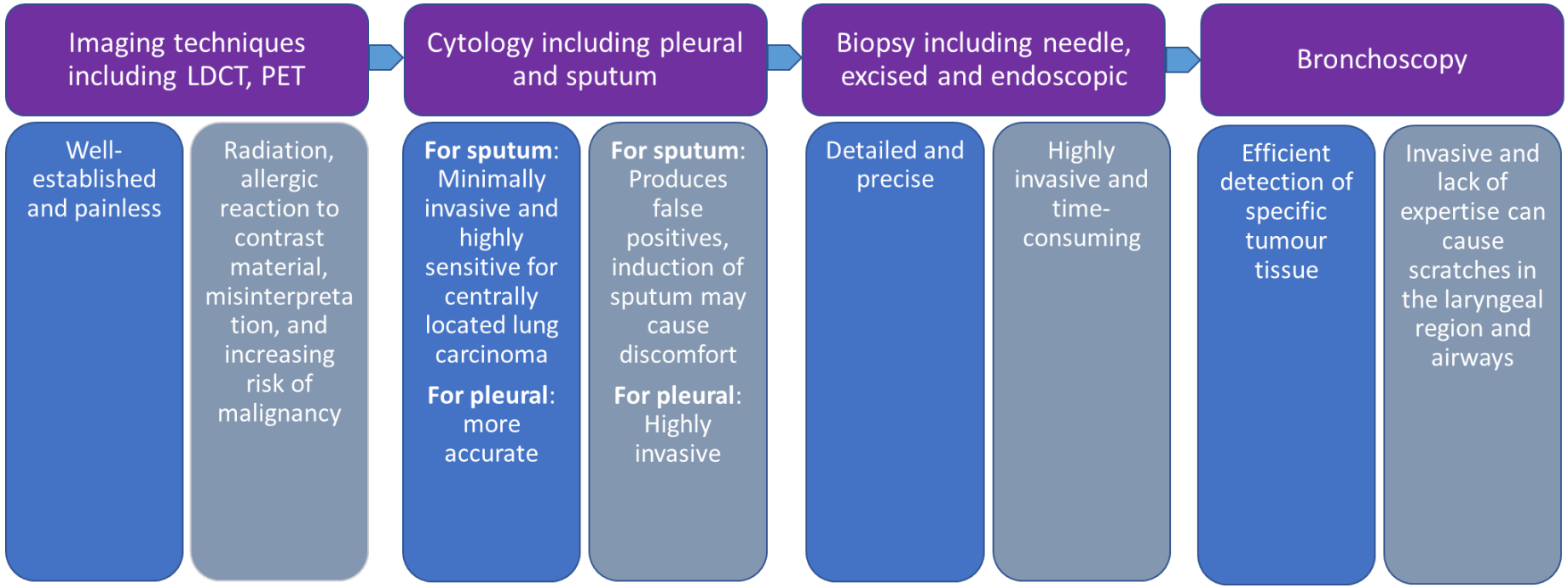


Figure 1.4 Conventional methods for LC diagnosis

LC diagnosis starts with imaging techniques to detect abnormality, if suspicious mass observed, then minimally invasive techniques such as bronchoscopy and cytology performed to confirm the presence of tumour. Data summarised from (Nooreldeen & Bach, 2021).

1.3 Adult B-cell acute lymphocytic leukaemia (aB-ALL)

aB-ALL is a haemopoietic malignancy characterised by impaired differentiation of progenitor cells both in the bone marrow and the extra-medullary organs such as spleen. It has many types which are classified according to duration and progenitor cell type: mainly myeloid and lymphoid cells. Regarding the onset of the disease, leukaemia is categorised as acute myeloid leukaemia and chronic myeloid leukaemia and same principle applied to lymphoid leukaemia (Faderl et al., 2010). Acute disease refers to rapid development where the white blood cells divide very quickly abnormally and spread to other organs such as the spleen. This disease can lead to death within few weeks or months without treatment (Paul et al., 2016).

Globally, leukaemia is a rare disease with incidence less than 1 case per 2000 individuals. In the UK, around 10,100 patients are diagnosed with this disease as at year 2020. It ranks as one 12th most common type of cancer in females and the 10th most common in males with 4,000 and 5,800 patients, respectively. It accounts of 3% of all new cases. The incidence rate increased in the period of 2014 by 5% and expected to increase up to 19 patients per 100,000 individuals by 2035 (Cancer research UK, 2021).

B-ALL is more common in childhood and accounts for about 80% of all ALL (**Figure 1.5**). The disease is thought to originate from genetic abnormalities during pregnancy that lead to malignantly transformed lymphocyte progenitor cells followed by second mutation occurring early in childhood with incidence peaks at age between 2-5 years (Roberts, 2018). While adult ALL constitutes 20% of all the disease and is characterised by heterogenous genetic mutations and chromosomal translocations, around 60% of patients are diagnosed at age less than 20 years with the median age of 14 years. The adults of 45 years comprise ~25% of patients and the elderly group of 65 years and above constitutes around 11% of cases (Mohseni et al., 2018). Essentially, ALL is more predominant in early childhood and in >45 years old individuals.

ALL diagnosis is based on the presence of 20% or more lymphoblasts in the bone marrow.

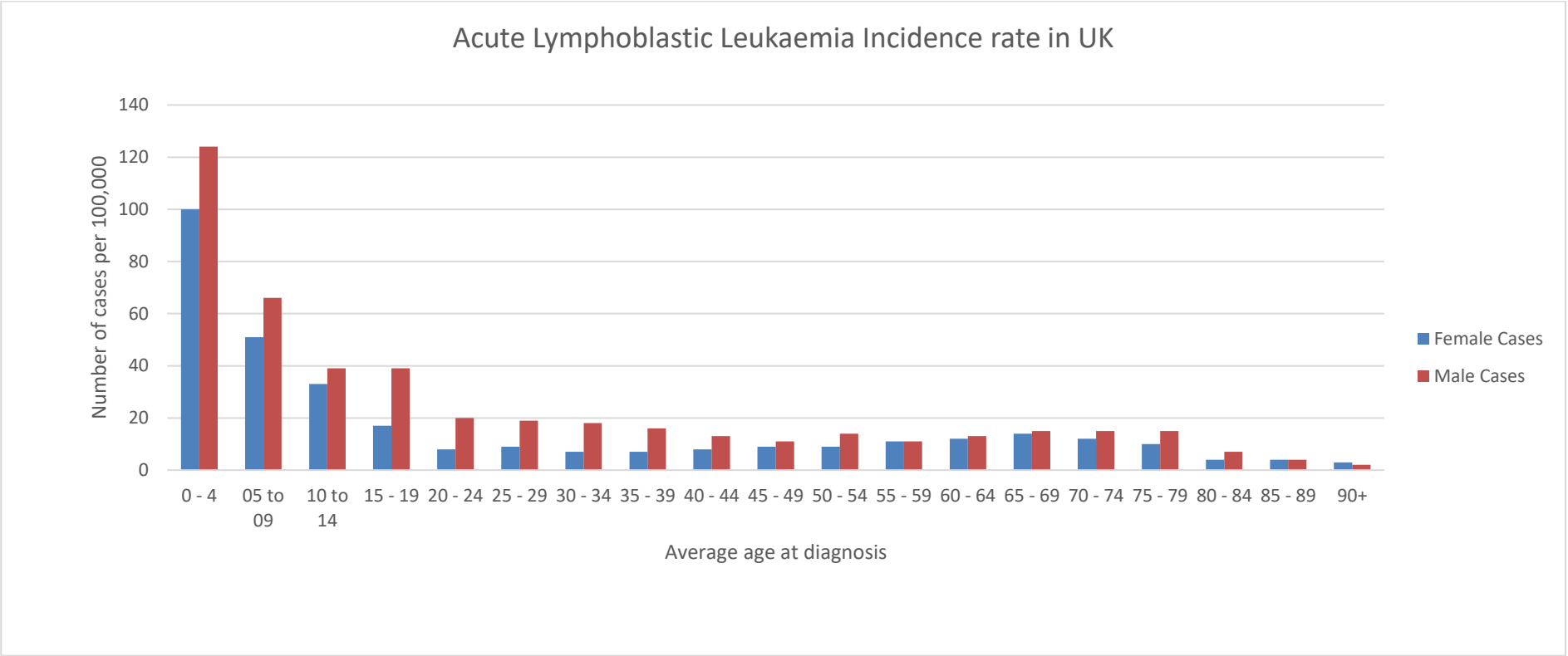


Figure 1.5 B-ALL incidence rate by age

B-ALL has a bimodal distribution with the first and largest peak observed in paediatrics in early childhood and the second one in adults, at the age of approximately 50 years. Data from (Cancer Research UK, 2020).

Historically, ALL is classified into L1, L2, L3 based on morphological criteria, including: cytoplasm, cell size, nucleoli vacuolation, and basophilia, according to the French-American-British (FAB) system (Lilleyman et al., 1986). Also, ALL is classified according to immunophenotypes and cytogenetic of blasts by World Health Organisation (WHO) includes recurrent genetic abnormalities and chromosomal rearrangement as shown in **(Table 1.1)** (Swerdlow et al., 2016).

Table 1.1 WHO classification for adult (a)B-ALL

B-cell lymphoblastic leukaemia not otherwise specified
B-cell lymphoblastic leukaemia with recurrent genetic abnormalities
B-cell lymphoblastic leukaemia with hypodiploidy
B-cell lymphoblastic leukaemia with hyperdiploidy
B-cell lymphoblastic leukaemia with t(9;22)(q34;q11.2)[*BCR-ABL1*]
B-cell lymphoblastic leukaemia with t(v;11q23)[*MLL* rearranged]
B-cell lymphoblastic leukaemia with t(12;21)(p13;q22)[*ETV6-RUNX1*]
B-cell lymphoblastic leukaemia with t(1;19)(q23;p13.3)[*TCF3-PBX1*]
B-cell lymphoblastic leukaemia with t(5;14)(q31;q32)[*IL3-IGH]*
B-cell lymphoblastic leukaemia with intrachromosomal amplification of chromosome 21 (iAMP21)
B-cell lymphoblastic leukaemia with translocations involving tyrosine kinases or cytokine receptors ('BCR-ABL1–like ALL')

1.3.1 Prognosis of ALL

Adult ALL constitutes ~20% where most patients are detected after 55 years, and this constitutes around the half (51%) of mortality rate of ALL. High death rate in adult ALL reveals that the old age has harmful impact on the pathophysiological outcome. The 5-year survival rate is poor around 20% while the survival decreases with increasing age (Sive et al., 2012). Poor outcome of adult ALL is attributed to the presence of comorbidities and other high-risk features increased with age. These risk factor may result in intolerance to chemotherapy and resistance. In addition, the prevalence of genetic and epigenetic mutations is higher in adult ALL when compared to children's patients. More alterations in repertoire leukemic cells is associated with adverse outcomes (Aldoss et al., 2019).

Molecular abnormalities of cytogenetic are also associated with prognosis of ALL (**Table 1.2**). Introduction of targeted therapy such as the BCR-ABL1 tyrosine kinase inhibitors (TKI), monoclonal antibodies against cell surface antigens including CD20 and CD22 and chimeric antigen receptor (CAR)-T-cell has resulted in improving the treatment of ALL (Rafei et al., 2019). The pathogenesis of ALL is due to the blockage in differentiation of lymphoid progenitors and results in abnormal proliferation and survival of B precursor cells with aberrant chromosomal numbers and structures including hyperploidy, and hypoploidy (Liu et al., 2016).

1.3.2 Risk factors of ALL

Considering the genetic factors, ALL is more common in twins of monozygotic and dizygotic and this indicates that the disease may be associated with genetic susceptibility (Aldoss et al., 2019). Moreover, ALL risk increases in patients with inherited syndromes that are characterised by massive chromosomal fragility such as Bloom syndrome and Fanconi anaemia and accounts for less than 5% of ALL cases, and chromosomal abnormalities, including Down's syndrome (Mertens et al., 1998; Chessells et al., 2001). Recently, the risk of ALL has been associated with the presence of polymorphism of methylenetetrahydrofolate reductase (MTHFR) gene in adult as well as in infant (Jabbour et al., 2005). Moreover, microbial infection contributes to the aetiology of ALL, example of such microbes are varicella and influenza viruses. As ALL proliferation is associated with pathogenic exposure, lymphoproliferative disorders are found to be more predominant in patients with human immunodeficiency virus (HIV) and Epstein-Barr virus are associated with mature B-cell ALL (Paul et al., 2016).

ALL risk increases with age and peaks in people over 70 years old. Although 25% of ALL patients have the median age of 45 years, and 11% have 65 years. Increasing age is associated with poor survival rates in the elderly patients (Goldstone et al., 2008). However, there is an inverse correlation between increasing age and survival with ALL. Furthermore, Caucasians in urban areas are at higher risk of developing ALL (Pui & Evans, 2006). Interaction between environmental factors with genetic variability has been investigated as a contributor to ALL development. Similarly, exposure to radioactive

radiation and carcinogenic chemicals increase the susceptibility to ALL (Jacobson et al., 2016).

1.3.3 Pathogenesis of ALL

Normal haemopoiesis starts in the bone marrow when the stem cell differentiates into myeloid and lymphoid stem cells (Baba et al., 2004; Seita & Weissman, 2010). Essentially, B-lymphopoiesis is complex and under the tight control of transcription factors (**Figure 1.6**). B-ALL occurs due to abnormal proliferation of common lymphoid progenitors in the bone marrow with hierarchical clonal expansion of abnormal blast and resistance to the negative selection and apoptosis. Leukaemogenesis is characterised by mutations of lymphoid transcription factors (IKZF1, PAX5, EBF1, and ETV6), cell cycle regulators and tumour suppressors (CDKN2A/B, TP53, and RB1), lymphoid signalling regulators (BTLA and CD200), and chromatin modifiers (CREBBP, SETD2, and WHSC1), and these are all common in B-ALL (Faderl et al., 2010). Table 1.2 shows the characteristic genetic abnormalities and chromosomal aberrations linked with ALL hallmark (Iacobucci & Mullighan, 2017). They include chromosomal translocations t(9;22), t(12;21) and less common t(1;19) as well as mixed linkage leukaemia (MLL) rearrangement. This translocation may lead to formation of fusion proteins with oncogenic properties. Aberration of chromosomal numbers such as hypodiploid (32-39 chromosomes) is also seen in part of cases and exhibited in Ras and PI3K signalling which may lead to aggressive ALL. Alteration of B-lymphoid transcription factors such as Ikaros, PAX5 include the deletions of variants in these factors (Faderl et al., 2010). These molecular abnormalities are crucial to be detected in diagnosis, patient's stratification and application of targeted therapy (Mullighan & Downing, 2009). Poor prognosis (Molina et al., 2021) is characterised by the adverse clinical features such as CNS involvement, high total count of white blood cells (more than 50×10^9/L) and low event-free survival rate.

Table 1.2 Cytogenetic abnormalities and their frequency in adults with B-ALL

ALL subtype	Frequency (%)	Genetic abnormalities	Prognosis	Reference
Ph-B-ALL (c-/Pre-B-ALL t(9;22))	15-30	Deletion of IKZF1 (70%), CDKN2A, CDKN2B, and PAX5, formation of BCR-ABL	Poor	(Faderl et al., 2003)
Ph-like B-ALL (c-/Pre-B-ALL no t(9;22))	20-25	50% CRLF2 rearrangement, 10% JAK-STAT, 2-8% RAS mutation, 3%-10% ERPO mutation	Poor	(Mullighan et al., 2009)
B-ALL with t(8;14)	1.7	C-MYC with various partners	Poor	(Angi et al., 2017)
Pro-B-ALL t(11q23)/MLL	3-7	KMT2A mutation, PI3K-RAS	Poor	(Malard & Mohty, 2020)
B-ALL hypodiploid	2-3	90% TP53, 13% IKZF2, 41% RB1, RAS	Poor	(Holmfeldt et al., 2013)
B-ALL t(1;19)	5	TCF3–PBX1	Intermediate	(Mohammadi et al., 2017)
B-ALL Hyperdiploid	25	RTK-RAS	Good	(Chilton et al., 2014)
B-ALL with t (12:21)	3	ETV6–RUNX1	Very good	(Moorman et al., 2007)

≠: Frequency within the population of aB-ALL patients

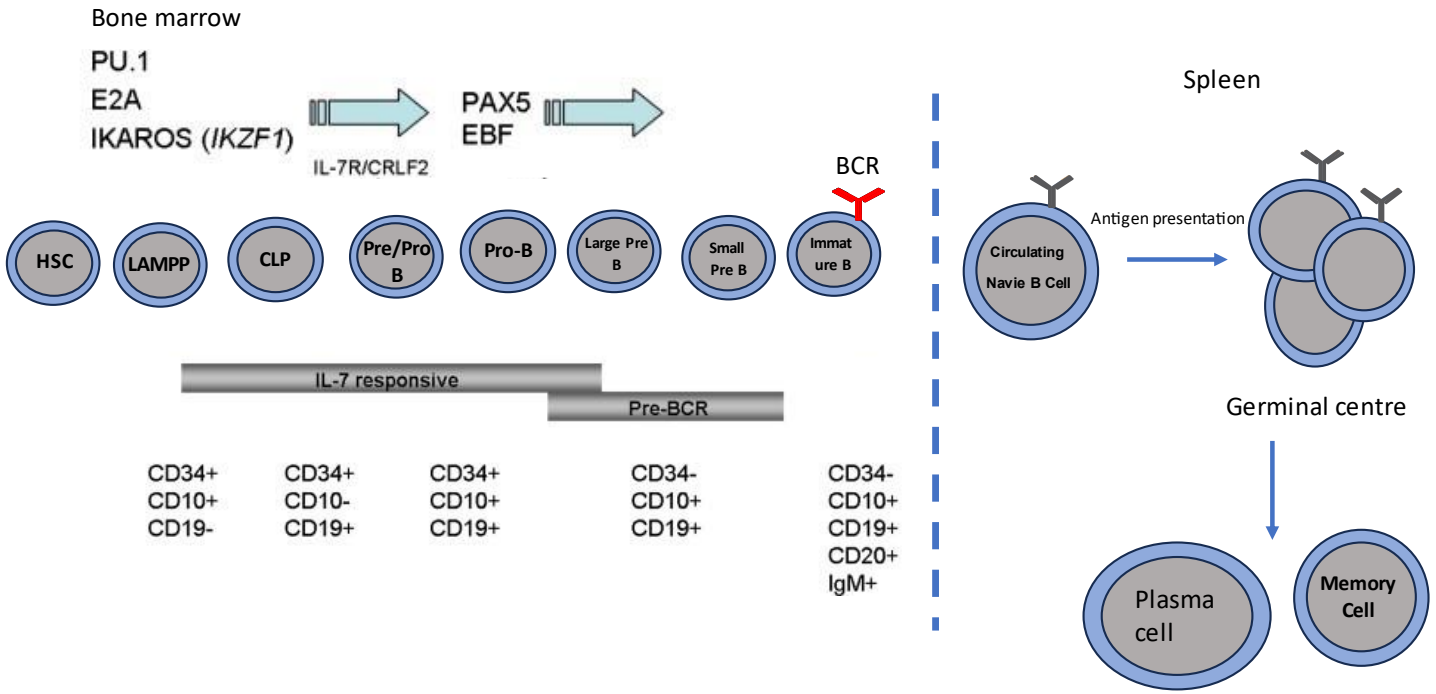


Figure 1.6 B-lymphopoiesis and B-ALL

B-ALL occurs due abnormal proliferation and differentiation of lymphoid progenitors in the bone marrow (mutation in IKZF1, reduced PAX5, high IL-7, high Pre-B signalling) leading to accumulation of immature blasts in bone marrow and other sites. Figure adapted from (Garcillán et al., 2018).

1.3.4 Treatment of B-ALL

B-ALL is a highly heterogenous disease and it is stratified based on cytogenetics abnormalities from good, intermediate to poor risk groups. The treatment (**Figure 1.7**) has four phases over 2-3 years including induction, consolidation, intensification, and long-term maintenance (Malard & Mohty, 2020). The induction phase comprises chemotherapy (including cyclophosphamide (CP), vincristine, doxorubicin (VAD), and etoposide, each of which can be used alone or in combinations) and glucocorticoids treatment. Intrathecal chemotherapy alone or with glucocorticoids is used for B-ALL with CNS involvement with or without irradiation (Liu et al., 2016). VAD (Liu-Dumlao et al., 2012) has been used infrequently because of their high toxicity and potential side effects such as cardiomyopathy, haemorrhagic cystitis, and cerebrovascular events. The failure of current chemotherapies such as CP, fludarabine, and alemtuzumab is associated with increased risk for relapse in patients with ALL. Minimal residual disease (MRD) is a crucial part of B-ALL treatment to determine number of blasts in bone marrow and the molecular and immunophenotypes of these blasts using PCR and flow cytometry. Low MRD is associated with favourable prognosis especially in consolidation and maintenance stage (Malard & Mohty, 2020). Introduction of tyrosine kinase inhibitors such as blinatumomab (Roberts et al., 2017; Tasian et al., 2017) has improved the B-ALL with Philadelphia and related subtypes which constitute 50% of aB-ALL.

In case of chemotherapy failure, B-ALL may be treated with various types of immunotherapy including allogeneic hematopoietic stem cell transplantation (HSCT), antibodies, and CAR-T cell (Malard & Mohty, 2020). HSCT has the survival rate in the first CR of 50%. Stem cell transplantation has many drawbacks due to comorbidities, severe infections make this not suitable especially for elderly B-ALL (Malard & Mohty, 2020). Elderly patients are particularly susceptible to the dose limiting toxicities of chemotherapies and are often excluded from allogenic stem cell transplantation on this basis (Liu et al., 2016). Also, therapeutic monoclonal antibodies (mAb) targeting cell surface antigens CD19, CD20, CD22 that are highly expressed on ALL blast. Rituximab is an example of developed antibodies against CD20 that CD20 is found in 30-50% of B-ALL and (Raponi et al., 2011).

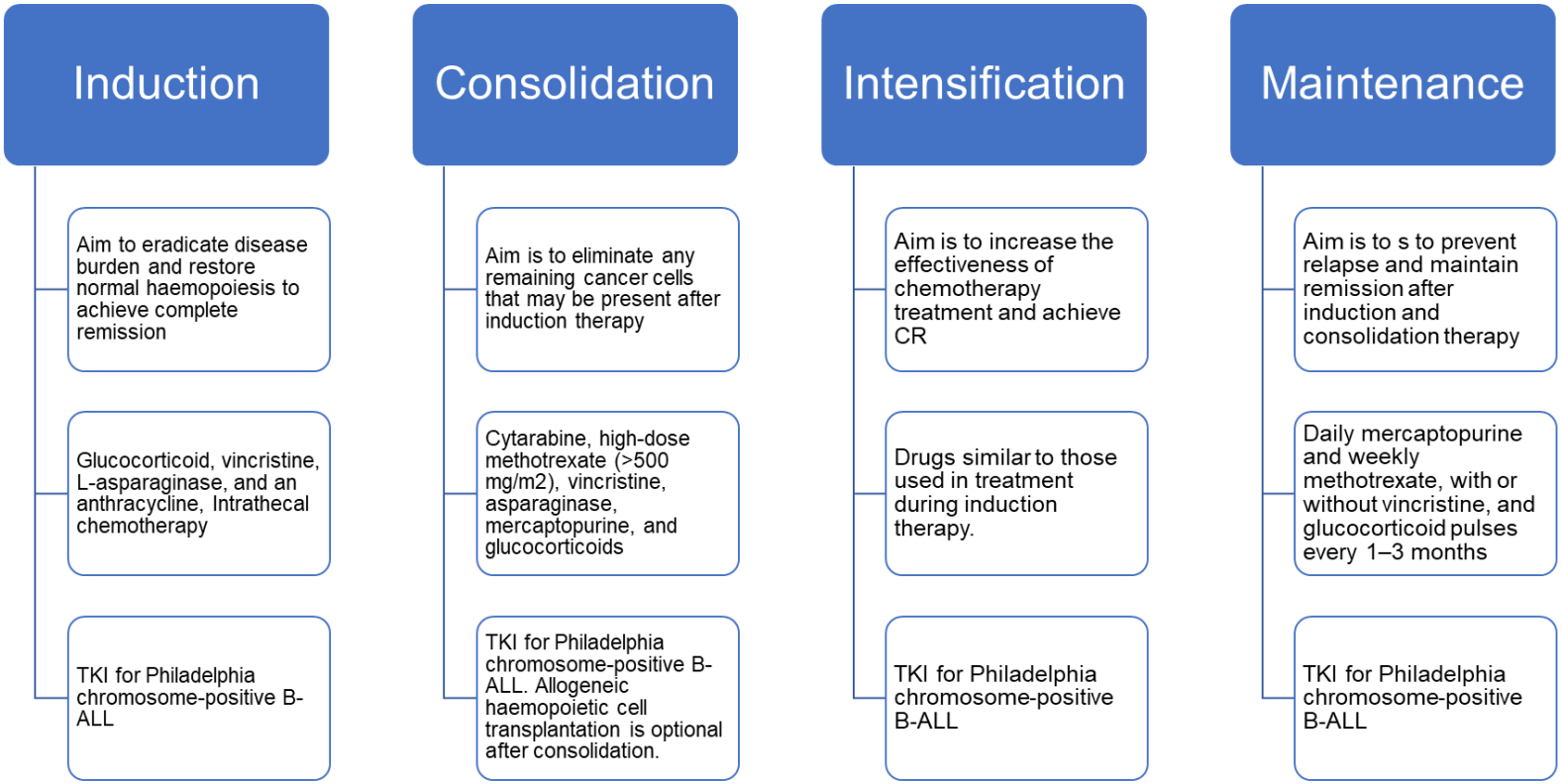


Figure 1.7 Phases of B-ALL treatment

Four phases of treatment aim to eradicate B-ALL cells and maximise the survival of the patients.

Rituximab has shown a statistically significant improvement with 2-years event free survival (EFS) with 65% in the rituximab against 52% in the control group. However, rituximab combined with chemotherapy are considered a standard for the treatment for B-ALL (Raponi et al., 2011).

Another cell surface antigen, CD22 is expressed on samples from 90% of B-ALL patients. This protein is internalised upon ligand binding. CD22 is indicated as an attractive target as it is not shed into the microenvironment, not recycled back to the cell surface. On ligand binding it undergoes endocytosis into B-cell, and is degraded in lysosomes (Piccaluga et al., 2011). Epratuzumab is an anti-CD22 mAb with limited therapeutic efficacy. Its combination with chemotherapy has not shown obvious benefit. When Inotuzumab ozogamicin (IO) is conjugated to antiCD22, it has shown efficacy against B-ALL *in-vitro*. Upon antigen binding, blasts internalize IO and release cytotoxin and kill the leukaemic cells (Piccaluga et al., 2011). Anti-CD19 combined with pyrolobenzodiapine (PBD) dimer containing toxin is known as ADCT-402, and it has shown potent efficiency against B-ALL both *in-vitro* and *in-vivo*. In addition, ADCT-402 is assessed in adult with relapsed/refractory B-ALL (R/R B-ALL) but it was terminated early due to its toxicity (Zammarchi et al., 2018).

Another attractive approach for B-ALL treatment is the use of bispecific antibodies, they have two different target epitopes linked to form a single chain antibody. Blinatumomab is the only BiSpecific T-cell engager (BiTE) approved by the FDA for clinical use (Casey et al., 2022): it consists of CD3 and CD19 for treating R/R B-ALL. When blinatumomab was evaluated against standard of care, it was found to improve complete remission (CR) rates from 34% to 16% ($p<0.001$) regardless of the percentage of blasts in bone marrow, previous therapy, age and in Ph status (Liu et al., 2016). CAR-T cells are derived from autologous T-cells from patients with B-ALL and genetically engineered and infused back to the patients. CAR-T cells contain extracellular domains that are responsible for antigen recognition. These are derived from antibodies, transmembrane domain, and intracellular domains ad are used to transmit an activation signal and co-stimulation (Ruella & Maus, 2016). The CD19 CAR-T has shown efficacy for R/R B-ALL with response rate of more than 80% but relapse rates are observed in patients without CD19, and short duration of therapy. In order to avoid relapse due to the CD19 negative

population, a dual CD19/CD20 CAR-T has been developed for the treatment of advanced R/R B-ALL and shown to be more effective in preclinical trials than the single CD targeting antibodies (Liu et al., 2016). CAR-T is considered more specific than conventional chemotherapy, but it is still associated with adverse effects including B cell aplasia, increased infection susceptibility due impairment of antibodies production, and more serious events such as cytokine release syndrome (CRS) leading to neurotoxicity and multiple organ dysfunction (Paul et al., 2016).

Novel targeted therapies offer the promise of effective anti-leukemic activity with reduced toxicity from off-target effects. Given the diverse molecular and genetic alterations occurring in ALL, it is unlikely that a single agent will be effective for all ALL patients (Rafei et al., 2019). However, with the ability to characterise the immunophenotype and genotype of each patient's leukaemia, targeted therapy can be expected to lead to improvements in remission and survival as part of individualised treatment strategies. Targeting signalling pathways such as JAK/STAT inhibitors (Whitlock, 2006) is also a promising approach, using small molecule inhibitors such as CHZ868 for patients with Cytokine receptor-like factor 2 (CRLF2) rearrangements. These agents, used in combination with conventional chemotherapy have improved disease-free survival (DFS). However, R/R B-ALL represents a challenge in terms of long-term remission (Mohseni et al., 2018). Tumour antigens are useful tools in terms of disease diagnosis and treatment. Search of antigens with high sensitivity and specificity is ongoing to improve the current methods for the disease diagnosis and treatment.

1.4 Tumour antigens

TSA includes neoantigens, oncoviral and endogenous retroviral elements. Neoantigens (Schumacher et al., 2019) are not found in normal tissues as they represent foreign proteins. These antigens may result from different non-synonymous genetic alterations such as gene fusions, frameshift mutations, insertions and deletions, and single point mutation. Melanoma ubiquitous mutated (MUM-1) is an example of a neoantigen derived from a single point mutation which can elicit a T-cell response (Coulie et al., 1995). In addition to neoantigens, the oncoviral antigens are epitopes derived from the

viruses mediating oncogenic transformation and they are not patient specific. For instance, EBV is associated with the development of many tumours such as B-cell lymphoproliferative disorders (Hollingsworth & Jansen, 2019). Endogenous retroviral antigens (ERVs) derive form retroviruses that infected the germ line cells of our ancestors and integrating their RNA and passed down through generations and now make up about 8% of the human genome (Bannert et al., 2018). Additionally, epigenetic mechanisms, such as methylation, play a role in suppressing the expression of ERVs in healthy cells. In the context of cancer, ERVs can be induced during malignant transformation or because of epigenetic therapy (Kassiotis & Stoye, 2016). This makes them potential targets for therapeutic approaches aimed at treating cancer. ERVH-5 is an example of these antigens and has been detected in various cancers such as bladder, colorectal, and lung squamous carcinoma (Gillison et al., 2000; Jiang et al., 2019).

Another group of TAAs include overexpressed, differentiation and CTAs. Overexpressed antigens are produced by genes that are excessively active in tumours and are wild type proteins. These proteins are only minimally expressed in healthy tissues but are consistently overexpressed in cancer cells. The proteins that are typically overexpressed play a crucial role in the survival of cancer cells. Differentiation antigens are proteins that are expressed due to the specific environment created in the target tissue. Prostatic acid phosphatase (PAP) and prostate-specific antigen (PSA) are examples of this group. It has been reported that two HLA-02-restricted peptides from PSA can elicit cytotoxic T lymphocytes (CTLs) responses (Olson et al., 2010). CTAs are overexpressed in tumours but are often not in normal tissues except in immune-privileged tissues such as placenta, testis, ovaries and trophoblasts at low levels in physiological conditions (Ward et al., 2016).

1.5 Tumour antigens as biomarkers for cancer

Tumour antigens can be used as indicators for determining abnormal cell functions present in cancer patients (Li et al., 2013a). Biomarkers can also be subdivided into diagnostic, prognostic and predictive markers based on their applications.

1.5.1 Diagnostic biomarkers

Diagnostic markers assist in detecting a particular disease before the onset of symptoms and they are used for screening purposes, such as the BRCA mutation (Li et al., 2013a) which is used for identifying risk, and predictive markers used for screening. Diagnostic markers are also used for early disease detection (Mayeux, 2004; Søreide, 2009; Li et al., 2013a) such as CA125 (Escudero et al., 2011) which is an approved biomarker for ovarian cancer but lacks specificity and can be indicative rather than an absolute indicator of disease.

1.5.2 Prognostic biomarkers

These markers allow the differentiation of "good outcome" cancer from "poor outcome" cancer guiding treatment and identifying the patients requiring an aggressive approach for disease management (Ludwig & Weinstein, 2005). Depending on the tumour microenvironment, prognostic biomarkers estimate the probability of disease recurrence after tumour resection. Identification of multi-gene expression may characterise the amount of residual cancer following surgical resection and determine the patient population that requires adjuvant therapy for the purpose of minimising the relapse risk. The prognostic markers assist in assessing disease recurrence (Mayeux, 2004). Examples include Kras overexpression in colon cancer (Lin et al., 2012).

1.5.3 Predictive biomarkers

These are markers identifying a sub-group of patients that will achieve benefit from certain drugs due to their molecular characteristics. For example, HER2 amplification in breast cancer will have better response when treated with trastuzumab (HER2 antibodies) compared to patients who receive tamoxifen only (Sawyers, 2008). In addition to response prediction, these markers may provide more information regarding treatment resistance using genotype-based analysis. For example, LC patients with distinct mutations in KRAS will not respond to the EGFR inhibitors such as erlotinib (Sharma et al., 2007).

1.5.4 Pharmacodynamic biomarkers

Pharmacodynamic cancer biomarkers assist in dose selection for new anticancer treatment in the early stage of clinical development by measuring the near-term

treatment effect of a drug on the tumour. Classically, the maximum tolerated dose is identified in a phase I clinical trial using a dose escalation study and after that this dose will be used in phase II clinical trials to determine anti-tumour activity of the new drug. This approach has one drawback for drugs that bind to a specific molecular target and thus it might be less relevant. Alternatively, target engagement studies involve determination of an appropriate dose by measuring the drug effect on its target using different doses (Shah et al., 2007). The dose selection for phase II trials, is based on the magnitude of target modulation. For example, imatinib mesylate inhibit the protein-kinase activity of BCR-ABL and enhances the clinical remission for chronic myeloid leukaemia patients at the same dose. Applying pharmacodynamics markers, the magnitude of BCR-ABL blocking is correlated with clinical outcome and could be used for the personalised drug dose selection (Shah et al., 2006).

Recently, biomarkers have become an area of interest due to various advantages such as effective detection at low concentration of biomarkers, the process is considerably faster and cost-efficient, well defined end-points and multiple biomarkers could be used for parallel detection in cancer management (Strimbu & Tavel, 2010).

1.5.5 Biomarker discovery phases

The Early Detection Research Network (EDRN) was developed by the National Cancer Institute (NCI) to encourage a systematic approach towards non-invasive cancer biomarkers development. The proposed biomarker discovery platform has five phases and involves: (1) identification of biomarkers (preclinical discovery); (2) validation of laboratories biomarkers (developing of assays and verification); (3) use of samples in clinical repositories from retrospective trials; (4) prospective screening programs; (5) Cancer Control (Pepe et al., 2001). A common biomarker discovery pitfall is when the biomarker does not represent the stage of the disease. Additionally, biomarker discovery presents variations in studies of selected populations, specimen sets, storage and processing (Ransohoff & Gourlay, 2010a). The EDRN developed a reference set which included a clear clinical application with defined specimens that represent the disease without bias but mainly matched the age, and sex to improve biomarker discovery. Phase one identifies the expression level of biomarkers in terms of transcript or protein levels or the presence of antibodies that recognise TSA, or other molecular

entities in tissue or body fluids such as ctDNA, miRNA (Mäbert et al., 2014). Biomarker identification starts with preclinical studies in which the marker expression is compared in cancer and non-cancerous tissues using mainly immunohistology chemistry (IHC) and Western blotting techniques. Introduction of high throughput techniques such as next generation sequence (NGS), and proteomics have revealed thousands of genes or proteins, that have differential expression. Usually, tissue biopsies are obtained at the time of diagnosis and before treatment, so that treatments do not interfere with the performance of the biomarkers. Tumour specimens are evaluated with respect to patient variability taking consideration of the fact that patients are from different geographical regions and have a complexity of genetic backgrounds. This is to be overcome as best as possible, using large patient cohorts to assess biomarkers – ensuring enough samples are used to assess sensitivity and specificity of each biomarker. Most specimens are obtained at a later stage of cancer while noncancerous tissue obtained from normal tissue adjacent to the tumour or abnormal tissue from related benign diseases (such as inflamed tissue) are used as control in the phase I EDRN studies (Pepe et al., 2001). Alternatively, blood-based, serum or plasma assays are widely used to determine protein or gene expression. Patients matched with controls should not differ for factors such as gender, age, race and lifestyle such as smoking wherever possible (He, 2006).

Ideal markers are expressed in cancer tissue only and are not found in healthy tissue from patients as well as healthy individuals. Assessing biomarker reliability and the reproducibility despite other factors (time of day, storage, age, gender, diet) is required as small changes in levels may obscure a promising marker (Pepe et al., 2001). This evaluation is based on assessing the sensitivity and specificity of marker in disease diagnosis (**Table 1.3**). Sensitivity is measured by number of true positives and the biomarker ability to distinguish the disease from other diseases or healthy individuals. Specificity is represented by the false-positive rate (FPR) in which control subjects that have positive expression of the markers. This binary system depends on many variables such as the sample size (Pepe et al., 2001).

Table 1.3 Definition of sensitivity and specificity as applied to biomarkers

	Positive	Negative
Patients	True positive = sensitivity	False negative
Healthy donors, related diseases, unrelated diseases	False positive	True negative = specificity

Generally, a larger population cohort can show a strong association either high or low sensitivity or specificity of the marker. Advances in statistics have developed a receiver operating characteristic (ROC) curve to overcome the problem of binary system associated with frequencies and the scale of raw-data measurement of different markers. ROC possesses the ability of quantifying true-and false-positive rates for markers. ROC calculation is a suitable for cancer screening with low FPR focusing on data analysis. Ranking markers based on statistical analysis is very common using ROC or the area under the curve (AUC) (Pepe, 2000; Pepe et al., 2001). Developing of statistical algorithms is ongoing to identify new promising markers (Pepe et al., 2001). However, heterogeneity of results is in phase I due to variation in samples obtained, collection, storage and statistical validation occur during data analysis. Study design is a key component in this process and depends on sample size of the participants, and specimen selection representing the disease. Markers variability is associated with the study objective and different factors are attributed to its variability, these include the number of promising candidates, the number and proportion of cancer subtypes in the studied sample, the markers capacity to distinguish different subtypes and the statistical algorithm (Baker, 2000). In addition, sample selection is confounded in small studies where random selection may result in disparity on some factors. These factors affect the biomarker value rather than the cancer itself, from perfectly matched subjects (Guyatt et al., 1986). Promising biomarkers undergo phase II for validation which includes assays for biomarker detection based on non-invasive specimen collection, commonly blood specimens. This phase focuses on estimating the true-positive rate (TPR), FPR and ROC for biomarker assay. In addition, biomarkers should be specific, sensitive, reliable, and reproducible by quantitative analytical methods as well as obtained through non-invasive methods. This phase involved optimizing assays for marker detection within the

laboratory and ensuring assay reproducibility at different laboratories. Several factors such as control matched in age, sex, lifestyle with studied individuals have affected assay optimization as in phase I. Assay optimization is an important step to determine marker thresholds as related to cancer stage or histology or other factors such as prognosis. Understanding of tumour initiating interaction with immune system and cancer microenvironment may assist in stratifying the patients with cancer and finding markers that diagnose cancer at early stages (Dunn et al., 2010).

Marker measurement is a non-invasive method, representing marker expression at the tumour tissue site. Thus, it is reflecting tumour microenvironment. Biomarker expression may be involved in disease pathways and may have clinical value. Evaluation of assay require adjustments for multiple comparisons and re-weighting statistical values estimating TPR, FPR in the targeted cohort. Marker for screening purposes include high risk patient without having cancer and require refining marker thresholds compared to other healthy participants and disease cases. Biomarkers act as surrogate and is a physical sign or a laboratory measurement correlated to a meaningful endpoint that has a substitute for a clinical value of disease progress, outcome, survival and prediction of therapy response (Fleming, 2005). Biological markers can be used as alternatives to surrogate endpoint as biomarkers are more easily measured and quantified in short time compared to other medical interventions and thus marker use may decrease the cost and length of disease diagnosis (Schatzkin & Gail, 2002). Translational research for biomarker is a challenging issue. Ideally promising marker identified in the basic research translated to the clinical use, reflecting the clinical endpoints in larger individual study cohorts and allowing relatively accurate conclusion (Fleming & DeMets, 1996).

Following the phase II, biomarkers undergo a retrospective longitudinal repository study. The phase aimed to identify the promising cohort for markers use. It includes the collection of clinical samples from cancer patients before clinical detection in comparison with patients who do not have cancer (control group). The comparison evaluates the marker capacity for detecting cancer pre-clinically. Marker expression has to show statistically significant difference from the control group in timeline of months or years prior the symptoms manifestation, consequently this marker has the capacity

for early detection. Biomarker discovery is based on investigating variables such as disease-characteristics, genotyping differences and its association with marker expression (He, 2006). It is important to use a well-defined protocol including target populations, all relevant control, sample collection, storage, laboratory assay and data analysis. Multiple sequential specimens may reduce the bias of analysis and provide more information about cancer and marker-interaction with immune system. Random and control case study are required with matched individuals as well as time follow-up length. Altered intervention may affect the estimated time and clinical value of new discovered marker. For example, if the study is not randomly controlled, it may result in misleading results with values of TPR and FPR. The advantage of this phase over phase II is that inter-individual variability may be assessed using longitudinal data with more significant comparisons of time-specific ROC curves (Pepe et al., 2001). Consequently, it increases the ability of distinguishing trend between control and case cohorts.

All mentioned three phases focus on retrospective studies, but phase 4 involves prospective study. It is aimed at detecting marker operating characteristics in terms of cancer nature and stage at the time of diagnosis. This phase describes the potential efficacy for the biomarker to detect early cancers as well as cancers with a slow growing nature or that undergo spontaneous regression. This phase involves the analysis of individuals who exhibit false positive results that require further follow-up procedures or who do not have cancer. This stage requires large cohorts as the prevalence of cancer is low, and the section of participants at this phase is not based on the disease status (Pepe et al., 2001). It therefore requires pilot studies with significant statistics and planning for people recruitment. Outcome of this phase may assess the feasibility of the assay implementation and compliance with work-up guidance (Ludwig & Weinstein, 2005).

The evaluation process of biomarkers involves a comprehensive assessment known as the Pivotal Evaluation of a Biomarker's Capacity (PEBC). This method is designed to accurately classify markers based on their correlation with a subject's outcome. It encompasses both prospective specimen collection and retrospective blinded evaluation (PRoBE) of the target population. During this phase, the analysis of samples and clinical data is conducted without knowledge of patient outcomes to validate the

biomarker's efficacy. Subsequently, patients and control subjects are randomly chosen from the cohort, and markers are evaluated in a blinded study. It is also crucial to note that specific considerations must be considered in biomarker studies. For instance, randomized clinical studies designed for predictive markers differ from those for diagnostic markers used as surrogate outcomes, considering marker cut-off values (Pepe et al., 2008).

The final phase aims to estimate the role of a marker in detecting cancer and its overall impact on the population's cancer burden. For example, prostate-specific antigen (PSA) has been shown to reduce prostate cancer mortality by 20% (Ross et al., 2000). This evaluation considers the cost of screening, work-up, and treatment in relation to the number of lives saved. It provides insights into protocol compliance in different settings, comparing costs and mortality rates (Lin et al., 1997).

The identification of a single biomarker detecting a cancer type with high sensitivity and specificity is still an unsolved challenge as cancer is heterogenous disease. However, a biomarker such as PSA has been proposed for prostate cancer. The PSA has high sensitivity and low specificity (Etzioni et al., 1999; Pepe et al., 2001). Specificity of the marker represents the percentage of negative samples as true negative value. Hence, low specificity may result in inaccurate diagnosis with unnecessary diagnostic tests and patient’s stress (Pepe et al., 2001). Thus, low false positive rate is required for the highly specific marker.

1.6 Tumour antigens as targets for immunotherapy

Neoantigens are appealing targets for immunotherapy since they are only present in tumours, which makes them extremely immunogenic (not necessarily all of them) and resistant to the effects of central tolerance (Schumacher & Schreiber, 2015). However, the heterogeneity within and between tumours may reduce its efficacy via inducing the negative selection of clones that are not expressed the targeted antigen. Mutational burden also plays a vital role in neoepitopes generations. Tumours with low mutational burden are difficult to identify because they may have less tumour epitopes to target them (Rooney et al., 2015; Vigneron, 2015; Hollingsworth & Jansen, 2019).

Oncoviral antigens are highly specific to tumours as they are not expressed in healthy cells and are commonly found in patients. Recently, EBV-seropositive infected tumour patients were treated with a vaccine that contained the full length of LMP2 and the C-terminal of EBNA1 proteins from EBV. This treatment resulted in a specific T cell response to LMP2 and/or EBNA1, indicating that it is possible to enhance the immune response against EBV in cancer patients (Taylor et al., 2014). However, further investigation is required to determine the extent of clinical benefits in a phase II clinical trial (NCT01094405). However, clinical application of oncoviral antigens (He et al., 2019) is limited to only 15% of cancers that have a viral cause.

One major obstacle is the difficulty of analysing the expression of these proteins in all types of tissues and under various physiological conditions. This makes it challenging to fully understand the safety profile of TAAs. There have been reports of potential risks associated with using these molecules in clinical treatments, including "on-target, off-tumour" toxicity and the onset of autoimmune diseases. For example, in patients with renal cell carcinoma (RCC), CAR-T cells designed to target carbonic anhydrase IX (CAIX) caused liver toxicity, leading to the discontinuation of treatment. Biopsies showed that CAIX was expressed in the bile duct epithelium, and T cells, including CAR-T cells, had infiltrated the area (Lamers et al., 2013). This is a clear example of "on-target, off-tumour" toxicity. Furthermore, the body central and peripheral tolerance mechanisms prevent T and B cells from recognizing self-antigens. To develop a TAA peptide-based vaccine, it is necessary to overcome this tolerance and stimulate the low-affinity and rare T cells that are still present in circulation (Hollingsworth & Jansen, 2019). This can interfere with the development of effective cancer therapeutic vaccines. One major drawback of using differentiation antigens as vaccines for cancer therapy is the risk of autoimmune toxicity. For example, when targeting melanoma-melanocyte antigens (Jäger et al., 2000), there have been reports of severe skin rashes and vitiligo lesions.

1.7 Prioritisation of tumour antigens as vaccine target

In 2004 Rosenberg *et al.* (Rosenberg et al., 2004) demonstrated that despite investment from funding agencies and a lot of time and effort, little improvement in patient outcomes had been made due to immunotherapy treatments in clinical trials. To try to improve this, the NIH asked the Cheever group and the Translational Research Working

Group (TRWG) of the top scientists in the field (Lepisto et al., 2008), to provide a mechanism for prioritising antigens that would dictate funding investment and ideally improve TAs translated to the clinical.

The Cheever group has proposed an antigen prioritization scheme to rank immunotherapeutic targets (Cheever et al., 2009). This model is based on Analytic Hierarchy Process (AHP) which is a structured method and mathematical model (Bhushan & Rai, 2007). They identified the ideal characteristics of an antigen that may be used in a vaccine and used weighted measures to rank them for further research investment by the NIH **(Figure 1.8)**. These characteristics were weighted using pairwise comparison process and resulted in the derived priorities according to the criteria and their relative weightings. Pair-wise comparisons create considerable variations in response and become explicit, resulting in a failure to achieve a consensus. After criteria weighting, sub-criteria may serve as the rating scales for each characteristic and provide a nonlinear relationship. The full value for each characteristic was given to the top sub-criteria and other sub-criteria, the level depending on the pre-determined weight, provided less value for the criteria. For example, "specificity" is weighed at 100% (when considered as absolute specificity, which is very unlikely), while "overexpressed in cancer" weighs only 35%. Thus, sub-criteria ranking may mislead antigen prioritisation (Silva et al., 2007b; Cheever et al., 2009).

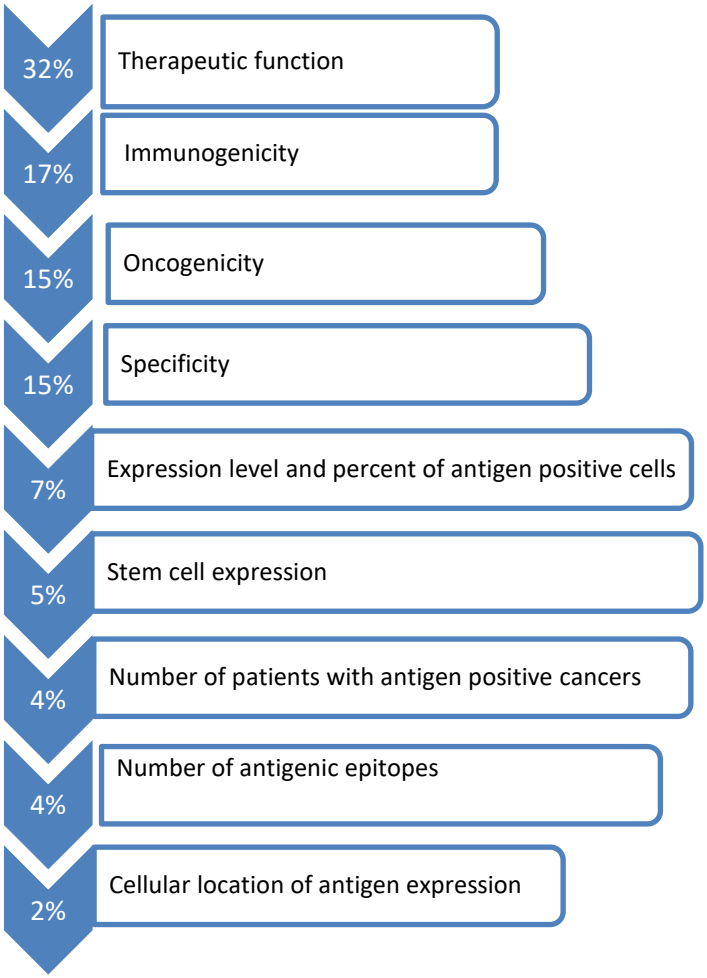


Figure 1.8 Characteristics of ideal antigen based on AHP in descending order as vaccine

The "therapeutic function" had the highest weight and weighed 32%. Then "immunogenicity" is represented by 17% of the weight, therefore therapeutic function was nearly twice as important as immunogenicity. Specificity and oncogenicity carries the same weight representation of about 15%. Figure modified from (Cheever et al., 2009).

Applying the same principle, cancer biomarkers should differentiate the disease condition from other disease or non-disease samples. Ideal makers have highly specific and sensitive for predicting disease diagnosis. In addition, reliable indicator of clinical outcome point, easily detected via non-invasive or easily invasive route such as blood, reproducibility of analytical detection method and cost effectiveness (Mäbert et al., 2014). Biomarkers translation into clinical practice is a challenging mission and even with approved markers such as carcinoembryonic antigen (CEA) for colon cancer, diagnosis still lacks ideal specificity and sensitivity as CEA is expressed in other cancers and non-malignant conditions (Ransohoff & Gourlay, 2010a). Cancer is a disease with a heterogenous and complex signalling pathways involving tumour immunity and host

response to tumour. The disease process involves changes in expression of intracellular and cellular genes, cell surface molecules and other mediators in complicated pathways in known and unknown ways (Schatzkin & Gail, 2002; Fleming, 2005). Although change in expression used as biomarkers may not reflect true clinical benefit, this has been associated with pathway modulation (Dunn et al., 2010).

1.8 Tumour antigens as biomarkers for the diagnosis of NSCLC

At the moment, the diagnosis of LC is widely based on using imaging techniques. In addition to aiding diagnosis, biomarkers could be useful tools for evaluate effective treatment, to monitor for recurrence after therapy and prognostic information prediction. The common serological markers (**Table 1.5**) investigated in LC include CA125, CEA and cytokeratin 19 fragments (CYFRA21-1) (Schneider, 2006; Patz Jr et al., 2007; Hanagiri et al., 2011). Although these markers are highly expressed in LC, they also have high expression in other benign lung diseases and thus have low sensitivity (Kulpa et al., 2002; Schneider et al., 2003). Therefore, because this of low sensitivity, tumour markers are not recommended as a tool for the early detection of LC. Tumour markers cut-off levels were adjusted to maximize the diagnostic yield of LC for checking healthy adults. Such cut-off levels have no benefit in imaging studies using method such as CT scan in patients suspected of LC. However, the cut off the biomarker has not been translated to the clinical meaningful end point, representing the gold standard of imaging finding. Considerably, the positive predictive value (PPV) of the CT has a greater value compared to tumour markers value at standard cut-off levels. However, markers PPV depend on the prevalence rates of lung carcinoma. High score of the PPV of tumour marker has been observed in patient populations with high prevalence rates (Okamura et al., 2013).

1.8.1 Carcinoembryonic antigen (CEA)

CEA is expressed in low levels specifically in foetal gastrointestinal epithelium, pancreas and liver under normal conditions (Grunnet & Sorensen, 2012). CEA production is mostly during foetal development and terminated before

Table 1.5 List of antigens known to be expressed in LC

Gene name (symbol)	Function	Healthy tissue	Expression in LC(s)	Reference(s)
Carcinoembryonic antigen (CEA)	Glycoprotein involved in cell adhesion and signal transduction	Low expression in colon, appendix	High expression all types in advanced stages	(Hammarström, 1999; Ayan et al., 2016)
Osteopontin (OPN)	Cell survival and angiogenesis	Gall bladder, placenta, brain	High expression associated with poor prognosis	(Berge et al., 2011; Ayan et al., 2016)
Cytokeratin 19 fragments (CYFR A 21-1)	Part of the cytoskeleton of epithelial cells	All epithelial cells	NSCLC mainly SCC. High expression associated with negative prognosis rather than advanced stages	(Yu et al., 2017)
Neuron specific enolase (NSE)	Glycolytic enzyme involved in inflammatory and neurotrophic activity regulating neuronal growth, differentiation, survival and death	Brain, adrenal, lung	Preferred for SCLC but also NSCLC and a marker of metastasis	(Isgrò et al., 2015)
Serum amyloid A (SAA)	Secreted during acute inflammation, transports cholesterol to liver, recruits immune cells to inflammatory sites	Housekeeping" role in normal human tissues	All types. High expression in late stages	(Urieli-Shoval et al., 2000; Biaoxue et al., 2016)
p53	Tumour suppressor gene	Low expression at cell stress	p53 mutation or overexpression was an indicator of poor prognosis, especially in patients with ADC	(Burotto et al., 2014)

birth. It is rarely detected in the blood of healthy people except smokers (Grunnet & Sorensen, 2012). CEA is a glycosyl phosphatidyl inositol-cell surface anchored glycoprotein (molecular weight of 180 kDa) playing a role in cell adhesion. The sialofucosylated glycoforms may serve as functional colon cancer to L-selectin and E-selectin ligands that play a crucial role in metastatic dissemination of colon carcinoma (Thomas et al., 2008; Thomas et al., 2009). CEA is already used as tumour marker in colon cancer diagnosis. CEA is also highly expressed in many cancer types such as gastric carcinoma and LC (Thomas et al., 2008; Grunnet & Sorensen, 2012). It was found that CEA expression in LC is higher than 3 ng/ml in blood, however this level is slightly higher in smokers compared to non-smokers. High level of CEA correlates with poor survival rate. Furthermore, serum CEA level could be a useful survival marker at the early stage of patient with NSCLC (Sawabata et al., 2002). It could also be a predictive marker for survival at early stage of NSCLC. There is evidence that high expression of CEA after surgical intervention is strongly associated with poor prognosis. CEA is commonly used as a biomarker for distinguishing malignant lesions from benign ones (Seemann et al., 1999). This worse prognosis may be associated with the tumour burden or failure to eradicate all lung carcinoma even after surgical resection. High levels of preoperative (CEA >50ng/ml) is associated with high frequency of metastasis even following surgical resection. CEA level may be an indicator of postoperative survival in patients with NSCLC. It has a sensitivity ranging from 40-70% in NSCLC and 30-65% in SLC. High concentrations of CEA have been reported in ADCs and large cell LC while squamous tumour have the lowest sensitivity and serum concentrations of CEA (Grunnet & Sorensen, 2012). CEA expression could be useful for monitoring the response to chemotherapy following surgery (Nonaka et al., 2004; Hsu et al., 2007). The significance of the preoperative serum CEA in females with LCs remains unclear and it is still unclear whether CEA expression has a prognostic value in males with SCC (Doseeva et al., 2015). This is perhaps because SCC is associated with a history of smoking more than ADC. As mentioned above, persistent high levels of CEA after complete resection for stage I NSCLC (confirmed pathologically) indicates poor prognosis than patients with normal preoperative serum CEA level. Poor prognosis at early stage of LC could undergo to be a specific group for therapies for subsequent survival benefit instead of observation alone in current guideline. The high preoperative CEA in non-smoking patients and poorer

survival are required to be considered in association to smoking status (Hsu et al., 2007). As CEA is identified in many carcinomas such as gastric cancer, colon cancer and breast cancer, it has been of limited value for use in diagnosis of LC due to its poor specificity. Efforts had been made focusing on CEA and other markers to detect LC (Doseeva et al., 2015).

1.8.2 CA19-9

CA19-9 was also known as a sialyl Lewis antigen (Schmiegel et al., 1985) and was firstly identified in colon cancer (Herlyn et al., 1982). CA19-9 contains a type I sugar chain with epitope of sialylated lacto-N-fucopentaose II, oligosaccharide related to the Lewis blood antigen. The location of Lewis gene is on chromosome 19q13.3. CA19-9 is found in normal epithelial tissues of many organs including bile duct, pancreatic duct, gall bladder, stomach, colon, rectum, uterus, prostate, bronchus, and salivary glands (Dietel et al., 1986). High serum concentration of CA19-9 release from cancer cells to blood (Magnani et al., 1983). It is a bulk moiety in the same form as high-molecular weight mucin-like molecules (Magnani et al., 1983). Interestingly, CA19-9 containing molecules are not found in individuals that have Lewis antigen-negative with congenital deficiency of a fucosyl transferase enzyme even if they have cancer. These individuals (Soejima & Koda, 2005) constitute 10% of Asians and Caucasians with the Lewis a/b–negative blood type (Magnani, 2004). CA19-9 is proposed to be specific for colon tumour, however, its serum level in pancreatic cancer patients is higher than in serum from those with colon cancer (Huang & Liu, 2014). CA19-9 was shown to be a biomarker for pancreatic cancers with specificity 0.80 (95% CI 0.77– 0.82) suggesting that it could be an important biomarker (Huang & Liu, 2014). High levels of CA19-9 have been reported in 44.2% of NSCLC patients especially the ADC subtype (56.5%). Moreover, its expression was associated with disease stage in positive CA19-9 ADC with expression of 30, 67, and 80% in stages I, II, and III, respectively. High concentration of CA19-9 is correlated with poor prognosis in preoperative patients. Serum CA19-9 level was linked to tumour CA19-9 concentration using immunohistochemistry; but Kawai *et al.* (Kawai et al., 1993) failed to demonstrate a correlation between postoperative survival and serum CA19-9 levels. CA19-9 combined with an increase in TPA provided a 49% sensitivity in patients with resectable NSCLC but limited sensitivity in SCLC when compared to the sensitivity of TPA

(44%) without significant differences in specificity (Toumbis et al., 1995). The Tsumatori group (Tsumatori et al., 1999) studied the serum E-selectin concentration and correlated it to the survival of NSCLC patients suggesting CA19-9 or the sialyl LewisX (sLeX) interaction enhanced metastasis. Elevated E-selectin expression was linked to significantly poorer survival rates in individuals compared to those with lower E-selectin expression, whether they exhibited carbohydrate antigens CA19-9 or sLeX. However, among patients lacking both CA19-9 and sLeX, there was no notable difference in survival rates between those with normal E-selectin levels (Tsumatori et al., 1999).

Apart from their expression in cancers, CA19-9 and sLeX are prominently present in interstitial pneumonia, serving as serum biomarkers. Despite the common use of KL6 as a biomarker for interstitial pneumonia (IP), CA19-9 exhibits a sensitivity of 42% and specificity of 94.3% in diagnosing interstitial pneumonia (Yokoyama et al., 1998). False positive results for CA19-9 predicted LC and suggest a concurrent interstitial pneumonia. Elevated CA19-9 levels are observed in various benign conditions, including pulmonary sequestration, bronchiectasis, chronic pancreatitis, and liver cirrhosis. In summary, serum CA19-9 positivity is relatively common in lung adenocarcinoma, potentially increasing with disease progression or stage. However, its predictive efficacy for post-operative survival seems to be limited.

1.8.3 SCC antigen

SCC antigen is a structural cytoplasmic protein that is found in higher levels in patients especially with squamous cell LC compared to other NSCLCs and it is also known as tumour-associated antigen or TA-4. Over-expression of SCC in blood indicates potential metastasis of the disease. It's sensitivity ranges from 15-55% for NSCLC (Kagohashi et al., 2008). SCC has a prognostic value when combined with NSE and CA125, showing that patients with high SCC had significantly shorter overall survival (OS) than squamous cell LC patients. However, poor prognosis is generally correlated with high levels of CA125 and neuron specific enolase (NSE) better than high SCC. Recently, a comparison study of six markers (CA125, CA19.9, NSE, CEA, CYFRA21-1, and SCC) showed that CYFRA21-1 serum levels was an independent prognostic factor of the 5-year survival rate for patients while other markers including SCC had no significance (Kagohashi et al., 2008). Furthermore, high concentrations of SCC and CYFRA21-1 have been reported in

30% and 59% of NSCLC patients, respectively (Yu et al., 2013a). SCC had low sensitivity of 13% in metastatic patients while the sensitivity of CYFRA21-1 was up to 74%. Due to poor sensitivity, SCC serum level is generally an inadequate tool for diagnosis but may be useful in monitoring recurrence for patients with NSCLC (Kagohashi et al., 2008).

1.8.4 Neuron Specific Enolase (NSE)

NSE is a glycolytic enzyme with three isozymes of enolase expressed by different genes. The isozymes are neuron specific and include enolase α which is ubiquitous; enolase β which is muscle-specific and enolase γ. Expression of NSE occurs at the late stage of neural differentiation in forms of dimer such as γγ- and αγ- which are used as an indicator for neural maturation (Barlési et al., 2004). NSE is expressed at low level in specific tissue such as neurons and peripheral neuroendocrine cell under normal circumstances. It is known as specific maker for neurons and it is high expression observed in malignancy proliferation (Barlési et al., 2004). Thus, it could be useful tool to aid in diagnosis, and staging of related neuroendocrine tumours. In addition, it may serve as target for treatment (Barlési et al., 2004). NSE is mainly tumour marker for SCLC in diagnosis, follow-up and prognosis. However, some reports have identified NSE expression in NSCLC (Altintas & Tothill, 2013). Its expression is associated with burden of tumour, metastatic sites number and treatment response (Altintas & Tothill, 2013).

In addition to LC, NSE is highly expressed in all stages of neuroblastoma and high expression found in late stages and metastatic disease. NSE is also expressed in neuroendocrine tumours (NETs) and gastroenteropancreatic (GEP)-NETs and might value in diagnosis. NSE could be used to determine the probability of neuroblastoma in newborns by testing the cord blood. Furthermore, high level of NSE has been reported in many cancers such as melanoma, renal cell carcinoma, malignant phaeochromocytoma, carcinoid tumours, and immature teratomas (Isgrò et al., 2015). NSE could improve the diagnosis and evaluation of outcome in many diseases, such as seizures, intracerebral haemorrhage, and ischaemic stroke. It is also useful in measuring cardiac arrest after cardiopulmonary resuscitation in comatose patients and brain injuries through quantitative evaluation of brain damage (Dittadi & Gion, 2013).

1.8.5 Serum amyloid A (SAA)

It is a cytokine-induced protein and the precursor in inflammation-associated reactive amyloidosis (AA-type) (Urieli-Shoval et al., 2000). It is expressed in the acute inflammatory phase in response to different insults. For example, it is found in liver but its physiological function is not fully understood while it can stimulate various cytokine production and play a role in acute immune response. SAA proteins found in normal histology tissues, as well as in inflammatory, Alzheimer, and malignant tissues (Biaoxue et al., 2016). SAA protein has binding sites for laminin, calcium, high density lipoproteins, and heparin/heparan-sulphate. Identification of adhesion motifs affected cell adhesion, proliferation, migration and aggregation (Biaoxue et al., 2016). This finding highlighted SAA role in various physiological and pathological processes such as AA-amyloidosis, rheumatoid arthritis, inflammation, atherosclerosis and malignancy. Recently, SAA protein may have "housekeeping" role in normal human tissues (Urieli-Shoval et al., 2000). New view of tumour progression is associated with chronic inflammation and different inflammatory factors may be useful as diagnostic or prognostic markers for specific tumours. It is a common believe that chronic inflammation enhances angiogenesis and cell proliferation, therefore it suggests having a role in tumourigenesis (Urieli-Shoval et al., 2000). Blood SAA protein has been identified in early stage of many cancers such as uterine, ovarian, renal, nasopharyngeal, LC and melanoma using immunochemistry and by proteomics methods. LC is considered as inflammatory and malignant development associated with different inflammatory mediators and cells factors. Thus, SAA is suggested as a marker for diagnosis (distinguish healthy individuals from patients with LC) and for prognosis prediction. Previous studies have shown the relationship between the expression of SAA and LC and shown its value for LC diagnosis (Moshkovskii, 2012; Biaoxue et al., 2016).

1.8.6 Human epididymis 4 (HE4)

It is a protein precursor of human epididymis protein, encoded by the whey-acidic-protein 4-disulphide core domain 2 (WFDC2) gene. WFDC2 gene is in chromosome 20q12–13.1 and this gene family has antibacterial and anti-inflammatory functions. HE4 is normally found in the epididymis, respiratory tract, and genital tract (Escudero et al., 2011). HE4 has been proposed by FDA as a biomarker for ovarian cancer diagnosis

combined with CA125 (Moore et al., 2011). Studies showed HE4 has more sensitivity and specificity than CA125 for ovarian cancer detection in premenopausal and postmenopausal females (Escudero et al., 2011; Moore et al., 2011). Interestingly, a negative predictive value of HE4 combined with CA125 would assist in differentiation between benign and malignant gynaecological diseases (Escudero et al., 2011). Moreover, it has a prognostic value in ovarian cancer and detection of recurrence at early stage. HE4 combination with CA125 had more sensitivity than using CA125 alone for endometrial ADC, suggesting this could be a prognostic marker of this carcinoma (Macuks et al., 2012).

Recently, high sera HE4 was reported in patients with lung ADC compared to matched healthy controls (Tokuishi et al., 2012). Abnormal expression of HE4 was firstly observed in tissue microarrays from SCC and lung ADC (Tokuishi et al., 2012). This observation is due to overexpression of WFDC2 gene in LC. Lung ADC cell lines showed moderate to high expression of different variants of HE4 (V1-4) with predominantly HE4-V3 variant and up 92.1% positive results (Tokuishi et al., 2012). In a clinical study, one third of patients expressed high level of HE4 indicating the possibility of its production by LC (Drapkin et al., 2005). High level of HE4 found in LC suggests its role in malignancy and prognostic value, but only there is limited date available in this respect. One study shows that combination of HE4 with CEA and CA125 could be helpful diagnostic markers in patients especially males with LC (Nagy et al., 2014). In addition to expression in ovarian and LCs, HE4 has been found in other carcinomas such as mesotheliomas, breast ADCs, and less frequently in, renal, gastrointestinal, and transitional cell carcinomas (Bingle et al., 2002; Drapkin et al., 2005).

1.8.7 Cancer testis antigens (CTAs) expression in NSCLC

CTAs are not expressed in healthy tissues except immunologically protected sites (that lack MHC class I). These are often reproductive tissues including testicles, ovaries and placenta (Gjerstorff et al., 2015). However, CTAs are often aberrantly expressed and at high(er) levels in many cancers including ovarian cancer, oesophageal cancer, and lung cancer (Fratta et al., 2011). About 50% of CTAs are encoded on the X chromosome and involved in proliferation of male germ cells and associated with inverted DNA repeats (Gjerstorff et al., 2015). The rest of CTAs are non-X CTAs that are found in later stages of

germ cell differentiation and encoded by loci dispersed throughout autosomes. Non-X CTAs are not associated with inverted DNA repeats (Gjerstorff et al., 2015). The X-linked CTA genes are more frequently activated in cancer cells, and specific gene families are simultaneously de-repressed in a tumour-specific manner. For example, in LC cells, NY-ESO-1 is often activated alongside MAGEA1 and MAGE-3. CTA (Gure et al., 2005; John et al., 2013) are further subdivided into four groups based on tissue expression: (i) testis-restricted transcripts such as BAGE (ii) germline genes expressed <3 non-gametogenic tissues such as MAGE (iii) differentially expressed CTA expressed in <6 non-gametogenic tissues such as XAGE1 (iv) ubiquitously expressed cancer/germline genes such as OY-TES-1 and found in juxtacentromeric regions (chromosomes 9, 13, 18 and 21). Tumours expressing CTAs could be divided into three groups based on number of CTAs and frequency of their expression (Gjerstorff et al., 2015). Tumours with high CTA expression i.e. >50% of the tumours have CTAs transcript found in >20% of the cells; examples of these cancers including melanoma and NSCLC. The second group consists of moderate CTA expressors, i.e. 30-50% of tumours express the CTA transcript at a frequency of >20% of the cells. The third group consists of low CTA expressors, <30% of the tumours express CTA transcripts in >20% of the cells. Breast and prostate cancers are examples of tumour with moderate CTA expression while leukaemia could be considered as low CTA expressor (Gjerstorff et al., 2015). However, 57% HAGE is frequently expressed in CML, and to a lesser extent in AML patient samples (Adams et al., 2002; Guinn et al., 2005). CTAs are highly immunogenic and hence are considered as a target for immunotherapy (Scanlan et al., 2004). CTA expression in NSCLC (**Figure 1.9**) ranges from 90% of SCC expressing at least one CTA, 62% of bronchiolo-alveolar cancer and 67% of ADC expressing at least one CTA. Expression of CTAs has been associated with advanced tumour and poor outcomes (Gure et al., 2005).

1.8.7.1 Melanoma-associated antigen gene (MAGE)

MAGE is a family of highly homologous proteins of approximately 200 amino acids length each sharing a conserved MAGE Homology Domain (MHD). These proteins regulate the ubiquitination of p53, MDM4, 5'-AMP-activated protein kinase catalytic subunit α-1 (AMPKα subunit), zinc finger transcription factors (KZNFs) (Weon & Potts, 2015). MAGE proteins are found in the germ line of male, placenta and many cancer

types such as melanoma, breast, prostate, brain, and LCs (Campagnolo et al., 2004; Krishnadas et al., 2013). MAGEs play an important role in cancer biology and have been increasingly reported as cancer biomarkers. The MAGE family includes the MAGE A, B, and C proteins. Expression of the MAGEs are associated with advanced disease and poor prognosis, as well as chemotherapy resistance (Gjerstorff et al., 2015). MAGE proteins increase the cancer cell survival either directly through interaction with tumour suppressor p53 or indirectly through regulation of E3 RING ubiquitin ligases activity (Doyle et al., 2010; Gjerstorff et al., 2015). It was also shown that MAGEs increase cancer metastasis by enhancing cell motility and thus increase invasive capacity to other tissue (Gjerstorff et al., 2015). MAGE-A1 is prevalent in 30% of NSCLCs by using reverse-transcription polymerase chain reaction, whereas it has 43% expression in NSCLCs by immunohistochemistry (IHC) (Fanipakdel et al., 2019).

MAGE A3 is a recombinant protein (Tyagi & Mirakhur, 2009), used as the adjuvant NSCLC immunotherapy in phase III MAGRIT trial. The use of immunotherapy did not improve the DFS. The CAR-T cells targeting the MAGE A3/A12 HLA-A0201 restricted peptide tested in 9 patients, but it resulted in severe neurotoxicity and subsequently 2 patients died. It was found that MAGE A12 is expressed in the brain and this may explain the neurotoxicity (Morgan et al., 2013). The use of MAGE A12 was cautioned as an adjuvant therapy.

1.5.2.2 XAGE-1 gene

The XAGE-1 gene has CTA like features and is known as a PAGE/GAGE-related gene on the X chromosome. Four alternative splice variants, XAGE-1a, b, c and d have been identified. XAGE-1b was dominantly expressed in lung ADC (Egland et al., 2002). 45% (14/31) mRNA expression of XAGE-1b has been observed in lung ADC, while only 6% (1/18) was found in other histological types of LC. XAGE-1b protein expression was found in most of the NSCLCs using immunohistochemistry (Nakagawa et al., 2005). Moreover, LC patients have shown an immune response to XAGE-1b protein (81 amino acids). XAGE-1b expression has been observed in hepatocellular and gastric carcinomas (Sato et al., 2007).

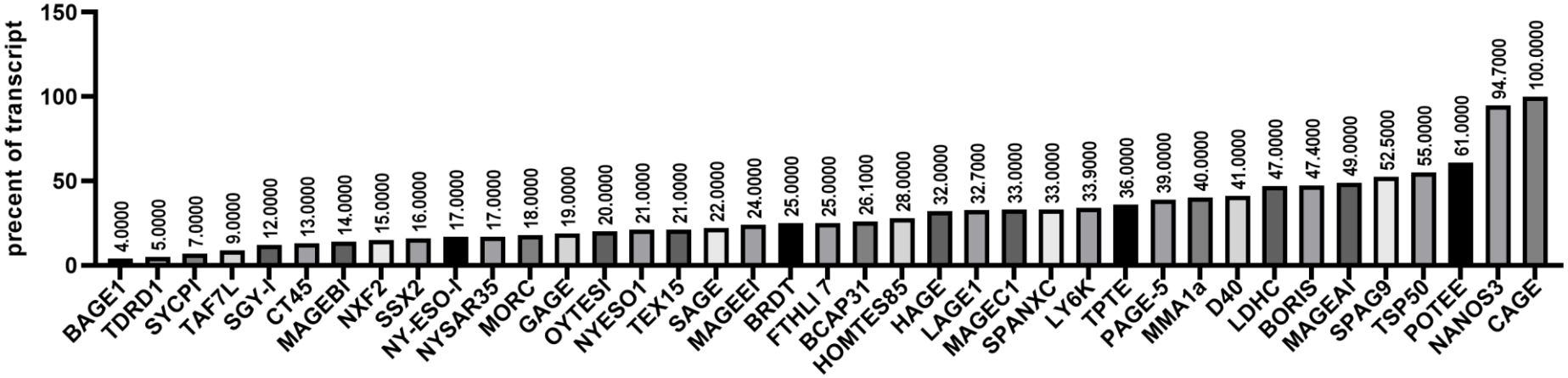


Figure 1.9 Expression of CTAs in NSCLC.

NSCLC is considered to be a tumour type with high expression of CTAs. CAGE/CT26 is widely expressed in most of NSCLC while CT34, CT8, CT12.1 have limited expression at less than 10%.

It has been observed that lung ADC patients with tumours expressing both XAGE-1b and HLA class I have prolonged survival post-operative (Nakagawa et al., 2005). However, patients with tumours expressing XAGE-1b may have shorter survival with down-regulation of HLA class I expression. XAGE-1b and HLA class co-expression may eradicate minimal residual disease via eliciting efficient CD8+ T-cell responses following surgery (Kikuchi et al., 2008). Consequently, XAGE-1b could represent a promising target for immunotherapy against NSCLC in an adjuvant setting.

1.9 Previous studies of LC/nodules diagnosis and the knowledge gap

The National Lung Screening Trial showed around 20% reductions in mortality were associated with LC when using low-dose CT screening, in known risk factors people. CT scans show a good efficiency in detecting small peripheral lesions particularly ADC (Field et al., 2016). In most of the cases, both benign and malignant nodules have high degree of similarity at early stages and different scanning errors occur due to frequent false positives in CT scan techniques (Nasrullah et al., 2019).

Despite of some successes, new challenges are encountered as the number of lung nodules detected increases and continues to rise. Many limitations of CT scans for the screening of LC include the limited reliability in distinguishing between benign and malignant LC due to overlapping radiological features. The nodules detected on CT scan screening lead to high false positive results in more than 95% of patients (MacMahon et al., 2017). Additional procedures consist of bronchoscopy, fine needle aspiration, transthoracic needle aspiration and surgical biopsy. As an alternative to CT scanning, FDG-PET and contrast CT are used to assess malignancy risk (MacMahon et al., 2017). While CT scans fail to detect preinvasive lesions that are centrally located, bronchoscopy and sputum cytology can identify 25% of lung malignancies that cannot be detected by imaging techniques. Positive screening of widespread cancers including breast, colorectal, cervical, prostate, and skin can be quickly followed up with a tissue biopsy at minimal extra-risk to patients (Huang et al., 2016; Welch et al., 2016; Bretthauer et al., 2017). However, this is not the case in LC as there is the requirement for invasive procedures with anaesthesia usage and this is associated with an increased risk including significant rates of pneumothorax (Huang et al., 2016). Furthermore, cost-benefit analysis showed that >40% of the total cost of LC management is attributed to benign

diseases with invasive approaches (Lokhandwala et al., 2017). Due to these invasive procedures being associated with morbidity, increased costs and delays in diagnosis, the development of non-invasive approaches is needed. Experience of biomarkers for evaluation in the fields of endocrine (HbA1C for diabetes) or infectious diseases (HIV viral load) are successful examples for clinical practice (Saenger et al., 2011).

Many studies have investigated non-invasive diagnosis of LC patients with indeterminate pulmonary nodules (IPNs) focusing on circulating biomarkers. The most attractive diagnostic method is represented by blood biomarkers mainly proteins and micro-RNA panels, given that blood sampling has low risks for patient and is easily accessible. Blood biomarkers have been proposed for distinguishing benign from malignant lung nodules (Kammer & Massion, 2020). For example, 552 patients have been studied (113 benign nodules and 339 malignant) using serum C-reactive protein (CRP) combined with CEA, in the presence or absence of nodule spiculation, calcification, and CT bronchus signs. It has been found that CRP correlates with inflammation while CEA is one of glycoproteins that can assist in cellular adhesion and is thought to be upregulated in many epithelial cancer including LC due to metastasis (Yonemori et al., 2007). Furthermore, a phospholipid hydrolase enzyme is known as secretory phospholipase A2-IIa (sPA2-IIa) that facilitates several precursors to eicosanoids release, regulating several mechanisms including immunity, inflammation, and carcinogenesis. The sPA2-IIa was found to be highly expressed in prostate cancer but could also assist in differentiating LCs from healthy individuals. However, it failed to discriminate between LCs and benign nodules with AUC of 0.68 (less accurate) (Kupert et al., 2011a).

Studies focused on miRNA as biomarkers in patients with lung nodules (Yu et al., 2010; Shen et al., 2011b). Around 65 patients with solid nodules (33 patients with benign nodules and 32 with malignant masses) were recruited as well as 92 smokers without any solid nodules used, as controls. Moreover, validation of the miRNA set used an independent group of 156 patients with solidarity pulmonary nodules (SPNs) consisting of two sets (76 patients with malignant masses and 80 patients with benign nodules). Studies have shown difference in expression of five miRNAs between LC and normal lung tissues with higher expression of miR-21, miR-210 and miR-375 in lung malignancy and lower expression of miR-126 and miR-486-5p in LC compared to healthy lung tissues (Yu

et al., 2010; Shen et al., 2011b). Therefore, these five miRNAs were examined in pulmonary nodules to determine if there was any expression difference between benign and malignant nodules. It has been found that there were higher levels of miR-21 and miR-210 expression in malignant nodules in comparison to both patients with benign nodules and smokers with no cancer evidence. In contrast, miR-486-5p expression was lower in patients who had malignant nodules compared to both groups of patients bearing benign nodules and smokers with no cancer evidence. Thus, these three miRNAs could assist in differentiating malignant nodules as potential plasma markers in this study. miRNAs expression in plasma was correlated to nodule size and smoking patterns only. However, the estimated correlations of three miRNAs were low and combing the three gene surpassed a single test used alone as miRNA expressions were complementary to each other. miRNA gene changes had no association was found between age, ethnicity, nodules histology, and gender. Moreover, no statistical difference was found between the three genes in terms of specificity and sensitivity in diagnosis among different stages of LCs. Regarding the validation set, in Shen *et al.* miR-21 and miR-210 were higher in patients bearing malignant nodules and lower expression of miR-486-5p in malignant nodules (Shen et al., 2011a). Optimal thresholds were used to determine miRNA diagnostic performance in the test setting. The conclusion was very similar to Huang et al., with no association of gene changes between age, sex, and ethnicity except smoking and size of nodules (Huang et al., 2009). Benign nodule formation was correlated to chronic smoking, infections, granuloma, etc. Some molecular changes have been observed in these non-malignant conditions (Minna et al., 2002; Belinsky, 2004). Not surprisingly, benign nodules had a higher degree of aberrant miRNA expression in comparison to tissue from healthy people without nodules. However, the highest expression of abnormal miRNAs was found in LC patients, compared to patients with lung nodules.

This finding resembles the results of Boeri *et al.* (2011) who showed that miRNAs could assist in disease diagnosis even before it's detection by using spiral-CT (Boeri et al., 2011). The sensitivity and specificity of three miRNA genes together were higher than the individual genes. Combinations of miRNA expression with CT may help increase the accuracy of diagnosis and assess the low sensitivity CT images in early diagnosis of lung

malignancy. Also, this combination can overcome the difficulties of determining tumour location that is the major challenge with circulating biomarkers. Interestingly, sputum-based genomic probes combined with CT images have supported this concept and increased the sensitivity of detecting stage I LC in comparison to the use of CT alone (Jiang et al., 2009). Thus, miRNA expression may assist in lung nodule management as it is believed that malignant masses are commonly bigger in size than solid nodules due to the finding that the level of miRNAs was statistically associated with nodule size and aberrant expression of miRNAs was identified in malignancy-related changes (Ost et al., 2003; Djebbi et al., 2022). Remarkably, there was no association between miRNA levels and histological subtypes. miRNA alterations were found in both advanced stage and early-stage LCs. Using miRNAs is cost-effective (less expensive, non-invasive) in differentiating malignant from benign nodules and could be applied to early LC (Djebbi et al., 2022). However, there are several limitations of miRNAs as biomarkers. Firstly, there are a limited number of miRNAs selected to be part of biomarker panels. Therefore, miRNA biomarkers do not possess high enough accuracy to be utilized in clinical settings (Shen et al., 2011a). Comprehensive miRNA microarrays could be used to overcome the challenge of detecting malignant nodules in patients, in comparison with patients with benign nodules. Genome-wide miRNA profiling will identify more information about miRNAs that could be used to detect malignant nodules and improve the diagnosis efficiency of plasma miRNAs as global biomarkers. Secondly, solid nodules were highly heterogenous with most malignant nodules being associated with primary LCs and the rest being caused by metastases from other organs. It remains unclear whether the three miRNAs described by Shen et al, 2011a could aid in the detection of lung metastatic cancer (Shen et al., 2011a).

Interestingly, more studies have focused on patients with lung nodules and have followed-up patients at least for 2 years and they remained cancer-free (using annual LDCT) included in Daly *et al.* study. Around 136 patients were divided into two cohorts: the first group included 69 patients who were pathologically diagnosed to have lymph-node negative LC and the second group consisted of 67 patients with benign diseases. The second group contained 35 patients with benign resected disease, 21 patients with granulomatous inflammation, 9 patients with non-specific inflammatory changes and

five patients with lung infections. The validation set included 81 patients split between two groups: 20 patients with lymph-node negative LC and 61 patients with benign diseases (Daly et al., 2013). Farlow et al., (2010) identified 17 markers that were differentially expressed between NSCLC and benign diseases (Farlow et al., 2010a). These seventeen markers were measured in lung nodules and included: CA125, cytokeratin 19 fragment 21-1, stromal cell-derived factor -1(α+β) (SDF-1α+β), osteopontin, interleukin-1 receptor antagonist (IL-1ra), sIL-2Rα, IL-6, IL-10, Eotaxin, monocyte chemotactic protein 1, macrophage inflammatory protein-1α, tumour necrosis factor alpha (TNF-α), soluble epidermal growth factor receptor (sEGFR) matrix metalloproteinase 2, and C-reactive protein (Farlow et al., 2010a; Daly et al., 2013). The group showed that nine of the 17 markers in the panel had an AUC greater than 0.60 while eight markers revealed a significance for identifying indeterminate nodules between the LC and benign diseases. Four markers (CA125, sE-selectin, IL-1ra, and IL-10) had strong significance (p<0.01) in distinguishing LC in indeterminate nodules, with all found to be significantly raised in LC patients. Using multivariate analysis, seven markers had 76.5% accuracy in differentiating NSCLC patients from benign cases and included IL-10, IL-6, IL-1ra, sIL-2Rα, SDF-1α+β, TNF-α, and MIP-1α. This panel had 35 cases of true negative, 32 false positives, 69 true positives and 0 false negative, therefore, the specificity was 52.2%, 100% sensitivity and a negative predictive value, and the AUC=0.91. In contrast, the validation cohort had shown 15 true negatives, 19 true positive cases, 46 false positives, and one false negative case. This is consequently lead to 24.6% specificity, 95% sensitivity, 93.8% negative predictive value, 0.676 AUC (Daly et al., 2013). The decrease in sensitivity and specificity may be due that validation set had 20 patients differentiated with LC from 61 cases with benign disease. Previous plasma markers panel including IL-1Rα, TNF-α, MIP-1α successfully aided in prediction of NSCLC versus benign diseases (Farlow et al., 2010a). Here, the plasma markers panel had 76.5% accuracy in differentiating LC and benign disease and included IL-6, IL-10, IL-1ra, sIL-2Rα, SDF-1α+β, TNF-α and MIP-1α. However, false-positive cases highly impacted the accuracy. Markers panel should have the clinical value of differentiating indeterminate nodules that proceed to LC when negative predictive value is 100% with no false-negative cases seen, thus the panel predicts benign disease. In validation cohort, this failed to distinguish LC from benign disease resulting in a substantial count

of 46 false-positive results. Despite this challenge, the false-negative instances persisted at a notable rate of 93.8%, and the panel capability to predict benign disease remained unaltered. This panel has accurately determined that 15.4% of patients had unnecessarily undergone surgical biopsy and were recommended for more conservative treatment, with serial LDCT screening. Regardless of high sensitivity in testing set (100%) and validation cohort 95%, the specificity was lowered to 52.2% and 23.3% for two sets respectively. The panel could lead to reduction in morbidity, mortality, and health care costs even with low specificity. Efforts continue in new biomarkers discovery using proteomics and immunoproteomic techniques to enhance the overall accuracy of plasma tests. Another attempt included SCC patients that expressing unique tumour antigen to increase the overall efficiency value to marker panel (Farlow et al., 2010a). Other blood tests have been developed for assigning clinical significance to indeterminate nodules including Early CDT-lung test manufactured by Oncimmune and a multianalyte serum biomarker panel by Bigbee *et al.* (Murray et al., 2010; Boyle et al., 2011; Bigbee et al., 2012). The Oncimmune test included test of autoantibodies against NY-ESO-1, p53, GBU4-5, annexin I, SOX2 examined and validated against early-stage LC and control patients. However, the sensitivity was low around 39% with specificity of 90% (Murray et al., 2010; Boyle et al., 2011). 11 markers panel were validated by Bigbee *et al.* for predicting cancer risk in high-risk individuals with indeterminate lung nodules. Markers in the panel included prolactin, transthyretin, sE-selectin, thrombospondin-1, C-C motif chemokine 5 (CCL5; RANTES), macrophage migration inhibitory factor (MIF), plasminogen activator inhibitor, tyrosine-protein kinase, erbb-2, CYRA 21-1, and serum amyloid A (SAA). The sensitivity of this panel was 73.3% and specificity of 93.3% (Bigbee et al., 2012). However, this panel had not changed the treatment plan and patients underwent subsequently invasive procedures, with modest clinical impacts. On the other hand, this panel had good negative predictive value of 77.8% in validation cohort and could aid in clinical management of screening population (Bigbee et al., 2012). The positive predictive value is required to be improved by exploring biomarker targets. Large-scale validation will allow marker panel in large prospective clinical trials to become highly efficient tools for clinical practice (Daly et al., 2013). In addition to predicting nodules propensity to malignancy obstacles, LC studies often referred to LC as one single disease; however, LC is heterogeneous group of diseases rather than one

entity. It is obvious that NSCLC patients respond differently to treatment because they suffer from a biologically heterogeneous group of LCs (Tufman et al., 2013). The role of intra-tumour heterogeneity (ITH) and genetic diversity within a single tumour remains unclear as well as their impact on the sensitivity of tumours to immune modulation (Castle et al., 2012; Robbins et al., 2013; Schumacher & Schreiber, 2015).

1.10 Antigenic targets for the treatment of B-ALL and knowledge gap

Molecular analysis of B-ALL subtypes has revealed genetic, epigenetic, and pathway alterations, yielding new insights for B-ALL management. TKI and JAK2 inhibitors have been proposed for B-ALL treatment. Thus, identification of antigens associated within B-ALL remains an attractive approach for B-ALL treatment. CTA have not been found to be frequently expressed in leukaemia (Adams et al., 2002; Khan et al., 2019) . RAGE-1 and NY-ESO-1 are not found in ALL (Greiner et al., 2004) while SCP-1 and SSX-2 are expressed in 29% of ALL (Niemeyer et al., 2003) and may be target for immunotherapy. CTA expression was varied from 65% in ALL of one CTA gene to 41% of cases with 2-5 CTAs the expression of which did not correlate with any subtypes (Niemeyer et al., 2003). HOM-TES-14/SCP-1 was found in 12% of ALL. SSX-1, HOM-MEL-40/SXX-2 were expressed in 29% of cases. Around 47% of ALL showed expression of SCP3a (Niemeyer et al., 2003). MAGE A genes are weakly expressed in ALL except MAGE A3. MAGE A1 was not found in ALL (Chambost et al., 2001) but it was detected as weak signal in 2 out of 53 ALL cases and, similarly, MAGE B2 was detected in one case. MAGE A3 is strongly expressed only in 38% of ALL (Martinez et al., 2007).

About 90% of aB-ALL patients achieved first complete remission (CR1), but more than half of these patients will relapse. Few of the patients who achieve CR1 will progress to CR2; however, relapse and long-term survival are challenging for subsequent treatment. B-ALL heterogeneity is another obstacle for the disease management. Although surface antigens are specific to B-cells and rarely found in other tissues, these antigens present in both normal and malignant B-cells leading to side effects (Gardner et al., 2016). This issue is particularly pronounced in early B cells exhibiting stem cell features, such as mixed lineage leukaemia (MLL). When treated with CD19, there is a notable loss of the B-phenotype, causing a shift in lymphoid cells toward the myeloid lineage, thereby excluding any further opportunity for CD19-targetted immunotherapy. This lineage shift

is not exclusive to CD19 but extends to other cell surface B antigens. Despite the promise of antibody-based therapy and its approval for B-ALL, this form of treatment has a limited duration of effectiveness. Its efficacy is contingent upon the density and percentage of antigen expression. Repeated immunotherapy sessions lead to the persistence of antigen expression, suggesting that antibody-based therapy alone may not suffice for the comprehensive treatment of B-ALL (Raponi et al., 2011; Gardner et al., 2016). Another challenge lies in the failure of activated T-cells to recognize tumour antigens, resulting in the elimination of CAR-T cells expressing the corresponding receptors, as exemplified by CD38 CAR-T, posing challenges in their production and effectiveness (Depil et al., 2020). Additionally, inadequate expansion of CAR-T cells, limited persistence due to the absence of memory T cells, and the presence of immunosuppressive signals from the tumour microenvironment may constrain the clinical advantages of this approach (Gauthier & Turtle, 2021). Antigen loss or downregulation due to the pre-existence of splice variants leads consequently to antigen escape mediated relapse. The utility of surface protein properties may affect the immunotherapy's efficiency. For example, CD22 is better served as a target for antibody-drug-conjugates as it is more efficiently internalized compared to CD20, while CD20 is more efficient for Fc dependent activation mechanisms and is optimised for use as a naked antibody due to its slow internalization and the long exposure of the Fc fragment (Gardner et al., 2016).

Boullosa *et al.* 2017 (Boullosa et al., 2017) showed that WT1 and BIRC5 were expressed in samples from B-ALL patients at the transcript level using qPCR, but transcripts were not found in the healthy control samples. Only BIRC5 levels were significantly higher ($p= 0.015$) and its protein was found only in B-ALL patient samples. Survivin is an antiapoptotic gene involved in cell cycle regulation, and it is found to have low levels of expression in terminally differentiated healthy tissues such as embryos. BIRC5 is upregulated in many cancers such as lung and breast cancer (Chang et al., 2020; Li et al., 2023b) due high expression of oncogenes and loss of tumour suppressor genes that leads to over-activation of JAK/STAT, Akt/PI3K and TCF-β-catenin pathways promoting tumour proliferation and survival. BIRC5 overexpression is correlated to chemotherapy resistance and tumour aggression (Li et al., 2023b). All of this suggests that BIRC5 may

be a novel target for B-ALL therapies. Serum profiling of B-ALL and age- and sex-matched healthy donors revealed three tumour antigens were differentially expressed in sera from patients compared to controls. These were bone marrow tyrosine kinase (BMX), dCTP pyrophosphatase 1 (DCTPP1), and vestigial like 4 (VGLL4) (Jordaens et al., 2020). The BMX is located in chromosome Xp22.2 and belongs to Tec kinase family which is a non-receptor tyrosine kinase involved in many signal transduction processes including the PI3K and TNF pathways. In addition to its expression in bone marrow cells, BMX is highly expressed in prostate cancer and healthy cells such as keratinocytes (Jiang et al., 2022a). Targeting of BMX may be achieved through the use of an epidermal growth factor receptor (EGFR) inhibitor, BMX-IN-1 and CTN06 a small molecule inhibitor of both BMX and BTK. This treatment suppresses tumour growth and migration via induction of autophagy and apoptosis. Another BTK target, ibrutinib has been targeted pre-BCR signalling and BMX (Liu et al., 2013). Ibrutinib is effective for the management of B-cell malignancies in clinical trials including R/R chronic lymphoblastic leukaemia (CLL) and B-ALL (Kim et al., 2017). DCTPP1 is an intracellular regulator of 5'-methyl dCTP metabolism and is involved in DNA hypermethylation. Decreased DCTPP1 is due to activation of PI3K/AKT/mTOR pathway and associated with poor prognosis and reduced overall survival in solid cancer (Marin et al., 2020). DCTPP1 is suggested as a biomarker of chemoresistance in gastric cancer as its expression is associated with increasing chemotherapy resistance (Xia et al., 2016). VGLL4 is a co-factor of TEA domain-containing transcription factors (TEADs) and is involved in tumour proliferation and migration (Zhang et al., 2017c). Thus, DCTPP1 and VGLL4 may be an attractive target for B-ALL and further studies are required to verify its expression in B-ALL samples and role in B-ALL pathogenesis. Alternatively, to the present immunotherapy targets, identification of novel antigens is important for survival in leukaemia especially for subtypes associated with poor prognosis and phenotypic plasticity such as MLL B-ALL. The search for antigens that are important and effective for B-ALL may provide novel future strategies for treatment.

1.11 Hypothesis

Early cancer detection has been widely demonstrated using different molecules in the literature. Tumour antigens are suggested for early diagnosis, they have low sensitivity

and specificity and more accurate for identifying the advanced diseases and could be targets for the treatment. Further research utilises advanced methodologies to identify signatures associated with the specific cancer because tumour antigens are found in normal conditions and non-cancer diseases may be a major challenge for tumour antigens analysis. The gold standard procedures for cancer diagnosis still include the invasive procedures and imaging techniques that requires skilled expertise, and expensive equipment. Identifying new biomarkers may assist in early diagnosis especially in the case of lung nodules that may proceed to LC within two years. Furthermore, tumour antigens could be used as biomarkers for prognosis as in B-ALL. B-ALL with different cytogenetic abnormalities have different outcomes and identifying the molecular abnormalities with flow cytometry and other techniques is still expensive. Identification of robust biomarkers which correlate with prognosis would assist not only the diagnosis but also it may offer new targets for treatment. The work presented in this thesis aims to address these limitations in the current techniques for diagnosis of LC and B-ALL.

1.12 Thesis Aims

This study aimed to address the need for early disease biomarkers and new targets for immunotherapy in two diseases with poor associated outcomes for patients. These were chosen based on sample accessibility and the interests of the lab I was working in.

The aims of this study were as follows:-

- To identify biomarkers that enable the earlier diagnosis of lung cancer through the systematic literature review and by analysing RNAseq focusing on cancer testis antigens
- To analyse adult B-ALL sera from patients at disease presentation and identify novel antigens for the immunotherapy of disease
- To identify enriched pathways in B-ALL and prioritise antigens using Cheever criteria

Chapter 2: Systematic Literature Review

This work has been submitted for publication:

Mohamed, E., García Martínez, D.J., Hosseini, M-S., Yoong, Y-S., Hart, S. & Guinn, B.A. (2023) Identification of biomarkers for the earlier detection of non-small cell lung cancer: a systematic literature review and meta-analysis.

The contributions of each author are indicated throughout the chapter.

2.1 Systematic Review process

The review was performed and reported according to the he Preferred Reporting Items for Systematic reviews and Meta-Analyses for Protocols (PRISMA) guidelines (Moher et al., 2015; Shamseer et al., 2015) to answer the research questions: which TAs/ tumour associated antibodies (TAABs)/RNA/miRNAs/ctDNA have been identified as possible biomarkers for the earlier detection of NSCLC? The review was conducted using the predefined protocol (Supplementary Data I) that was registered on the international prospective register of systematic reviews (PROSPERO: CRD42022336488). A search of the literature in the Cochrane library, PubMed, MEDLINE, CINAHL Complete, Scopus, Web of science, and Clinical trial.gov, provided limited evidence of other similar systematic reviews but did identify a review of autoantibodies as biomarkers for LC detected by ELISA (Yang et al., 2019). Development of the search strategy was based on index terms found in three to six sentinel articles that an initial PubMed screen of the literature identified. The population, intervention, comparison, outcomes (PICO) framework provided structure for the search for eligibility criteria (human studies, NSCLC patients compared to healthy donors and other inflammatory diseases). García Martínez, D.J and Eithar Mohamed performed all searches for antigens in LC using MeSH term as follows:

(cancer* or tumour* or tumour* or neoplasm* or carcinoma* or malignancy*) N2 (lung* or pulmonary) AND (antigen* OR protein* OR RNA* OR miRNA* OR cell surface marker* OR inflammatory cell*) AND (early detection OR early diagnosis OR early

biomarker OR early marker). Research terms used in different databases such as Medline, Scopus, Web of science stated (Appendix I). Following the protocol, duplicates were removed using Endnote. Then articles were screened by title and abstract and remove not relevant studies. Assess of eligibility using the exclusion and exclusion criteria was addressed. The inclusion criteria (Table 2.1) included primary research on human studies, number of participants >10 individuals while the review excludes cell line and animal studies and studies not reporting the sensitivity and specificity of markers. The selected studies were assessed for risk of bias using Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) (Appendix I).

A systematic literature review (SLR) was performed to identify the tumour antigens that had already been found to be expressed in LC. We combined these data and performed meta-analysis to determine which antigens provided the most promising targets for future therapies and/or biomarkers for disease. These antigens would provide the basis of future studies, providing controls and comparators by which we can compare the novel tumour antigens identified in Chapter 4.

2.2 Results

The database searches identified 7025 articles in total and 2398 duplicates were removed. 4627 articles were excluded based on title and abstract. After evaluating 148 full-texts, 79 articles were included in this systematic review (Figure 2.1, Table 2.1). Studies were published by groups in China (40 studies), USA (17 studies), Korea (3 studies), Canada (1 study), France (1 study), Italy (1 study), Spain (1 study), Taiwan (1 study) and multiple countries (two studies). The sample sizes varied across the included studies, ranging from 18 to 1479 LC cases. 28 studies investigated either single antigens or antigen panels and reported sensitivities ranging from 48-95% and most studies investigated blood biomarkers. 19 studies investigated autoantibodies, 21 studies focused on miRNAs and RNA and 11 studies explored circulating tumour cells (CTCs) in early-stage NSCLC. Six studies (Table 2.2) identified biomarkers that had a sensitivity and specificity of more than 90%. ctDNA and CTC had the highest values of sensitivity, and a high specificity, with the lowest Standard Deviation of all of the groups suggesting these were the best options for the early (minimally invasive).

2.2.1 Risk of bias (ROB)

74 studies had high ROB (Figure 2.2; Supplementary Table 9.1) most commonly due to the use of case-control study designs, causing the "patient selection" domain to score highly. Three other domains also scored poorly, most notably "flow and timing". Applicability concerns were low for all studies. Different methods of detection were used in each study and could have impacted the robustness of the results obtained. ELISA was the most common technique applied to analyse this early LC biomarkers, mainly for antibodies and antigens while RT-PCR was used for miRNA and ctDNA detection.

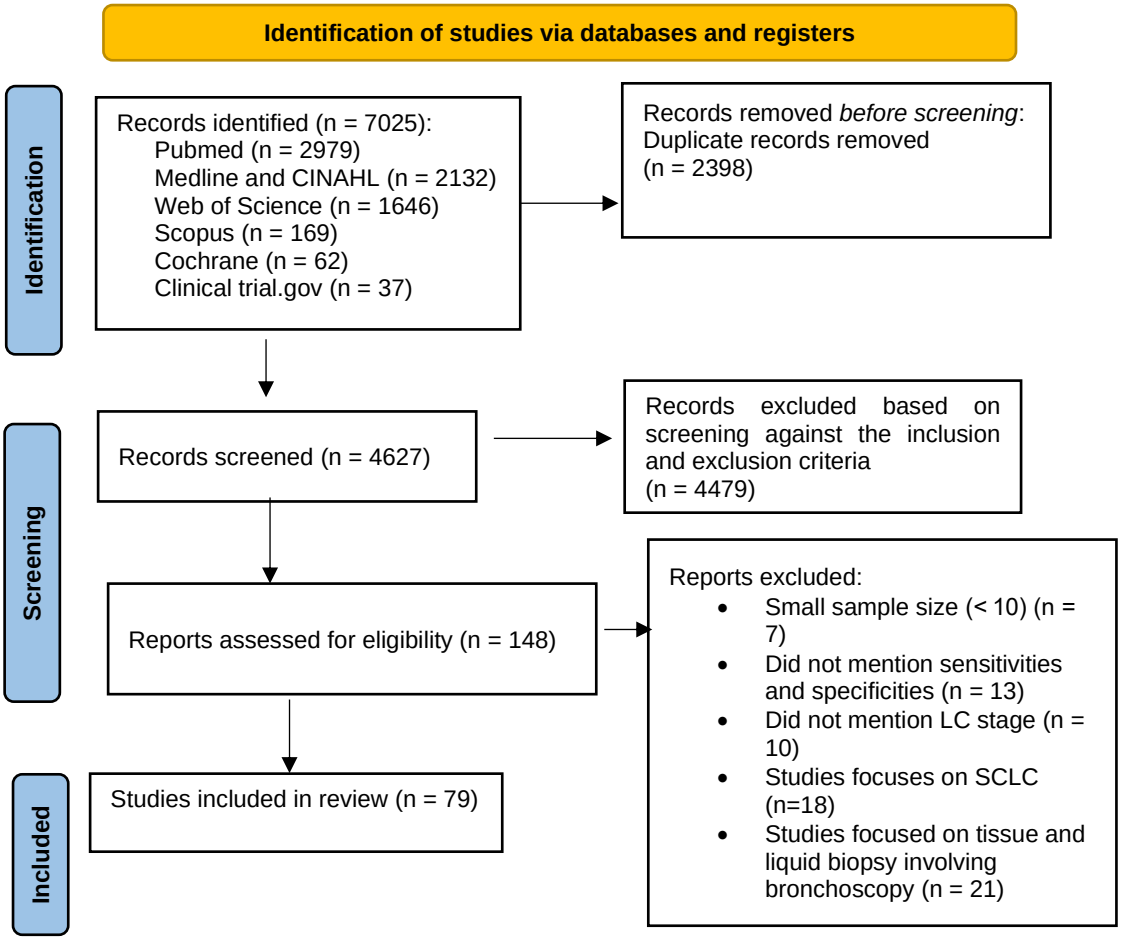


Figure 2.1 PRISMA flow diagram for systematic review for biomarkers for early NSCLC detection

Articles were identified from different databases, then duplicate references were removed using the Endnote. The articles were screened by title and abstract and not relevant studies removed (n=4,479). Then studies were assessed for eligibility using inclusion and exclusion criteria and full-text studies (n=79) were used for quantitative analysis.

2.2.2 Synthesis of the results - Meta-analysis

Twenty-three studies reported adequate data to enable the pooling of AUC, and random effects meta-analysis found that the pooled AUC was 0.86 (95% CI 0.82-0.90) (**Figure 2.3**), indicating that the diagnostic performance of biomarkers for early NSCLC were excellent. However, the heterogeneity was also considerable (I^2 = 98%, p < 0.00001). Sensitivity analysis found that the pooled AUC remained consistent, indicating that the results were robust. Subgroup analysis found that there was no significant subgroup difference (I^2 = 62.2%, p = 0.05) based on the type of biomarker used. Among the four types of biomarkers, pooled AUC for the autoantibodies subgroup was the lowest (pooled AUC = 0.77, 95% CI 0.68-0.87). Subgroup analysis based on the type of control used in the study showed that there was a significant subgroup difference in diagnostic performance (I^2 = 58.3%, p = 0.03). The biomarkers performing the least accurately were those differentiating early NSCLC from benign lung diseases (pooled AUC = 0.71, 95% CI 0.60-0.82) (**Figure 2.4**). The findings of this meta-analysis should be interpreted with caution as each study investigated different individual biomarkers. There were limited studies with data suitable for meta-analysis, hence this meta-analysis was not representative of all studies encompassed by the systematic review. However, we provide preliminary evidence that current NSCLC biomarkers can generally be expected to perform well diagnostically.

Table 2.1 Characteristics of included studies

	Author	Sample size	Comparison groups	Name of protein(s) evaluated	Sensitivity %	Specificity %	AUC 95% CI
Antigens	(Ajona et al., 2021)	78	NSCLC/indeterminate nodules	C4c, CYFRA 21-1, and CRP	82	95	0.9
	(Bigbee et al., 2012)	56/30	NSCLC/indeterminate nodule	Prolactin, transthyretin, thrombospondin-1, E-selectin, C-C motif chemokine 5, macrophage migration inhibitory factor, plasminogen activator inhibitor, receptor tyrosine-protein kinase, erbb-2, cytokeratin fragment 21.1 and serum amyloid A	77.10	76.2	NG
	(Fahrmann et al., 2022)	1,299	High risk	A four-marker protein panel (4MP) consisting of CA125, CEA, SPA, CYFRA21-1	91.5	45.4	0.79
	(Farlow et al., 2010a)	90/43	NSCLC/non-cancer	TNF-α, CYFRA 21-1, IL-1ra , MMP-2, MCP-1 and sE-selectin	99	95	0.979
	(Gasparri et al., 2023)	46/41	NSCLC/high-risk	ARSA, PRKCA, ACTR3B, and CD59	94.83	93.56	0.8
	(Goebel et al., 2019)	1,479	NSCLC/HC	CA-125, CEA, CYFRA21-1, EGFR/HER1/ErBB1, Gro-Pan, HGF, IL-10, IL-12p70, IL-16, IL-2,IL-4, IL-5, IL-7, IL-8, IL-9, Leptin, LIF, MCP-1, MIF, MIG, MMP7,MP9, MPO, NSE, PDGF-BB, Rantes, Resistin, sFasL, SAA,sCD40-ligand, sICAM-1, TNFRI and sTNFRII.	80	95	0.96
	(Higgins et al., 2012)	35/170/160	LC/inflammatory diseases	Ciz1	95	74	0.96
	(Jeong et al., 2021)	70/16	NSCLC/HC	Exosomal GCC2	90	75	0.84
	(Joseph et al., 2012)	1,182	NSCLC/nodules	OPNV	80	88	0.88
	(Jung et al., 2017)	200/150	LC/control group	EGFR1, MMP7, CA6, KIT, CRP, C9 and SERPINA3	75	91.70	0.82/0.77
	(Kupert et al., 2011b)	145	NSCLC/BN/HC	secretory phospholipase A2-IIa	48-67	86	0.68-0.86
	(Lai et al., 2022)	201/112/94	NSCLC/HC/Nodules	CEA, CYFRA21-1, CST1	88.4	89.1	0.92
	(Li et al., 2023a)	37/11	NSCLC/HC	MDK, WFDC2, and CXCL14	NG	NG	0.96
	(Li et al., 2022)	98/100	NSCLC/BLD	CA153 + CA125 + CEA + TNF − alpha + hs – CRP	66.82	93.51	NG
	(Ma et al., 2021)	318/239	NSCLC stage I/HC	MDH2	70.13	66.11	0.77
		769/493			68.92	58.22	0.72
	(Meng et al., 2023)	60/15	NSCLC/HC	EpCAM and CEA	93.3	86.7	0.92
	(Nolen et al., 2011)	172	LC/high risk	MIF, TTR, THSP, sVCAM-1 and tPAI-1	70/74	90/93	0.85/0.89

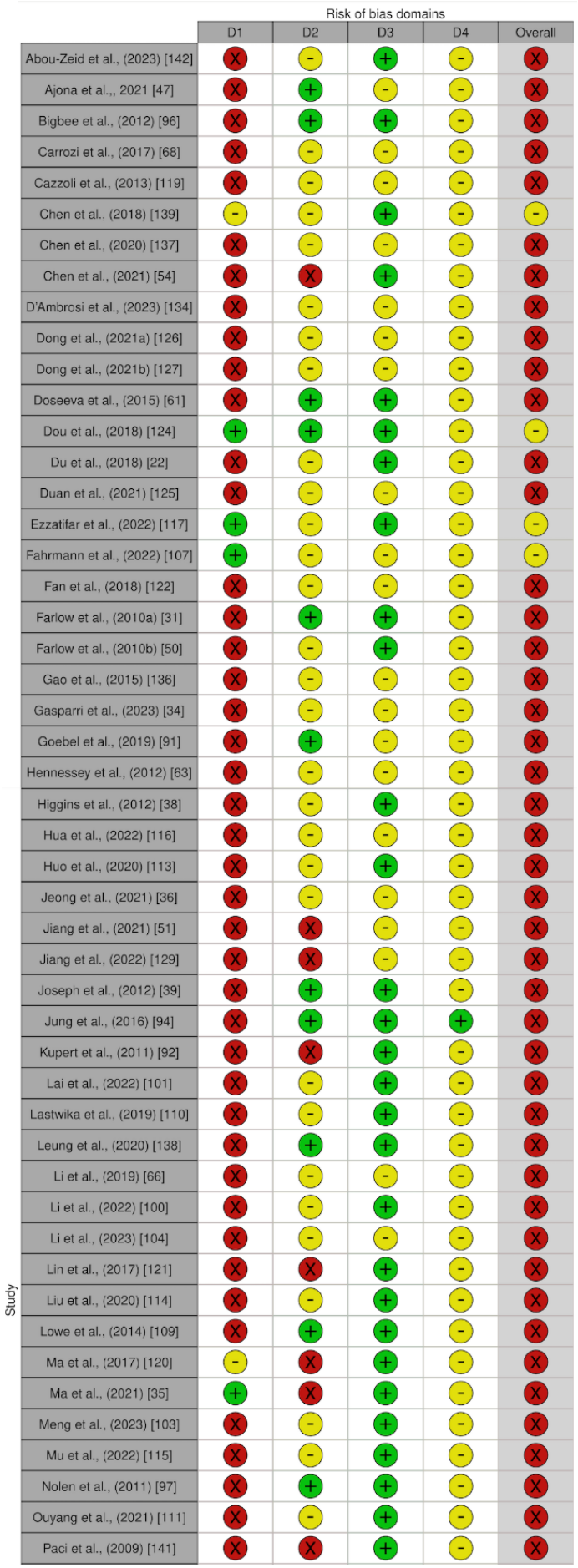
	(Pakvisal et al., 2022)	76/12/53	NSCLC/BLD/HC	C5AR1, CLEC4A and NLRP3 specific to CD3	71.5	70	
	(Sun et al., 2020a)	1223	NSCLC/BPC/OC/ HC	IDH1	63.3/55	86.8/86.3	0.91/0.79
	(Song et al., 2022)	30/15	NSCLC/HC	p53-anti-p53-autoAb complex	81.6	93.3	NG
	(Wang et al., 2017)	350/411	NSCLC/control (BLD, HC)	MIC-1, CYFRA21-1, CA125 and CEA	84.40	90	0.96
	(Wang et al., 2013)	132/48/92	LC/BLD/HC	NSE + CEA + CYFRA21-1	75.76	89	0.63
	(Wieskopf et al., 1995)	161/97	LC/BD	CYFRA 21-1	59	94	0.85
	(Wu et al., 2020a)	102/84	ADC/HC	Beta-1,4- galactosyltransferase 1, CD44, eukaryotic initiation factor 4A-I, galectin-1, mucin-16, protein disulfide-isomerase A3 and vimentin	97.2	61	0.76
	(Yang et al., 2020b)	370/110	NSCLC stage I/BLD	Ferritin, CA125, CEA, NSE and CYFRA21-1	92.97	90	0.95
	(Yu et al., 2023)	513	Nodules	ACSL4	65.1	90.2	0.76
	(Yuan et al., 2022)	175/160	LC/HC	HSP90α, CEA	95.63	99.97	0.996
	(Zhang et al., 2022b)	78/44	NSCLC/BLD	CEA, CYFRA21-1, miR3149 and miR-4769.3p	88.46	81.82	0.90
Autoantibodies	(Chen et al., 2021a)	458	NSCLC/ nodules/HC	MAGE A1, PGP9.5, SOX2, and TP53	71.8	89	0.89
	(Doseeva et al., 2015)	230/150	NSCLC/BLD	One autoAb marker (NY-ESO-1) and 3 Ags (CEA, CA-125, and CYFRA 21–1)	74/77	80/80	0.81/0.85
	(Du et al., 2018)	397	LC/nodules	Seven TAAs (p53, PGP9.5, SOX2, GAGE7, GBU4-5, CAGE and MAGE A1)	56.53	91.60	NG
	(Ezzatifar et al., 2022)	190/30	NSCLC/HC	Nucleolin	85	96.67	0.948
	(Farlow et al., 2010b)	16/196	NSCLC/COPD/non malignant nodules/NC	IMPDH, phosphoglycerate mutase, ubiquillin, Annexin I, Annexin II, and HSP70-9B	94.8	91.1	0.964
	(Hua et al., 2022)	83/26	NSCLC/BLD	7-TAAbs (P53, PGP9.5, SOX2, GAGE7, GBU4–5, MAGE A1 and CAGE)	55.44	87.5	0.65
	(Huo et al., 2020)	121/34/100	NSCLC/HC/nodules	7AAb (GAGE7, CAGE, MAGE A1, SOX2, GBU4-5, PGP9.5, and p53)	45.5	85.3	0.66
	(Lastwika et al., 2019)	20/10/250	LC/nodules	IgG: EPB41L3, ANKRD36B, FGCR2A, and LINGO1; IgM: S100A7L2	50	70	0.74/0.78
	(Liu et al., 2020)	211/200	NSCLC/HC	CD25-MUC1-VEGFR1	49.6	95	0.883
	(Lowe et al., 2014)	600	AAH & SCD	AAH: LTBP1*, BMI1*, GAGE7*, AGBL5 HES1*	86	78	0.81/0.88
	(Jiang et al., 2021)	150	LC /HC/BLD	7TAAb (TP53, NPM1, FGFR2, PIK3CA, GNA11, HIST1H3B, and TSC1)	94.4	84.9	0.897
		744			89.4	78.2	0.838
	(Mu et al., 2022)	633/147	NSCLC/BLD	7-TAAbs+SCCA+CYFRA21-1	37.76	81.84	0.648
	(Ouyang et al., 2021)	443	NSCLC /HC/BLD	7 TAAB, CEA, CYFRA 21-1	52.26	77.46	0.686

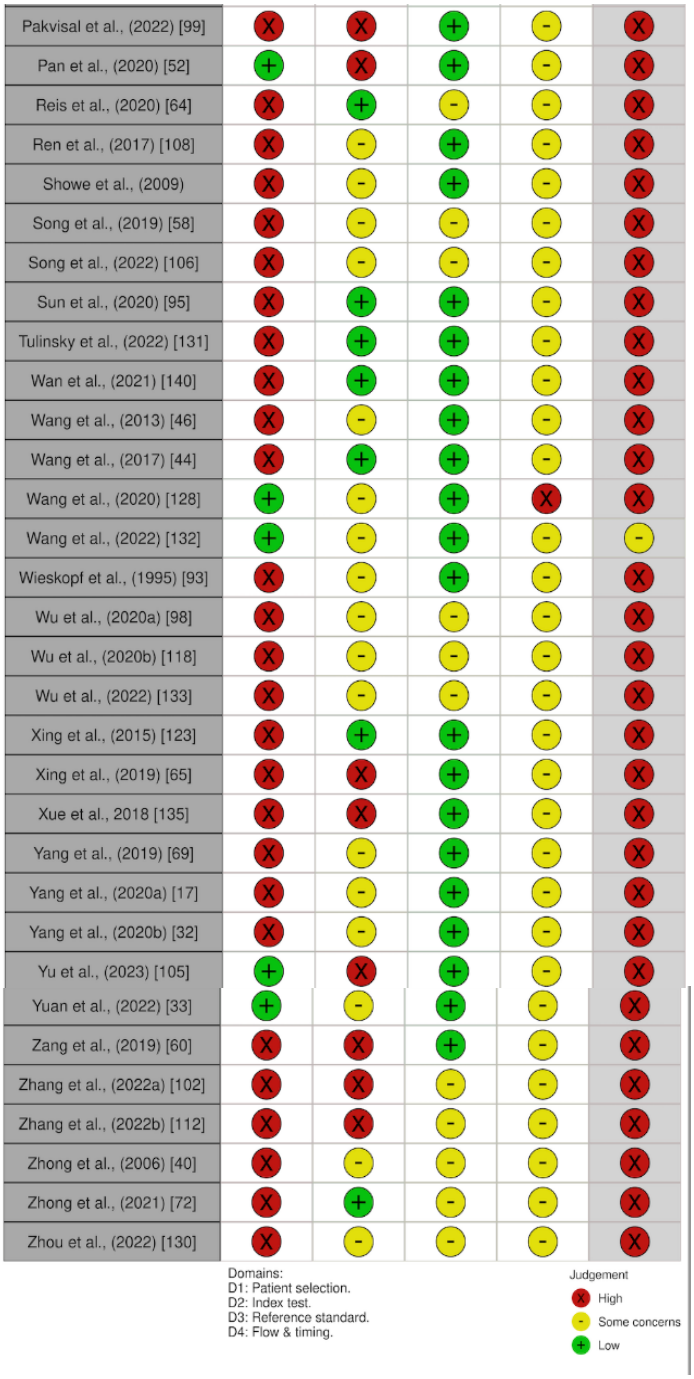
		569			44.02	83	0.668
	(Pan et al., 2020)	69/30/25	NSCLC stage I/HC/BLD	6 autoantigens BCL7A, TRIM33, MTERF4, CTAG1A, DDX4 and MAGEC2	73.5	>85	0.503
		88/36/18			68.2	87	0.673
	(Ren et al., 2018)	2008	LC/patients (GGNs) and/or solid nodules	p53, GAGE7, PGP9.5, CAGE, MAGE A1, SOX2 and GBU4-5	59/62	90	0.781
	(Song et al., 2019)	170	NSCLC/HC	CYFRA 21-1–anti-CYFRA 21-1 autoAb immune complex (CIC) and free CYFRA 21-1	76	87.5	NG
	(Yang et al., 2020a)	(42) 61/24/29	LC/BLD/HC	HE4	54.76	96.23	0.848
	(Zang et al., 2019)	176/140	LC/HC	CEA, CA125, Annexin A1-Ab, and Alpha enolase-Ab	86.5	82.3	0.897
	(Zhang et al., 2022a)	68/68	ADC/HC	CEA, 5 IgM AAB (TSHR, ERBB2, survivin, PIK3CA, and JAK2)	56.63	93.98	0.744
	(Zhong et al., 2006)	46	Stage I NSCLC & risk-matched control	PXN, SEC15L2, BAC clone RP11-499F19, XRCC5, and MALAT1	100	95.7	0.99
		102			91.3	91.3	
miRNA and RNA	(Cazzoli et al., 2013)	30/105	LC/BD/HC	miR-151a-5p, miR-30a-3p, miR-200b-5p, miR-629, miR-100, and miR-154-3p	97.5/96	72/60	0.76
	(D'Ambrosi et al., 2023)	30/27/3	NSCLC/HC/Nodules	2 circRNAs (circSLC8A1 and circCHD9) and 3 mRNAs (PSMB9, RUNX1, and LILRB1)	85	86	0.96
	(Dong et al., 2021a)	300	NSCLC/HC	CEA, miR-1247-5p, miR-301b-3p, and miR-105-5p	88.4	64.7	0.815
	(Dong et al., 2021b)	290/105	NSCLC/HC	CEA, TEP SNORD55	66.3	90	0.828
	(Dou et al., 2018)	50/35/29 44/32/51	ADC I,II/BLD/HC	hsa-miR-101-3p/hsa-miR-126-5p	81.1/70.4	78.1/72.7	0.82/0.742
	(Duan et al., 2021)	12/120	NSCLC/HC	miR-492, miR-590-3p, and miR-631	86.7	71.7	0.828
	(Fan et al., 2018)	128/193	NSCLC/BPD	Five miRNA ratios: miR-15b-5p/miR-146b-3p, miR-20a-5p/miR-146b-3p, miR-19a-3p/miR-146b-3p, miR-92a-3p/miR-146b-3p, and miR-16-5p/miR-146b-3p	70	90	0.79
	(Hennessey et al., 2012)	50/130	NSCLC/HC	miR-15b and miR-27b	100	84	0.98
	(Jiang et al., 2022b)	35/15	NSCLC/HC	miR-152-3p and miR-1277-5p	73.3	86.7	0.79
	(Li et al., 2019a)	(33) 64/40	NSCLC /HC	CEA+Exo-GAS5	89.06	90.00	0.919
	(Lin et al., 2017)	135 126	Indeterminate nodules	miR-21–5p miR-103a-3p miR-126–3p miR-135a-5p miR-145–5p miR-141–3p miR-193b-3p miR-200b-3p miR-205–5p	89.90 73.5	90.90 75.5	0.91
	(Ma et al., 2017)	1272 111	Indeterminate nodules	miRs-19b-3p and -29b-3p	80.30 72.6	89.40 81.9	0.91

	(Reis et al., 2020)	54/40	Early NSCLC/HC	miR-16-5p, miR-92a-3p, miR-451a	84	100	0.87
	(Tulinsky et al., 2022)	60/60	NSCLC/HC	miR-126, miR-143, miR-145, let-7a and let7g	75-85	75-85	0.90-0.93
	(Wang et al., 2022a)	165/118	NSCLC/HC	SNORD42B and SNORD111	61.8	77.1	0.719
	(Wang et al., 2020b)	82	Pulmonary nodules - benign & malignant	miRNA-17, miRNA-146a, miRNA-200b, miRNA-182, miRNA-155, miRNA-221, miRNA-205, miRNA-126, miRNA-7, miRNA-21, miRNA-145, and miRNA-210	50	92.9	0.896
	(Wu et al., 2022)	100/100	NSCLC/HC	miR-340 and miR-450b-5p	78.33	77.5	0.862
	(Wu et al., 2020b)	48/48/32	NSCLC I/II/HC/ LBL	Four serum miRNAs including miR-21-5p, miR-141-3p, miR222-3p, and miR-486-5p, and 2 serum exosomal miR-146a-5p and miR-486-5p	85.42	92.50	0.96
	(Xing et al., 2015)	122/136/155	Indeterminate solid nodules	miR205/miR708/ miR375/miR200b/ miR182/ miR155/ miR372 miR143 (miRs21, 31, and 210)	82.93/ 82.09/ 80.52	87.84/88.41/86.08	0.919
	(Xing et al., 2019)	17/534	NSCLC/control (BN/HC)	ITGA2B	92.8/91.2	78.6/56	0.892
	(Zhou et al., 2022)	15	ADC	SNORD60	74.2	75.3	0.828
DNA and CTC	(Abou-Zeid et al., 2023)	25/25	NSCLC/HC	HOXA9, SOX2, HV2	88	100	0.958
	(Carozzi et al., 2017)	1356	LC/smokers/ex-smokers	MSI/LOHs loci, with the loci 1 to 5 (3p14.2, 3p21-p23, 3p26.1, 3p13, 5q15) and 7 to 9 (9p22-p23, 9p21, 13q12.3)	90	71	NG
	(Chen et al., 2020a)	161	Nodules	CDO1, SOX17 and HOXA7	90	71	NG
	(Chen et al., 2018b)	41/10	NSCLC/HC	EpCAM and Folate receptor alpha (FRα)	75.61	90	NG
	(Gao et al., 2015b)	89	Nodules	APC, RASSF1A	56.9	90.3	0.81
	(Leung et al., 2020)	211	NSCLC/HR	ctDNA (EGFR, KRAS, and TP53 mutation)	75	89	NG
	(Paci et al., 2009)	151/79	NSCLC/HC	Amplification of hTERT	85.8	46.8	0.79
	(Wan et al., 2021)	48	NSCLC	NOTCH1, IGF2, EGFR and PTCH1	65.85	62.5	NG
	(Xue et al., 2018)	(31) 72/26	NSCLC/control	FR+CTC	74.19	73.08	0.8221
	(Yang et al., 2019b)	50	Nodules	Methylation of 8 genes (CDH13, WT1, CDKN2A, HOXA9, PITX2, CALCA, RASSF1A, and DLEC1)	72	91	NG
	(Zhong et al., 2021)	18	Solid nodules	(CEP8) CTC, CA125 or NSE	83	100/83 with NSE	NG

AAH: Atypical adenomatous hyperplasia, ADC: adenocarcinoma of the lung, ARSA: Arysulfatase A, AUC: area under curve, B: blood, BN: benign nodules, BPC: benign pulmonary condition, BLD: benign lung diseases, Bmi-1: B-lymphoma Moloney murine leukaemia virus insertion region-1, C4c : complement-derived fragment, CI: confidence interval, CRP: C-reactive protein, CTC: circulating tumour cells, Ciz1: nuclear matrix-associated DNA replication factor, CXCL14: C- X-C motif chemokine ligand 14, FR: folate receptor, FOXL2: fork-head box L2 gene, HC: healthy control, HE4: Human epididymis secretory protein 4, HES1: mammalian hairy and Enhancer-of-split homologues 1, IDH1: isocitrate dehydrogenase 1, ILra: interleukin-1ra, LTBP1: Latent Growth Factor Beta Binding Protein, MCP1: monocyte chemotactic protein-1, MDH2: malate dehydrogenase 2, MDK: Midkine, MIC-1: Macrophage inhibitory cytokine-1, MIF: macrophage migration inhibitory factor, MMP2 : matrix metalloproteinase-2, MSI/LOH: genomic instability loss of heterozygosity/ microsatellite instability, ncRNA: non-coding RNA, NG: not given, OC: other cancers, OPV: OPN velocity, P: plasma, PTGER4: prostaglandin E receptor 4 gene, QMSP: real-time quantitative methylation-specific polymerase chain reaction, SCD: squamous cell dysplasia, S: serum, SHOX2: methylation of short stature homeobox 2 gene, sVCAM-1: soluble vascular cell adhesion molecule, T: tissue, TAAB: tumour associated autoantibodies, TC: Training cohort, TB: tuberculosis, THSP: thrombospondin, TNF-α: tumour necrosis factor α, tPAI-1: tissue plasminogen activator inhibitor 1, TTR: transthyretin, VC: Validation cohort, WFDC2: WAP four-disulphide core domain 2. The number between the parenthesis represents NSCLC studies that include stages I and II.

(A)


Risk of bias domains
D1
D2
D3
D4
Overall
Study
Abou-Zeid et al., (2023) [142]
Ajona et al.,, 2021 [47]
Bigbee et al., (2012) [96]
Carrozi et al., (2017) [68]
Cazzoli et al., (2013) [119]
Chen et al., (2018) [139]
Chen et al., (2020) [137]
Chen et al., (2021) [54]
D'Ambrosi et al., (2023) [134]
Dong et al., (2021a) [126]
Dong et al., (2021b) [127]
Doseeva et al., (2015) [61]
Dou et al., (2018) [124]
Du et al., (2018) [22]
Duan et al., (2021) [125]
Ezzatifar et al., (2022) [117]
Fahrmann et al., (2022) [107]
Fan et al., (2018) [122]
Farlow et al., (2010a) [31]
Farlow et al., (2010b) [50]
Gao et al., (2015) [136]
Gasparri et al., (2023) [34]
Goebel et al., (2019) [91]
Hennessey et al., (2012) [63]
Higgins et al., (2012) [38]
Hua et al., (2022) [116]
Huo et al., (2020) [113]
Jeong et al., (2021) [36]
Jiang et al., (2021) [51]
Jiang et al., (2022) [129]
Joseph et al., (2012) [39]
Jung et al., (2016) [94]
Kupert et al., (2011) [92]
Lai et al., (2022) [101]
Lastwika et al., (2019) [110]
Leung et al., (2020) [138]
Li et al., (2019) [66]
Li et al., (2022) [100]
Li et al., (2023) [104]
Lin et al., (2017) [121]
Liu et al., (2020) [114]
Lowe et al., (2014) [109]
Ma et al., (2017) [120]
Ma et al., (2021) [35]
Meng et al., (2023) [103]
Mu et al., (2022) [115]
Nolen et al., (2011) [97]
Ouyang et al., (2021) [111]
Paci et al., (2009) [141]



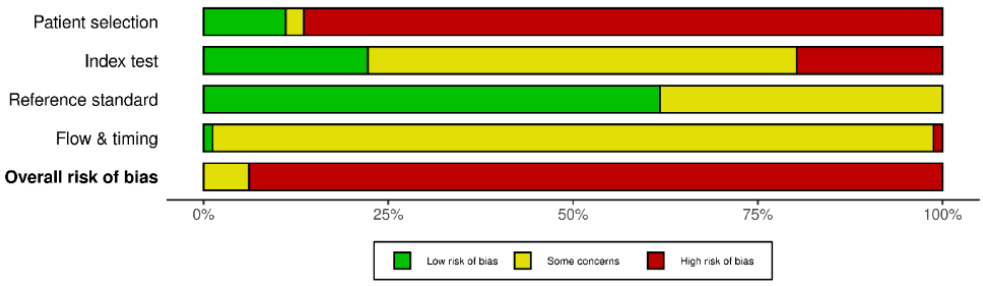


Figure 2.2 Risk of bias assessment of the selected studies by QUADAS- 2

(A) Risk of bias rating for each study and **(B)** for each domain across all studies. This assesses the risk of bias of a study across four domains: participant selection, index test, reference standard, and flow and timing, as well as an overall assessment of risk of bias (detailed in Supplementary Table 3). Patient selection has high risk of bias in most studies.

Table 2.2 Studies with sensitivity and specificity > 90%

TC: Training cohort; VC: Validation cohort

Group	Study	Sensitivity TC/VC	Specificity TC/VC	AUC	Biomarker panel
Antigens	(Farlow et al., 2010a)	99	95	0.979	TNF-α, CYFRA 21-1, interleukin-1ra, MMP-2, monocyte chemotactic protein-1 and sE-selectin
	(Yang et al., 2020b)	92.97	90	0.95	Ferritin, CA125, CEA, NSE and CYFRA21-1
	(Yuan et al., 2022)	95.63	99.97	0.996	HSP90α and CEA
	(Gasparri et al., 2023)	94.83	93.56	0.98	ARSA, PRKCA, ACTR3B and CD59
Autoantibodies	(Zhong et al., 2006)	100/ 91.3	95.7/ 91.3	0.99	Paxillin, SEC15L2, BAC clone RP11-499F19, XRCC5 and MALAT1
	(Farlow et al., 2010b)	94.8	91.1	0. 964	IMPDH, phosphoglycerate mutase, ubiquillin, Annexin I, Annexin II and HSP70-9B

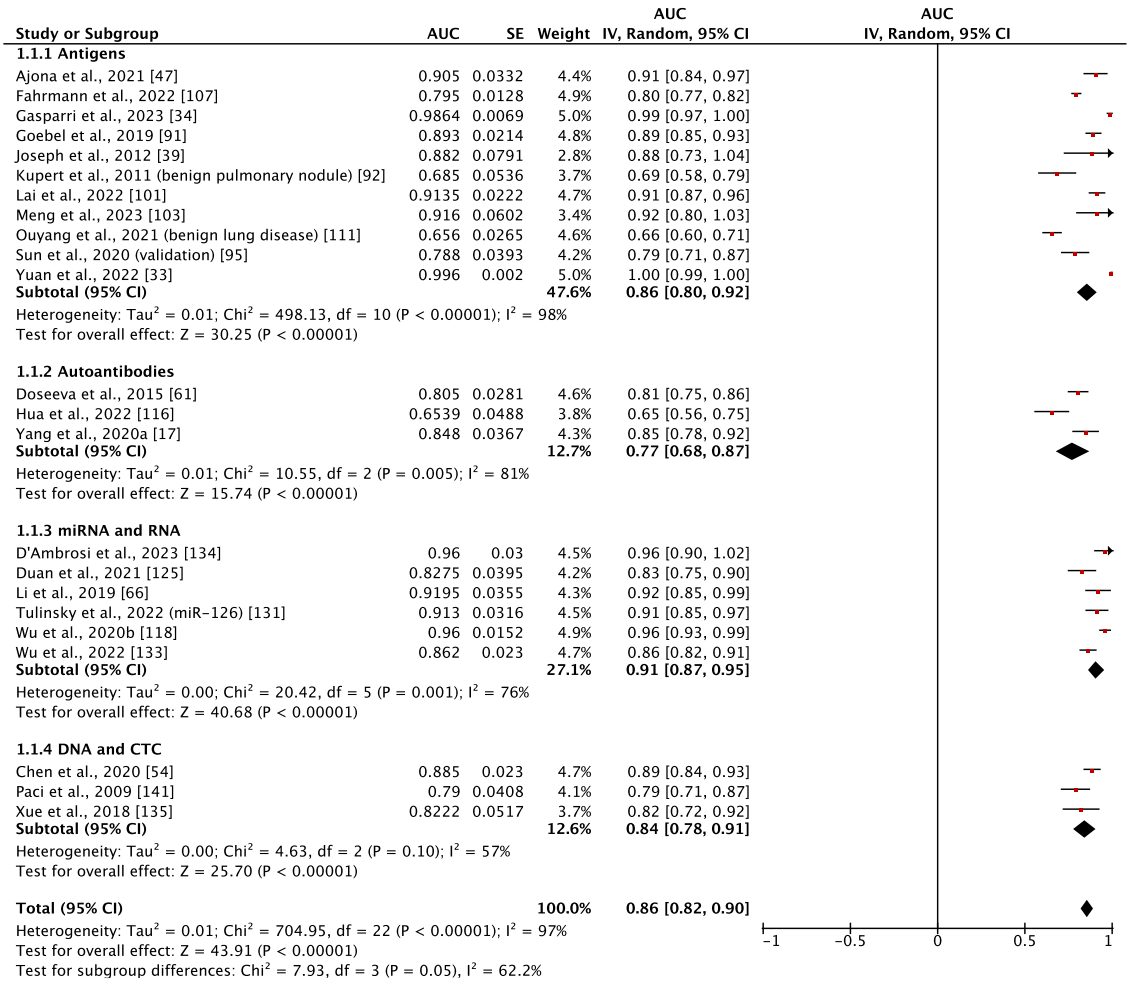


Figure 2.3 Forest plot for meta-analysis of AUC and 95% CI, and subgroup analysis based on type of biomarker

The y axis shows the name of each study and the x axis (the horizontal axis - the point estimation) shows the study's (based on the type of study and the reported scale: -1 to +1), 1 means we are observing a positive relationship of diagnostic performance based of a given biomarker type (listed in bold), SE: standard error, red dot which shows the weight of the study in the analysis (most importantly the sample size of the study which means the study with bigger population has a bigger impact on the overall analysis). Also, each rectangle and a line crossed through, which shows the 95%CI for the used scale. Black diamond shows heterogeneity of studies among the groups (Chi2, the value of Chi-square test for heterogeneity). Big diamonds mark high heterogeneity between studies.

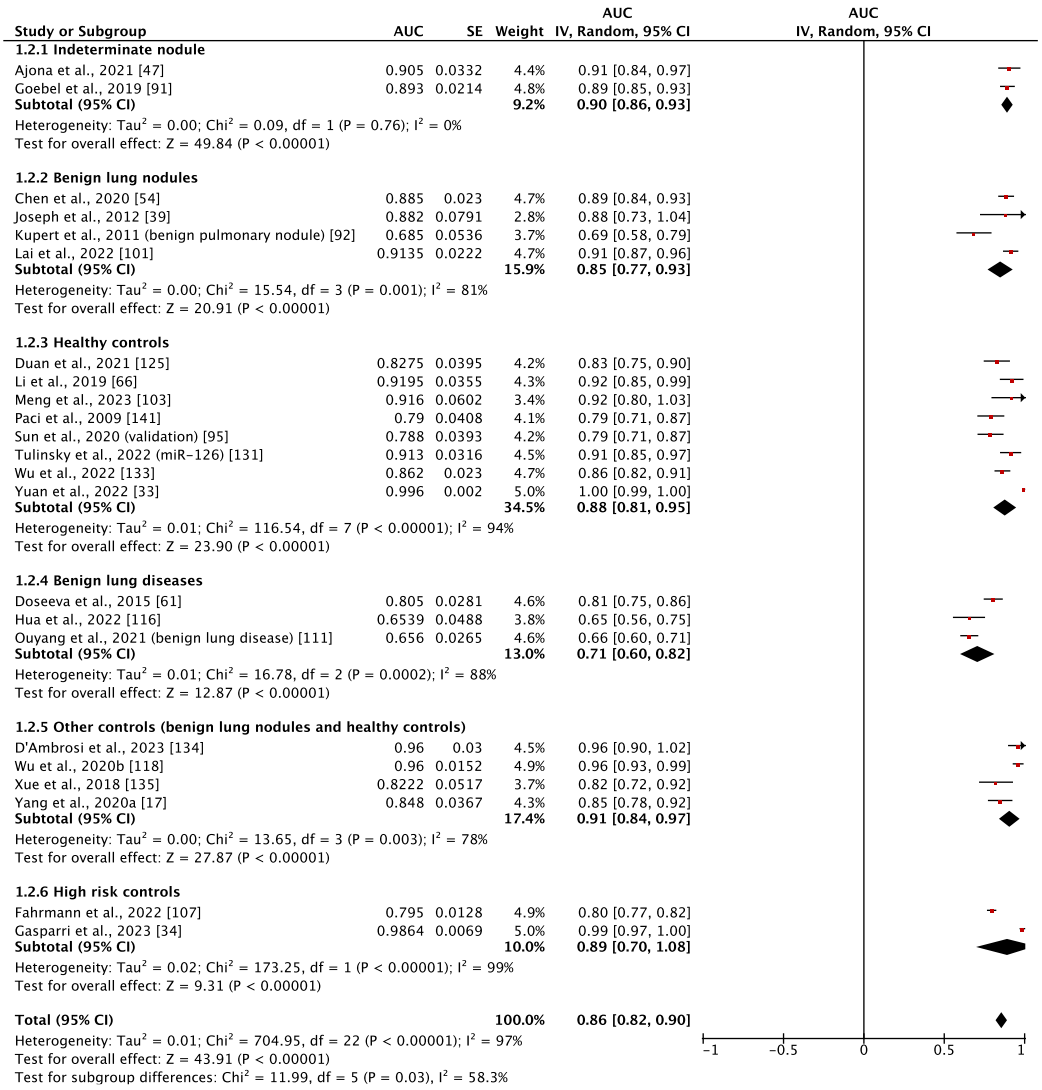


Figure 2.4 Forest plot for meta-analysis of AUC and 95% CI, and subgroup analysis based on type of control

The study or subgroup shows the name of each study and the x-axis (the horizontal axis - the point estimation) shows the study's (based on the type of study and the reported scale: -1 to +1), 1 means we are observing a positive relationship (AUC) between the biomarker efficiency for LC relative to each control type (listed on bold). The biomarkers performed the least accurately in differentiating early NSCLC from benign lung diseases (pooled AUC = 0.74, 95% CI 0.67-0.81). There was also no significant subgroup difference based on the source of biomarker (I2 = 0%, p = 0.95). Red dot which shows the weight of the study in the analysis (most importantly the sample size of the study which means the study with bigger population has a bigger impact on the overall analysis). Also, each rectangle and a line crossed through, which shows the 95%CI for the used scale. Black diamond shows heterogeneity of studies among the groups (Chi2, the value of Chi-square test for heterogeneity). Big diamonds mark high heterogeneity between studies.

2.3 Discussion

The included studies assessed a wide range and combination of biomarkers, which commonly were not validated in multi-centre studies hence we were unable to make firm conclusions on their diagnostic accuracy, nor conduct a meta-analysis for each biomarker. Our review found that well-performing single biomarkers for early NSCLC diagnosis include Ciz1 (sensitivity: 95%) and exosomal GCC2 (sensitivity: 90%) with a slightly lower specificity of 71% for CIZ1 and 75% for exosomal GCC2 respectively. Tumour-educated blood platelets (ITGA2B) also had high sensitivities in both the training (92.8%) and validation cohort (91.2%) but low specificity. In contrast, CYFRA 21-1 and anti-HE4 had high specificity for LC (95% each). OPNV had a sensitivity of 80% and a specificity of 88% as a biomarker for early LC. Biomarker panels (Table 2.2) had high sensitivity and specificity (greater than 90%). Combined biomarkers are more likely to help early detection of NSCLC, especially when antigens or autoantibodies are combined with miRNAs. Early detection is the holy grail of NSCLC diagnoses as it offers the opportunity to significantly increase survival rates, aid in the management of the disease, and reduce overall healthcare costs. CYFRA21-1 is a prognostic biomarker for advanced NSCLC as it is predominately found in lung tissues and correlates with tumour size, lymph node involvement and the stage of the disease (Edelman et al., 2012). Lower baseline levels of CYFRA21-1 were associated with both longer overall survival and failure free survival ($p<0.0001$ and $p=0.0003$) (Edelman et al., 2012). Wang *et al.* (Wang et al., 2018b) found that serological levels CYFRA21-1 combined with other markers have different sensitivities and specificities, depending on the sample size. A commonly used combination of CYFRA21-1, CEA and NSE for LC detection revealed a very low sensitivity of 31% in contrast to its very high specificity of 96%. These assays have the drawback of low sensitivity especially in the early stages of LC, whilst the same panel of biomarkers have shown high sensitivity at advanced stages (Wang et al., 2018b). Therefore, this combination could not be recommended for use in early detection in clinical practice. However, the high concentration of these biomarkers in body fluids/levels in tissues are poor prognostic indicators. Thus, these three biomarkers could be used to predict relapse before the onset of clinical symptoms as their concentration can be used to monitor therapy response/resistance.

Blood represents one of the most suitable mediums for the analysis of biomarkers for NSCLC detection. It has been suggested that cancer elicits a complex immune response (Nardi-Agmon & Peled, 2017) that can be detected in the peripheral blood, which is less invasive than a bronchoscopy. Thus, changes in the immune response could be detected in blood before clinical symptoms appear. Different molecules such as proteins and miRNA have been shown to be very sensitive biomarkers, which is cost effective and safer compared to imaging techniques, such as CT scans associated with an increased risk of cancer due to radiation, requiring a lot of specialist training and expensive, compared to blood tests. Blood is one of the most analysed tissues for disease diagnosis especially when compared to other body fluids such as urine and sputum (Nardi-Agmon & Peled, 2017). Blood biomarkers are more stable and reliable for LC detection than breath (Nardi-Agmon & Peled, 2017). For example, miRNAs in blood have a higher sensitivity for LC diagnosis compared to miRNAs in sputum as the oral cavity contains many enzymes such as those that degrade these small molecules (Kammer & Massion, 2020). However, miRNAs lack specificity for early LC diagnosis as they are expressed in many cancers and healthy tissues. Due to their low specificity and reduced stability, miRNAs are not suggested for clinical use (Kammer & Massion, 2020). In contrast, protein antigens are frequently used as markers for disease diagnosis with the aid of imaging techniques (Thomas et al., 2008; Thomas et al., 2009).

Biomarkers can be diagnostic, predictive for a treatment response, or indicative of a disease prognosis (Voon & Kong, 2011; Nalejska et al., 2014). Biomarker discovery is largely dependent on an analytic validation for measuring biomarkers in body fluids. Blood is mostly used to detect molecular changes associated with LC after depleting the abundant proteins leaving the biomarkers of interest that are usually present in very low concentrations. The stability of biomarkers is a crucial factor, affecting the reproducibility and analytic validation procedures (Schatzkin & Gail, 2002; Fleming, 2005). In addition to study design, population selection should also be considered. Sample size should be statistically valid as a low number of participants exaggerates diagnostic performance of biomarkers. The required sample sizes should be calculated to achieve 95% confidence levels and 80% power for purpose of testing the validity of the biomarker (Hajian-Tilaki, 2014). Moreover, age and gender match of controls should

be considered. For example, postmenopausal women have less oestrogen than childbearing females and this difference may leads to false positive results (Ransohoff & Gourlay, 2010b). Ideally, biomarkers must be highly sensitive and specific for cancer diagnosis. However, there is no marker in clinical practice that possesses both 100% sensitivity and specificity. The use of biomarkers have been proposed in addition to imaging techniques, which would have greater benefit-to-risk ratio compared to using either markers or imaging alone (Koscielny, 2010). An example of the successful use of biomarkers in clinical practice is human epidermal growth factor (HER2) to detect HER2 positive breast cancer. Patients with this biomarker receive targeted therapy, which has improved the overall survival and reduced the side effects of multiple treatments in cancer patients (Piccart-Gebhart et al., 2005).

Biomarker research should be optimised by developing a common workflow. Identifying the optimal cut-off point of biomarkers is required for their application in clinical settings. Most biomarkers in this review utilised retrospective designs and samples from tissue banks. Ideal biomarker studies should have a prospective design such as randomised controlled trials, with a large sample size ensuring that the study is able to achieve adequate precision following the Standards for Reporting of Diagnostic Accuracy Studies (STARD) guidelines and examine populations with disease and compare them to age- and sex-matched controls (Cohen et al., 2016). This would reduce false positives associated with CT results and thus overtreatment and side effects from unnecessary interventions (Mazzone et al., 2017). Although a change in biomarker expression may not reflect true clinical benefit, this may associate with pathway modulation (Dunn et al., 2010). Biomarker translation into clinical practice is a challenging mission and even with approved markers such as CEA for colon cancer diagnosis, it's sensitivity is still not ideal as it is expressed in other cancers and in non-malignant conditions (Ransohoff & Gourlay, 2010b).

Sensitivities and specificities are dependent on the biomarker selected and the LC types studied. Biomarker assays also require both robustness and reproducibility to be applied for clinical use (Pass et al., 2013). Studies with validation cohorts are more robust than studies with only a testing group (Mehan et al., 2014). For example, Xing *et al.* (Xing et al., 2019) showed that variations in results were due to a difference in the number of

participants and controls with a range of non-malignant conditions being used to determine the specificity of the biomarkers. Goebel *et al.* (Goebel et al., 2019) examined 21 candidate biomarkers including antigens and cytokines using a multiplex immunoassay but many were excluded even with >80% sensitivity and >95% specificity, as the assay lacked reproducibility and was difficult to perform using such a large number of biomarkers. Developing an optimal multiplex test is required to validate the findings of this study and to examine its functionality and clinical use (Goebel et al., 2019).

This systematic review has several limitations. We only included articles in English and some quantitative studies could not be included as they did not adequately report the diagnostic performance of the biomarkers investigated e.g. sensitivity and specificity, which in this study was considered crucial information for the evaluation of a diagnostic biomarker/biomarker panel. There was also considerable variability across studies in terms of timing, participants and control groups, sampling, and biomarker detection methods. Included studies assessed a combination of biomarkers, which commonly were not validated in multi-centre studies, hence we were unable to make firm conclusions on their diagnostic accuracy, nor conduct a meta-analysis for each biomarker. Future studies should report their findings following the STARD guidelines for the construction of 2-by-2 tables for diagnostic meta-analysis, and minimally by including the 95% CI of diagnostic effect measures (Cohen et al., 2016). NSCLC biomarker diagnosis should emphasise the validation of biomarkers so that they can be translated into clinical use and impact patient treatment and care.

Chapter 3: Materials and methods

3.1 Bioinformatics on NSCLC

GSE81089 was downloaded from the Gene Expression Omnibus (GEO) databases and reanalysed to identify differentially expressed genes (DEGs) with aqp-value <0.05 and an average log fold change (avg log_2FC)>1 for upregulated genes and (avg log_2FC)<-1 for downregulated between NSCLC compared to normal tissues that may act as biomarkers for the early NSCLC. The Bigomics software depends on DESeq2 package, which is part of the 'Bioconductor' package (Love et al., 2014), was used to identify the differentially expressed genes (DEGs). Gene set enrichment analysis (GSEA) was performed by Enrichr. Fresh frozen tumour tissue from 199 individuals diagnosed with NSCLC and surgically treated at Uppsala University Hospital in Uppsala, Sweden, from 2006 to 2010, as well as 19 associated normal lung tissues, were studied (Mezheyeuski et al., 2018). Clinical information was obtained from a regional LC registry. Sample characteristics values represent (Figure 3.1).

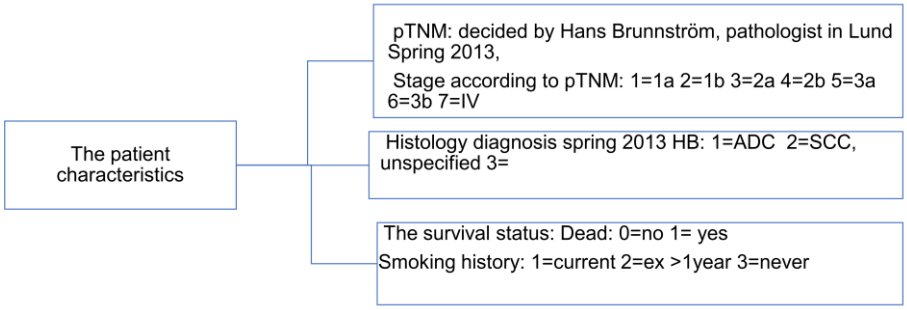


Figure 3.1 Patient characteristics

Tumour staging pTMN and histology as determined by Hans Brunnström, a pathologist in Lund Spring in 2013. Staging pTMN indicates by numbers patient Ia=70, Ib=45, IIa=25, IIb=23, IIIa=3, IV=3 and gene expression correlated stages (analysis performed figure 4.4). Hist_1 is ADC and number of patients=108, Hist_2 is SCC and patient number=67, Hist_3 is non-specific histology and patient number=24 and gene expression correlated to histology (analysis performed figure 4.5).

RNA was isolated from frozen tissue samples that contained more than 10% cancer cells. The samples were prepared using the Illumina TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, CA, USA) with polyA selection. The Uppsala University Hospital Review Board approved the trial, and all patients provided informed permission in accordance with the Helsinki Declaration. The sequencing was done multiplexed with five samples per lane on Illumina HiSeq2500 machines (Illumina) using the standard Illumina RNAseq methodology with a read length of 2100 bases. Bigomics was used (https://bigomics.ch/) for the analysis (Akhmedov et al., 2020). Data downloaded from GSE81089 and analysed for Chapter 4. Normalisation was performed using the limma package, DESeq2 (Wald, LRT) package was used to identify DEGs when comparing levels between tumour and healthy tissues, the p and q values were calculated using the Fisher's test (Akhmedov et al., 2020).

3.2 Samples from patients with incidental lung nodules and LC

Ethical approval was received from the Faculty of Health Sciences Ethics Committee for the study 'Identification of novel biomarkers for the early detection of LC' (FHS283), sponsorship approval has been secured and the application was approved by the NREC, IRAS: 273139. Ethical approval was sought for the collection of sera from patients with LC and non-malignant lung diseases. Tissue samples are being collected from NSCLC patients who are undergoing invasive procedures such as tissue biopsies **(Figure 3.2).** Samples were also collected from healthy volunteers to be used as controls to examine TA expression in physiological conditions.

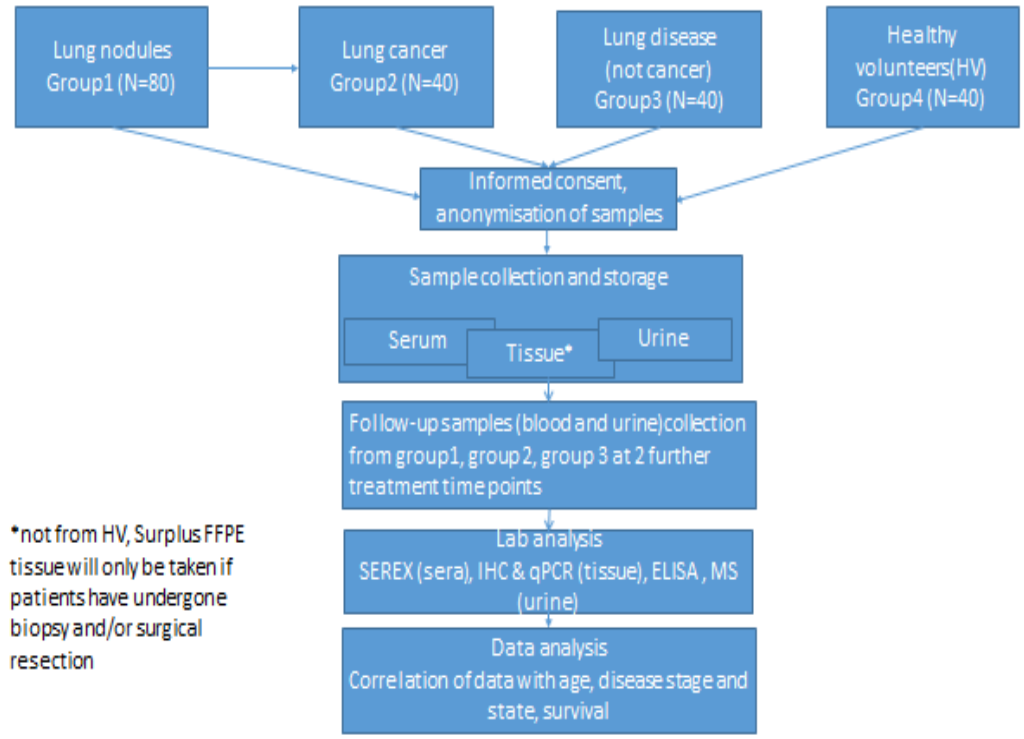


Figure 3.2 A schematic overview of the trial

The plan was that informed consent would be obtained, samples of urine and blood would be collected from patients with LC, non-malignant diseases and healthy volunteers as well as tissue biopsy samples from patients who had undergone surgery for disease confirmation. Samples will be analysed to identify markers related to the disease stage and survival.

3.3 Analysis of aB-ALL samples

DEGs (as before) were identified through the analysis of previously published microarrays. GSE38403 included 215 aB-ALL patients prior to treatment in comparison to 12 healthy donor controls (Geng et al., 2012) and GSE13204 compared 205 aB-ALL patient samples to 74 healthy controls (Kohlmann et al., 2008). IPA is a software from Qiagen used to predict the upstream regulators that are either activated or inhibited and ranked according to their p-value and the genes involved from the uploaded list (SEREX, protoarray, LAA, GSE38403, and GSE13204). Enrichr is a web-based tool for analysing gene sets and returns any enrichment of common annotated biological features. Enrichr is used to show pathway enrichment. The core genes in the selected pathways were examined for their interrelatedness using STRING. Szklarczyk et al. (2019) developed the Search Tool for the Retrieval of Interacting Genes (STRING)

analytical program (Szklarczyk et al., 2019). STRING analysis can group genes using a variety of techniques, including co-occurrence, co-expression, text mining, fusion, experimental, neighbourhood, and database evidence. This aided in gaining a rudimentary understanding of the roles of these genes and the genes to which they were linked. The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (http://string-db.org) was used to determine protein partners and their functional roles in the cell. It was used to build and analyse protein-protein interaction networks of the signature gene. The enriched Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway were used to determine the variations in biological functions between the two groups. Genes associated with survival were analysed using MILE, a multi-laboratory database with approximately 3000 whole genome microarray analysis results (Kohlmann et al., 2008), organised by the European Leukaemia Network and financed by Roche Molecular Systems, Inc. then these were clustered in Bloodspot, a public dataset of gene expression. 649 B-ALL patients were selected compared to the healthy bone marrow samples (74). Genes with Spearman's correlation false detection rate (FDR) corrected $p < 0.05$ were filtered out and remaining genes were arranged in accordance with decreasing Spearman's r-value, thus creating a ranked correlation file. It contains gene expression correlated to survival with a Kaplan-Meier analysis. It also displayed a hierarchical tree depicting the relationship between different cell types in the database. Then, the antigens identified were prioritised using Cheever *et al.* criteria. The criteria in descending order of importance (followed by weighting in parenthesis) were: (a) therapeutic function (0.32); (b) immunogenicity (0.17); (c) role of the antigen in oncogenicity (0.15); (d) specificity (0.15); (e) expression level and percent of antigen-positive cells (0.07); (f) stem cell expression (0.05); (g) number of patients with antigen-positive cancers (0.04); (h) number of antigenic epitopes (0.04); (i) cellular location of antigen expression (0.02) (Cheever et al., 2009). None of the antigens were expected to have all the characteristics of an ideal antigen, but had one been perfect it would have scored 1.0.

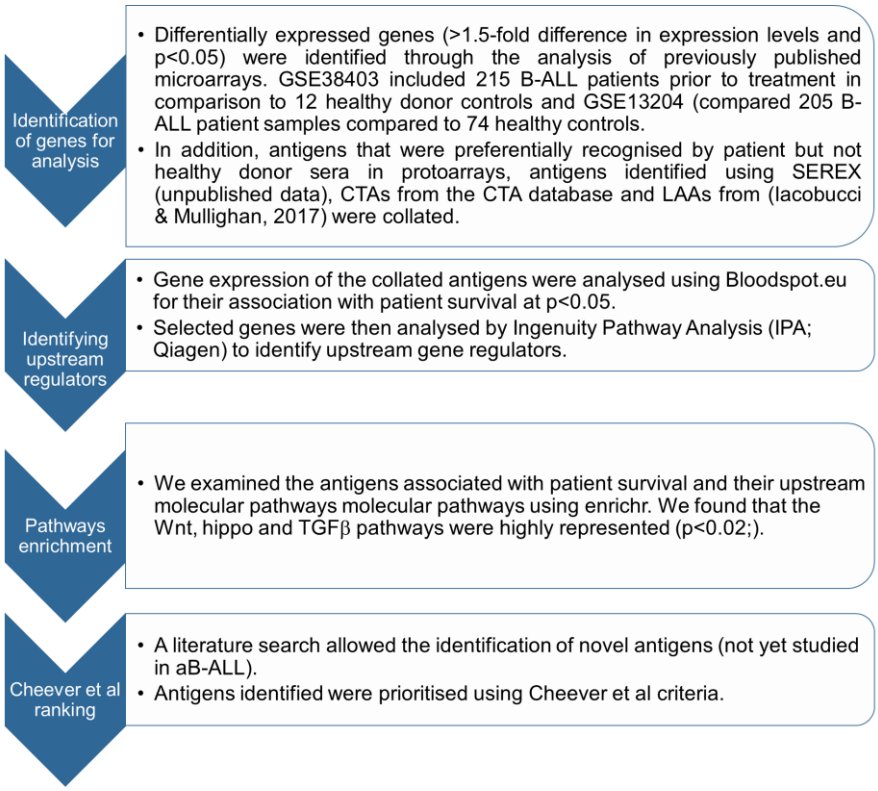


Figure 3.3 Workflow for the identification of cancer vaccine targets

Integration of six methods to identify DEGs that are key for B-ALL. Then the pathway enrichment was performed eliciting key pathways (Wnt, hippo, and TGFB) mainly associated with stem cell-like properties. Antigens that had not been studied before in aB-ALL were ranked using Cheever *et al.* model.

3.3.1 Ethical consent for aB-ALL and healthy donors

Adult patients with B-ALL who were attending the Departments of Haematology at University Hospital Southampton NHS Trust, Portsmouth Hospitals NHS Trust and the Royal Devon and Exeter Foundation Trust were invited to engage in this study following informed consent and local ethical approval (REC 07/H0606/88).

3.3.2 SEREX

3.3.2.1 Principle

SEREX is an immunological technique (Sahin et al., 1995) that is used to identify TAs that can elicit a high IgG immune response in autologous patient sera. It was established by Sahin et al. in 1995 (Sahin et al., 1995) through the modification of a method called autologous typing. SEREX was used to detect the immune-recognition of cancer proteins using in our study a testes cDNA library subclones into MRF' recombinant expression cloning and phage display to present them to antibodies in pre-cleared sera (**Figure 3.4**).

The testes cDNA library used as part of this project was generated by Dr Viktoriya Boncheva-Henderson (Boncheva, 2013). The first step of the construction of the new cDNA expression library were the ligation of the cDNA inserts into the ZAP Express Vector according to the protocol provided by Stratagene UK using reagents from the ZAP Express cDNA Synthesis Kit.

Once PBK-CMV phagemids are produced, the cDNA library was transfected into *Escherichia Coli* (*E.coli*), allowing the production of recombinant proteins from the cDNA insert. The bacteriophage carrying the antigenic cDNA can infect E. coli and initiate their lytic life cycle, resulting in high-level production of recombinant protein expressed as polypeptides on the phagemid surface. Plaque lifts, of the polypeptides produced by the phage, as well as from *E.coli* and any in the media are transferred onto nitrocellulose membrane.

To prevent non-specific binding of antibodies in sera to polypeptides from the *E.coli*, bacteria media or contamination on the membranes, the sera was pre-cleared (Section 3.3.2.3) and the non-specific binding sites were blocked using 5% non-fat dried milk (Marvel) in Tris-buffered saline with 0.05% Tween 20 (TBST). Once membranes with bound polypeptides were washed with TBST, then they were incubated with serum samples diluted 1:100 in TBST overnight at 4°C (**Figure 3.5).**

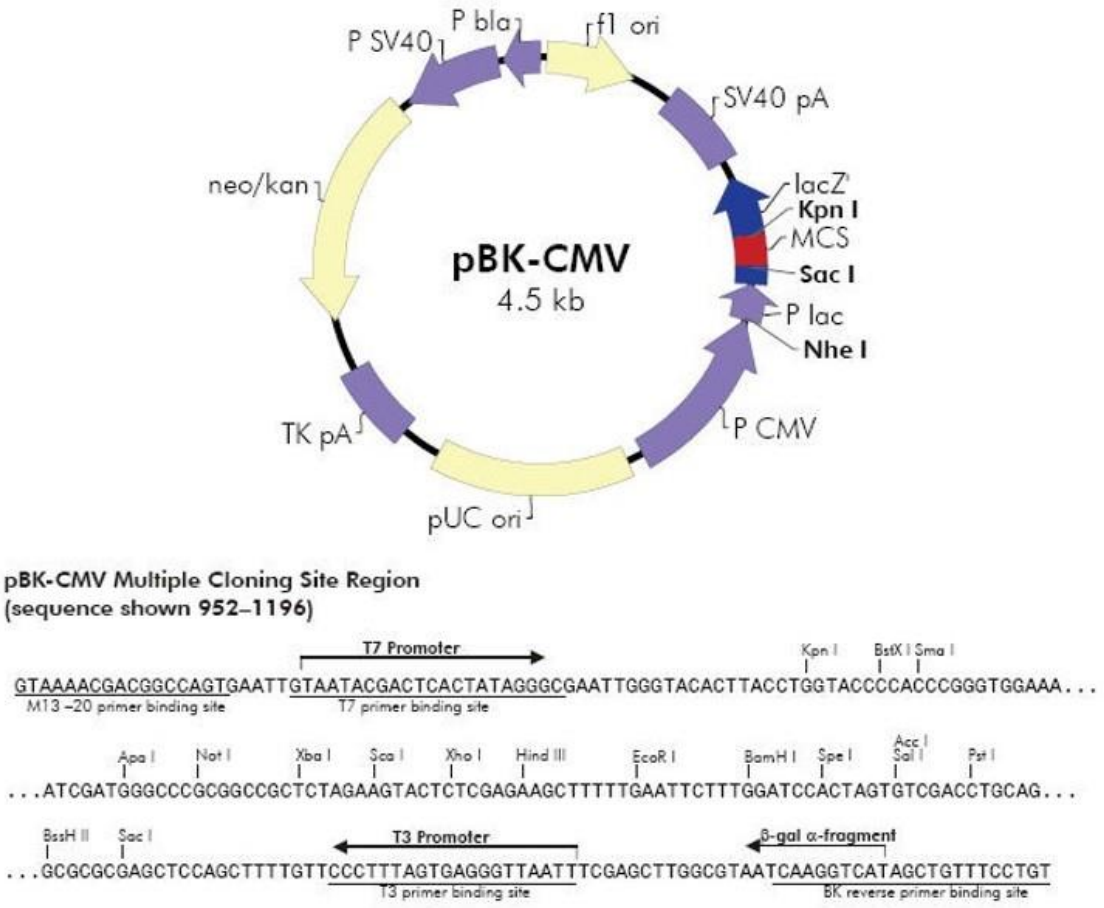


Figure 3.4 pBK-CMV vector map

Sequences are transcribed and translated from the CMV promoter on the same side of the multiple cloning site as the T3 primer binding site. The protein recognised by patient sera is therefore likely to be a protein read in the T3-T7 direction. T7 sequencing was used to confirm the sequence obtained from the T3 primed sequencing and showed that both (in the 5'-3' orientation) were the same.

3.3.2.2 Patient samples

Peripheral blood samples using for the immunoscreening of the testes cDNA library were collected from aB-ALL patients mainly at diagnosis (ALL001, 002, 003 and 004) and one following allo-transplant (ALL005). Part of the sample was collected in tubes without anti-clotting agents and allowed to form a clot for 30 mins at 37°C. The clot was removed, by centrifuging the blood for 7 mins at 1500 rpm. The straw-coloured sera were collected, aliquoted and frozen at -80°C.

3.3.2.3 Preclearing of patient sera

Preclearing can significantly reduce background noise and non-specific antibodies, but it may also remove specific antibodies of interest (Kavran & Leahy, 2014). This method had been carefully optimization by the original SEREX authors (Sahin et al, 2005) and subsequent iterations were achieved by Dr Geng Li, Nottingham Trent University who taught the Guinn Lab the SEREX method (Guinn et al, 2002; Liggins et al, 2005). The aim of preclearing was to effective removal of non-specific antibodies while retaining specific ones relevant to SEREX screening. CNBr-activated Sepharose is a resin that contains cyanogen bromide groups, which can covalently bind to primary amines. In this case, the Sepharose resin is activated with CNBr to allow binding of proteins, including antibodies. With regards to this study sera preclearing was performed by Dr Hannah Wickenden as follows. Sera was precleared using Cyanogen Bromide (CNBr)-Activated Sepharose TM 4B beads (Merck) to prepare the following (1) lytic column; (2) mechanical column; (3) lytic membrane.

A lytic column was prepared as follows:

A single colony of XL1 Blue MRF' *E.coli* bacteria, grown on an LB AMP plate was inoculated into 3ml LB media (Sigma-Aldrich Co Ltd) and incubated overnight in a Multritron shaker stack (Infors HT, Switzerland) (250 rpm**,** 37^{o}C). The cells were then pelleted and resuspended in 7ml LB media with 10 mM $MgSO_4.7H_2O$. 200 µl of the resuspended cells was mixed with 7ml LB media, 10 mM $MgSO_4$, 7.5 µl of a 12.5 µg/ml tetracycline, and supernatant from a single blue phage. *E.coli* was incubated with the XL1 Blue MRF' phage for 15min 37^{o}C and then and incubated for 4 hours (250 rpm, 37^{o}C). The bacterial cells were then lysed by freeze-thawing and stored at -80^{o}C. The thawed solution was sonicated to ensure the release of as many proteins from the mixture as possible. CNBr-activated sepharose beads are washed with 1mM HCl and added to the lysed bacteria along with coupling buffer. The mixture was rotated to couple bacterial/phage proteins to the beads. The matrix was pelleted 2000xg for 10 minutes with 30 ml of coupling buffer, blocked with 30 ml of 0.1MmTris-HCl pH8, washed again in a Multritron shaker stack (Infors HT, Switzerland) (250 rpm**,** 37^{o}C) with 0.1 M NaOAc and 0.1 M Tris-HCl buffers, and finally washed with Tris buffered saline (TBS)/ 0.1% Na azide solution at 2000xg for 10 minutes. Serum was added to the matrix at a 1:10

dilution and rotated 250 rpm overnight at 4°C. The matrix was then pelleted by centrifuging at 2000xg for 10 minutes, once completed the supernatant was saved and matrix was discarded.

Mechanical columns:

The mechanical columns were prepared through the inoculation of a single *E.coli* colony that had not been infected with phage, swiped using a sterile loop into 3ml LB media supplemented with 0.2% maltose and 10 mM $MgSO_4.7H_20$. The cells were then pelleted 2000xg for 10 minutes, resuspended in 5 ml Phosphate buffered saline (PBS), and lysed using freeze-thaw and sonicated eight repeats of 5 second pulses. CNBr-activated sepharose beads were washed with 200 ml 1mM HCl over a sintered glass filter to remove any additives and resuspended 7 ml 1 mM HCl, 5 ml of coupling buffer was added to the lysed bacteria as well as 4 ml of the was sepharose beads. The matrix was pelleted, blocked 30 ml 0.1 M Tris-HCl pH 8.0 and left at room temperature for 2 hours. The matrix was washed three times (30 ml wash Buffer 1 followed by 30 ml wash Buffer 2) and then washed with 50 ml 1x TBS/0.1% sodium azide. The matrix was combined with the 1:10 diluted serum from the lytic column procedure, the falcon tube was sealed and rotated overnight at 4°C. The matrix was pelleted at 2000xg for 10 minutes and the supernatant was saved.

Lytic membrane:

The lytic membrane method involved several steps. First, NZY agar plates were inverted at 37°C to remove excess moisture. Meanwhile, the top agar was melted and cooled in a water bath at 55°C. XL1-Blue cells are grown in LB media until they reach a specific optical density (0.5-0.7). The culture was then kept on ice for 1 hour. The supernatant was saved. The phage and cells were combined and incubated at 37°C for 15 minutes. 2-3 ml Top agar and 12 µl 0.5M isopropyl β-D-1-thiogalactopyranoside (IPTG; Sigma-Aldrich) were added to the cells, mixed, and poured onto the NZY agar plates. The plates are left at room temperature for 10 minutes before being incubated overnight at 37°C. After incubation, a nitrocellulose membrane was placed on top of the plate and inverted for 4 hours in a 37°C incubator. A block solution was prepared by sterilizing a bottle, adding TBS with low-fat dried milk, boiling it, and cooling it to room temperature with the addition of Tween-20. The membrane is carefully removed from the plate, washed

in TBS-T on a shaker, washed again in TBS on a shaker, covered with block solution, incubated for 1 hour on a shaker, washed four times in TBS-T on a shaker (changing petri dish after the third wash), washed once more in TBS for 5 minutes, covered with serum, incubated overnight on a shaker. After incubation, the membrane was discarded and the serum collected into a Falcon tube for storage long-term temperatures. This lytic membrane procedure is repeated twice more to ensure clean serum storage**.**

Preclearing sera: Each sera (5 ml) were defrosted (Table 3.1), diluted 1:10 in TBS/ 0.1% Na-Azide, incubated and rotated 250 rpm overnight at 4^{o}C with each of the following (1) lytic column, (2) mechanical column and (3) lytic membrane in sequential order. After each step, the column or membrane was discarded and the serum collected and moved to the next step.

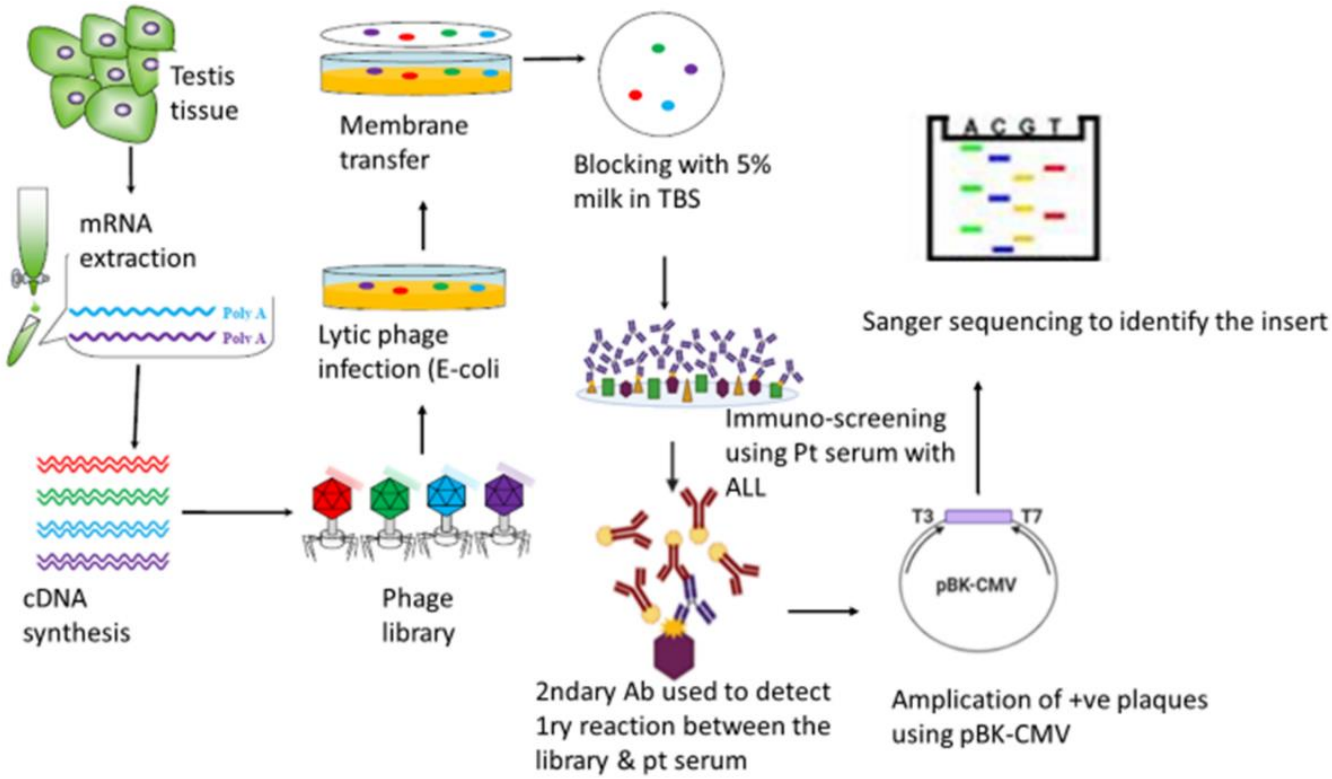


Figure 3.5 Diagrammatic representation of the SEREX technique.

The SEREX technique is based on autologous typing and is used to identify antigens recognized by antibodies in patient sera. It involves the immunoscreening of a cDNA library. The lambda vector allows any cDNA insert to form a phagemid that is expressed as a polypeptide on the capsid surface. The phage lyses the E-coli as part of its lytic life cycle. Each clearing represents a single cDNA insert. These polypeptides are transferred to nitrocellulose membrane and non-specific binding sites are blocked. After incubating the membrane with patient's sera overnight, the membranes are washed in T-TBS and TBS, and secondary antibody (rabbit anti-human IgG-alkaline phosphatase conjugated) forms a blue/violet colour in the presence of 5-bromo-4-chloro-3-indoyl phosphate p-toluidine salt (BCIP) and Nitroblue tetrazolium chloride (NBT) when mixed with alkaline phosphate (AP) colour development solution. The membrane is aligned to the NZY plate of phage and positive plaques are isolated and placed in saline magnesium (SM) buffer. This process is repeated with each potentially positive plaque and 2-3 negative/blue plaques. PCR was performed using primers that bind to the T3 and T7 flanking regions. After PCR amplification and purification of the insert, Sanger sequenced were performed and genes identified following comparisons to the NCBI database.

Precleared sera was defrosted and diluted 1:10 in 1 x TBS/0.1% sodium azide to a final dilution of 1:100 and used for primary and secondary immunoscreening. Sera often works for up to 6 months after defrosting and we were mindful that each defrosting step reduced activity of the antibodies by approximately 50%. This serum was frozen on collection, defrosted for preclearing and defrosted finally for immunoscreening. Initially sera gave high background on membranes, with rapid colour development in the final step taking 3-5 mins. By the time sera was depleted colour development would take 25-30mins when it was discarded.

3.3.3 Phage cDNA library efficiency

The optimised protocol of Guinn et al. (Guinn et al., 2002) was used as follows. Library efficiency was determined using different volumes of testes cDNA library in phagemid to transfect the XL1 Blue MRF' *E.coli* to determine the recombinant rate by virtue of blue/white colony screening. Blue plaque where there is no cDNA insert and the X-galactosidase gene remains intact. Plaque clearing where bacteriophage have cleared the XL1 Blue MRF' *E.coli* as part of their lytic life cycle represents cDNA insert. The recombinant rate was determined using the following equation:-

$$\frac{White\ plaque\ count}{Total\ plaque\ count} \text{ x100\%}$$

Transfection was performed by adding different volumes of (0, 1, 2, 5µl) phage cDNA library to 600µl of XL1 Blue MRF' culture and incubating the mixture in a water bath for 15 min, 37°C for achieving optimal transfection. 8-9 ml of molten but cooled top agar (NZY media prepared by 22g NZY Broth (Sigma-Aldrich Co Ltd) was weighed, dissolved in 1 L and pH adjusted to 7.0.) , 0.7% bacteriological agar, 10% filtered maltose which is requisite for phage to bind the surface of E.coli, 5% filtered sterile H_20), 120 µl 0.5M isopropyl β-D-1-thiogalactopyranoside (IPTG; Sigma-Aldrich), 40µl 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal; Melford) dissolved in dimethylformamide (DMF; Sigma-Aldrich Co Ltd) were added to each tube prior to pouring of the top agar onto NZY plates for blue/clear plaques determination.

The plates were allowed to set and then incubated overnight at 37°C, inverted on a completely flat rack in the incubator to ensure even coverage of plaque and bacteria on

the plate. The plaques were counted using the light box to determine the plaque forming units (pfu)/µl of each cDNA library tested.

3.3.3.1 Primary immunoscreening

Primary screening aims to identify potential positive plaques using sera from patients with adult B-ALL **(Figure 3.6)**. Day 1: after growing the *E.coli* to an OD600 of 0.5-0.8 and placing the cells on ice, the large plates were warmed for one-hour prior to transfection. If plates were particularly wet then excess moisture was drained and/or the plates warmed with the lid partially offset. Transfection was performed by adding 1, 2 or 5 µl of library phage to 600ul E-coli, which was then incubated at 37°C for 15 minutes. 6-8ml of molten 0.7% NZY. The polypeptides from the phage surface, along with proteins from *E.coli* and the media in the top agar, were then transferred onto nitrocellulose membranes (PALL Corporation, New York, USA) for 2-4h at 37^{o}C, with 3h being used most commonly. The membranes were pricked with a sterile 18G needs and peeled off to leave the top agar intact and on the NZY plates. Pricking with a sterile needle allows later orientation to find the potentially positive plaques.

Membrane washing: The membranes were lay face up and washed in T-TBS to remove any residual top agar and bacteria, by smoothing gloved hand over the membrane. Membranes were then placed inverted in empty 132mm petri dishes and washed in 10-15ml TTBS a further 3 x, 5 mins per wash at 1000 rpm. The membranes were then moved inverted into a new empty 132mm petri dish, washed for 5 min in TTBS at 1000 rpm and then a final wash in TBS for 5 mins at 1000 rpm.

Membranes were then blocked in boiled 5% low-fat milk (Marvel) in TBS (20 mM Tris, 137 mM NaCl. Once cooled a final volume of 0.05% Tween-20-pH 7.6 was added. Membranes were shaken at 1000 rpm for 2hrs at room temperature. The membranes were washed as described above and incubated overnight with precleared serum at a 1:100 dilution. After washing, the membranes were incubated with an alkaline phosphatase (AP) buffer conjugated antihuman IgG antibody Fc fragment specific and reactive plaques were visualized using 100 µl of Nitro-blue tetrazolium chloride (NBT, Melford) which was prepared by dissolving 60mg of NBT/ml in 70% dimethylformamide (DMF; Sigma-Aldrich Co Ltd) and 100 µl of 5-bromo-4-chloro-3-indoyl phosphate p-

toluidine salt (BCIP-30mg/ml in 100% DMF) were mixed thoroughly placed in this mixture and incubated in a dark container to develop colour within 30 minutes.

After drying the membrane for at least 2 hrs at room temperature on white paper towels (to avoid colour bleeding from the paper towels into the membrane), the needle pricks placed in the membrane previously were aligned with the needle pricks in the NZY plate. Using a compass, the three needle pricks were used to identify the location of each potential positive plaques, and on excision with a scalpel approximately three negative plaques were removed as well and placed in the Eppendorf tube containing 500 µl of saline magnesium (SM) buffer (1.16g NaCL, 10ml 1M Tris HCL (pH 7.5) (Fisher Bioreagents), 0.4g $MgSO_4.7H_20$, and 1ml 2% W/V gelatin (Sigma-Aldrich Co Ltd). The tube was wrapped with a para-film, rotated in the cold room at 1000 rpm overnight. The next day the samples were removed from the rotator, briefly centrifuged and 20 µl of chloroform were added to prevent bacterial contamination. The samples were vortexed, microfuged briefly and placed at 4°C for up to 6 months and -80°C for longer term storage.

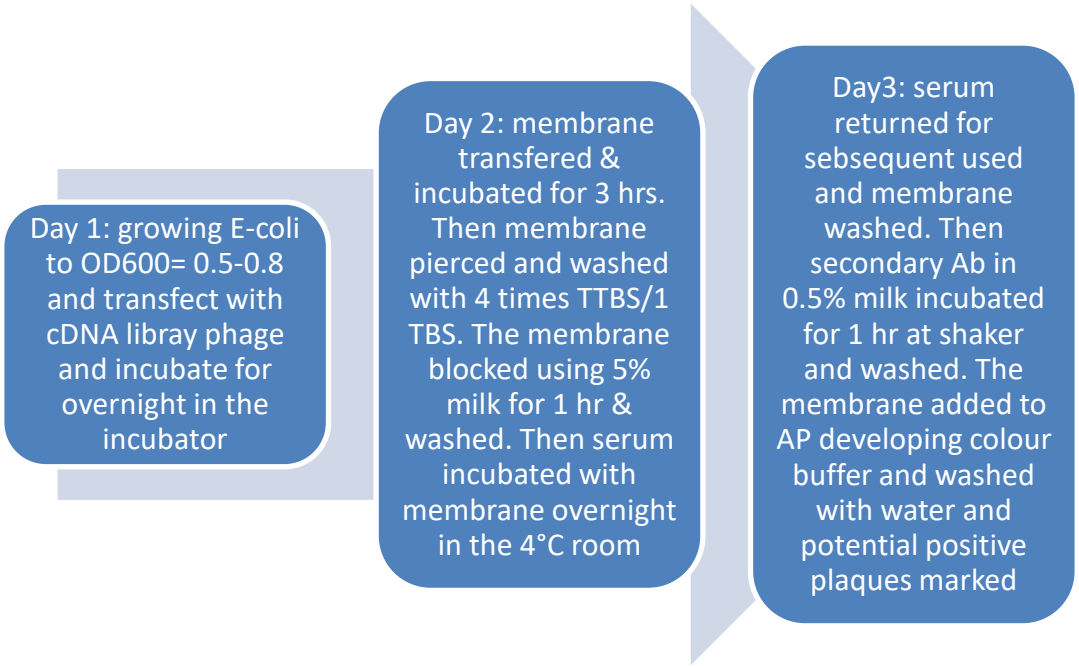


Figure 3.6 SEREX summary with timeline.

Transfection is performed on day 1, XL-blue are grown in LB media until they reach an OD600 of 0.5. Phage containing the cDNA library are then used to infect the *E.coli* and added to top agar and incubated overnight. On the 2nd day, nitrocellulose membranes were added to the plates and incubated for 3 hrs in the incubator. In the last day, the membranes were incubated for 1

hr with 2ndary Ab after washing from serum. The plaque was marked due to reaction of AP buffer and aligned to the original plate and cut for the subsequent analysis.

3.3.3.2 Secondary immunoscreening

It used to confirm the potential positive plaques. It included three days similar to primary screening. The membrane has 82mm 0.22µm nitrocellulose transfer membranes (PALL, USA)

3.3.3.3 PCR for SEREX

Positive plaques were eluted, rescreened, and confirmed by comparison to phages without a cDNA insert as controls. pBK-CMV phagemids were removed. The identified cDNA inserts were amplified via polymerase chain reaction (PCR) using primers that bind the T3 and T7 promoter regions of the pBK-CMV plasmid. A kit from Sigma-Aldrich Co Ltd was used for the PCR. Briefly, 12.5µl ReadyMix Taq PCR with $MgCl_2$ (Sigma Aldrich, UK; Cat #P4600), 9.5µl dH_20, 1µl T7 primer (5'-TAATACGACTCACTATAGGG-3'), 1µl T3 Primer (5'-GCAATTAACCCTCACTAAAGG-3') were mixed in each tube with 1µl of the confirmed plaque. Negative and positive control were prepared by adding 1µl SM buffer and 1µl 1 x pBK-CMV/GKT-AML-8 to the master mix, respectively. The positive control, by virtue of there being a 1.1Kb cDNA insert in the multiple cloning site was first identified by SEREX (Guinn et al, 2002) and renamed ZNF465 in according to HUGO convention when characterised (Collin et al., 2015). PCR tubes were vortexed to mix all contents thoroughly, microfuged and placed in 2720 thermal cycler (Applied Biosystems, California, USA) (Table 3.2).

Table 3.2 2720 Thermocycler Protocol

Denaturation and activation of hot start Taq	5mins	95°C	
Denaturation	60 sec	95°C	
Annealing	30 sec	°C	x 40
Extension	30 sec	72°C	
Final extension	7 min		

The PCR products were stored at 4°C. One gram agarose powder (Sigma Aldrich Co Ltd) was added in 100ml 1x Tris-acetate-EDTA (TAE) buffer (Fisher Scientific, Invitrogen, California, USA), microwaved till dissolve and cooled to 55°C. Before pouring 1% agarose to gel tray and allowed to set for 20 min at RT, 3µl SYBR safe DNA gel stain (Invitrogen) was added. 3µl of the 6-x bromophenol blue loading dye were added to each PCR product. Samples were loaded into the gel and the ladder of 5µl HyperLadder™ 100bp (BioLine, London, U.K) was loaded into one well. Then gel was run for 40-45 min at 100V and after running to the three quarter of the gel, the samples are visualized using UV light on a Bio-Rad gel documentation system (9.1 MP).

3.3.3.4 Sequence analysis of cDNA inserts

PCR products were gel purified via a QIAQuick gel extraction kit following the manufacturer's instructions (Qiagen Ltd., Manchester, YORK, UK; Cat #28704) and sent for Sanger sequencing (the Department of Biochemistry at the University of Cambridge). Applied Biosystem Sequencing Analysis Software was used for generating nucleotide sequences. The produced sequences were compared with known sequences in the gene, expressed sequence tag (EST) and protein databases, including the National Centre for Biotechnology Information (NCBI) BLAST.

3.3.4 Cell culture

K562 cells were culture in media as shown in **Table 3.3**. Media was prepared using 500ml DMEM (Lonza, Verviers, Belgium) or RPMI (Sigma-Aldrich Company Ltd., Dorset, UK), 10% foetal bovine serum (FBS) (ThermoFisher Scientific, Leicestershire, UK) and 1% of penicillin/streptomycin (PS; 10,000 U Penicillin/ml, 10,000U Streptomycin/ml) (ThermoFisher Scientific). K562 was brought from ATCC while A549 was bought from Sigma-Aldrich Co. Ltd. Cell lines were defrosted from liquid nitrogen and partially defrosted and added to warm media. The mixture was spun at 800g for 5 min and incubated in a humidified incubator at 37°C with 5 % CO_2. Cell lines were used to optimise techniques and as positive controls.

Table 3.3 Cell lines, disease and original source details

Cell lines	Cell lines	Cancer subtype & patient details	Adherence	Media	Citation
Leukaemia	K562	Chronic myeloid leukaemia	Non-adherent	RPMI	(Lozzio & Lozzio, 1975)
LC	A549	Derived through explant culture of lung carcinomatous tissue from a 58-year-old Caucasian male.	Adherent	DMEM	(Giard et al., 1973)

A549 cells were adherent and when confluent were washed in 1x sterile PBS, treated with trypsin for 3 mins at 37^{o}C and then resuspended in DMEM media.

3.3.5 Cell counting

Cell counting was performed by 10 µl of 0.4% Trypan Blue (Thermofisher). After mixing 10 µl of the cell suspension was placed onto a Neubauer Improved Haemocytometer Counting Chamber (Hawksley). Clear (live) cells were counted within four squares and the count determined using Average cell count per square x dilution factor x 10^4 cells/ml.

3.3.6 Preparation of cell lines for qPCR:

Cell lines were grown until confluent with >95% viability and along with primary cells were counted (Section 3.3.4). Aliquots of 10^6 cells in media were placed in 15 ml falcon tubes and centrifuged at 12000 rpm for 7 mins. The excess supernatant was decanted and the cell pellet dried using paper towel/pipette tip. Cells were immediately subjected to RNA extraction as described in Section 3.3.7 and analysed as in 3.3.8 and 3.3.9. Where excess cells were available, aliquots of cell pellets were prepared and placed at –80^{o}C for later use.

3.3.7 RNA extraction from aB-ALL samples

aB-ALL samples and HV (whole white blood cells from both blood and BM) from Table 3.1 and 3.2 were taken from -80°C and RNA extraction was performed using the RNeasy Mini Kit® (Qiagen) following the manufacturer's handbook. Briefly, 350 µl lysis buffer RLT were added to the sample and homogenized by vortexing for 1 minute. 350µl of 70% ethanol were added to the homogenized lysate and mixed well by pipetting. 700µl the sample was transferred to a RNeasy spin column placed in a 2 ml collection tube, centrifuged for 15 seconds (s) at > 8000xg, and the flow-through discarded. 700 µl buffer RW1 added to the RNeasy spin column, centrifuged for 15 s at

Table 3.1 patient information

ID	Disease stage	Cytogenetics	Age≠	Sex	Sample type
ALL001*	Diagnosis	Ph + ALL: t(9;22)	39	M	PB
ALL002*	Diagnosis	46, XY, t(1;7)(p36;p15)	19	M	PB
ALL003*	Diagnosis	t(1;19)	26	F	PB
ALL004*	Diagnosis	No result	19	M	PB
ALL005*†	¶Post-allotransplant	No result	46	M	PB
ALL006	Diagnosis	Normal karyotype. FISH failed	19	M	PB
ALL007†	Diagnosis	Loss of one copy of ETV6 (12p13) and gain of one copy of ABL1 (9q34) by FISH;	24	M	PB
ALL008†	Diagnosis	46XY 5,del(5)(q15q33),dic(9;16)(p11;q11),del(13)(q12q14)	19	M	PB
ALL009	Diagnosis	46,XY,t(1;7)(q25;q3?5),add(3)(p1?3)	19	M	BM
ALL010	Diagnosis	Complex including t(4;11)	64	M	PB
ALL011	Diagnosis	No result	19	F	PB
ALL012	Diagnosis	t(11;14)(q24;q32)	33	M	BM
ALL014	Diagnosis	47,XY,+2,add(2)(p1)[3]/46,XY[47].nucish (CRLF2)x2[100],	56	M	BM
ALL015	Diagnosis	Gain of one copy of CRLF2 (Xp22.3/Yp11.3) and loss of one copy of CSFR1 (5q32) and EBF1 (5q33.3) detected by FISH.	20	F	PB
ALL016	Diagnosis	Hyperdiploid; 56-57 XX +X, +4, +6, +9, +10, +14, +17, +18, +21, +marker	27	F	BM
ALL020†	Diagnosis	46,XY, t(1;7)(q25;q3?5), add(3)(p1?3) TCF3 ex16-PBX1 ex3 fusion transcript detected	56	F	PB/BM

*: samples used for SEREX immunoscreening; †: also used in (Boullosa et al., 2018); ¶: 6 months; ≠ age at sampling; BM: bone marrow; PB: peripheral blood

Figure 3.2 Healthy volunteers information

HV control*	Age ≠	Sex	Sample type
HV008	40	F	PB
HV010	22	M	PB
HV012	46	F	PB
HV021	34	M	PB
HV043	NK	M	PB

HV: healthy volunteer

> 8000xg, and the flow-through discarded. 500 µl buffer RPE added to the RNeasy spin column, centrifuged for 15s at >8000xg, and the flow-through discarded. 500 µl buffer RPE added to the RNeasy spin column, centrifuged for 1min at >8000xg, and the flow-through discarded. The RNeasy spin column was placed in a new 1.5 ml collection tube. 30µl RNAse-free water added to the spin column and centrifuged for 1minute at >8000xg to elute the RNA. RNA was aliquoted to the PCR tubes and stored -80°C for cDNA preparation for B-ALL and healthy donors with high quality of RNA yield.

3.3.8 Reverse transcription first strand synthesis

Genomic DNA elimination mix for each RNA sample was prepared in a sterile RNAse and DNAse free PCR tube according to the manufacturer's instructions (**Table 3.4; MBI Fermentas, Cat. 330404)**. The reagents were mixed gently by pipetting up and down, and then centrifuged briefly. The mixture was incubated for 5 min at 42°C and placed immediately on ice for 1 min.

Table 3.4 Genomic DNA elimination mix

Components	**Volume of reaction**
RNA	25 ng – 5 µg
Buffer GE	2 µl
RNase-free water	Variable
Total volume	**10 µl**

The reverse-transcription mix was prepared according to **Table 3.5** and added in the order shown. The RT mix was added to the genomic elimination mix and incubated for

42°C for 15 min (to activate reverse transcriptase) and immediately incubated at 95°C for 5 min (to deactivate it).

Table 3.5 First strand synthesis mix

Components	Volume of reaction
5x Buffer BC3	4 µl
Control P2 1	1 µl
RE3 Reverse Transcriptase Mix	2 µl
RNase-free water	3 µl
Total volume	10 µl

3.3.9 QPCR

Quantitative PCR (qPCR) was performed using the RT² SYBR® Green Master-mix, RT² qPCR Primer Assay and cDNA synthesis reaction (Qiagen) was used. Primers from Qiagen that detected each transcript SRY-Box Transcription Factor 4 (SOX4-PPH01950A), Rho associated coiled-coil containing protein kinase 1 (ROCK1; PPH01966C), Yes-associated protein (YAP1; PPH13459A), TEA Domain Transcription Factor 4 (TEAD4; PPH10558A-200), SMAD family member 3 (SMAD3; PPH01921C), and T cell receptor 4 (TCF4; PPH02770A) and BIRC5 (PH00271E) primers (all Qiagen). The samples were run in triplicate by mixing the master-mix, primers, cDNA from the sample until the volume 25 µl in total. The plate was sealed, centrifuged, put in the thermocycler. Comparative CT method and StepOne software v2.0 (Applied Biosystems) were used (**Table 3.6**):

Table 3.6 Thermocycler protocol

Initial denaturation	5 min	95°C	
Denaturation	10 sec	95°C	} x 40
Annealing/Extension	30 sec	60°C	

3.3.9 Statistical Analysis of qPCR

The StepOne software v2.0 (Applied Biosystems) and the comparative CT method (Livak& Schmittgen, 2001) were used to analyse the qPCR data using the relative quantification approach. GraphPad Prism was used to perform the statistical analysis, including normality Kolmogorov-Smirnov and Shapiro-Wilk tests. Data are presented as

the mean ± SD. In case of comparisons between two groups, means were evaluated using Student t-test or Mann-Whitney U test, for parametric and non-parametric data, respectively. If more than two groups were considered, one-factor or two-factor analysis of variance (ANOVA) or Kruskal-Wallis test (for parametric and non-parametric data, respectively). p-value ≤0.05 was considered significant. When comparing antigen expression in B-ALL to healthy controls, the results were normalised with the TBP1 and PRKG1 reference genes (ΔCT = CT reference genes–CT gene of interest) as well as combining two reference genes using the average of CT values.

3.3.10 Preparation of cells for ICC:

Cells were counted, centrifuged, any excess media removed, the cell pellet flicked and resuspended in 1 x TBS to achieve 5 x 10^6 cells/ml. 5 µl of cells were spotted in each of 2 sites on clean microscope glass slides and allowed to dry for 4-6 h.

3.3.11 Immunocytochemistry (ICC)

Cells which had been air-dried onto glass microscope slides were taken from storage at -20°C and defrosted for 20 minutes at room temperature before carefully removing the saran wrap. The cell buttons were marked using hydrophobic ink (Abcam ab2601) and the cells were fixed in cold 100% methanol for 15 minutes and washed with TBS three times. Immunolabeling of SMAD3 and TEAD4 were conducted using monoclonal rabbit anti-human against TEAD4 (1D10) and SMAD3 (2C12) (ThermoFisher Scientific, Leicestershire, UK) at a concentration of 3.5 µg/ml. K562 was used as a positive control. Cells were washed in TBS 3 x using a circular motion to ensure the cells were washed but not directly impacted by the flow of buffer. The primary antibody concentration was optimised and used at a 1:100 dilution for each antibody, diluted in TBS. The samples were incubated for 1hr at room temperature in a humidity chamber. Controls that were used – no primary, isotype (MOPC-21) and cells only as well as actin (ACTN05(C4)). Immunolabeling by primary and secondary antibodies were detected using the EnVision®+ Dual link system (DAB+), which involved HRP-labeled anti-rabbit polymer (Agilent Technologies Deutschland GmbH, Waldbronn, Germany; Cat# K4063), following the manufacturer's instructions. Mayer’s haematoxylin: Lillie’s modification (Dako Cytomation) was added to as a counterstain and washed with water to remove all excess stain. Slides were mounted in Faramount aqueous mounting medium (Dako

Cytomation*)* for analysis with an Olympus CX21 light microscope and analysed at 40x magnifications, and images were recorded. A Histo(H)-Score was used to characterise the expression of TEAD4 and SMAD3 based on the intensity of brown colour. Staining intensities were scored according to a five-tiered scale described originally by (Biesterfeld et al., 1996) as follows:- 0 = negative; 1–29:weak; 30–143: moderate (mod) and 144–228: high; >228: very high. The percentage of positively stained cells was based on the cell count of stained cells per microscopic view and represented on a five-tiered scale (0: 0 %; 1: 1–10%; 2: 11–50%; 3: 51–80%; 4:>80%). The final immunoreactivity score was obtained by multiplication of the percentage of positive stained cells scored over ≥5 microscopic views by the value for staining intensity within the same (Deng et al., 2014).

3.3.11 Statistical Analysis of ICC data

GraphPad Prism was used to perform the statistical analysis, including normality Kolmogorov-Smirnov and Shapiro-Wilk tests. Data are presented as the mean ± SD. If more than two groups were considered, ANOVA with p-value ≤0.05 was considered significant.

Chapter 4: Identification of antigens in NSCLC that may act as biomarkers of early disease using RNAseq data

4.1 Introduction

Two-thirds of NSCLC cases are diagnosed at the advanced stages and <20% of cases have >5-year survival rate. Late-stage detection of NSCLC has a significant and detrimental impact on the patient's survival rates (Coakley & Popat, 2020). The main challenges in achieving earlier detection include the difficulty in accessing the sites where tumours originate and the multiple locations from which tumours can develop (Coakley & Popat, 2020). NSCLC diagnosis must be focused on early detection to increase the survival rate and reduce suffering and cost to the health service. The principal role of screening is to identify the high-risk patients and detect asymptomatic patients who have enhanced survival rates of LC at early stages (Pastorino et al., 2019). Although screening can lead to false positive or misdiagnosis, causing harm to individuals, missed early diagnosis can be fatal. The current methods sometimes fail to detect the presence of early disease due to the overlap in benign and malignant nodule characteristics which often require a follow-up to confirm a diagnosis of LC (Field et al., 2016). TAAs such as CEA, CYFRA21-1 and SCC have been studied as biomarkers to facilitate NSCLC diagnosis, but they remain unsuitable for early diagnosis due to their low sensitivity (Kulpa et al., 2002; Schneider et al., 2003).

As CTAs are not generally found in normal tissues (except immune privileged sites), these may act as very specific biomarkers for NSCLC. NSCLC is considered as a tumour with high CTA expression (Gure et al., 2005), with upregulation detected in 10-50% of NSCLC samples and associated with poor survival and advanced disease stages (Gure et al., 2005). CTAs are involved in gametogenesis and spermatogenesis (Li et al., 2020c). Germ cells are similar to trophoblasts exhibiting many features akin to cancer cells, as immature spermatogonia continues their proliferative capacity until they differentiate into spermatocytes (Old, 2007). The spermatocytes undergo meiosis, comparable to chromosomal changes found in most cancers (Old, 2007). Furthermore, germ cells colonise in the gonads similar to cancer progression and metastasis (Old, 2007). This is

supporting the hypothesis that activation of embryonic or gametogenic programs are normally active in the germ line may play a crucial role in tumourigenesis. This is supported by the fact that many germ cell and placental proteins, including CTAs, are aberrantly expressed in cancer (Silva et al., 2007a). CTAs have been found to have various functions in cancer cells. They can promote cell proliferation and survival, inhibit apoptosis, enhance invasion and metastasis, and modulate immune responses (Silva et al., 2007a). Overall, the role of CTAs in the germ line is complex and still not fully understood.

4.2 Aims

To determine if CTAs are highly expressed in early NSCLC

4.3 Results

Following data normalisation, 199 NSCLC samples from GSE81089 RNAseq (Mezheyeuski et al., 2018) and their clinical characteristics provided (Figure 3.1), F**igure 4.1** shows the hierarchical cluster analysis demonstrating the pair-wise correlation between all 199 NSCLC cases based on stages; gene expression at each of the four main stages (Ia, Ib, IIa, IIb, III and IV – top row) and grouped into four core gene expression clusters labelled S1, S2, S3, and S4. It revealed that cancer histology is by far the most dominant factor for gene expression differences and responsible for clustering. Differential expression genes between tumour (T) and normal tissues (N), of 402 genes with a FC>2 for upregulated and FC<-2 for downregulated ones and a qp-value<0.05. Among these DEGs, 184 were upregulated and 218 were downregulated. **Figure 4.2** displays the top upregulated genes, including COL11A1, SPP1, FAM83A, COL10A1, GREM1, MMP1, MMP13, CTHRC1, KRT6A, CYP24A1 and GJB2. Conversely, the downregulated genes are SFTPC, CLDN18, SLC6A4, SFTPA1, SFTPA2, AGER, FABP4, FCN3, ADH1B, CYP4B1, TMEM100, AQP4, AGTR2, FAM107A, LGI3, GPD1, CA4 and GKN2. COLL11A1 had the highest $\log_2$FC as it represented a low expression of 0.7165 in normal tissues compared to 6.571 in NSCLC. In contrast, SFTPC was downregulated in NSCLC (6.375) compared to the expression levels in normal tissues of 12.79.

To identify potential biomarkers for cancer prediction, the machine learning algorithms, including LASSO (linear regression), elastic nets (elastic net regularization), random forests, and extreme gradient boosting (linear regression and logistic regression) were used providing the top 40 features according to cumulative ranking with accordance of a variable importance score (Akhmedov et al., 2020). **Figure 4.3** shows an aggregated score calculated as the cumulative rank of variable importance. Among the upregulated genes identified in our analysis is collagen type XI alpha 1(COL11A1). Zhu *et al*. has shown that COL11A1 (Zhu et al., 2022) is an oncogene upregulated in NSCLCs, particularly ADC. It plays a role in cancer progression and its knockdown in NSCLC cell lines reduces colony formation units and enhances cancer apoptosis. Additionally, COL11A1 contributes to the tumour microenvironment by increasing immune cell infiltration and promoting tumour escape (Zhu et al., 2022). Interestingly, COL11A1 (Jia et al., 2016) has also been found to be upregulated in other cancers such as pancreatic cancer. Another gene of interest is surfactant protein C (SFTPC), which is involved in pulmonary protection as part of a protein-lipid complex on alveolar surfaces. Our analysis revealed downregulation of SFTPC in NSCLC suggesting its role as a tumour suppressor (Zhu et al., 2022). Laminin subunit alpha 1 (LAMA1) is correlated to stages and its expression increases from stage I to most upregulation in IV with log2CPM **7 (Figure 4.4)**. LAMA1 belongs to the laminin family which is a family of adhesion molecules found in the extracellular matrix (Yoshimura et al., 2020). It has been previously found upregulated in NSCLC as well as other cancers including oesophageal carcinoma and melanoma (Zhou et al., 2021b).

In terms of histology correlation genes were identified by combining several machine learning logarithms (ML), including LASSO (spls.da), elastic nets (glmnet.a1 and glmnet.a0), random forests (randomForest), and extreme gradient boosting (xgboost and xgboost.lin). S100 calcium binding protein A2 (S100A2) was found to be mostly associated with ADC (**Figure 4.5);** it may be worth to be investigated if its high expression may differentiate ADC from other NSCLC subtypes as well as its targeting potential. As indicated by the SLR, the antibodies generated in cancer patients recognise S100A analogues and were among a panel of antibodies (IgG: EPB41L3, ANKRD36B, FGCR2A,

LINGO1; IgM: S100A7L2) with low sensitivity of 50% and 70% specificity (Lastwika et al., 2019).

Furthermore, the expression difference between normal and cancer tissues was examined for CTAs. Although PRAME had the highest log_2FC value of 3.5 among CTAs (**Figure 4.6**), it did not appear in the top differentially expressed gene list. The expression patterns of CTAs in various normal tissues including pancreas, liver and brain were plotted (accessing 10.5281/zenodo.8419232).

Our pathway analysis revealed that most of the differentially expressed genes in tumour compared with healthy tissues were associated with various cancer-related pathways **(Table 4.1)**. Notably, the top upregulated pathway identified was gastric cancer, while complement activity was found to be the most downregulated pathway in NSCLC.

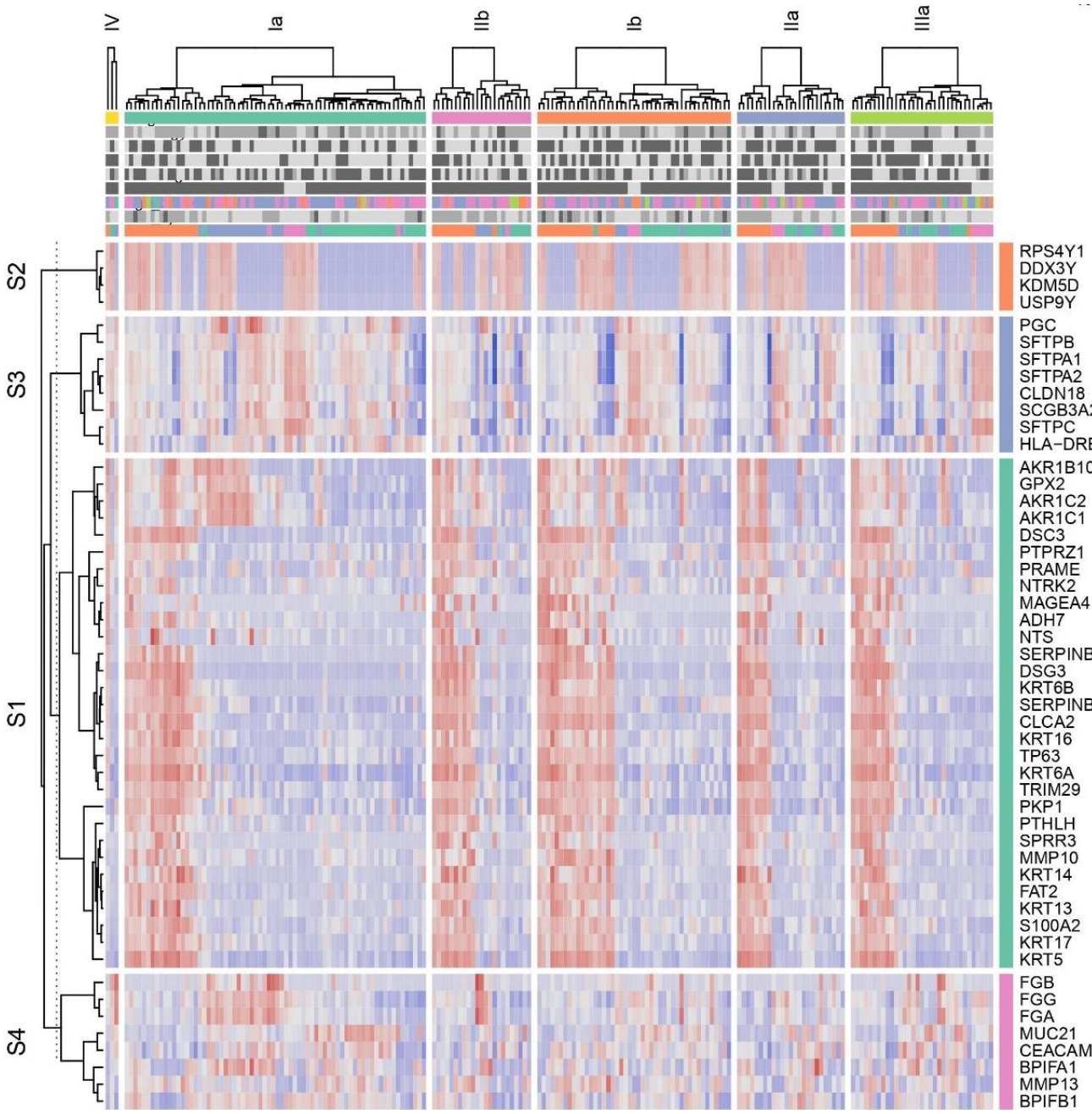
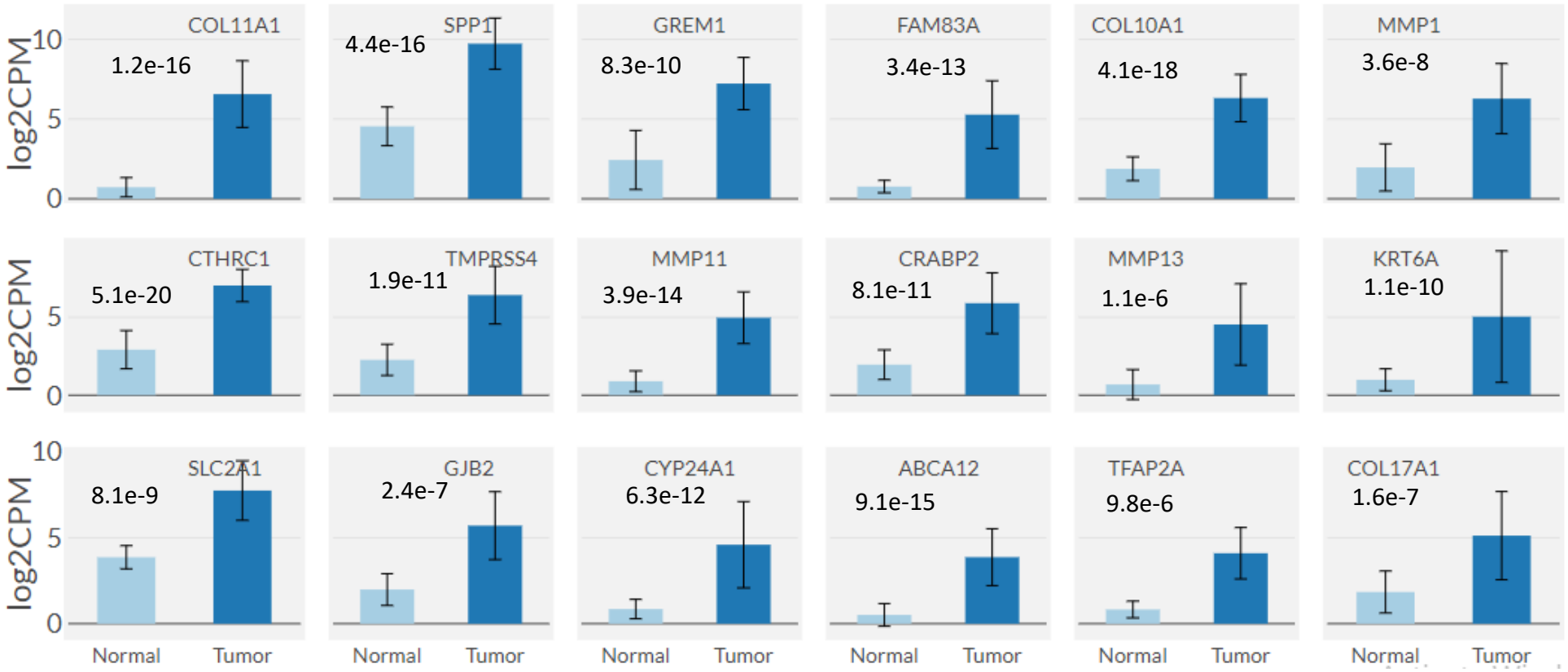


Figure 4.1 Phenotype clustering on NSCLC stages

GSE81089 RNAseq data and clustered hierarchically based on gene expression levels and organised by NSCLC stage. A multi-dimensional scaling plot was generated to visualise the differences in gene expression at each of the four main stages (Ia, Ib, IIa, IIb, III and IV – top row) and grouped into four core gene expression clusters labelled S1, S2, S3, and S4 on the y-axis.

(A)


COL11A1
1.2e-16
SPP1
4.4e-16
GREM1
8.3e-10
FAM83A
3.4e-13
COL10A1
4.1e-18
MMP1
3.6e-8
CTHRC1
5.1e-20
TMPRSS4
1.9e-11
MMP11
3.9e-14
CRABP2
8.1e-11
MMP13
1.1e-6
KRT6A
1.1e-10
SLC2A1
8.1e-9
GJB2
2.4e-7
CYP24A1
6.3e-12
ABCA12
9.1e-15
TFAP2A
9.8e-6
COL17A1
1.6e-7
log2CPM
10
5
0
Normal
Tumor

(B)

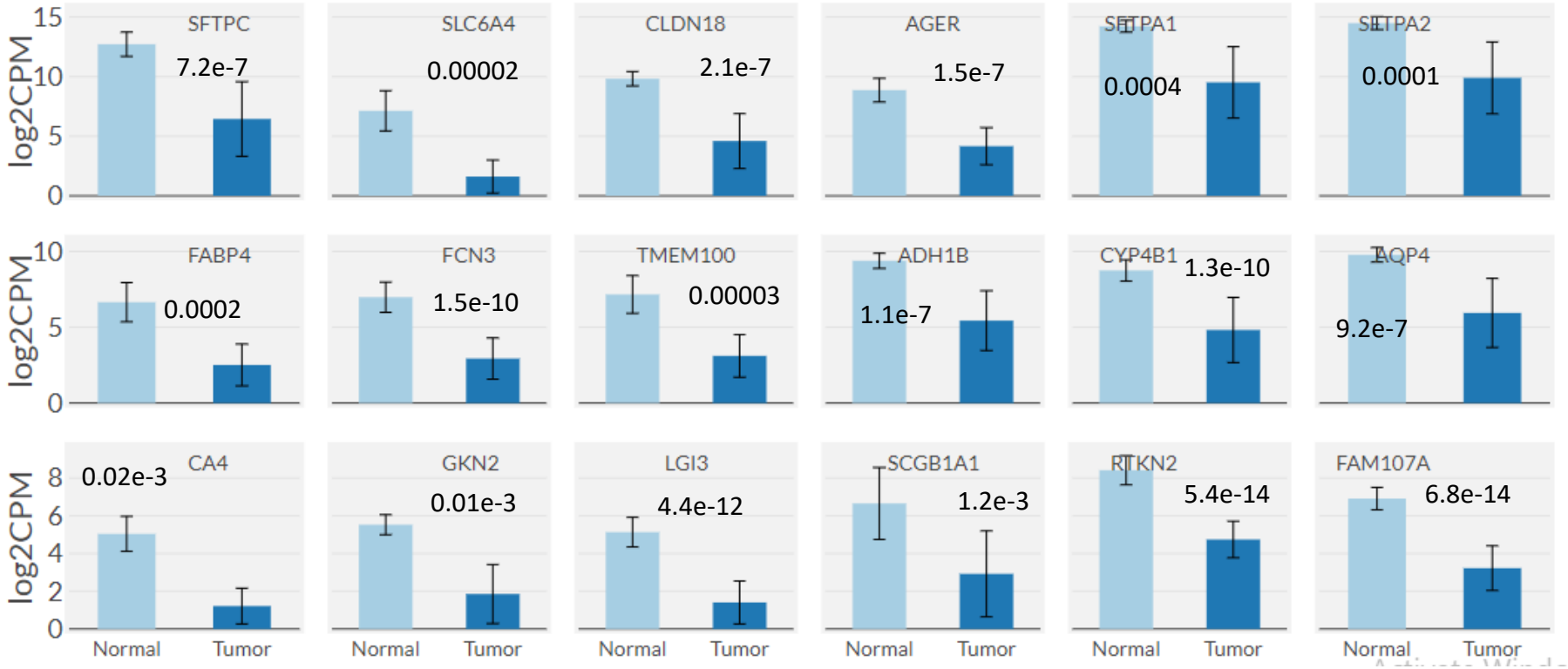


Figure 4.2 Top DEG identified by log FC >2 and significant q values (q < 0.05)

The 402 DEGs identified following analysis of GSE81089 RNAseq data and the comparison of gene expression in tumour (T) and normal tissues (N) were organised into those which were the 18 most A: upregulated and B: downregulated DEGs. Average expression plots were prepared based on the 199 NSCLC samples (Tumor) and 19 HV (Normal) samples in GSE81089. q-values were calculated and are shown below each gene plot. Expression was determined using DESeq2 and represented as log2 counts per million (CPM) as shown on the y-axis.

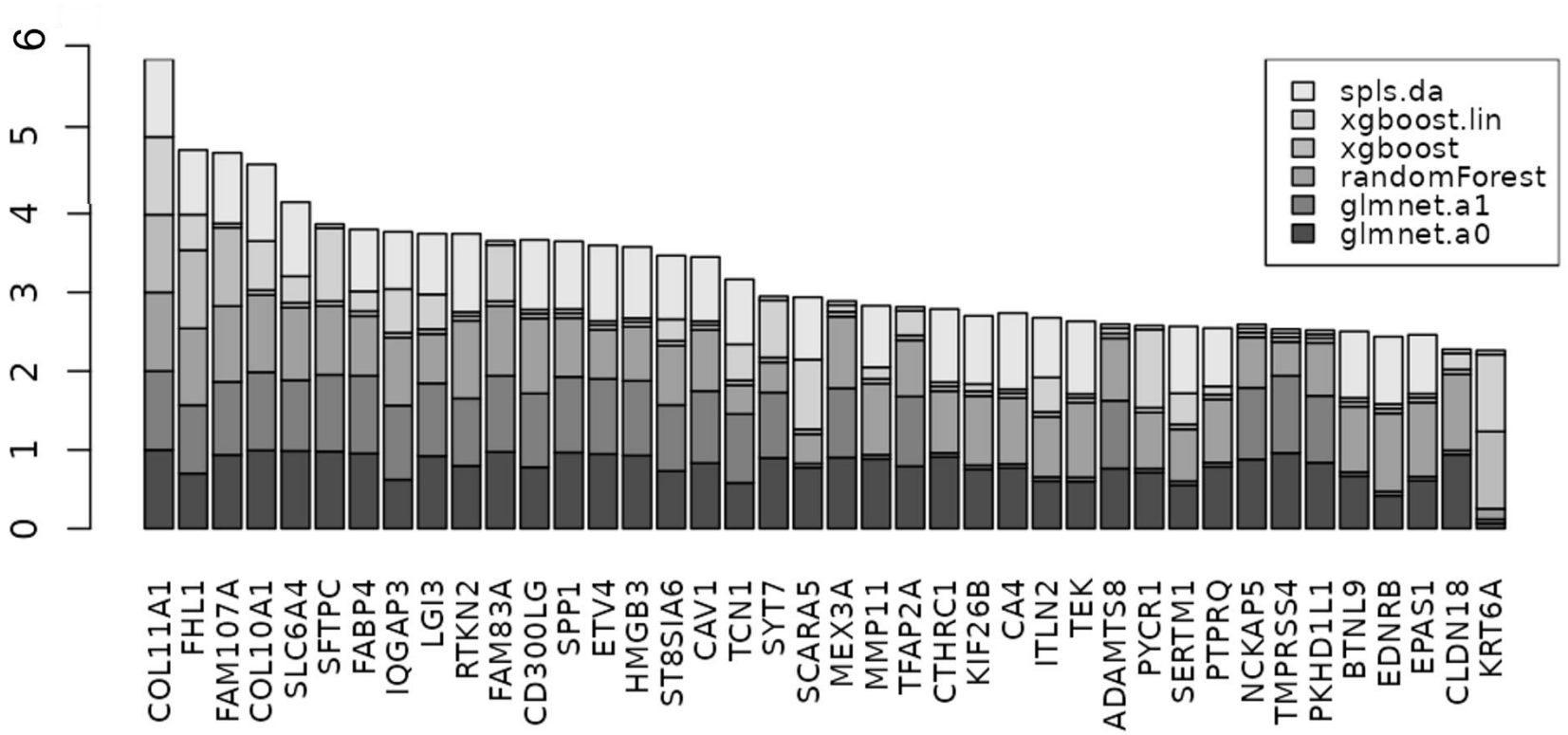


Figure 4.3 The cumulative weight of the top 40 biomarkers based on the mathematical modelling

By combining several machine learning logarithms (ML), including LASSO (spls.da), elastic nets (glmnet.a1 and glmnet.a0), random forests (randomForest), and extreme gradient boosting (xgboost and xgboost.lin), genes were identified that may act as biomarkers through the calculation of variable importance (0-6; y-axis). COL11A1 has the highest cumulative weight as determined by the ML.

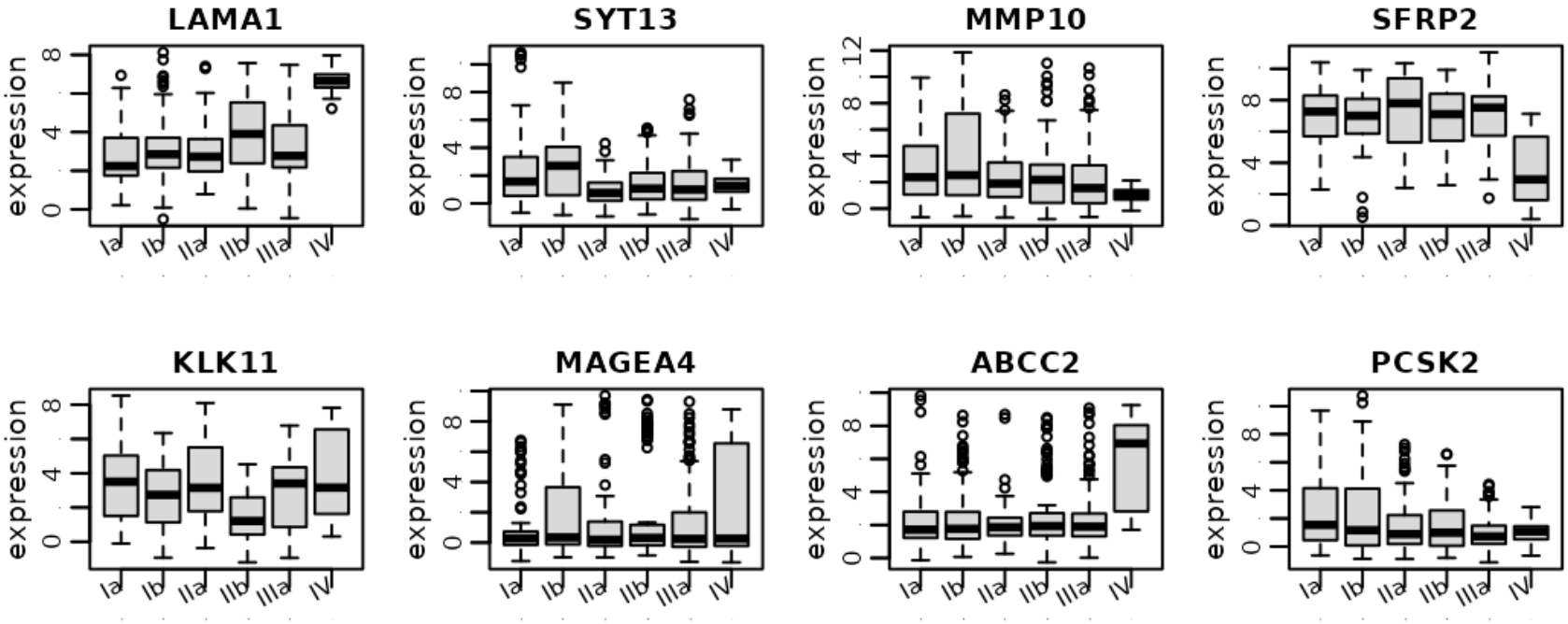


Figure 4.4 Average gene expression ± standard error of mean is shown for each NSCLC Stage

The expression (y-axis: log2CPM) of each gene in the 199 NSCLC samples were divided by stage. Patient numbers Ia=70, Ib=45, IIa=25, IIb=23, IIIa=3, IV=3 as by the description of Figure 3.1. These eight genes each showed expression that was associated with disease stage and were identified using the intersection of ML algorithms. LAMA1 expression is significantly associated with stage, although MAGE A4 is correlated with stage but its expression was not significant. The Box-Whisker plots show upper and lower quartile range (25% to 75% by the box), mean (thick horizontal line across the box). Circles indicate outliers.

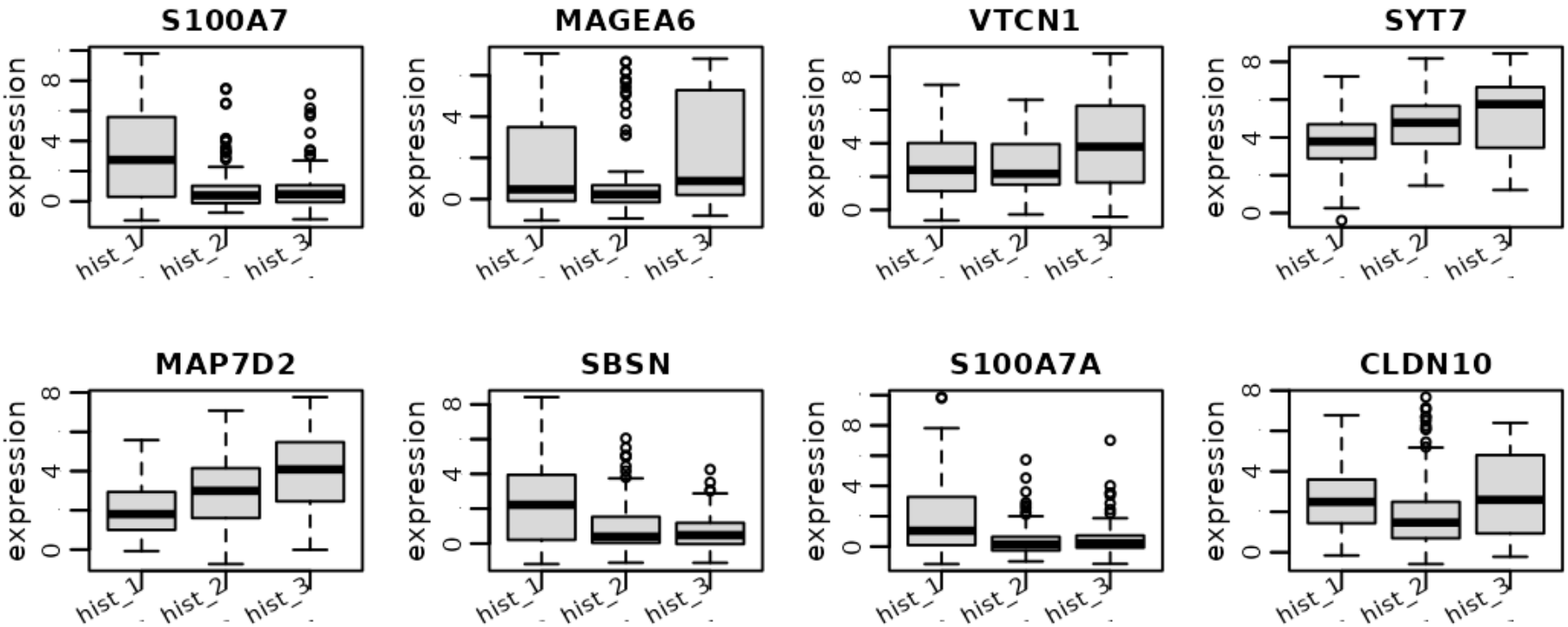


Figure 4.5 Average gene expression levels expression ± standard error of mean is shown for NSCLC histology

The expression (y-axis: log2CPM) of each gene in the 199 NSCLC samples. Hist_1 is ADC and number of patients= 108, Hist_2 is SCC and patient number=67, Hist_2 is non-specific histology and patient number=24 from GSE81089 as by the description of Figure 3.1. S100A7, SBSN and S100A7A expression are significantly associated with ADC. The Box-Whisker plots show upper and lower quartile range (25% to 75% by the box), mean (thick horizontal line across the box). Circles indicate outliers.

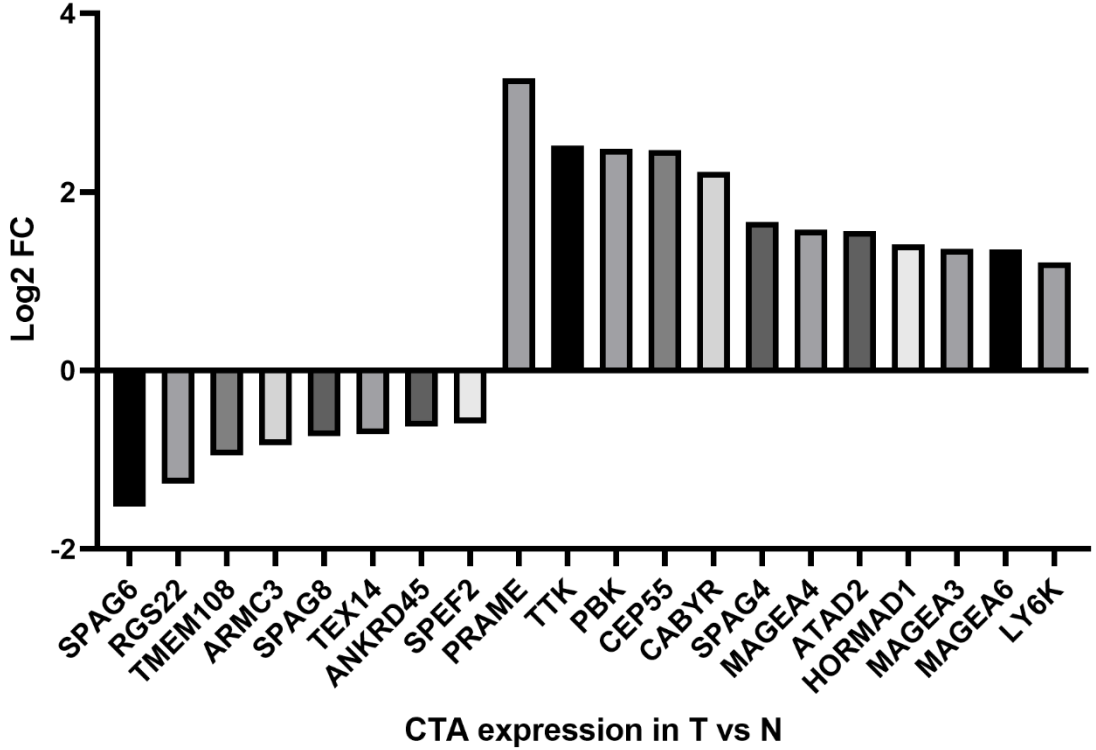


Figure 4.6 Differential CTA expression in this dataset between tumour and normal different tissues

The expression (y-axis: log_2FC) of CTAs in the 199 NSCLC samples (T) versus 19 normal tissues (N) from GSE81089 analysis. The leading log2FC, which represents the average of the largest log-fold changes between sample pairs, was calculated. PRAME and TKK were the most upregulated CTAs while SPAG6 and RG22 were the most downregulated CTAs in tumour samples compared to normal tissues.

Comparing the results of RNASeq analysis in this Chapter and the SLR performed (Chapter 2), CEA and HSP90α had both high sensitivity and specificity. In the RNAseq CEA analogues, CEACAM1 had log_2FC 0.93 while CEACAM21, CEACAM4 and CEACAM8 were downregulated with log_2FC of -1.19, -0.67, and -0.62 respectively when comparing diseased and healthy tissues. HSP90AA1 had a log_2FC 0.55. Having shown high sensitivity in the SLR, ITGA2 was found to be upregulated in diseased tissue with log_2FC (1.49). The strength of this RNAseq data was that the normal samples have been taken from different tissues which indicated their expression not only in normal lung tissues but in also in the other tissues showing an average expression in physiological conditions. Interestingly, MAGE genes were found to be expressed in low levels in LC. Due to the large number of samples analysed (199 with NSCLCs), the data had very high background noise reflecting heterogeneity of NSCLC. Although the number of NSCLCs were 199, only four patients were stage IV showing COL11A1 was upregulated as NSCLC developed.

Table 4.1 Enriched pathways in NSCLC

Pathway	**Log FC**	**qP**
Upregulated		
Gastric cancer network	1.37	0.0007
Regulation of sister chromatid separation at the metaphase-anaphase transition	1.23	0.0031
DNA replication	0.96	0.0007
Retinoblastoma gene in cancer	0.90	0.0007
Cell cycle	0.90	0.03
Downregulated		
Complement activation	-0.85	0.0008
Cells and molecules involved in local acute inflammatory response	-0.84	0.001
Platelet-mediated interactions with vascular and circulating cells	-0.81	0.005
GPCRs, class B secretin-like	-0.76	0.002
Lipid metabolism in senescent cells	-0.65	0.04

4.4 Discussion

Our results align with previous finding that CTAs are overexpressed in advanced NSCLC (Gure et al., 2005; Zhou et al., 2021a). To help identify the genes whose expression was specific to NSCLC, we examined the expression of a range of CTAs. To ensure these antigens were important to NSCLC pathogenesis we examined the relationship between

the gene expression (above and below median levels) and patient survival (**Table 9.2**). Genes whose expression were associated with patient survival were more likely to play a key role in disease pathogenesis, and less likely to be bystanders impacted by changes in the expression of other key proteins involved in disease initiation and progression. Most CTAs are oncogenes involved in tumour proliferation, epithelial mesenchymal transition (EMT) and metastasis via upregulation expression of oncogenes and cell signalling activation (Mecklenburg et al., 2017) (**Figure 4.7**). For example, SSX2 is a protein that binds to DNA and interacts with chromatin, thereby influencing its structure (Greve et al., 2015). Its regulatory role extends to the modulation of Polycomb Group proteins, which are crucial epigenetic regulators involved in gene expression and implicated in cancer progression (Greve et al., 2015). Consequently, the present findings suggest that SSX proteins facilitate the proliferation of cancer cells by modulating gene expression. SSX2 knockdown results in a significant reduction of proliferation in melanoma cells (Greve et al., 2015). Further SSX2 is involved in activating several proliferative pathways such as Wnt and MAPK (D'Arcy et al., 2014). As SSX2 is involved in activating of B-catenin and other key proteins such as SNAIL, it may play a role in EMT and resulting in tumour metastasis (D'Arcy et al., 2014).

MAGE A represents an interesting biomarker for NSCLC as it is expressed in 70% of SCC with high tumour specificity and immunogenicity (Karimi et al., 2012). Patients with MAGE A expression levels of 0.2% or higher in at least one sample of bone marrow or blood during tumour surgery had significantly lower overall, cancer-free, and distant metastasis-free survival rates compared to patients with MAGE A levels below 0.2% in all samples. There was no significant difference in locoregional recurrence-free survival between the two groups. The hazard ratios for death, cancer-related death, and development of distant metastasis were also higher in patients with MAGE A levels of 0.2% or higher (Mecklenburg et al., 2017). The five-year Kaplan-Meier estimates showed a lower distant metastasis-free survival rate (43%) in patients with MAGE A levels of 0.2% or higher when compared to those with MAGE A levels below 0.2% (87%) (Mecklenburg et al., 2017). Specifically, we examined the activation frequencies of MAGE genes in primary NSCLCs.

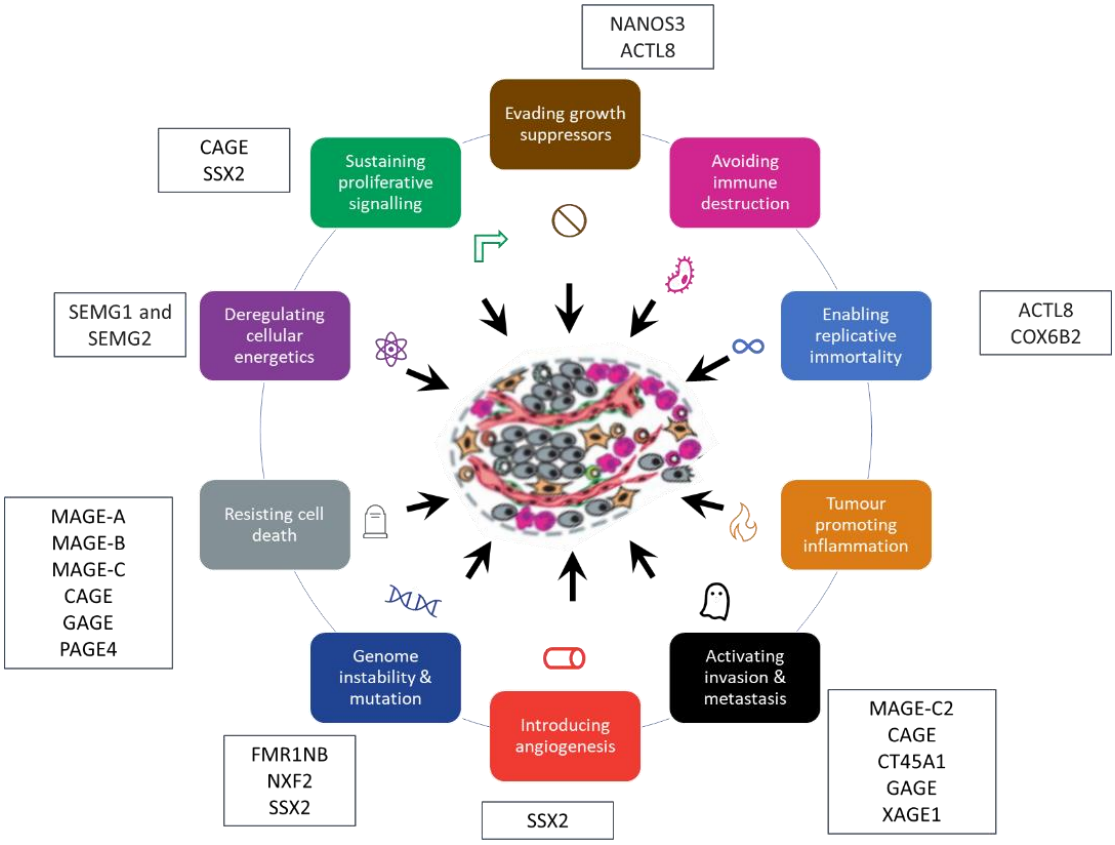


Figure 4.7 CTA function in LC acting as oncogenes

The whole cancer hallmarks are indicated in the circle with different colours and CTAs involved in hallmarks found in boxes aligned. Oncogenic CTAs promote tumour initiation and proliferation such as MAGE-A and some CTAs are involved in more than one hallmark of cancer such as SSX2.

We focused on three MAGE genes (MAGE A1, A3, and B2) that have previously been found to be expressed in LCs (Mecklenburg et al., 2017). Our findings revealed that beyond MAGE in LC is expressed in normal lung tissues adjacent to cancers and bronchial epithelial cells from former smokers without LC, in agreement with (Mecklenburg et al., 2017). This suggests that MAGE gene activation can occur at an early stage of lung carcinogenesis (Mecklenburg et al., 2017). Unfortunately, MAGE gene expression was present in normal-appearing lung tissues, which was unexpected (Mecklenburg et al., 2017). This raises the possibility that MAGE gene activation may not be limited to cancer cells alone. Furthermore, 35-60% of bronchial epithelium samples from former smokers without LC showed expression of the MAGE genes (Bhutani et al., 2011). This supports the idea that the activation of MAGE is a common occurrence in bronchial epithelium exposed to carcinogens. Additionally, we observed that the frequencies of expression for MAGE A1 and MAGE A3 were significantly higher in tumour sections compared to

bronchial brushes obtained from former smokers without LC. This difference is likely due to all bronchial brushes being collected from individuals without detectable LC. Notably, when we analysed a cDNA library derived from the lungs of a nonsmoking 17-year-old female, we did not detect any expression of the three MAGE genes.

MAGE antigens have been found to play a crucial role in promoting carcinogenesis and cancer cell survival. One of the key mechanisms through which MAGE CTAs contribute to tumour growth is by binding and regulating the function of the tumour suppressor protein p53 (Yang et al., 2007). By doing so, they enhance the survival of cancer cells (Yang et al., 2007). Yang et al., have shown that when MAGE A, MAGE B, and MAGE C proteins are knocked down in melanoma cells, the complex formation between p53 and its co-repressor KAP1 is inhibited (Yang et al., 2007). This leads to increased p53 activity and ultimately triggers apoptosis or programmed cell death. Additionally, MAGE A proteins directly interfere with the interaction between p53 and chromatin, inhibiting its role as a transcriptional regulator. They also recruit transcription repressors (histone deacetylases) to p53-regulated genes, further down-regulating its function (Monte et al., 2006). Furthermore, members of the MAGE family have been identified as regulators of E3 RING ubiquitin ligases. These ligases facilitate the degradation of various proteins, including p53 (Doyle et al., 2010). Therefore, by influencing these ligases activity, MAGE proteins contribute to the proteasomal degradation of p53. MAGE A2 may be involved in cellular senescence which is a state of irreversible growth arrest that acts as an important tumour suppressor mechanism. However, little is currently known about how cells bypass this response during tumourigenesis (Smith & Kipling, 2004). Peche *et al.* found that MAGE A2, a protein found in human fibroblasts, has been shown to promote cell proliferation in response to the expression of RasV12 by limiting the senescence response to this oncogene. This effect is believed to be due to MAGE A2 inhibition of p53 function, suggesting that MAGE A2 may play a unique role in the early progression to malignancy by interfering with p53 function (Peche et al., 2012). This interference may prevent the activation of the senescence program, which is a crucial defence mechanism against cell transformation (Peche et al., 2012). Further research should prioritize investigating the potential cooperation between CTAs that interfere with cell cycle regulation and oncogenic signalling. Apart from MAGE proteins, other CTAs like

GAGE and PAGE proteins have also been implicated in promoting apoptotic resistance in cancer cells. For example, GAGE7 prevents apoptosis triggered by different stimuli such as interferon-γ, while knockdown of PAGE4 induces cell death and attenuates tumour growth (Zeng et al., 2011). B cell receptor-associated protein 31 (BCAP31) is a novel CTA and its expression is associated with histological grade and p53 status in NSCLC (Wang et al., 2020a). High expression of BCAP31 combined with calreticulin, glucose-regulated protein 78, and glucose-regulated protein 94 predicts the poor overall survival (Wang et al., 2020a). The integration of this test with imaging techniques like LDCT, has the potential to significantly enhance the accuracy and precision of LC diagnosis. It can also aid in evaluating recurrence during clinical assessments and ultimately improve patient outcomes. If a reliable, precise, and rapid alternative test had been available, patients in data set 1 with benign nodules could have avoided unnecessary lung resections. Although our study on remission and recurrence samples is limited in size, the results are highly promising, suggesting that AKAP4 could potentially serve as a sensitive marker for monitoring remission and an early indicator of recurrence (Gumireddy et al., 2015). While further validation is necessary, AKAP4, straightforward test that has the potential to contribute to the advancement of more personalized patient care strategies. The findings mentioned above indicate that several CTAs may support tumour growth and also suggest that CTAs may be important in determining treatment responses to cytotoxic or growth inhibitory anti-cancer drugs. It is clear that any CTA that enhances cell survival may reduce the effectiveness of treatment with cytotoxic agents. This has been demonstrated with MAGE A, MAGE C, GAGE, PAGE-4, and CAGE proteins, which make cells resistant to DNA damage-inducing drugs commonly used in clinical settings like etoposide and paclitaxel (Atanackovic et al., 2010; Weeraratne et al., 2011). They have also been shown to confer resistance to other cytotoxic drugs used in cancer treatment. In breast cancer, MAGE A proteins have been implicated in the development of tamoxifen resistance, as they are up-regulated in tamoxifen-resistant clones. Knockdown of MAGE A2 has been found to sensitize cells to tamoxifen (Wong et al., 2014). Additionally, there is a significant association between MAGE A expression and reduced overall survival in oestrogen receptor-positive, tamoxifen-treated breast cancer patients (Wong et al., 2014). Therefore, CTAs may serve prognostic and predictive markers. The role of CTAs in tumour progression is

crucial, particularly in the context of metastasis. While these antigens are rarely expressed in benign neoplastic lesions, their expression becomes more frequent in primary melanoma and even more so in distant metastases (Lüftl et al., 2004). This suggests that CTAs may directly contribute to the complex process of metastasis, which involves various steps such as local invasion, intravasation, survival in circulation, extravasation, and colonization. One significant characteristic of cells with metastatic potential is their increased motility and invasive capabilities. Interestingly, several CTAs, including MAGE C2, GAGE, XAGE1, CAGE, and CT45A1, have been found to enhance these phenotypes (Caballero et al., 2013; Yang et al., 2014a). This implies that the presence of these antigens can promote the movement and invasiveness of cancer cells. Another important aspect of tumour progression is EMT, a process where epithelial cells lose their adhesion and polarity while gaining migratory and invasive properties similar to mesenchymal cells. EMT has been shown to play a crucial role in the metastatic progression of various cancers, particularly melanoma. Molecular analysis has revealed that the expression of MAGE C2 in breast cancer cells leads to signs of EMT such as reduced E-cadherin and cytokeratin levels, increased vimentin levels, and increased fibronectin levels (Yang et al., 2014a). Hence, CTA association with EMT further underscores their significance in facilitating the spread of cancer. Similarly, the proteins CAGE and CT45A1 have been shown to play a role in regulating the function of key proteins involved in EMT, such as beta-catenin, SNAIL, and TWIST (Kim et al., 2009b; Shang et al., 2014). This regulation may contribute to the development of a metastatic phenotype. In contrast to melanocyte differentiation antigens MART and GP100, which are down-regulated during EMT, CTAs like CT45A1 have been found to be up-regulated or unchanged in expression. This aligns with previous observations of increased CTA expression in metastatic cancers compared to primary cancers (Woods et al., 2014). It suggests that targeting CTAs could block metastatic progression or specifically target cancer cells in established metastatic lesions. Among the CTAs, GAGE proteins have been identified in both migrating primordial germ cells and trophoectodermal cells, which are known for their high motility and invasiveness (Gjerstorff et al., 2008). Knockdown experiments have shown that reducing GAGE protein levels significantly impairs migration and invasion of melanoma cells lines (Gjerstorff et al., 2008). Additionally, studies have found that GAGE proteins are highly up regulated in

metastatic clones of breast and gastric cancer models, further supporting their potential involvement in metastasis formation. However, direct evidence for the role of GAGE proteins in metastasis is still lacking (Lee et al., 2015). Therefore, it is crucial to further investigate the impact of GAGE proteins and other CTAs on the formation of metastases using *in-vivo* models.

The role of CTA in genomic instability is a subject of critical examination. Genomic instability in cancer cells leads to mutations, chromosomal rearrangements, and changes in chromosome number (Mecklenburg et al., 2017). These alterations contribute to the development of mutant genotypes that provide selective advantages to specific cell subclones, ultimately supporting the growth of tumours. While there is controversy surrounding the extent to which different mechanisms contribute to genomic instability, it is widely accepted that DNA double-strand breaks and abnormal segregation of chromatids during mitosis play a role (Keeney et al., 1997). Furthermore, the meiotic process involves the generation of double strand breaks during genetic material exchange between sister chromatids and the subsequent pairing and segregation of chromatids. This suggests that activation of meiotic programs in cancer cells may contribute to genome instability. Meiosis-specific CTAs, such as SPO11, SCP1, and HORMAD1 (Chen et al., 2005), have not identified in this dataset. SPO11 plays a crucial role in meiotic chromatid exchange by creating double-strand breaks. It is possible that SPO11 may promote chromosomal rearrangements in cancer cells through a similar mechanism. SCP1 and HORMAD1 are involved in chromosome pairing during meiosis, and their presence in somatic cells could disrupt normal mitotic processes. Interestingly, meiotic proteins have also been linked to reducing polyploidy in cancer cells (Kalejs et al., 2006). They may serve to maintain a balance between increased genome instability driving genetic variation and decreased genome instability necessary for the propagation of malignant clones. In conclusion, understanding the role of CTAs in genomic instability is vital to comprehend the mechanisms underlying tumour development. The involvement of meiotic proteins suggests potential avenues for therapeutic interventions aimed at targeting genomic instability in cancer cells. Several CTAs and other proteins that are predominantly expressed in the testis have been found to play a role in facilitating successful cell division in cancer cells. For example, when

FMR1NB, NXF2, STARD6, and FSIP1 were depleted in LC cells, it resulted in an increased occurrence of mitotic arrest and micronucleation when exposed to Paclitaxel or Nocodazole treatment, which induces mitotic stress (Cappell et al., 2012). This suggests that certain CTAs are crucial for maintaining accurate cell division and resistance to chemotherapy drugs in cancer cells. Interestingly, overexpression of NXF2, STARD6, FSIP1, and SSX2 CTA has also been linked to defective cell division and genomic instability (Whitehurst et al., 2010; Cappell et al., 2012). These findings indicate that multiple CTAs and testis proteins play important roles in delicate processes within cancer cells, and disruptions to these processes can lead to abnormal cell division (Cappell et al., 2012). Therefore, targeting these CTAs could potentially be a promising approach for anti-cancer therapy. Apart of oncogenic CTAs, several other CTAs may act as tumour suppressors being identified as having the ability to inhibit cancer cell proliferation, angiogenesis, and metastasis. One such CTA is testis-specific gene antigen 10 (TSGA10), which has been found to suppress cancer development in various malignant tumours (Amoorahim et al., 2020). TSGA10 achieves this by inhibiting the expression of HIF-1, reducing tumour cell metastatic capability, and decreasing metabolic activity in breast cancer. Additionally, TSGA10 has been shown to diminish the angiogenesis of human vascular endothelial cells. However, it is important to note that TSGA10 expression is significantly reduced in cancer patients. This downregulation of TSGA10 is associated with high levels of VEGF, which promotes tumour angiogenesis and cancer metastasis (Hoseinkhani et al., 2019). Furthermore, miR-10b-3p and miR-23a have been found to reduce the expression of TSGA10, thereby promoting cancer progression (Zhang et al., 2019). Furthermore, G-protein signalling 22 (RGS22) has been identified as a tumour suppressor in liver cancer, pancreatic ADC. It functions by inhibiting tumour cell invasion and metastasis (Hu et al., 2015). Similarly, MAGE-A4 has pro-apoptotic activity and promotes tumour suppression by binding to RING E3 ligases, p21Cip1, Miz1, and p53, leading to DNA damage. In low-invasive LC cells, the expression levels of sperm protein associated with the nucleus on the X-chromosome family members A (SPANXA) are significantly higher compared to high-invasive cells. Overexpression of SPANXA inhibits tumour cell invasion and metastasis both *in-vitro* and *in-vivo* by suppressing the c-JUN-SNAI2 axis and increasing E-cadherin levels (Hsiao et al., 2016). These findings suggest that CTAs have a potential role in tumour suppression.

However, more research is required to explore tumour-suppressive CTAs could provide a promising avenue for developing novel cancer therapeutics.

The failure of CTA therapy lies in its focus on targeting only a single antigen in anti-cancer vaccination trials (Walter et al., 2012). Recent data suggests that including several antigens in immunotherapy can have a greater clinical impact (Gordeeva, 2018). By co-targeting biologically connected proteins, such as multiple CTA antigens involved in cancer cell survival, in a multi-epitope setting, the magnitude and flexibility of the vaccine-induced anti-tumour response can be increased. This approach may also prevent tumour cells from escaping through the selective loss of single target antigens. While multiple CTA antigens are up regulated in many cancers, the amount of each antigen varies within individual cancer cells and between tumours of different patients with the same cancer type. Therefore, simultaneously targeting these proteins may be more effective than targeting them individually. However, it is currently unknown how the redundancy of many MAGE A proteins or additional co-expressed CTA antigens in regulating survival pathways may affect tumour responsiveness to anti-MAGE A3 treatment. To address this uncertainty, ongoing trials are evaluating the effect of vaccines targeting multiple antigens that affect cancer cell growth and survival. In a completed but unreported study, several MAGE family members (MAGE A1, MAGE A3, MAGE A4, MAGE A10, and MAGE C2) were simultaneously targeted in melanoma patients using a peptide vaccine (Gjerstorff et al., 2015). The outcome of these trials will provide valuable insights into the potential benefits and limitations of targeting multiple antigens in cancer immunotherapy.

Chapter 5: Analysis of the expression of leukaemia associated antigens (LAAs) in aB-ALL using SEREX

5.1 Introduction

This chapter focuses on the identification of LAAs that are recognised by B-ALL patient sera and may act as targets for immunotherapy. Current immunotherapy strategies have mostly been utilised to treat B-ALL by targeting CD19, CD20, and CD22 which are only expressed on B cells (pre-B-cells to plasma cells) and not on hematopoietic stem cells or other tissues (Malard & Mohty, 2020). Such therapy can kill not just malignant but also normal B cells, resulting in hypogammaglobulinemia, which can be treated with intravenous or subcutaneous immunoglobulin injections (Malard & Mohty, 2020). Although most B-ALL cases achieved first remission, treating relapsed cases remains a therapeutic challenge. As a result, innovative approaches that target the oncogenic programme behind eukaemia cells are of special interest. The SEREX technique was used to identify novel antigens expressed by B-ALL cells in adults and are recognised by autologous humoral responses. Using a testis cDNA library increased the chances of finding CTAs, providing a wider range of gene transcripts for immunoscreening due to global promoter hypomethylation in the spermatogonia (Fratta et al., 2011). CTAs are predominantly expressed in cancer cells and immunologically protected sites such as the testes and placenta (Fratta et al., 2011). As such they provide promising targets for immunotherapy as the immune response would only be stimulated to target diseased cells and avoid attacking the immune privileged healthy tissues.

5.2 Aims

To determine the reactivity of antibodies in aB-ALL patient sera against polypeptides present on the surface of a phage display library developed from testis tissue. To prioritise the identified genes from SEREX using criteria described by Cheever *et al.* 2009.

5.3 Results

5.3.1 Determining the recombinant rate

The library (labelled T2 28.07.11 Clear) was taken from -80^{o}C and defrosted. Storage for up to 6 months was at 4^{o}C. Chloroform was regularly added to the library so that it

formed a small visible phase beneath the phage library. Before use the library was vortexed and centrifuged for 2-3 mins at 300 x g. A range of volumes of the library were plated from 0.5 - 5µl (**Table 5.1**) to determine the optimal concentration required for complete coverage of the library and to assess the proportion of non-recombinant plaques generated. The final volume used was 1.8 µl of library as this gave individual plaques at the highest density to immunoscreen the library as rapidly as possible but dilute enough to maximise the identification of positive plaques. Blue/white colony screening (**Figure 5.1**) was used to show the recombination frequency of the library. The library labelled T2 28.07.11 Clear © was selected as it consistently had a higher proportion of phage containing cDNA inserts in the testis library.

Table 5.1 Identifying the optimal pfu density for plating cDNA testes library for SEREX

Library ID	**Volume (µl) per 90mm plate**	**Blue plaques (pfu)**	**White plaques (pfu)**	**Total plaques (pfu)**	**Recombination rate (% white colonies)**
T2 28.07.11 Clear	0	-	-	200	NA
T2 28.07.11 Clear ©	1	70	101	171	59.06
T2 28.07.11 Clear	2	390	824	1,214	67.87
T2 28.07.11 Clear	5	476	1,845	2,321	79.49

5.3.2 Primary and secondary immunoscreening

The primary screening identified immunoreactive plaques. These were isolated with two nearby non-reactive plaques and secondary screening was performed to confirm the positivity by virtue of having one third of the plaques appearing positive. Two-three positive plaques from each immunoreactive clone following secondary immunoscreening were isolated and eluted in SM buffer for PCR amplification **(Figure 5.2)**. 310 sero-positive plaques were identified by primary immunoscreening, 134 plaques were confirmed by secondary screening (**Table 5.2**) and these were sequenced.

Table 5.2 Total number of plaques and positive plaques identified during primary and secondary immunoscreening

Number of plaques screened from primary	**Number of potential positives identified**	**Confirmed positives plaques**	**Confirmed negative**
10^6	310	134	176

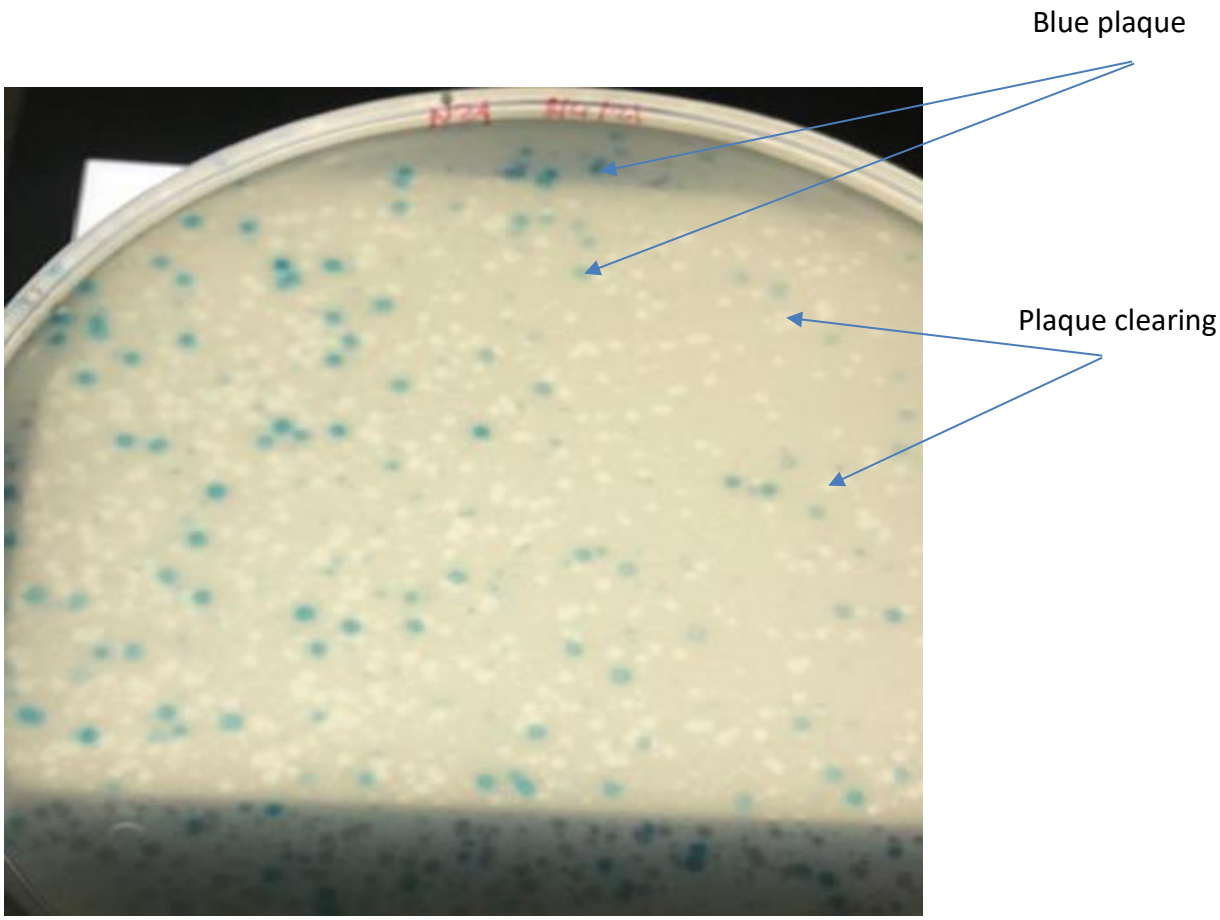


Figure 5.1 Recombination frequency was determined by plating different volumes of phage supernatant onto a 182mm petri dish

A lawn of *E.coli* was created in 0.7% NZY top agar which had been transfected with phage library at different pfus. Clear and blue plaques were counted and the recombination frequency was determined. Blue plaque where there is no cDNA insert and the X-galactosidase gene remains intact. Plaque clearing where bacteriophage have cleared the XL1 Blue MRF' *E.coli* as part of their lytic life cycle represents cDNA insert.

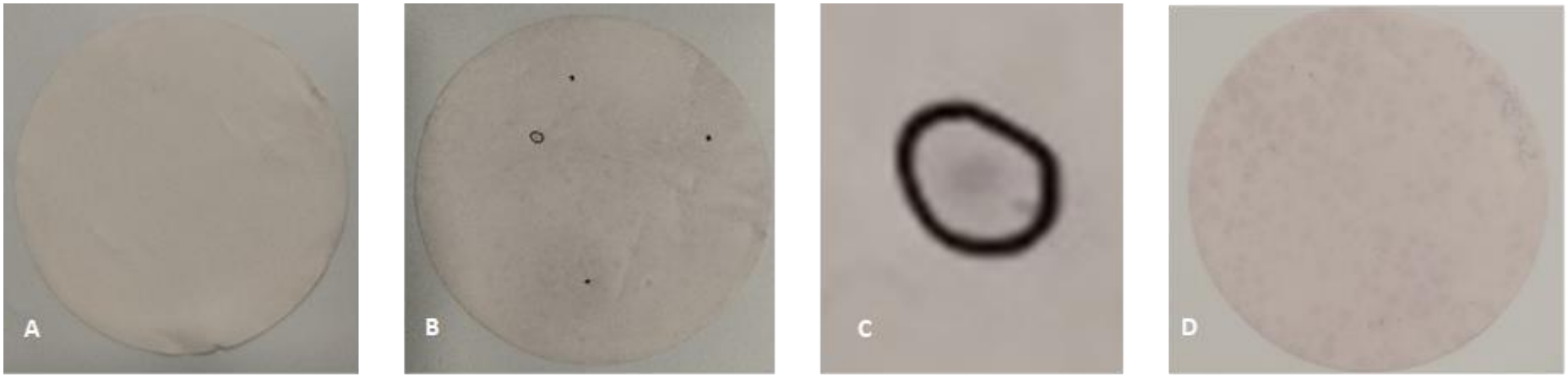


Figure 5.2 Representative membranes following primary and secondary immunoscreening

A were negative primary immunoscreening. B: shows one plaque that was positive on the nitrocellulose membrane primary immunoscreening. C: the plaque was cut out and subjected to secondary immunoscreening. D: represents the results secondary immunoscreening that was negative. E: represents a positive sencodary screening with almost ¼ of the plaques were found to be positive and were isolated for PCR

5.3.3 PCR Result

The plaques that were confirmed as containing cDNAs encoding for immunogenic epitopes during secondary immunoscreening were amplified using PCR (**Figure 5.3**). The predicted pBK-CMV multiple cloning site (MCS) size (when amplified with the T7-T3 primers) was 244bp, only bands larger than 44bp were excised to increase the likelihood that the MCS contained cDNA inserts.

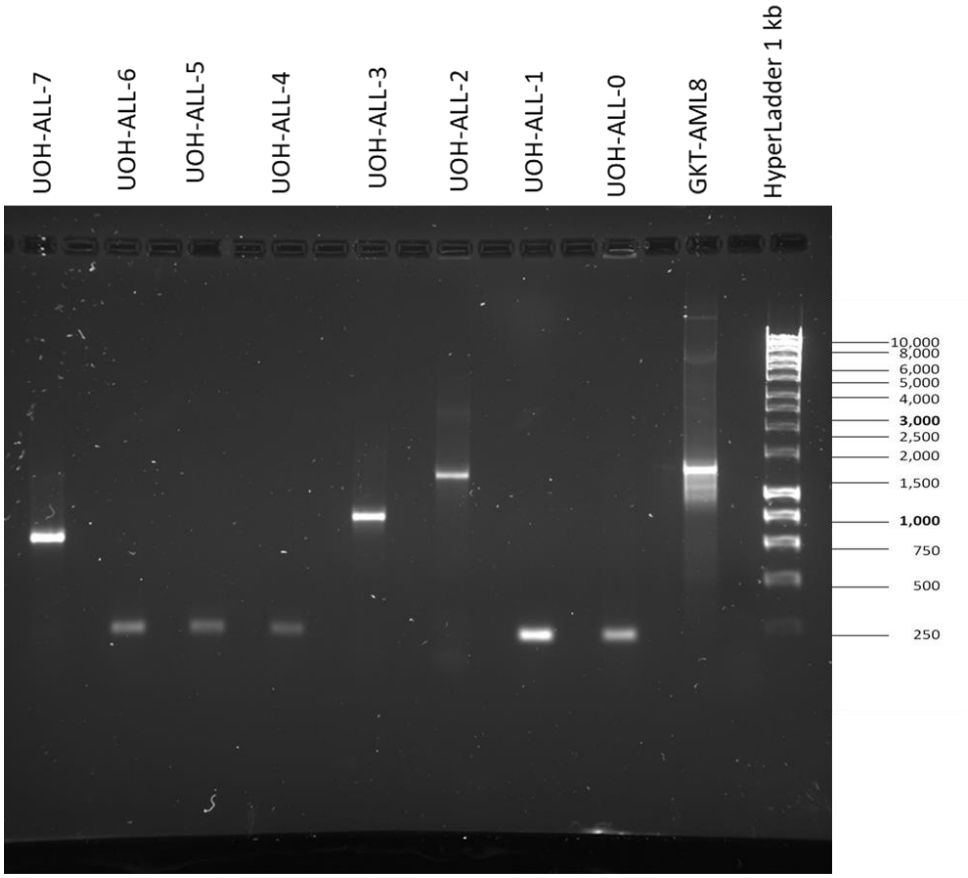


Figure 5.3 PCR for confirmed secondary plaques

The PCR products were isolated from the gel for purification and sequencing. Amplicons from UOH-ALL-2 around 1500bp and UOH-ALL-3 around 1000 bp and UOH-ALL-7 around 800 bp. UOHALL-0 is the empty MCS has an amplicon size of 244bp. UOHALL01 appears to be an empty vector. GKT-AML8, the positive control, by virtue of there being a 1.1Kb cDNA insert in the multiple cloning site was first identified by SEREX (Guinn et al, 2002) and renamed ZNF465 in according to HUGO convention when characterised (Collin et al., 2015).

5.3.3 Sequencing of cDNA inserts

The results of sequences were analysed and ensembled via NCBI using databases of BLASTN (for nucleotides and cDNA sequences). The identity of the cDNA insert was

made based on significant nucleotide identity from the result of BLASTN (Table 5.3). Sequencing identified 72 independent genes from 134 cDNA inserts; to date, three of them (UOH-ALL-104, UOH-ALL-105, UOH-ALL-106) remain novel and as yet do not correspond to annotated genes. The number of genes identified were smaller than number of plaques that means more than more plaques recognised the same gene. Antigens identified by SEREX included genes with different functions such as CUL1 and ROCK1, transcription factors such as ZNF676, signalling molecules such as DKK3, and RAB5C, enzymes such as IDI1, AASDH1 and adhesion molecules such as HNRNPLL. In addition, non-coding RNA such as LINC00261 was identified.

Table 5.3 Characterisation of cDNAs sequenced from phagemids following secondary immunoscreening

SEREX id	# times found *	Gene name	GenBank number Sequence ID	Percent of identity ≠	Size (bp)	Chromosomal localisation
UOH-ALL-1	1	Tropomyosin 3 (TPM3) transcript variant Tpm3.1, mRNA	NM_153649.4	100	3177	1q21.3
UOH-ALL-2	1	Transmembrane and coiled-coil domains 3 (TMCO3)/ "C13orf11"	NM_001014283.2	99	66237	13q34
UOH-ALL-3	1	Chromosome 10 open reading frame 82 (C10orf82), transcript variant 1, mRNA	NM_144661.4	100	990	10q25.3
UOH-ALL-4	1	Ubiquitin conjugating enzyme E2 D2 (UBE2D2), transcript variant 1	NM_003339.3	100	2530	5q31.2
UOH-ALL-5	1	Myosin light chain 12B (MYL12B), transcript variant 1	NM_001144944.1	100	1041	18p11.31
UOH-ALL-6	1	Translation machinery associated 7 homolog (TMA7), transcript variant 1, mRNA	NM_015933.6	97.75%	561	3p21.31
UOH-ALL-7	3	Protein tyrosine phosphatase non-receptor type 1 (PTPN1), transcript variant 1, mRNA,	NG_012119.2	97.12	81198	20q13.13
UOH-ALL-8	2	Isolate CHM13 chromosome 16	CP068262.2	97.34	963	16
UOH-ALL-9	1	col-3 genes for 18S rRNA, ITS1, 5.8S rRNA, ITS2, 28S rRNA	AB665697.1	87.08	690	-
UOH-ALL-10	2	Wnt2b mRNA for wingless-type MMTV integration site family member 2b, complete cds	AB205148.1	94.85	2427	-
UOH-ALL-11	1	Vacuolar H(+)-ATPase subunit	NM_015933.6	90.79	1078	-
UOH-ALL-12	1	Isolate BMF2-1/1 mitochondrion	KJ801454.1	97.20	16569	-
UOH-ALL-13	1	Isolate H7_91_CI-9_FU_P3S2_Heavy immunoglobulin variable region mRNA, partial cds; and IGHV3-15*07, IGHJ5*01, and IGHG2 mRNAs, complete sequence.	MW176666.1	97.16	1563	-
UOH-ALL-14	1	Chromosome 5 clone RP11-2O17	AC026746.6	95.62	179144	5
UOH-ALL-15	2	DNA methyltransferase 1 (DNMT1), RefSeqGene (LRG_362)	NG_028016.3	100	104941	19p13.2
UOH-ALL-16	1	Ataxin 10, mRNA (cDNA clone MGC:4152 IMAGE:3030062)	BC007508.2	99.70	1952	-
UOH-ALL-17	1	mRNA for Six2 protein	AJ316542.1	93.44	1254	-
UOH-ALL-18	1	Isolate 026 haplogroup N9 mitochondrion	MH553652.1	96.49	16570	-
UOH-ALL-19	1	Wnt ligand secretion mediator (WLS), transcript variant X2,	XM_003807869.4	100	2741	1
UOH-ALL-20	1	Tubulin alpha 3c (TUBA3C), mRNA Syno "bA408E5.3; TUBA2"	NM_006001.3	100	1521	13q11
UOH-ALL-21	1	Seizure related 6 homolog like 2 (SEZ6L2), transcript variant 2	NM_201575.4	100	3804	16p11.2
UOH-ALL-22	1	DNA, chromosome 18, nearly complete genome.	AP023478.1	100	77846715	18
UOH-ALL-23	1	Transcription factor 7 like 2 (TCF7L2), transcript variant X35, mRNA	XM_032177767.1	87.19	3876	10q25.2
UOH-ALL-24	2	Stromal RNA regulating factor mRNA, alternatively spliced (SRRF)	AY236962.1	97.87	1931	2p22.1

UOH-ALL-25	1	Myosin VI (MYO6), transcript variant 2, mRNA "DFNA22; DFNB37"	NM_001300899.2	100	8546	6q14.1
UOH-ALL-26	1	Haplogroup H-T152C mitochondrion	ON688208.1	100	16570	-
UOH-ALL-27	2	Filamin B (FLNB), transcript variant 4, mRNA, ABP-278; ABP-280; AOI; FH1; FLN-B; FLN1L; LRS1; SCT; TABP; TAP	NM_001164319.2	96.91%	9369	3p14.3
UOH-ALL-28	1	Chromosome 16 clone RP11-89D3, complete sequence	AC022168.6	96.93%	177322	16
UOH-ALL-29	1	TIMP metallopeptidase inhibitor 1 (TIMP1), mRNA	NM_003254.3	99.38%	769	Xp11.3
UOH-ALL-30		Atlastin GTPase 3 (ATL3), RefSeqGene on chromosome 11	NG_033985.1	99.80%	54893	11q13.1
UOH-ALL-31	1	cDNA DKFZp686O033 (from clone DKFZp686O033)	AL833551.1	97.41%	6513	3
UOH-ALL-32	2	Kinesin family member 5B (KIF5B), mRNA	NM_004521.3	98.64		
UOH-ALL-33	2	Aminoadipate-semialdehyde dehydrogenase (AASDH)-ACSF4; LYS2; NRPS1098; NRPS998"	NG_046885.1	99.12	58732	4q12
UOH-ALL-34	1	Translocase of outer mitochondrial membrane 20 like (TOMM20L), transcript variant X1, mRNA	XM_011536742.4	100	631	14
UOH-ALL-35	1	Nucleoside-triphosphatase, cancer-related (NTPCR)	NM_032324.3	97.94	6324	1q42.2
UOH-ALL-36	1	Keratin 31 (KRT31), mRNA	NM_002277.3	99.23	1615	17q21.2
UOH-ALL-37	2	Isolate E5_53_SR-6_LA_P2S23_Heavy immunoglobulin variable region mRNA, partial cds; and IGHV3-23*01, IGHJ5*02, and IGHG2 mRNAs	MW176520.1	98.33	1028	-
UOH-ALL-38	1	Isolate C114_Kampong_Cham_F1a1a1 mitochondrion, complete genome Sequence	KT587463.1	98.52	16567	-
UOH-ALL-39	1	TNF receptor associated protein 1 (TRAP1), transcript variant 2	NM_001272049.2	100	2064	16p13.3
UOH-ALL-40	1	Isolate ACAD11197 haplogroup A2 mitochondrion, complete genome Sequence	KU523267.1	99.77	16568	-
UOH-ALL-41	1	LJ45347 fis, clone BRHIP3011082, highly similar to ubiquitin-protein isopeptide ligase (E3)	AK127280.1	98.85	5196	-
UOH-ALL-42	2	Serine incorporator 3 (SERINC3), transcript variant 1, mRNA	NM_006811.4	100	4396	20q13.12
UOH-ALL-43	1	Testis expressed 43 (TEX43), mRNA- C5orf48; Tseg7	NM_207408.3	100	478	5q23.2
UOH-ALL-44	1	FER tyrosine kinase (FER)	NG_011445.2	99.69	456020	5q21.3
UOH-ALL-45	2	cDNA DKFZp686C15213 (from clone DKFZp686C15213) Sequence	BX640874.1	97.88	1732	-
UOH-ALL-46	1	MCL1 apoptosis regulator, BCL2 family member (MCL1), transcript variant 1, mRNA	NM_021960.5	99.37	3950	1q21.2
UOH-ALL-47	1	Ankyrin repeat domain 17 (ANKRD17), transcript variant 3, mRNA	NM_001286771.3	95.99	10358	4q13.3
UOH-ALL-48	1	cDNA fis, A-COL04217, highly similar to Homo sapiens mitochondrion, NADH dehydrogenase subunit 2 Sequence	AK026903.2	99.49	1054	-
UOH-ALL-49	1	Coiled-coil domain containing 89 (CCDC89)	NM_152723.3	95.8	2348	11q14.1
UOH-ALL-50	1	Ubiquitin C, mRNA (cDNA clone MGC:14624 IMAGE:4076286) (HMG20)	BC039193.1	97.73	2222	-
UOH-ALL-51	1	FLJ21112 fis (full insert sequence), clone CAS05418, highly similar to AF116692 Homo sapiens PRO2207 mRNA	AK024765.1	100	1405	-

UOH-ALL-52	3	Thymopoietin (TMPO), transcript variant 2	NM_001032283.3	93.96	4136	12q23.1
UOH-ALL-53	1	Homeobox D8 (HOXD8), transcript variant 3 (HOX4; HOX4E; HOX5.4)	NM_001199747.2	100	1597	2q31.1
UOH-ALL-54	1	Isolate PG0247 mitochondrion	MN687198.1	86.07	1656	-
UOH-ALL-55	2	Isolate CHM13 chromosome 15	CP068263.2	99.70	9975	15
UOH-ALL-56	1	Tumour protein, translationally-controlled 1 (TPT1), transcript variant 2	NM_003295.4	96.99	4548	13q14.13
UOH-ALL-57	1	MIP mRNA for major intrinsic protein	NM_012064.4	100	2614	12q13.3
UOH-ALL-58	1	Long intergenic non-protein coding RNA 661 (LINC00661)	NR_026828.1	86.07	3350	19p13.12
UOH-ALL-59	1	Prostate cancer associated transcript 6 (PCAT6), transcript variant 1	NR_046325.1	99.11	764	1q32.1
UOH-ALL-60	1	High density lipoprotein binding protein (HDLBP), transcript variant X14	XM_047444076.1	100	8546	2
UOH-ALL-61	1	Isolate 9_T haplogroup U2e1 mitochondrion	KY670862.1	99.25	16579	-
UOH-ALL-62	1	Sad1 and UNC84 domain containing 1 (SUN1), transcript variant 71	NM_001367708.1	99.10	3841	7p22.3
UOH-ALL-63	1	Myosin heavy chain 11 (MYH11), transcript variant SM2B	NM_001040113.2	94.27	6940	16p13.11
UOH-ALL-64	1	Clone VMRC53-239I22 from chromosome 1	AC275668.1	95.63	1677	1
UOH-ALL-65	1	Rho-associated coiled-coil containing protein kinase 1 (ROCK1),mRNA	NM_005406.3	91.01	9446	18q11.1
UOH-ALL-66	1	B9 domain containing 1 (B9D1), transcript variant 2, mRNA	NM_015681.6	99.86	913	17p11.2
UOH-ALL-67	1	RAB5C, member RAS oncogene family (RAB5C), transcript variant 1, mRNA	NM_201434.3	99.83	1757	17q21.2
UOH-ALL-68	1	RAB34, member RAS oncogene family (RAB34), transcript variant 6, mRNA	NM_001256277.2	97	1344	17q11.2
UOH-ALL-69	2	Adaptor protein, phosphotyrosine interacting with PH domain and leucine zipper 1 (APPL1), mRNA	NM_012096.3	99.42	6069	3p14.3
UOH-ALL-70	1	mRNA for hypothetical protein, complete cds	AB353305.1	100	1431	-
UOH-ALL-71	1	Chloride intracellular channel 4 (CLIC4), mRNA	NM_013943.3	99.48	4253	1p36.11
UOH-ALL-72	1	TNF receptor associated protein 1 (TRAP1), transcript variant 2, mRNA	NM_001272049.2	96.34	2064	16p13.3
UOH-ALL-73	1	Isopentenyl-diphosphate delta isomerase 1 (IDI1), transcript variant 1, mRNA	NM_004508.4	97.76	2739	10p15.3
UOH-ALL-74	1	Isolate 167 mitochondrion, complete genome	MF437201.1	98	585	-
UOH-ALL-75	1	Isolate HGDP00779 mitochondrion, complete genome	KF451299.1	80.45	1050	-
UOH-ALL-76	1	TIMP metallopeptidase inhibitor 1, mRNA	BC007097.1	91.67	841	-
UOH-ALL-77	1	Ubiquitin protein ligase E3C (UBE3C), mRNA	NM_014671.3	83.69	5214	7q36.3
UOH-ALL-78	1	Zinc finger protein 676 (ZNF676), transcript variant X1	XM_047438352.1	92.45	2859	19p12
UOH-ALL-79	1	PTPN23 transcript variant 2, mRNA	NM_001304482.2	97.84	5109	3p21.31
UOH-ALL-80	1	Surfeit 2 (SURF2), transcript variant 1, mRNA	NM_017503.5	90.69	833	9q34.2
UOH-ALL-81	1	KIF1Bbeta mRNA for kinesin family member 1Bbeta isoform IV, (KIF1B)	AB088213.1	96.09	8572	1p36.22
UOH-ALL-82	1	Testis expressed 38 (TEX38), transcript variant 3, mRNA	NM_001300864.2	90.58	689	1p33
UOH-ALL-83	2	Tubulin polymerization-promoting protein family member 2(TPPP2), mRNA (cDNA clone MGC:47825 IMAGE:5169309), complete cds	BC038970.2	90.86	791	14q11.2

UOH-ALL-84	2	BCL2-associated transcription factor 1 (BCLAF1), mRNA (cDNA clone IMAGE:6501081), partial cds	BC063846.1	95.58	2277	6q23.3
UOH-ALL-85	1	Epididymal protein pseudogene (LOC338963), non-coding RNA	NR_034139.1	94.36	1381	15q25.2
UOH-ALL-86	2	F-box protein 22 (FBXO22), mRNA (cDNA clone IMAGE:3349806)	BC008762.1	94.56	3768	15q24.2
UOH-ALL-87	1	UVRAG divergent transcript (UVRAG-DT), long non-coding RNA	NR_144531.1	92.79	689	11q13.5
UOH-ALL-88	1	A-kinase anchoring protein 1 (AKAP1), transcript variant X8, mRNA	XM_047436869.1	88.74	2081	17q22
UOH-ALL-89	2	Eukaryotic translation initiation factor 3, subunit A, mRNA (cDNA clone IMAGE:4519449) (EIF3A)	BC020285	99.79	1721	-
UOH-ALL-90	1	MYH10 variant protein, clone: hf00991	AB210026.1	100	7693	-
UOH-ALL-91	1	Clone FLB8503 PRO2286 mRNA, complete cds	AF130085.2	94.77	443	-
UOH-ALL-92	1	Ribosomal protein L28, mRNA (cDNA clone MGC:20081 IMAGE:4054251), complete cds (RPL28)	BC010173.2	99.40	511	-
UOH-ALL-93	1	Morf4 family associated protein 1 (MRFAP1), transcript variant 2	NM_001272053.2	97.46	1455	4p16.1
UOH-ALL-94	1	Isolate PNG53 haplogroup P2 mitochondrion, complete genome	MN849846.1	100	16568	-
UOH-ALL-95	1	Protein tyrosine phosphatase, non-receptor type 23, mRNA (cDNA clone IMAGE:4111514), partial cds	BC027711.2	99.79	4234	-
UOH-ALL-96	1	FLJ21112 fis, clone CAS05418, highly similar to AF116692 Homo sapiens PRO2207 mRNA (PARL)	AK024765.1	100	1405	-
UOH-ALL-97	1	Cullin 1 (CUL1), transcript variant 3, mRNA	NM_001370661.1	98.86	3198	7q36.1
UOH-ALL-98	1	Long intergenic non-protein coding RNA 251 (LINC00251), long non-coding RNA (C8orf25)	NR_038901.1	100	1058	8q13.1
UOH-ALL-99	1	Chromosome 18 open reading frame 32 (C18orf32), transcript variant 2, mRNA (GPIBD25)	NM_001035005.4	100	5548	18q21.1
UOH-ALL-100	1	Ribosomal protein S15, mRNA (cDNA clone MGC:70657 IMAGE:6053159), complete cds (RIG, DKK3)	BC064908.1	100	539	11p15.3
UOH-ALL-101	2	Testis specific protamine 1 (P1) mRNA, complete cds (PRM1)	AY651260.1	100	385	16p13.2
UOH-ALL-102	1	cDNA clone IMAGE:5802820	BC042915.1	100	1934	-
UOH-ALL-103	1	Clone FLB8503 PRO2286 mRNA, complete cds	AF130085.2	100	2359	-
UOH-ALL-104	1	-	-	-	917	-
UOH-ALL-105	1	-	-	-	700	-
UOH-ALL-106	1	-	-	-	600	-

*: found when immunoscreening the testes cDNA library; ≠: with genes from NCBI Blast

5.3.4 Analysis of genes identified using Bloodspot

To understand the role of the identified genes in cancer (Table 5.4), their function and distribution in normal tissue was determined through a comprehensive search of the literature. Interestingly, UOH-ALL-3, UOH-ALL-20, UOH-ALL-83, and UOH-ALL-101 were identified as corresponding to the following CTAs - C10orf82, TUBA3C, CT152 (TPPP2) and PRM1 while three UOH-ALL-58, UOH-ALL-85, and UOH-ALL-98 were non-coding RNAs (LINC00661, LOC338963, and LINC00251) that had testis-restricted expression. 69/72 of the genes identified had already been shown to be involved in tumour pathogenesis. There were five genes with no literature on their role in cancer. These were TEX43, CCDC89, LOC338963, C18orf32 and ATL3. LOC338963, TEX43 and CCDC89 were highly enriched in the testis and involved in spermatogenesis while their functions in cancer are poorly understood.

The SEREX-identified genes were examined for their expression in different B-ALL subtypes in comparison to healthy bone marrow using the MILE study dataset (Kohlmann et al., 2008). Upregulated genes (SIX2, ATL3, DNMT1, ANKRD17; $p<0.001$) and downregulated ones (WLS, KIF1B, HNRPLL, TMCO3; $p<0.05$) were identified (Table 5.5). The association with survival was determined for each gene. 10 genes were found to be associated with survival with a p value of<0.05 (**Table 5.4**).

5.3.5 Assessment of SEREX-associated B-ALL LAAs as targets for cancer vaccines

Cheever *et al.* described the prioritisation of antigens based on prior information focussing on the capacity of an antigen to be translated into clinical studies using pre-weighted criteria. This provided a significant advantage in the ranking to antigens that met those criteria. Here we implemented the criteria described by Cheever *et al.* (see section 1.5) and used it to prioritise the antigens identified in aB-ALL by SEREX for their potential to act as targets for immunotherapy using cancer vaccines. The top genes **(Figure 5.4)** with the highest score out of 1.0 were (in descending order) UOH-ALL-65, UOH-ALL-84, UOH-ALL-100, UOH-ALL-79, UOH-ALL-92 corresponding to ROCK1 (0.41), BCLAF1 (0.36), DKK3 (0.32), PTPN23 (0.32) and RPL28 (0.30), respectively. These LAAs had a low ranking compared to WT1 (0.81) and BIRC5 (0.55). The highest scoring

identified SEREX genes have been involved in molecular functions mainly involving binding and catalytic activity **(Figure 5.5)**.

Table 5.4 The association between gene expression and survival in B-ALL compared to healthy bone marrow using MILE study (Kohlmann et al., 2008)

Antigen	Probe set	Highest expression	Survival
CUL1	238509-at	ALL with t(1;19)/ ALL hyperdiploid - p<0.01	0.0422
	207614-s-at	ALL with t(1;19)/ ALL hyperdiploid - p<0.01	0.0161
DKK3	202196-s-at	ALL t(8;14)/ Pre-B-ALL no t(9;22) -NS	0.0233
HMG20	211296-x-at	Pre-B-ALL t(9;22) / Pre-B-ALL no t(9;22) -NS	0.0002
IDI1	208881-x-at	All types of B-ALL p<0.01 except ALL t(8;14)- NS	0.0007
	204615-x-at	All types of B-ALL p<0.01 except ALL t(8;14)- NS	p<0.0001
KIF5B	201992-s-at	ALL t(12;21)/ Pre-B-ALL no t(9;22)/hyperdiploid p<0.001	0.0066
PARL	218271-s-at	Healthy bone marrow- p<0.01	0.0017
PTPN23	223149-s-at	All types of B-ALL p<0.001	0.0378
ROCK1	213044-at	All types of B-ALL p<0.05 except ALL t(8;14)/ ALL t(1;19)/ Pre-B-ALL t(9;22)- NS	0.0014
	214578-s-at	All types of B-ALL p<0.05 except ALL t(8;14)/ ALL t(1;19)/ Pre-B-ALL t(9;22)- NS	0.0057
SERINC3	221472-at	Pro-B-ALL t(11q23)/MLL/ Pre-B-ALL t(9;22) p<0.05	0.0246
WLS	228949-at	Healthy bone marrow- p<0.01	0.0045

The identified SEREX genes (TCF7L2, WNT2B, DMNT1, MCL1, TRAP1, ROCK1, and CUL1) were involved in mostly the Wiki pathway (Table 5.5) as shown by Enrichr. In addition, the two envelope proteins (TCF7L2, WNT2B) that are essential to maintaining the structure and function of the nucleus mainly correlated with Emery-Dreifuss muscular dystrophy (EDMD) (Kuleshov et al., 2016). From Table 5.5 MSigDB Hallmark 2020 database, the mitotic spindle had a high p-value of 0.001 as this pathway plays a crucial role in cell division and chromosomal separation. The mitotic spindle had a significant the q-value (an adjusted p-value) taken into consideration the FDR which is significant <0.05. The Cadherin pathway was detected as being involved with a high proportion of the identified UOH-ALL antigens (14.3%) when using PANTHER.db (**Figure 5.6**). From STRING analysis (**Figure 5.7**), SEREX identified genes that were involved in ubiquitination included UBC, UBE2C, UBE2D2, CUL1, RBX1 and SKP2. Genes involved in regulating the intrinsic apoptosis process were BCLAF2, PARL, SERINC3, TRAP1, TPT1, MCL1 and PTPN1. Genes involved in formation of a pool of free 40S subunits, and elongation factors included EIF5, EIF3A, EIF3D, C18orf32, RPL28 and TPT1

Table 5.5 Pathways identified involving members of the SEREX identified genes

Term	p-value	q-value	Overlap_genes	Pathway
Envelope proteins and their potential roles in EDMD physiopathology	0.0007	0.0892	TCF7L2, WNT2B	Wiki Pathway
Hematopoietic Stem Cell Gene Regulation	0.001	0.064	DNMT1, MCL1	
TGF-beta Signalling Pathway	0.015	0.229	TRAP1, ROCK1, CUL1	
Mitotic Spindle	0.001	0.038	ROCK1, KIF5B, FLNB, KIF1B, MYH10	MSigDB Hallmark 2020
Notch Signalling	0.007	0.117	TCF7L2, CUL1	
Apoptosis	0.026	0.286	ROCK1, TIMP1, MCL1	

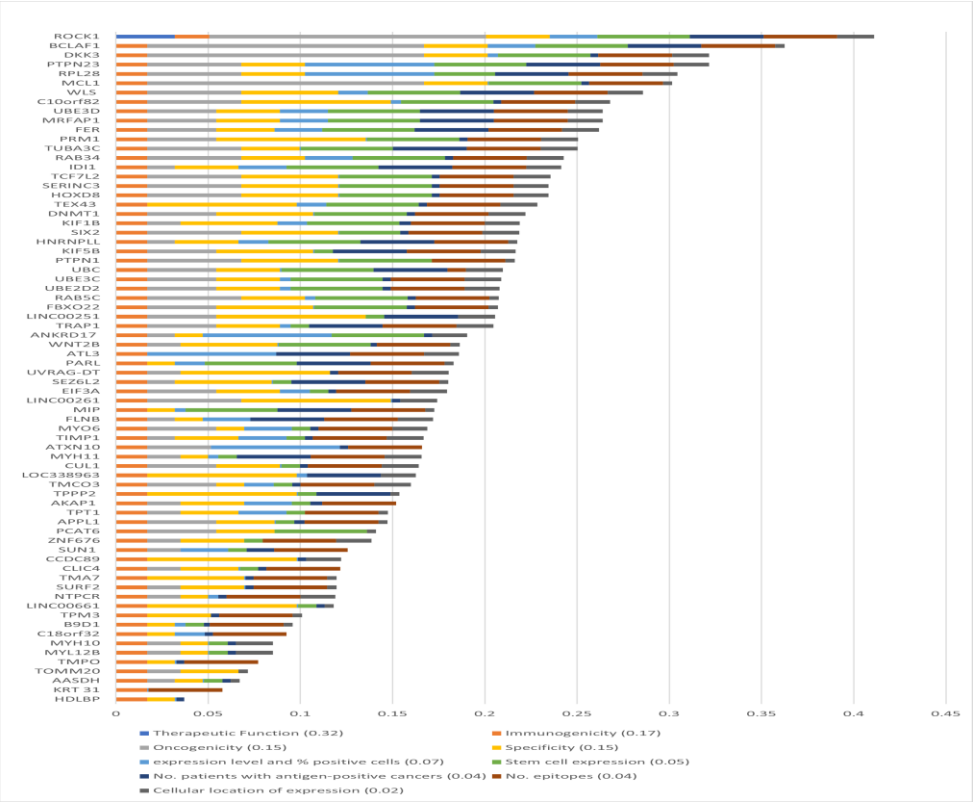


Figure 5.4 Ranking of tumour antigens based on their potential to act as vaccine targets

72 genes identified from SEREX were prioritised by Cheever criteria based on therapeutic function (blue) accounting 0.32 followed immunogenicity (0.17 indicated by orange). The oncogenicity (grey) and the specificity (yellow) account of 0.15. The key shows the evaluation criteria and the maximum attainable score in parenthesis. ROCK1(0.41), BCLAF1 (0.36), DKK3 (0.32), PTPN23 (0.32), and RPL28 (0.30) had the highest accumulated score and as such show the most promise as vaccine targets.

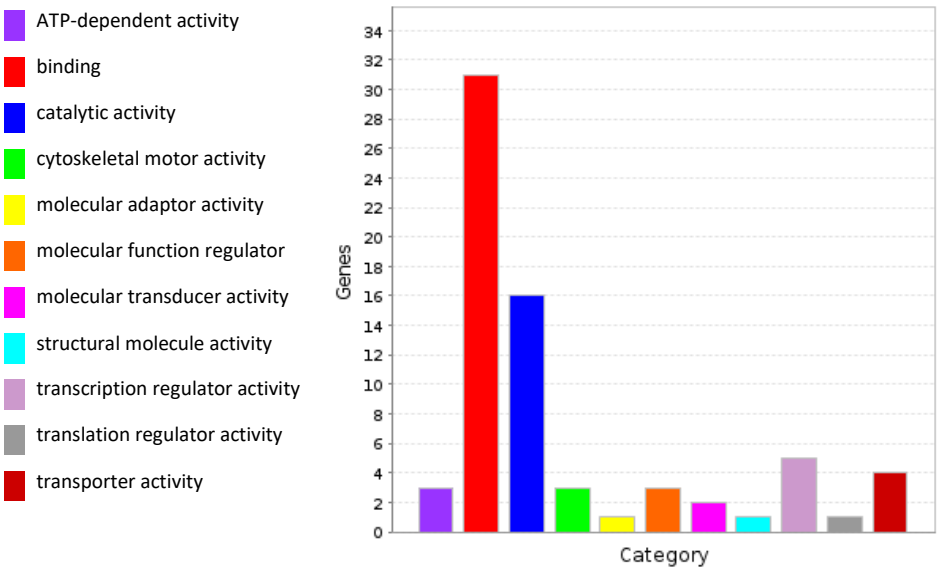


Figure 5.5 The molecular functions of identified LAAs

Using PANTHER.db different molecular functions identified by colour code, around 31% of 72 SEREX identified LAAs were involved in binding (indicated by the red bar). The blue bar shows 16% of LAAs were involved in catalytic functions. 5% of LAAs were involved in transcription regulator activity (lilac bar).

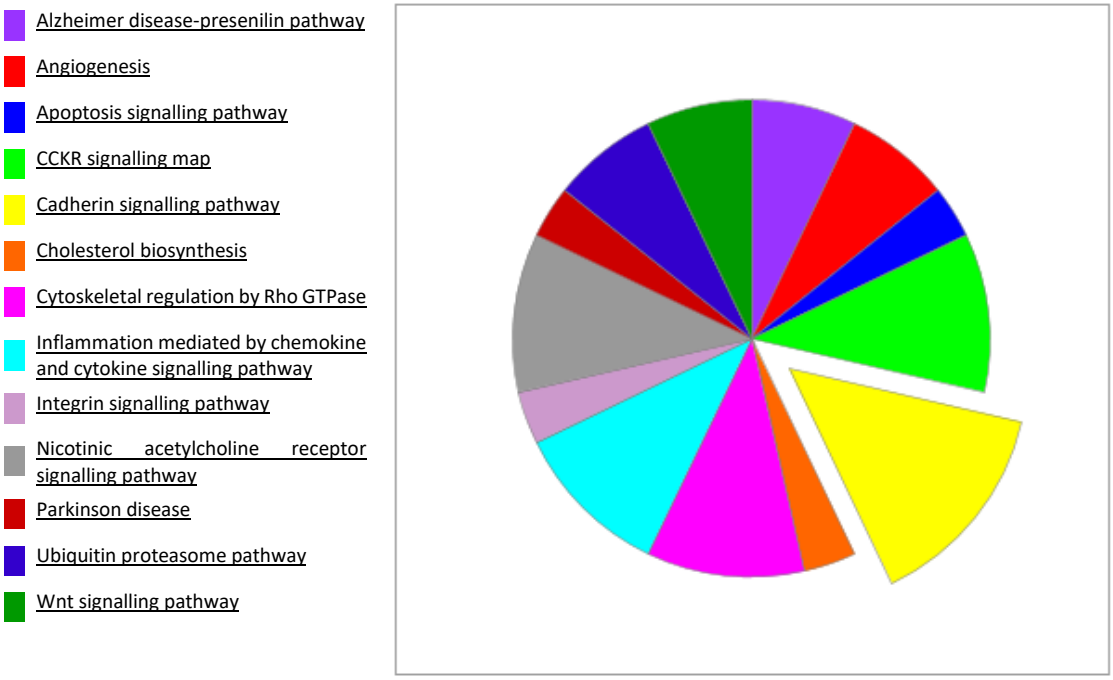


Figure 5.6 LAAs and their pathways involvement

From PANTHER.db, different pathways identified by colour code using 72 of identified SEREX. The most enriched pathway was the Cadherin signalling pathway (14.3%) and the participant genes included PTPN1, TCF7L2, FER and WNT2B (indicated by yellow section). AKAP1, ROCK1, MCL1 are involved in Cytoskeletal regulation by Rho GTPase (lilac), inflammation (light blue)

mediated by chemokine and cytokine signalling, and CCKR signalling (light green) that are secondly enriched weighed 10.3% as well as nicotinic acetylcholine receptor signalling (MYO6, MYH10, MYH11) indicated by grey section.

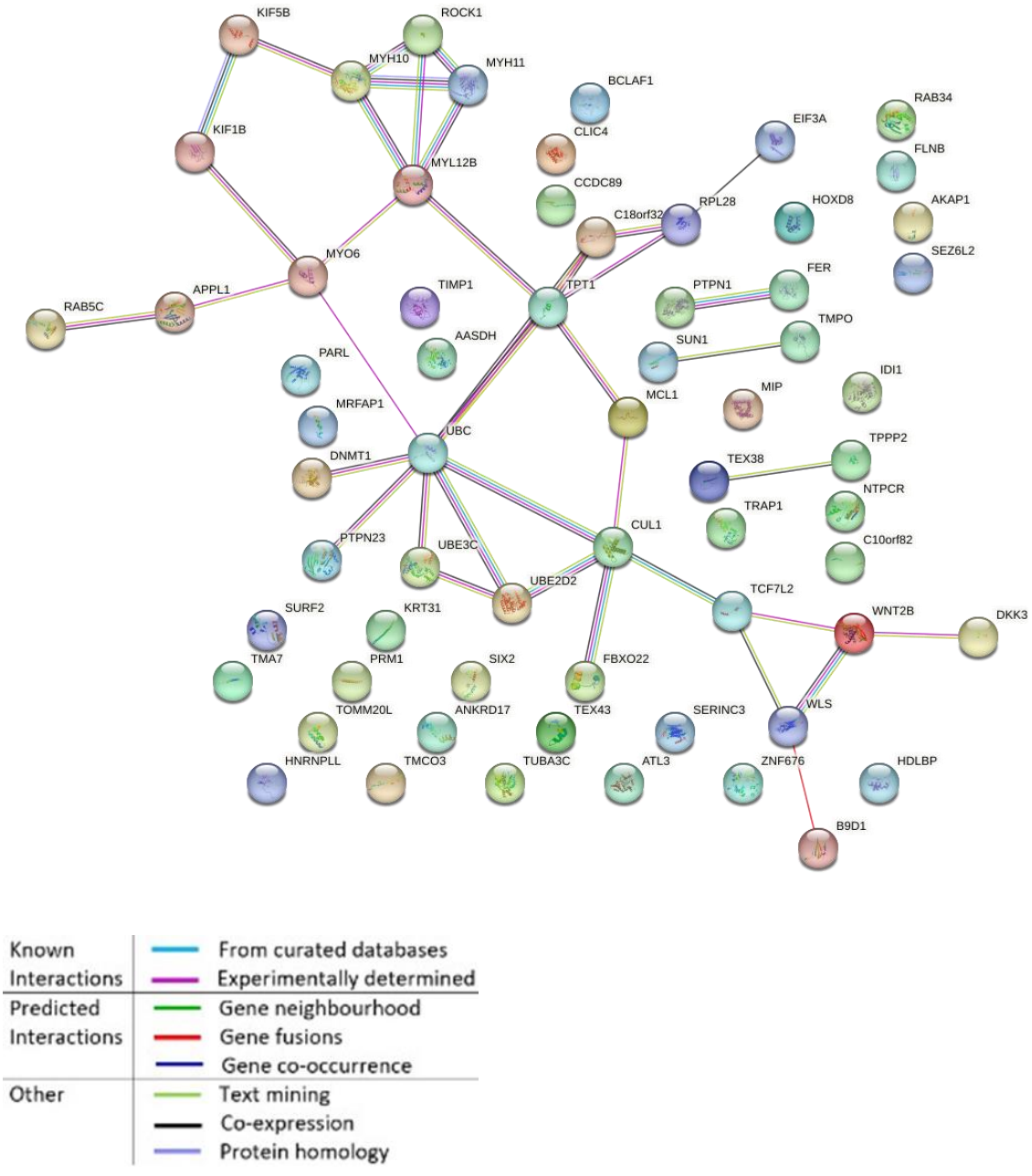


Figure 5.7 Protein-protein interaction of LAAs using STRING

The key shows different sources of interactions identified by curated databases and experiments in aqua and purple colour clustered as known interactions using STRING database. Clusters of proteins involved in ubiquitination (UBC, UBE2C, UBE2D2, CUL1, RBX1 and SKP2), intrinsic apoptosis (BCLAF2, PARL, SERINC3, TRAP1, TPT1, MCL1 and PTPN1) and the formation of a pool

of free 40S subunits, and elongation factors (EIF5, EIF3A, EIF3D, C18orf32, RPL28 and TPT1) are shown. The lack of interactions among other genes in this signature are likely explained by a variety of biological pathways utilised within B-ALL.

5.4 Discussion

72 out of 134 clones identified through the immunoscreening of a testis cDNA library with sera from five aB-ALL patients corresponded to known genes with different functions of cellular components. 59 of the UOH-ALL sequences were previously uncharacterised providing a novel source of genes with unknown function for future analysis outside the scope of this PhD. Three of the SEREX identified antigens (UOH-ALL-104, UOH-ALL-105, UOH-ALL-106) were novel and were identified by Sara Goodman (2020, MSc) but were not identified as ORFs via sequencing analysis previously. SEREX is widely used as a source of new antigens discovery with one third of identified clones corresponded to uncharacterised genes (Scanlan et al., 2002). Our results identified a high percentage (44%) of uncharacterised genes reflecting the heterogeneity of B-ALL. The number of identified antigens was less than the number of positive clones as the same genes were identified by different clones. In addition, some antigens may reflect recombinant β-gal fusion proteins expressed in lytic plaques by E. coli, possibly with a native conformation that are recognised by human sera (Seliger & Kellner, 2002).

CTAs represent an attractive target for immunotherapy due to their restricted expression in immune-privileged sites in healthy tissues, while being expressed in a range of cancers. Four UOH-ALL-3, UOH-ALL-20, UOH-ALL-83, and UOH-ALL-101 corresponded to (TUBA3C, C10orf82, CT152, and PRM1), each of which are CTAs. The presence of autoantibodies suggested a reactive immune response in some of patients with B-ALL and may be a therapeutic target for B-ALL. The MILE study was used to identify CTA expression in B-ALL compared to the normal healthy bone marrow, PRM1 and CT152 were mainly found in ALL t (8;14) without any significant expression compared to normal bone marrow as well as TUBA3C is found in Pro-B-ALL t(11q23)/MLL. While C10orf82 was downregulated in B-ALL (high expression in healthy bone marrow p-NS). Boullosa *et al.* investigated the following CTAs (HAGE, NY-ESO-1, SSX2 and PASD1) in aB-ALL that found SSX2, the binding partner of SSX2IP have been only detected by ICC but not RT-qPCR (Boullosa et al., 2018). This suggests that CTAs are

not widely expressed in B-ALL and their targets may not have clinical benefit. A broad expression of CTA in B-ALL subtypes was not fully elicited. Future studies are required to identify which subtype of aB-ALL may have higher expression of CTA as well as to identify their effect if they targeted.

PRM1 transcripts were detected in 11/41 (26.8%) of CLL patients and around 50% of patients had a high titre of PRM1 antibodies (P = 0·0001). These were not found in healthy donor controls (Meklat et al., 2009). C10orf82 is a CTA whose overexpression has been found to be associated with the overall survival in ovarian cancer (Almutairi et al., 2022). Another CTA, TUBA3C is upregulated in breast cancer and is a good biomarker for taxane sensitivity as its level decreases with chemo-resistances (Nami & Wang, 2018). CT152 (TPPP2) is found in hepatocellular carcinoma and may play role in cancer progression and metastasis (Xu et al., 2023).

Gene hypermethylation in hematologic neoplasia, with some lineage specificity, provides a strong rationale for developing demethylating agents for the treatment of these diseases. Demethylated agents such as decitabine (DEC) have shown a transient CR in 23% patients with cB-ALL in phase I trial. DEC combined with chemotherapy has led to CR in 52% of relapsed cB-ALL (Yanez et al., 2009). However, Roolf *et al.* found that DEC delayed leukaemic proliferation in an xenograft model and did not eradicate B-ALL (Roolf et al., 2018) as well as DEC is more effective in B-ALL with MLL rearrangements. Further investigation is required to validate these results and identify the B-ALL subtype that would benefit most from this therapy. DEC is used to treat patients with AML and its mechanism of action depends on upregulating p73, an upstream regulator of p21 and leads to its reactivation (Hoang & Rui, 2020). It increases SOCS3 expression inhibiting JAK/STAT signalling in AML. The use of demethylated agents synergises with cytotoxicity of CTA immunotherapy, as DEC combined with CTA-specific immunotherapy (Kang et al., 2022) induces high expression of NY-ESO-1 promoting its effect in an AML xenograft mouse model. CTA-specific immunotherapy may be combined with demethylating agents to manage B-ALL.

Cheever *et al.* provided a criteria to prioritise antigens for immunotherapy. It was based on a mathematical model ranking TAs based on a predefined criteria (See Section 1.8) to facilitate their transition into clinical trials. The top TAs based on Cheever criteria are

UOH-ALL-65, UOH-ALL-84, UOH-ALL-100, UOH-ALL-79 and UOH-ALL-92 corresponding to ROCK1 followed by BCLAF1, DKK3, PTPN23 and RPLA28 respectively. ROCK1 encodes a serine/threonine kinase protein that binds to a GTP-bound form of Rho upon activation (Julian & Olson, 2014). It functions in regulating cell polarity and cytoskeletal organization, adhesion and motility. It is expressed in bone marrow, lung, urinary bladder and adipose tissues (Julian & Olson, 2014). ROCK1 is involved in regulating normal haematopoiesis via negatively regulating erythropoietic stress and inflammation (Mali et al., 2014). Our finding that ROCK1 was recognised by B-ALL sera suggests a role in B-ALL pathogenesis. As ROCK1 (Hu et al., 2019) is a well-recognised oncogene in NSCLC and AML (Liu et al., 2019), understanding its role in different cancers may assist in predicting its function in B-ALL. ROCK1 (Figure 5.8) promotes NSCLC tumour growth, proliferation, and angiogenesis by inhibiting the PTEN/PI3K/FAK pathway (Hu et al., 2019). Furthermore, ROCK1 acts as non-canonical Wnt activator promoting breast cancer proliferation, stemness and migration (Mohammadi-Yeganeh et al., 2016). ROCK1 was found to be overexpressed in AML and was associated with poor survival ($p < 0.01$). ROCK1 knockdown of Kasumi-1 and AML-193 cell lines enhanced tumour cell apoptosis and significantly inhibited blast proliferation ($P < 0.05$). It has been suggested that miR340-5p downregulation is the main cause of ROCK1 upregulation in AML (Liu et al., 2019).

ROCK1 may be an attractive oncogene for treatment and different inhibitors have been investigated, such as GSK269962A (Pan et al., 2022), a selective ROCK inhibitor, which had inhibited tumour growth, apoptosis and clonogenicity in an AML mouse model, significantly prolonging survival. GSK269962A has antiproliferative activity through inhibition of cell cycle kinase (CDK6), inducing cell cycle arrest. It also induced apoptosis via increased expression and phosphorylation of p53 as well as decreasing the expression of antiapoptotic genes including survivin, Bcl-xL, and induced cleavage of PARP in AML (Pan et al., 2022). Interestingly, some treated mice remained disease-free for at least 140 days. ROCK1 knockout in is involved in AML mouse model enabled survival for more than 4 months (Pan et al., 2022). Taking all results together, ROCK1 may be a key therapeutic target in AML playing an important role in AML pathogenesis. To our date, ROCK1 expression is not verified in B-ALL and so we will examine its

expression in our available B-ALL samples (see section 6.4) to identify new targets for B-ALL.

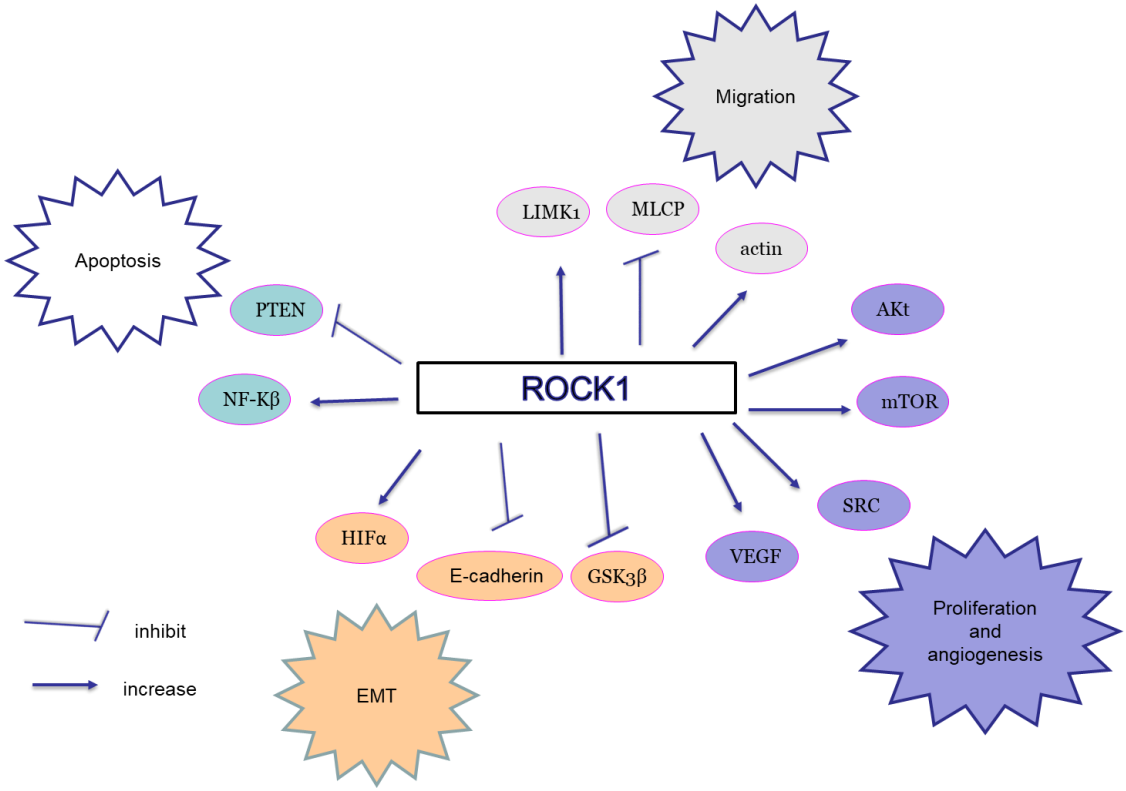


Figure 5.8 ROCK1 pathways involved in carcinogenesis showing cancer hallmarks

ROCK1 functions as an oncogene, promoting cell proliferation, migration and invasion as well as inhibiting apoptosis. ROCK1 supports migration and actin cytoskeleton reorganisation by activating RhoA/ROCK pathway downstream targets such as MLC, LIMK and cofilin. It also promotes tumour growth by activating the PI3K/Akt/mTOR pathway (Julian & Olson, 2014; Hu et al., 2019). ROCK1 is involved in maintaining Wnt activation by inhibiting GSK3 and NF-kB activity and interacting with TAK1-binding protein 2 (TAB2), thereby boosting tumour survival (Mohammadi-Yeganeh et al., 2016). ROCK1 regulates several cellular pathways implicated in carcinogenesis, making it an appealing therapeutic target for cancer treatment methods.

The second TA based on Cheever criteria prioritisation is BCLAF1 which is a nuclear protein that interacts with BCL2 and induces apoptosis activating p53. BCLAF1 has a dual function in carcinogenesis depending on the tumour type (White et al., 2018). BCLAF1 acts as an oncogene and is upregulated in AML and associated with shorter patient survival. miR-194-5p binds to the 3' UTR of BCLAF1 and inhibits its translation and consequent blast differentiation (White et al., 2018). While BCLAF1 acted as an oncogene in AML (White et al., 2018), it showed the activity of a tumour suppressor in

multiple myeloma (MM) and diffuse large B-cell lymphoma (DLBCL) inducing apoptosis of tumour cells via NF-κB inhibition (Figure 5.9). BCLAF1 triggers the pro-apoptotic proteins Bak and Bax and inhibiting Bcl2 (antiapoptotic protein). It also promotes DNA repair by interacting with DDR proteins such as BRCA1 and is implicated in RNA splicing regulation. BCLAF1 was identified by our SEREX library screen and its expression from the MILE study to be upregulated in B-ALL subtypes ($p<0.05$), except ALL t(12;21) and ALL t(1;19), suggesting that it may act as oncogene in those subtypes. However, BCLAF1 may act as a tumour suppressor in ALL t(12;21) and ALL t(1;19), and verification of BCLAF1 expression may be of interest to determine its function, providing elucidation.

BCLAF1 emerges as a compelling target for novel cancer therapies, given its ability to promote apoptosis, facilitate DNA repair, and regulate RNA splicing (Al-Odat et al., 2022). The upregulation of BCLAF1 has been observed following the administration of histone deacetylase inhibitors (HDACi) such as LMK-235 in diffuse large B-cell lymphoma (DLBCL), where it inhibits NF-κB and enhances the apoptotic rate of malignant B-cells (Jariwala et al., 2019). By investigating BCLAF1 expression and functionality in B-ALL, it may serve as a potential target in B-ALL, warranting further assessment of its efficacy.

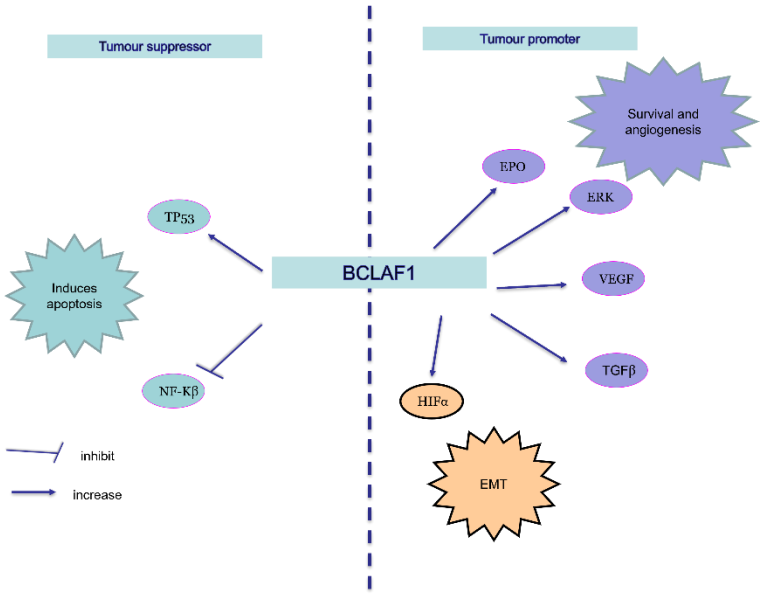


Figure 5.9 BCLAF1 key interactions and pathways involved tumourigenesis

BCLAF1 has emerged as a novel key player in cancer, with dual functions based on the type of cancer. It promoted carcinogenesis via increasing EPO, ERK, VEGF, TGF and HIF transcripts levels (Wen et al., 2019), thereby increasing tumour survival and

proliferation. However, BCLAF1 can also act as a tumour suppressor inducing apoptosis by activation of p53 and inhibiting NF-kB.

The third TA based on Cheever criteria prioritisation is PTPN23, a non-receptor phosphatase that belongs to the endosomal sorting complexes required for transport (ESCRT) driving cell surface receptor degradation and reducing oncogenic signalling leading to tumour suppressor functions (Gurzov et al., 2015). Phosphatases have functions opposed to those protein tyrosine kinases (PTK), involved in regulating different cellular homeostasis including apoptosis, haematopoiesis, cell proliferation and differentiation (Gurzov et al., 2015). PTPN23 was found by SEREX and has never been explored in B-ALL before. Analysis using data from the MILE study shows it may be upregulated in B-ALL. To understand the role that PTPN23 plays in aB-ALL we searched it's known function in healthy tissues and other cancers as phosphatases significantly affect AML, reviewed in (Liu et al., 2023). Phosphatases play a crucial role in tumour progression, dephosphorylating proteins such as tyrosine kinases that can activate oncogenic pathways depending on cancer type (Veenstra et al., 2019). They are associated with chemotherapy and radiotherapy resistance; hence targeting phosphatases may enhance anti-tumour efficacy in colorectal cancer (Kerr et al., 2021). Phosphatases may act as a tumour suppressors in breast cancer (PTPN2, PTPN6) (Veenstra et al., 2019) while PTPN6 (Mok et al., 1995) acts as an oncogene in ovarian cancer, enhancing tumour cell proliferation. PTPN23 acts as a tumour suppressor and the hemizygous PTPN23 deletion increases the predisposition of B-cell lymphoma and activating the progression MYC-driven lymphoma. PTPN23 deletions (Manteghi et al., 2016) in 50% of B-cell lymphoma are associated with poor prognosis, suggesting it may be a new haploinsufficiency tumour suppressor gene. PTPN23 deletion enhances focal adhesion kinase (FAK) activation and integrin mediated cell survival and invasion. However, PTPN23 (Jariwala et al., 2019) enhances bladder carcinoma by promoting MAPK and SRC activation and thus increasing angiogenesis and tumour proliferation (Figure 5.10). Verification of PTPN23 expression in B-ALL is worth further investigation. Future knockdown in B-ALL cell lines and mouse models will help investigate its effect on leukaemogenesis.

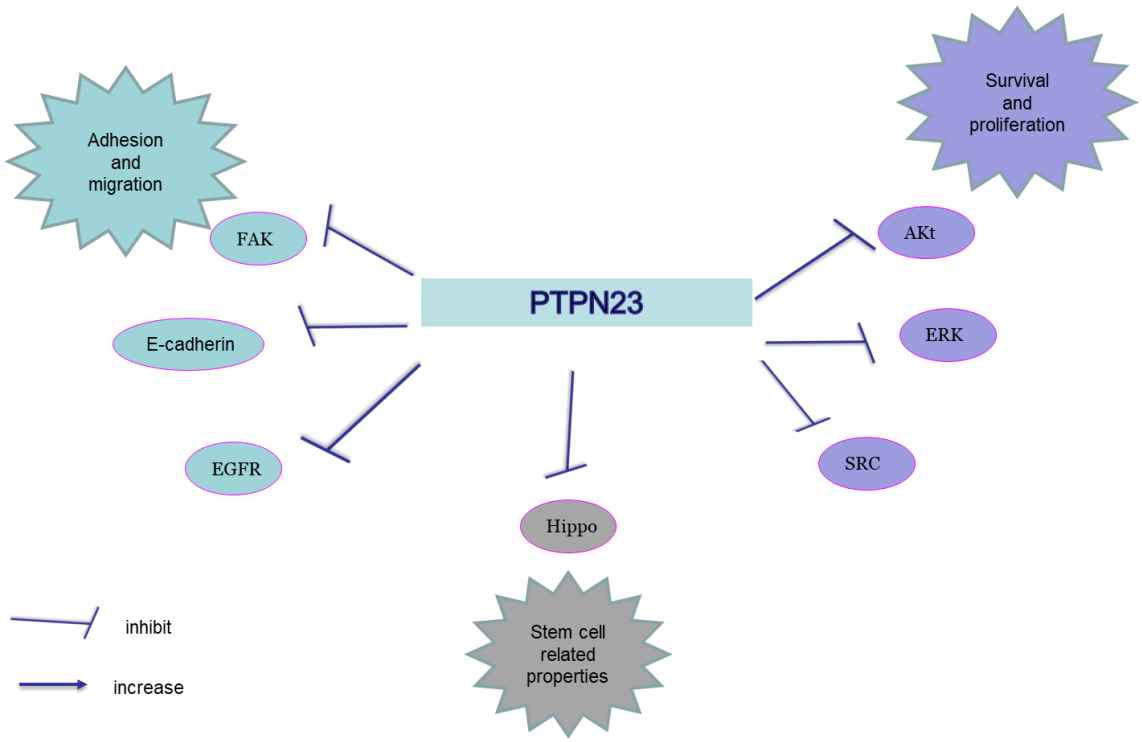


Figure 5.10 Key interactions of PTPN23 and its pathways

PTPN23 acts a tumour suppressor via inhibiting PI3K/Akt/mTOR pathway leading to suppressing of growth and proliferation of tumour. It also dephosphorylates ERK1 inhibiting MAPK/ERK pathway (Jariwala et al., 2019). It suppresses β-catenin translocation and regulates E-cadherin expression decreasing cell migration. It regulates Hippo pathway via dephosphorylation of YAP1 inhibiting its translocation and transcriptional activity (Hendriks & Böhmer, 2016).

Dickkopf-3 (DKK3) has the same score as PTPN23 according to the Cheever model. DKK3 is a Wnt inhibitor that is decreased in aB-ALL. The reduced expression of DKK3 may be due to promoter hypermethylation caused by overexpression of miR-708. Methylation of DKK3 (Veeck & Dahl, 2012) correlated with a higher relapse rate (58%) in comparison to patients with unmethylated DKK3 in aB-ALL cells (35%). DKK3 methylation was associated with low DFS (10.5%) and OS 15.1% for hypermethylated patients (P=0.001 and 0.09) compared to 49.8 and 45.6% for patients with non-methylated DKK3. Consequently, DKK3 methylation was proposed as an independent prognostic marker for ALL. High expression of miR-708 was associated with poor prognosis and may be a reliable biomarker for relapse in B-ALL (Zhang et al., 2017b). DKK3 restoration decreases glycogen synthase kinase 3β (GSK3β) which is a negative regulator of β-catenin in the NALM-6 cell line (Dandekar et al., 2014). Inactivation of DKK3 was also correlated to

chemotherapy resistance and increased gene expression was associated with relapse (Zhang et al., 2017b). Using the demethylating drug decitabine or 5-aza-cytidine restored chemo-sensitivity and reduced miR-708 expression as well as increased DKK3 (Dandekar et al., 2014). miR-708 inhibitor restores DKK3 independent promotor methylation and inhibits the Wnt pathway reducing tumour growth (Dandekar et al., 2014). This result suggested a new strategy for targeting post-translational modification and managing B-ALL by targeting an oncomiRNA. However, DKK3 (Katase et al., 2020) acts as an oncogene in head and neck carcinoma (HNC) enhancing tumour proliferation via Akt activation (Figure 5.11). In this context, its knockdown results in a significant growth inhibition.

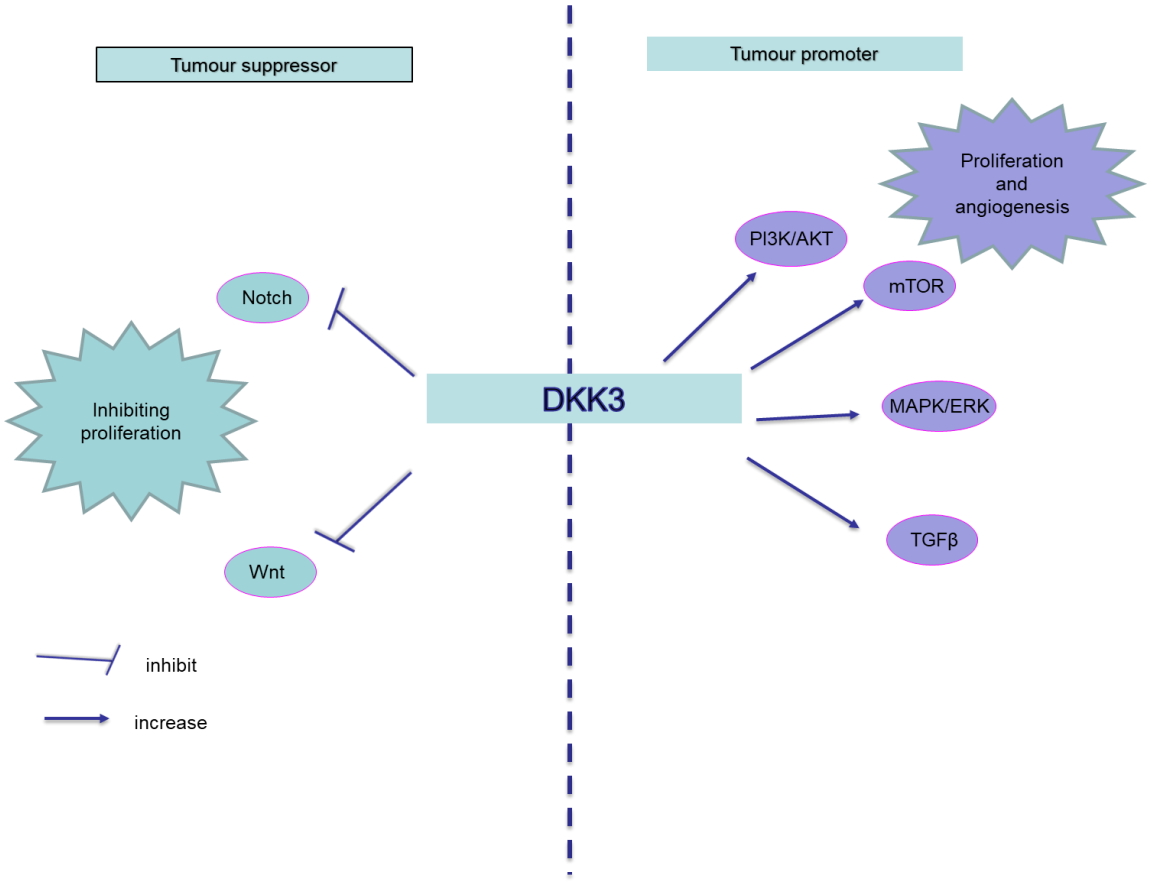


Figure 5.11 Key interactions of DKK3 and its pathways

DKK3 acts a tumour suppressor inhibiting Wnt and Notch pathways leading to suppressing of growth and proliferation of tumour. However, it promotes oncogenesis in HNC through Akt activation.

The fifth TA is RPL28 that belongs to the ribosomal protein L28E family and is overexpressed in breast cancer and associated with metastatic stages (Warner & McIntosh, 2009). Its upregulation is linked to reduced lamin A. Inhibition of RPL28 increases lamin A expression promoting breast cell differentiation and restoring cell growth regulation. RPL28 (Hounguè et al., 2019) has been suggested as a biomarker for monitoring the sensitivity of histone deacetylase inhibitor drugs and analogues because of its dynamic expression and it is induced by extracellular toxic reagents. RPL28 (Capo-chichi et al., 2022) may be inhibited by chemotherapy including doxorubicin and cisplatin decreasing the ectopic ribosomal proteins. From MILE study, RPL28 is upregulated in all B-ALL subtypes and further investigations are warranted to verify RPL28 expression and role in B-ALL.

The signal transduction pathways activated by mutations, gene amplification, or other leukaemia treatment may be an appropriate approach to identify new targets restricting leukaemia growth and drug resistance (Steelman et al., 2011). Pathways enriched in B-ALL identified from SEREX genes included Cadherin signalling pathway (p-value of 0.0008), hematopoietic stem cell gene regulation pathway (p-value 0.001) and apoptosis pathway (p-value 0.01).

From our results, the identified genes UOH-ALL-7, UOH-ALL-23, UOH-ALL-44, and UOH-ALL-10 correspond to (PTPN1, TCF7L2, FER, WNT2B) and are involved in cadherin pathway. PTPN1 is a growth factor regulating phosphatase and is involved in activation of RAS/MAPK. PTPN1 is associated with different cytogenetic abnormalities including Pax5, DUX4, ZNF384 fusions and hyperdiploidy (Cavé et al., 2016). RAS/MAPK overexpression is correlated to relapse and refractory B-ALL in comparison to first complete remission and this suggests MAPK play a crucial role in chemoresistance especially glucocorticoids resistance. MAPK inhibition sensitises B-ALL response to steroids (Cavé et al., 2016). Nonetheless, subtypes of hyperdiploidy are not resistant to steroids, and a correlation between hyperdiploidy and complex cytogenetic subtypes implies that more research is necessary to confirm Ras activation and identify patients who may experience relapses and benefit greatly from intensive therapy (Cavé et al., 2016).

The initiation and maintenance of B-ALL (Shu et al., 2015) are largely dependent on leukaemia-initiating cells (LICs). Both DNA methyl transferase (DNMT1) and MCL1 are involved in the hematopoietic stem cell gene regulation pathway and have been identified in our results corresponding to UOH-ALL-15 and UOH-ALL-46. DNA methylation (Guillamot et al., 2016) is common in haematological malignancies due to global hypomethylation. Aberrant methylation is correlated early leukaemogenesis. DNMT1 upregulation is found in cB-ALL promoting leukaemic stem cell maintenance. DNMT1 polymorphism increases the risk susceptibility of children to B-ALL in (Luo et al., 2015). Furthermore, hypermethylation of tumour suppressor genes (such as CDKN2B and PTEN) is predominant in 77.3% B-ALL around and is associated with poor prognosis, with low DFS of 37.2%, compared to 75.5% for patients with non-methylated genes (Roman-Gomez et al., 2004). Thus, methylation profiles may have clinical relevance and may be attractive new biomarkers for predicting outcomes of B-ALL.

In acute promyelocytic leukaemia, the oncogene PML-RAR recruits DNMT1 and DNMT3B to their target promoter inducing methylation and gene silencing (Tan et al., 2021). DNMT1 interacts with STAT3 and HDAC1 leading to inhibition of Src homology domain 2-containing tyrosine phosphatase-1 (SHP-1). Silencing of DNMT1 by using antisense peptide results in re-expression of SHP-1 restoring its functions in regulating haematopoietic cell growth via reduction of erythropoietin and FLT3 signalling (Tan et al., 2021).

MCL1 is found in B-ALL, especially patients with Ph+ subtype, and is associated with acquisition of stem cell features. MCL1 up-regulation is associated with TKIs resistance, and its downregulation sensitise B-ALL increasing the cytotoxicity of TKI. MCL1 inhibitor, S63845, induces cell apoptosis activating caspase-3 and PARP cleavage. S63845 combined with doxorubicin synergistically increase the caspase-dependent apoptosis (Ebrahimi et al., 2022).

UOH-ALL-65, UOH-ALL-29 and UOH-ALL-46 are ROCK1, TIMP1 and MCL1 all involved in apoptotic pathways given that p53 mutation is less frequent in leukaemia compared to solid tumours (Comeaux & Mullighan, 2017) and is found in hypodiploidy and relapses from all B-ALL subtypes. High expression of MCL1 and BCL2 is associated with poor prognosis in B-ALL, as upregulation of anti-apoptotic pathways blocks Bak and Bax and enhances chemo-resistance (Karlsson et al., 2013). Human homolog of murine double

minute 4 (HDM4), which is a negative regulator of p53, was detected in 80% (39 of 49) adults with B-ALL and was correlated with Ph+ subtype, while HDM2 was found in 26% of cases and not related to Ph+ presence (Han et al., 2007). This data suggests HDM4 may be a target for B-ALL and further studies are required to identify the mechanism of action and its variant splicing for a selective inhibition.

Wnt components (WLS, Wnt2b, TCF7L2, and DKK3) are frequently reported in our results. This result agrees with Khan *et al.* that Wnt family expression is found in B-ALL. Wnt activation is also associated with bone marrow stromal survival enhancing leukaemia stem proliferation (Khan et al., 2007).

Wnt pathway is involved in several pathways including stem cell self-renewal, cell growth, lymphopoiesis(Chiarini et al., 2020). Wnt activation may accumulate β-catenin in the nucleus and enhance the transcriptional activating TCF/LEF. Wnt signalling is regulated by the extracellular factors including Wnt inhibitory factor 1 (WIF1), DKK1–4, and secreted Frizzled related protein (sFRP)1–4 (Chiarini et al., 2020). Wnt aberration is associated with B-ALL; DKK3 is downregulated in B-ALL. Wnt 16 is especially activated by E2A-Pbx1 fusion and upregulated in B-ALL (1;19), promoting tumour cell survival. The Ph subtype has shown lower DKK3, SFRP, WIF1 due to hypermethylation. Low expression of Wnt inhibitors may result in upregulation of β-catenin, LEF1, TCF1, Fzd3, Wnt16 as well as suppression of cyclin D1, reversed by 5-aza-2'-deoxycytidine treatment (Román-Gómez et al., 2006). 25% of B-ALL cases had LEF1 overexpressed suggesting to be an independent prognostic factor of poor prognosis (Chiarini et al., 2020). Furthermore, overactivation of Wnt activates the BM stroma to increase cystic fibrosis transmembrane conductance regulator (CFTR), enhancing the inhibition of GSK3β. This suggests epigenetic regulations leading to overactivation of Wnt pathway in Ph subtype and associated with poor survival (Román-Gómez et al., 2006). According to Nygren *et al.* (2009), elevated β-catenin is linked to cadherin activation, which promotes and sustains cell to cell interaction and tumour microenvironment (Nygren et al., 2009). This indicates demethylating agents reduce Wnt activation and target epigenetic dysregulation. Wnt ligand secretion mediator (WLS) enables Wnt component secretion and binding. It is overexpressed in cB-ALL and associated with poor prognosis and relapse. It is involved in proliferation and anti-apoptotic activity via regulating GSK3β

activation (Chiou et al., 2014). Antibody targeting WLS has inhibited gastric cancer in PDX model. WLS expression needs to be verified in aB-ALL compared to healthy controls. Further its mechanism of action in the leukaemia cell lines needs to be defined to evaluate for future treatment. Retinoid acid and vitamin D have been suggested to help treat Wnt dependent carcinomas (González-Sancho et al., 2020). Their mechanisms induce DKK1,4 which are Wnt inhibitors, inhibiting tumour growth and enhancing apoptosis. Wnt antagonist, XAV939, has shown antileukemic activity enhancing chemotherapy-induced apoptosis in mouse model of B-ALL and overall survival (Yang et al., 2013). XAV939 accelerates β-catenin breakdown by blocking tankyrase and disrupting the BM niche protective driving chemoresistance (Yang et al., 2013). Further studies are required to correlate the epigenetic regulation of Wnt pathway with B-ALL microenvironment changes that would enable to identify B-ALL subtypes (beneficial for therapy). Moreover, β-catenin-dependent transcription inhibitor, iCRT14, inhibited Wnt signalling and induced cytotoxicity in ALL cell lines, relapsed ALL samples, and thus restored chemosensitivity (Dandekar et al., 2014). Inhibition of β-catenin reverses resistance to TKI in advanced stages regardless of the presence of BCR-ABL1 mutations. Targeting downstream genes in the Wnt pathway has been investigated for relapsed cB-ALL, for example by targeting BIRC5 with a novel survivin (mRNA) antagonist for phase I clinical trial (Raetz et al., 2014). However, dose related toxicities and haemorrhage combined with chemotherapy limited its use. Future critical studies are required to verify how the Wnt pathway is involved in the hematopoietic stem cell maintenance and differentiation to identify the safety of selective inhibition. BCL9 was found in B-ALL and other cancers including CRC to modulate β-catenin interaction with TCF and e-cadherin. As BCL9 was not found in normal tissues, this suggests blockade of BCL9/BL9/ β-catenin selectively targets tumours and avert damage to normal cells. Stapled α-helical peptidomimetic of the BCL9-HD2 domain, SAH-BCL9 peptide (Takada et al., 2012) has shown great efficiency in disrupting interaction of β-catenin, blocking invasiveness, metastasis, and angiogenesis of CRC, by significantly downregulating C-MYC, BIRC5, cyclin D1, and VEGF. SAH-BCL9 peptide demonstrated anticancer activity without disturbance of normal homeostasis; however, presence of carboaromatic rings may have high albumin binding, interfering with their location and destruction (Tanton et al., 2022). Further studies are required to modify the pharmacokinetic properties and

enhance their translation into clinical setting. Targeting protein interactions seems to be an attractive approach for overactivated Wnt pathway related carcinogenesis and stem cell targeting.

Interesting, we identified UOH-ALL-73 and UOH-ALL-60 (IDI1 and HDLBP) which are correlated to cholesterol synthesis. IDI1 (Xu et al., 2020) is involved in the regulation of the mevalonate-isoprenoid biosynthetic (MIB) pathway, which is critical to sustain cancer stem cells. It is found that venetoclax and azacytidine resistance (Wang et al., 2023) is due to the upregulation of fatty acid oxidation (FAO) enhancing AML proliferation by remodelling bone marrow adipocyte and lipolysis. FAO inhibition by using very long-chain acyl-CoA dehydrogenase (VLCAD) may suppress AML survival and proliferation (Wang et al., 2023). This finding suggests metabolic fatty acids contribute to AML pathogenesis. In addition, HDLBP is overexpressed in hepatocellular carcinoma and is involved in tumour growth and proliferation. Knockdown of vigilin leads to inhibition of carcinogenesis and sensitisation to cisplatin (Yang et al., 2014b).

We identified three UOH-ALL-58, UOH-ALL-85 and UOH-ALL-98 lncRNA (LINC00251, LINC00261 and LOC338963) as targets of B-ALL patients sera. Targeting lncRNA has become an area of interest as this was previously thought as undruggable due to lack of protein translation. LINC01257 is overexpressed in AML t(8;21) associated with poor prognosis and survival. siRNA-loaded lipid nanoparticle (LNP)(Connerty et al., 2021) has been used to inhibit lncRNA LINC01257 in AML. LNP-si-LINC01257 have shown safety in animal models providing that targeting lncRNAs increases the specificity for AML1-ETO subtype (Connerty et al., 2021). However, the expression of LINC00251, LINC00261, and LOC338963 should be verified using qRT-PCR and before these could be targeted using siRNA-LNP.

SEREX suffers from several drawbacks: 1) antibodies profiling requires rigorous validation as autoantibodies may be found in healthy individual; 2) the number of biological samples used was limiting; 3) specificity of the antibodies required further evaluation in patients with different types of leukaemia to determine the specificity for B-ALL versus other malignancies; 4) the bottleneck of library construction and protein identification displayed in the phage (Scanlan et al., 2002). Future directions include tertiary screening to compare B-ALL versus AML and CML antigens to healthy age and

sex-matched individuals for identifying LAAs that are for B-ALL. SEREX strategy fails to detect antigens with post-translational modification using phages. It is a complex method with many steps and associated with false positive results (Scanlan et al., 2002).

Lastly, Cheever *et al.* ranking has not been updated since 2009 and has the disadvantage of giving a high weighing to some antigens (Cheever et al., 2009). Thus, antigens further away from the development may be disadvantaged to known antigens with high scores especially in terms of therapeutic function and immunogenicity. Scoring different criteria may have different endpoints, different patient selection methods and characteristics that may not be precise in reflecting fairness and reliability of the data (Silva et al., 2007b; Cheever et al., 2009). In-depth analyses would be required to ensure reproducibility of the data. Further, the oncogenicity criteria was restricted in the definition; thus antigens such as FOS-related antigen-1, carbonic anhydrase IX do not count as oncogenes, even if they are associated with the poor survival and involved in malignancy (Cheever et al., 2009). Maintaining a cancer phenotype may be more relevant than oncogenic criteria of the antigens. Furthermore, stem cell criteria may need improvements by incorporating genes related to the stemness properties of cancer from more recent knowledge. Criteria related to stem cell expression are biological and would not change over time; instead, others, such as therapeutic efficacy will change by modifying vaccine formulations, and dose. The clinical efficacy may lack superb data due to many reasons including patient selection, or inadequate trial design, or vaccine formulations (Cheever et al., 2009). These drawbacks may be overcome by more clinical design ensuring the standards of reproducibility and accuracy. The Cheever model provided a systematic approach for antigen rankings, yet, none of ranked antigens have been approved by FDA for the clinical use so far.

Chapter 6: Identification of enriched pathways, and essential genes involved for B-ALL

6.1 Introduction

This chapter focuses on identifying the enriched pathways and genes that may be involved in pathogenesis and act as new targets for aB-ALL treatment(s). As treatment of haematological malignancy has improved, especially B-ALL in the paediatric population, the 5-year survival rate has improved by up to 85% (Elsallab et al., 2023). aB-ALL remains a significant obstacle with poor survival (Sive *et al*., 2012). Around 50% of patients with the first remission experienced relapse, which leads to a 5-year overall survival rate of approximately 40% for adults (Liu *et al*., 2016). Current treatment outcomes are not sufficient to improve patient survival. Consequently, it is crucial to thoroughly investigate the mechanisms behind leukaemia deterioration and relapse following treatment to enhance adult B-ALL patient survival. It is crucial to understand the complex biology of B-ALL in-depth to identify new targets (Liu *et al*., 2016). Despite B-ALL being characterised by low-burden mutations, loss, or downregulation of HLA class I molecules is uncommon. However, recognition of blast by cytotoxic T lymphocytes is significantly impaired due to tumour suppressive microenvironment (Rovatti *et al*., 2020). Natural killer (NK) cells also appear to play a role in immunosurveillance against ALL. However, leukaemia development impairs the immune system ability to respond to cancer, particularly in patients who do not respond well to treatment or are at a stage of relapse (Jiménez-Morales *et al*., 2021).

As B-ALL is a heterogenous malignancy, 134 cDNA clones were identified using SEREX (Table 5.3). Of these 72 known genes were prioritised with the Cheever *et al.* model focusing on TA with relatively high accumulative scores for further studies. ROCK1, the top TA, has been identified to examine its expression in B-ALL samples. Differentially expressed genes (DEG) and pathway enrichment may be attractive approaches for identifying key genes in B-ALL. Targeting this heterogeneity requires TAs to be broadly expressed in patients with B-ALL, avoiding negative selection and immunotolerance. Furthermore, enriched pathways could be targeted by both immunotherapeutic

approaches and small molecule inhibitors. Antigens from six methods (SEREX (5.34), protoarray (Jordaens et al., 2020), LAA (Iacobucci & Mullighan, 2017), CTAs from (http://www.cta.lncc.br/), GSE38403 (Geng et al., 2012), and GSE13204 (Kohlmann et al., 2008)) have been pooled together to identify enriched pathways (**Figure 6.1**). We determined which genes/antigens were associated with patient survival and their molecular pathways using Enrichr (Kuleshov et al., 2016). The core genes in selected pathways were examined for their protein-protein interaction networks using the Search Tool for the Retrieval of Interacting Genes (STRING) analytical program (Szklarczyk et al., 2019). We then priority-ranked the list of cancer antigens based on predefined and pre-weighted objective criteria developed using the analytical hierarchy process by Cheever criteria (Cheever et al., 2009). The top-ranking antigens were validated by qPCR and ICC. These included DEGs identified from GSE38403, GSE13204 (MILE), and protoarray (Jordenset al. 2020) were listed in (**Figure 6.2**). GSE38403 focused on methylation profiles of 215 aB-ALL compared to 12 normal preB samples suggesting epigenetic regulation in B-ALL. IL2RA(CD25) was upregulated in most B-ALL due to hypomethylation and was associated with poor outcomes. In addition, overexpressed BCL2 and FLT3 were found in the MLL subtype. Prof Ken Mills performed gene expression analysis and identified 16 upregulated and 14 downregulated genes. The median survival of patients over 10 years following diagnosis depended on whether gene expression was above or below the median as determined using BloodSpot (Table 9.5). Six genes (FLT4, CYTL1(C17), EMP1, PLVAP, MRC1, BMP-2, DPPA4) were associated with poor survival, and their expression in cancer was discussed (Table 9.14). GSE13204 contains 205 BALL patient's vs 74 healthy controls. Six genes (LIN7A, LILRA3, VSTM1, P2RY2, PTGS2, CYP4F2) were associated with poor survival (Table 9.15). Both LIN7A and VSTM1 are tumour suppressor genes and downregulated in AML due to promoter hypermethylation while P2RY2 is upregulated in AML via the activation of PI3k/Akt (Lin *et al*., 2022).

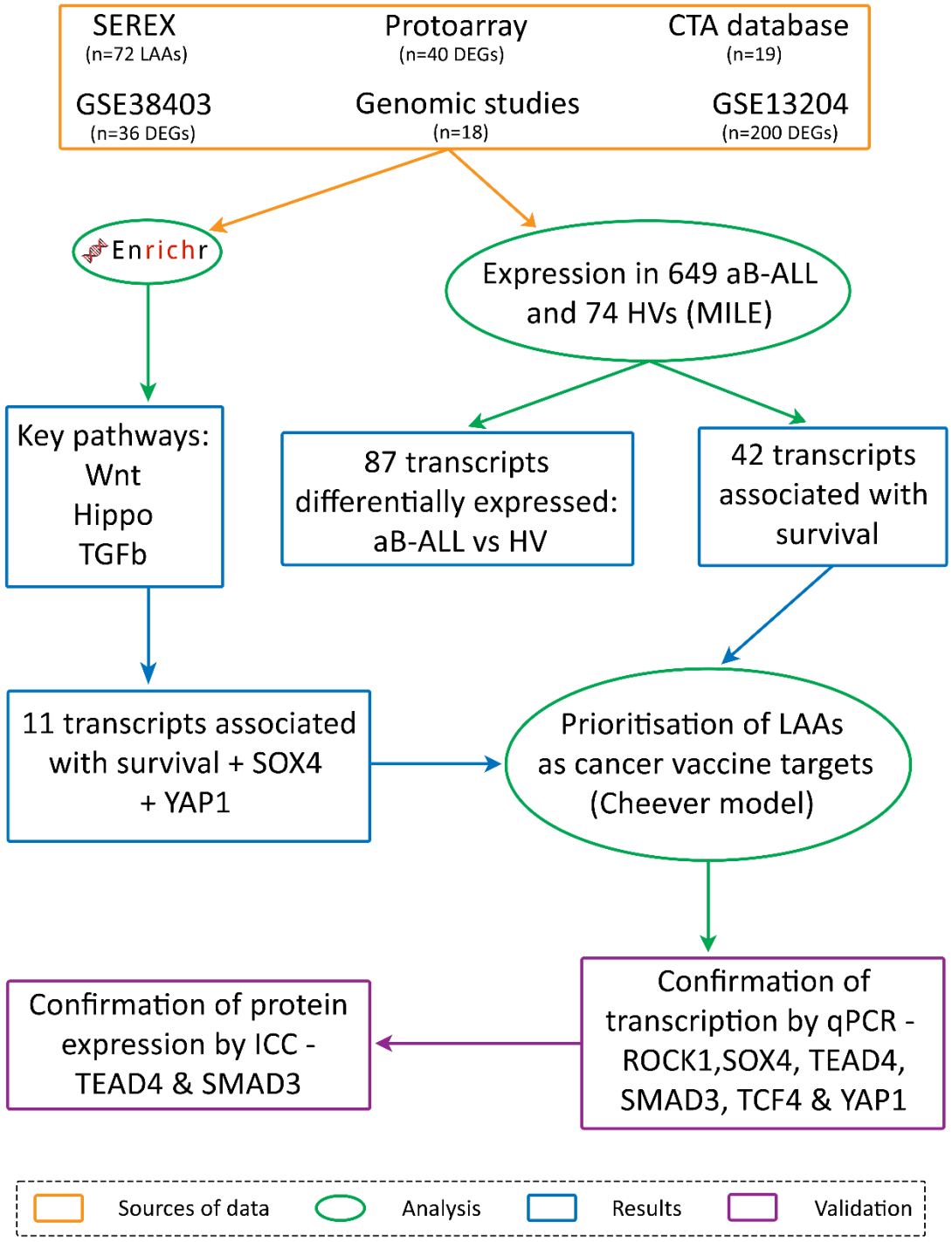


Figure 6.1 Method used for identification of TAs as targets for immunotherapy

N represents number of genes identified from each method. Combined antigens were identified from six methods and pulled by IPA to identify upstream regulators. We examined the antigens associated with patient survival and their upstream molecular pathways using Enrichr. We found that the Wnt, Hippo and TGF-β pathways were highly represented ($p<0.02$). A literature search allowed the identification of novel antigens (not yet studied in aB-ALL). Antigens identified were prioritised using Cheever *et al.* Criteria. Image prepared by Danny Fletcher.

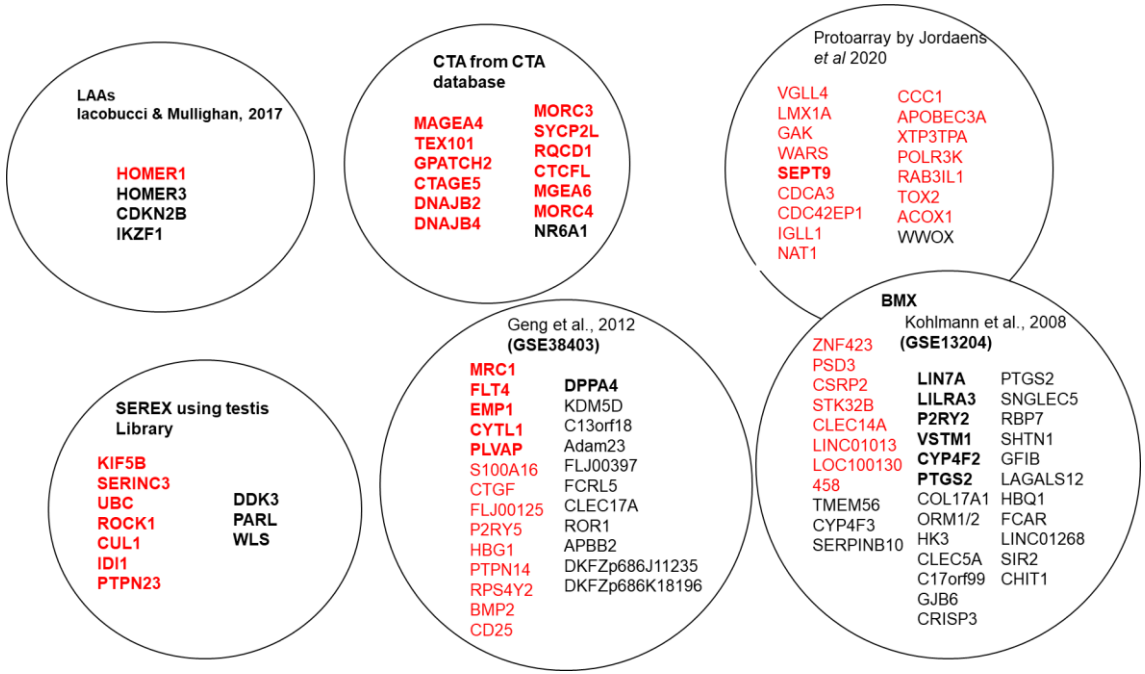


Figure 6.2 DEGs identified from the six sources used for identification of pathways enriched

Relative expression of aB-ALL associated genes from six sources (SEREX, protoarrays, microarrays (GSE38403, GSE13204), the CTA database). Red font are upregulated antigens while downregulated genes represented by black font, genes in bold are associated with survival $p<0.05$. Only BMX is common between the protoarray and MILE study while all other antigens were only identified once.

6.2 Aim

To identify new targets for B-ALL, suggesting TA targets as crucial candidates in pathways enriched.

6.3 Results

6.3.1 Bioinformatic analysis

Enrichr analysis showed that TGFβ, Wnt and Hippo pathways were highly represented by the identified antigens (**Table 6.1**). The combined score was calculated from the p-value and the deviation from the expected rank providing a compromise between both methods and the best rankings compared to the other scoring schemes. TGFβ is a multifunctional cytokine that is crucial in regulating the immune system. It is produced by various cell types, including haemopoietic and lymphoid cells, and can have pro- and anti-inflammatory effects depending on the context (Shen *et al*., 2018).

Table 6.1 Pathway enrichment in B-ALL

Name	P-value	Adjusted p-value	Odds ratio	Combined score
TGF-β pathway	0.001	0.19	8.3	52.44
Hippo pathway	0.01	0.14	14.02	63.81
Wnt pathway	0.01	0.14	66.54	272.01

In haemopoietic cells, TGFβ has been shown to regulate the differentiation and function of various cell types. As per Li *et al.* (2019), it has the potential to stimulate the development of regulatory T cells (Tregs), an essential cell type for preserving immunological tolerance and averting autoimmunity (Li *et al*., 2019). However, TGFβ can also prevent some types of hematopoietic cells from differentiating, including B cells (Shen *et al*., 2018). Given that TGFβ is a potent inhibitor of common progenitor development and that TKI-resistant hematopoietic stem cells cause complete remission of B-ALL, TGFβ is thought to be engaged in the tumour microenvironment and is essential for the maintenance of these stem cells (Vicioso et al., 2019).

In lymphoid cells, TGFβ has been shown to have diverse effects depending on the context. For example, in the presence of IL-6, it can promote the differentiation of Th17 cells, which is essential for host defence against certain pathogens but can also

contribute to autoimmune disease (Korn *et al*., 2009). However, TGFβ can also inhibit the proliferation and activation of other lymphoid cell types, such as CD8+ T cells (Thomas and Massagué, 2005). Overall, the balance of TGFβ signalling in haemopoietic and lymphoid cells is complex and context dependent. The TGF-β1 pathway is involved in various cellular functions, including inhibiting cell proliferation, and is inactivated in leukaemia. The intracellular mediators of TGF-β1 are SMAD proteins which are divided into three groups: receptor-activated Smad (R-SMAD), common-partner SMAD (Co-SMAD), and inhibitory Smad (anti-SMAD). R-Smad complexes with Co-SMAD enhance their translocation to the nucleus regulating gene transcription of targets (Shen *et al*., 2018). SMAD4 is a Co-SMAD mutated in AML (Shen *et al*., 2018). TGF-β1 interacts with other pathways, such as Bone morphogenetic proteins (BMPs) regulating pluripotent stem cells and their differentiation. BMP signalling regulates SMAD 1,5,6 rather than SMAD 2,3 (Korn *et al*., 2009). Dysregulation of this pathway has been implicated in various diseases, including cancer and autoimmune disorders (Travis *et al*., 2020). TGF-β1 is upregulated in cB-ALL. Elevated phospho-SMAD2/3 expression is used by leukaemia blasts to inhibit NK cytotoxicity (Rouce *et al*., 2016). In addition, TGF-β1 RI/II are overexpressed along SMAD2/3 in cB-ALL correlating with FOXO1 overexpression and downregulation of Runx1/3 (Liu *et al*., 2018). ALL blasts cultured alone or with NK cells had significantly higher levels of TGF-β1 in the supernatant than NK cells alone. Although the study did not directly measure TGF-β in patient sera (Liu *et al*., 2018), other studies have reported elevated levels in patients with lymphoid malignancies (Rouce *et al*., 2016). The use of TGF-β blocking monoclonal antibodies reversed ALL-mediated NK cell dysfunction (Rouce *et al*., 2016). In contrast, adding exogenous TGF-β1 to healthy donor NK cells induced an inhibitory phenotype similar to ALL-NK cells. This suggests a direct role for TGF-β in modulating NK cell function and phenotype through activating the SMAD signalling pathway (Rouce *et al*., 2016).

The Wnt pathway is involved in maintaining the process of normal hemopoiesis. It depends on a tight balance between extracellular factors such as Wnt ligands, agonists, antagonists, cell surface receptors, cytoplasmic components like adapters and destruction complex components, and nuclear factors (Chiarini et al., 2020). A small amount of Wnt is required to maintain normal hematopoiesis; high Wnt components

could lead to hematopoietic disorders (Chiarini et al., 2020). Wnt pathway dysregulation has been found in our SEREX result (Section 5.4). The polypeptide that makes up Noggin (NOG) suppresses TGFβ superfamily signalling as well as genes involved in development and differentiation, including Wnt3A, Wnt1, ID1, ID2, BCL2, P27, and MSX1. Nonetheless, Wnt pathway activation is still present overall (Khoury et al., 2006). Additionally, enhanced cystic fibrosis transmembrane conductance regulator (CFTR) and Wnt overactivation interact with the BM stroma, strengthening the inhibitory GSK3β. This suggests epigenetic regulations leading to overactivation of Wnt pathway in Ph subtype and associated with poor survival. Targeting Wnt ligands and receptors has been found useful in reducing leukaemic stemness. Knock-down of FZD6/ Wnt10B could be used as a target for leukemic stem cells (Cassaro *et al*., 2021). However, non-canonical pathways of Wnt increase the leukaemic burden. Table 9.16 shows different Wnt components studied in cancer specially leukaemia.

The Hippo pathway was also found to be dysregulated. This plays a crucial role in regulating stem cell differentiation and organ size (Zhu *et al*., 2015). It acts through a kinase cascade that includes Mst and Lats kinases and their co-factors WW45 and Mob (Zhu *et al*., 2015). This pathway is activated when cells proliferate in high density. Mst1/2 forms a complex with Sav1, which then phosphorylates Lats1/2. This activation leads to YAP phosphorylation on Ser127, leading to its cytoplasmic retention and degradation via the ubiquitin-proteasome pathway (Zhu *et al*., 2015). However, the Hippo pathway is often inactive in malignant tumours, leading to ineffective YAP phosphorylation and degradation. Unphosphorylated YAP enters the nucleus and binds to transcription factors, altering gene expression related to cell proliferation and apoptosis. YAP are crucial for various cancer cell survival processes, such as sustained proliferation, metastatic progression, and drug resistance (Han, 2019). Our results suggest the Hippo pathway be highly enriched (p=0.02), potentially involved in leukaemogenesis. Only Lats2 loss has been found in B-ALL and is correlated to promoter region methylation in leukaemic cells (Jimenez-Velasco *et al*., 2005). KIBRA is an upstream regulator of the Hippo pathway, frequently silenced and hypermethylated in 70% of cB-ALL. KIBRA methylation is not found in epithelial cancers and is linked to B-ALL, especially t (12;21) leukemogenesis (Hill *et al*., 2011). Loss of KIBRA is related to diagnosis and not found in

remission (Hill *et al*., 2011). KIBRA has been identified as a promising tumour suppressor gene in both Drosophila and mammalian cells. Its significance in B-cell leukemogenesis highlights a potential as a therapeutic target for clinical applications, mainly due to the reversible nature of promoter hypermethylation, for example 5-aza-2'-deoxycytidine (Hill *et al*., 2011). Further research is needed to determine if KIBRA is methylated in other B-cell leukaemia and lymphoma and to elucidate its role in B-cell development and differentiation (Hill *et al*., 2011). These findings contribute to the growing evidence indicating dysregulation of the Hippo tumour suppressor pathway in haematological malignancies (Jimenez-Velasco *et al*., 2005). The expression of the main component of Hippo, Wnt and TGFβ pathways in B-ALL and their association with the survival were shown (Table 9.9 and Table 9.10). Table 9.17 shows different Hippo components studied in cancer specially leukaemia.

6.3.1.1 Identifying upstream regulators using Ingenuity Pathway Analysis (IPA)

Somatic mutations in driving oncogenic processes and leukemogenesis in ALL, may result in signal pathways perturbation, such as the AKT pathway, sustaining the tumour survival (Barbosa et al., 2013; Nepstad et al., 2020). The control of genes synthesis occurs through various mechanisms, such as epigenetics, regulation of transcription and mRNA export, stability, modifications after transcription, and precise adjustment of the translation process involving ribosomes and initiation factors. Notably, the interplay between transcription and translation is regulated by a complex system that involves micro-RNA (miRNAs) dynamics (See section 7.2).

Therefore, upstream regulators were investigated suggesting that they may involve in leukaemogenesis and may act as targets for B-ALL. For example, low expression of Wnt inhibitors (upstream regulators such as DDK family) may result in upregulation of β-catenin, LEF1, TCF1, Fzd3, Wnt16. 25% of B-ALL cases had LEF1 overexpressed and had been suggested as independent prognostic factor of poor prognosis. β-catenin binds cadherins involved in tumour microenvironment and sustained LSC (Chiarini et al., 2020).

IPA was used to identify upstream regulators (**Table 6.2**) that are involved in gene regulation. Table 9.13 shows the upstream studied in leukaemia especially B-ALL. The median survival of patients over 10 years following diagnosis depended on whether

gene expression was above or below the median as determined using BloodSpot (Table 9.21).

Table 6.2 The upstream regulator identified by IPA

Upstream regulator	Predicted activation value	P-value
VEGF	2.41	2.27E-03
IL-6	2.39	1.86E-02
P38MAPK	1.92	1.62E-02
ErbB2	1.91	3.36E-02
IFN-α	1.68	1.61E-02
KLF6	1.45	2.04E-05
CD3	1.34	2.35E-02
Notch1	1	6.09E-03
Wnt3a	0.8	2.15E-03
PGE2	0.7	1.82E-02

6.3.1.2. Target identification

Prioritising TAs using the Cheever model (Cheever et al., 2009) is based on identifying antigens with high value and could be targets for immunotherapy (Section 3.3). Applying this to our gene list (Figure 6.2) to rank these antigens. Figure 6.3 shows ranking of TAs by Cheever model, SOX4 (0.89), TGFβ1 (0.81) and ROCK1 (0.41) had the top TAs and their expression verified from MILE Study. Both SOX4 and ROCK1 are upregulated in B-ALL $p<0.001$ except t(8;14) and t(1;19) for ROCK1 where TGFβ1 is downregulated in B-ALL $p<0.001$ when compared to healthy bone marrow. Figure 6.4 describes the three pathways and selected targets for qPCR analysis. SOX4 and ROCK1 act as oncogenes while TGFβ1 is downregulated and acts as a tumour suppressor. As miRNAs are master regulators of gene expression (Ho et al., 2022), identification and targeting of oncoMIR may be an attractive approach to restoring tumour suppressor gene functions (Section 7.6). In the case of TGFβ1, miR4286 overexpression may inhibit TGFβ1 and promotes EMT (Ho et al., 2022).

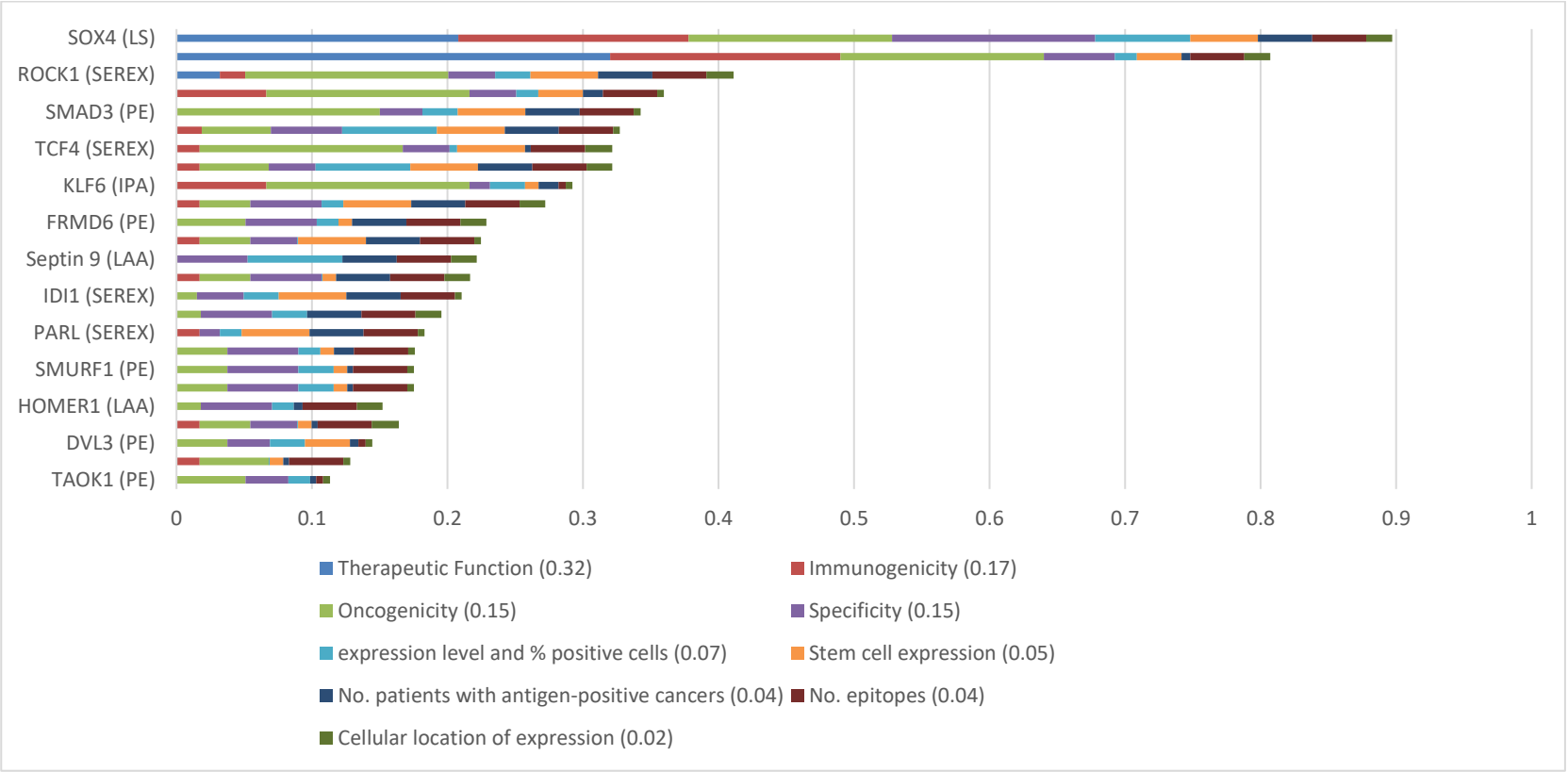


Figure 6.3 Ranking of tumour antigens based on their potential to act as vaccine targets

Antigens were found using SEREX, pathway enrichment (PE), literature search (LS), CTA database and microarrays (LAA). Genes were prioritised by Cheever criteria based on therapeutic function (blue) accounting 0.32 followed immunogenicity (0.17 indicated by red). The oncogenicity (green) and the specificity (purple) account of 0.15. The key shows the evaluation criteria and the maximum attainable score in parenthesis. SOX4 (0.89), TGFB1 (0.81) and ROCK1 (0.41) had the highest accumulated score.

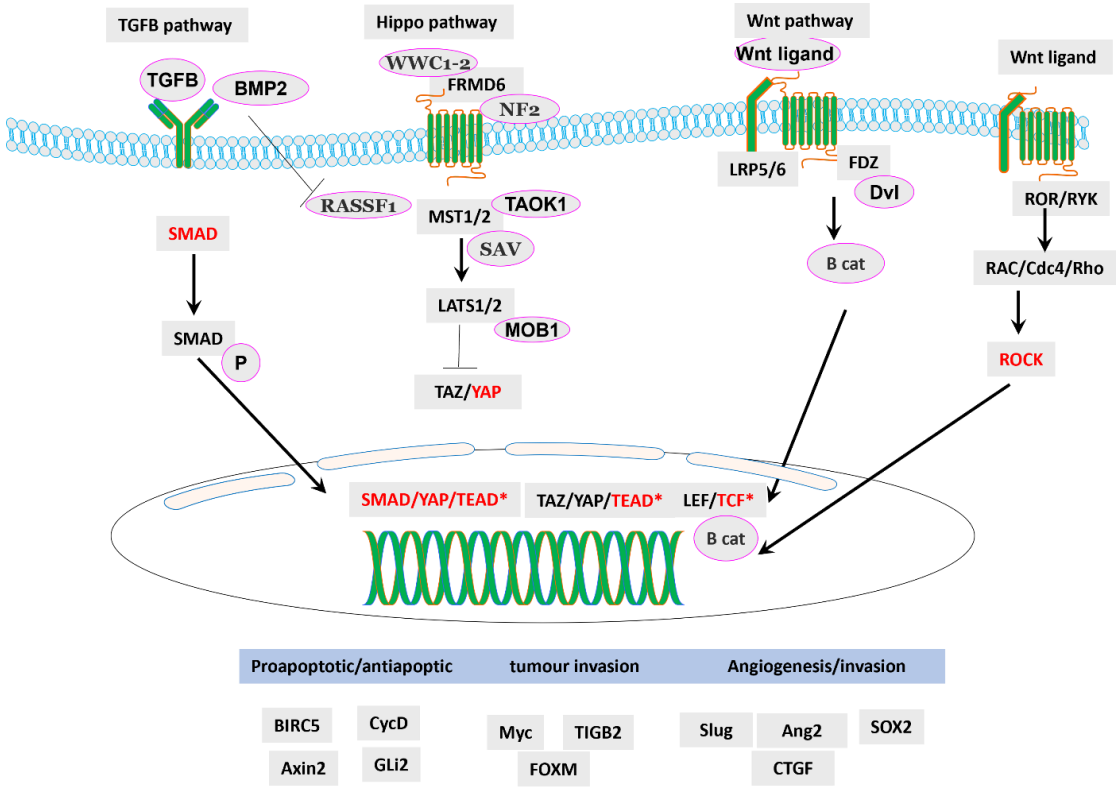


Figure 6.4 Pathways enriched and targets in B-ALL

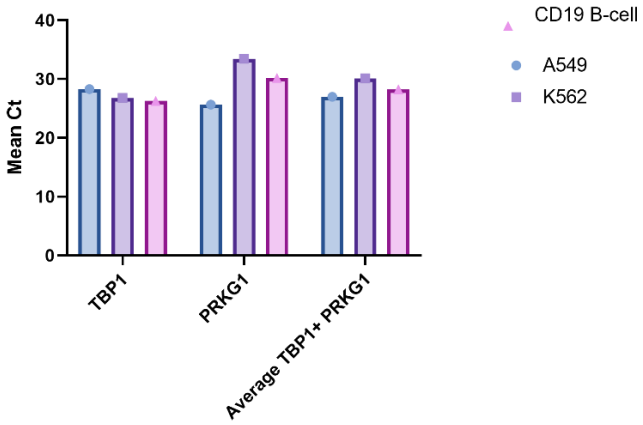
Hippo, TGFB and Wnt pathways involved in promoting leukaemogenesis via upregulation of anti-apoptotic genes and pro-proliferative genes that causes blasts to continue growing and resist to apoptosis. Red colour genes that have high rank from Cheever et al. 2009 and have been selected for verification of their expression in our B-ALL samples versus normal donors.

6.3.2 Genes expression verification in patients' samples and healthy control

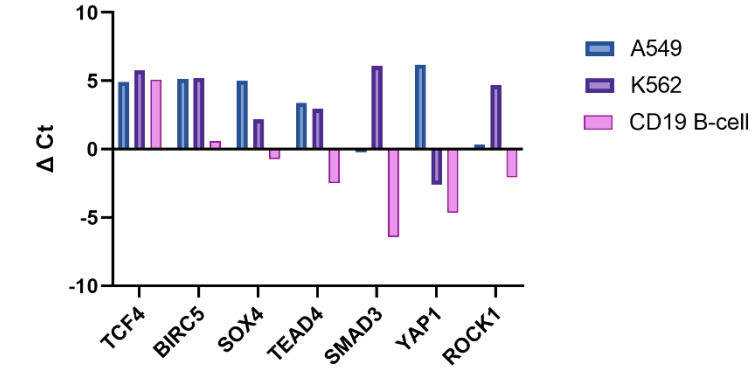
RTq PCR was used to quantify SOX4, ROCK1, TCF4, TEAD4, SMAD3 and YAP1 mRNA levels in aB-ALL samples from peripheral blood (PB) and bone marrow samples (BM). TBP1 and PRKG1 were used as reference genes in qPCR. TBP and PRKG1 were chosen in keeping with MIQE guidelines on qPCR protocol (Lossos *et al*., 2003). Furthermore, RNA input in each reaction was equal to 160 ng, using this as an internal control for quantifying the gene expression.

The K562 cell line (**Figure 6.5**) was used as a positive control for qPCR experiments using TBP1 and PRKG1 primers. ΔCT was calculated (ΔCT = CT reference genes–CT gene of interest) as well as the Ct was averaged for TBP1 and PRKG1. K562 was regarded for their Ct values with low variance between expressions for BIRC5, TCF4, SMAD3, ROCK1 relative to TBP1 expression. A549 was highly expressed YAP1 and had low variance between BIRC5, TCF4, SOX4, and TEAD4. CD19 B-cell were used to represent normal B-cell as pre-normal B cell was not available. The CD19 B-cell had high TCF4 expression and other genes were relatively less expressed. The results showed similar trends when normalisation was done relative to the average of two housekeeping genes. Out of the two housekeeping genes PRKG1 normalisation showed a general trend to increased expression of any targets probably because it is a less abundant gene in K652 and CD19 B cells than in A549 cells, where trends appeared more consistent between the two housekeeping controls.

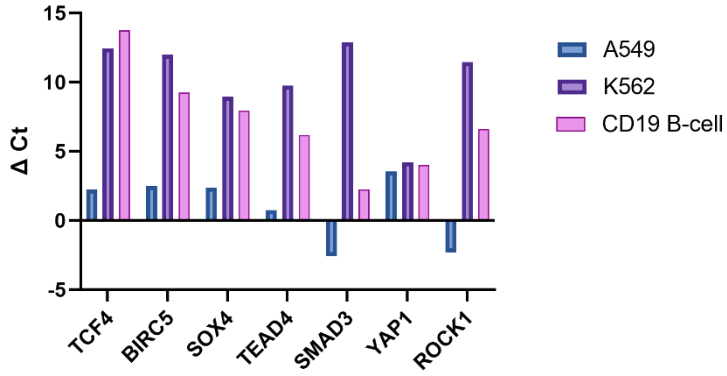
A



B



C



D

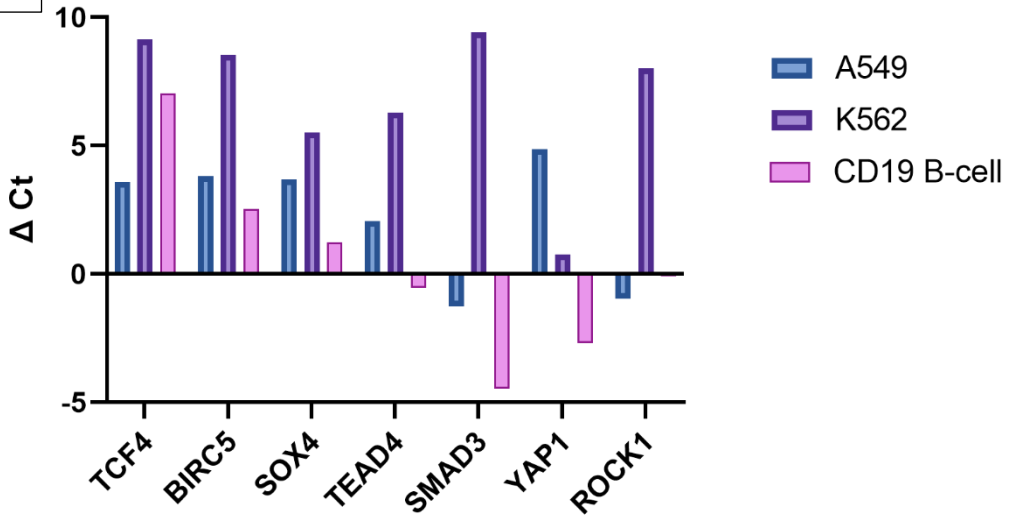


Figure 6.5 Expression of antigens of interest in control cell lines relative to TBP1 and PRKG1

(A) Reference genes (TBP1 and PRKG1 as well as the average of two) were examined in control cells and TBP1 was found to be expressed consistently in K562, A549 and CD19 B-cells. (B) Relative expression to TBP1. K562 is positive for all genes except YAP1 where YAP1 was highly found in A549. (C) Relative expression to PRKG1 seems positive in all cell lines (C) except SMAD3 and ROCK1 in A549. (D) Relative expression to both TBP1 and PRKG1 in all cell lines. The experiments were performed as three technical repeats, and the average of the three repeats are shown.

After creating positive control cell lines, both reference genes are expressed at common abundance with lower variability among patient samples and healthy volunteers **(Figure 6.6)**. The expression of targeted antigens was analysed through qPCR experiments. The analysis included aB-ALL samples from PB and BM compared to Normal PB. The results were analysed relative to TBP1 **(Figure 6.7),** relative to the PRKG1 (**Figure 6.8)** and relative to the average of both TBP1 and PRKG1 **(Figure 6.9)**. Average Ct values obtained for TBP1 and PRKG1 genes ranging from 24.7-31.7, 30.2-36.5 respectively. The mean Ct ranges from 23-35 for BIRC5, 20-33 for ROCK1, 22-36 for SMAD3, 17-37 for YAP1, 35.9-18 for TEAD4, 18-35 for TCF4, 31-20.9 for SOX4. The ΔCt values were obtained by subtracting the Ct value of the reference gene TBP1 or PRKG1 from that of each antigen. A higher ΔCt value indicates a higher expression of the antigen in comparison to the endogenous control, while a minus sign indicates the genes low expression.

BIRC5 was used as a positive control as it has been previously shown to be upregulated in aB-ALL by our group (Boullosa et al., 2018). Comparing aB-ALL samples to those from healthy controls, the relative expression to TBP1 (Figure 6.7), TEAD4 and SOX4 have shown to be significantly upregulated in B-ALL with p value <0.05 and $p<0.01$, respectively. TEAD4 expression is also significant relative to PRKG1 (Figure 6.8). Figure 6.9 shows the relative expression of BIRC5, SMAD3, and TEAD4 were significant to the average of two reference genes (TBP1 and PRKG1). Using multiple comparisons, the expression of these antigens between samples from BM to normal PB showed a significant difference for SMAD3 and TEAD4 **(Figure 6.9)**. In contrast, genes expression between BM to PB-ALL is not significant. Although threshold values were higher than expected for positive control samples, data from TBP1 tests did not suggest any issues with the positive control cell lines. Although PRKG1 has been suggested as the suitable reference gene for studying lymphoid malignancies, Green *et al.* considered PRKG1 as a poor normaliser for snap-freeze lymphoid samples (Green et al., 2009). This may partially explain its poor performance compared to TBP1.

By comparing the relative expression our result of genes from 15 samples B-ALL containing four samples from bone marrow to the GSE13204 (Kohlmann et al., 2008) that contains 205 aB-ALL patient samples to 74 healthy controls, ROCK1 is found to be

upregulated in All types of B-ALL with p<0.05, except ALL t (8;14)/ ALL t(1;19)/ Pre-B-ALL t(9;22)- NS. TEAD4 is highly expressed in B-ALL with t (8;14), t(1;19), c-/Pre-B-ALL t(9;22) -p<0.001. YAP1 has different expressions depending on the probe set. It is downregulated in B-ALL (healthy bone marrow- p<0.01 from 224894-at and not significant from 213342-at, while it is expressed in ALL t (1;19)-NS from 224895-at). These suggest no difference in YAP1 expression between the B-ALL and normal samples. TCF4 is upregulated in ALL subtypes of B-ALL-p<0.001. SOX4 is upregulated in ALL subtypes of B-ALL- p<0.001 except ALL t (8;14). SMAD3 is upregulated all subtypes of B-ALL- p< 0.05.

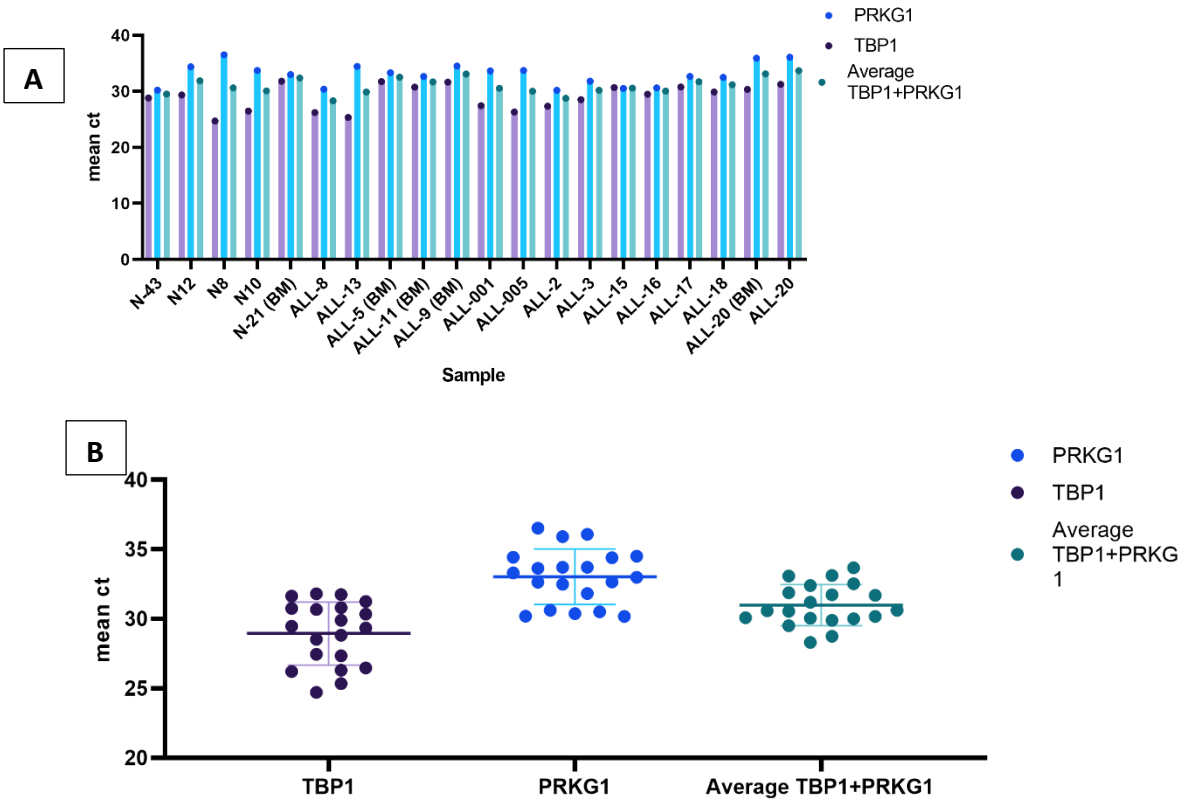


Figure 6.6 Mean Ct expression of the reference genes TBP1 and PRKG1 in aB-ALL and healthy donor samples

The mean Ct of TBP1 and PRKG1 were determined for each patient sample, TBP1 was found to be more consistent in its levels among aB-ALL patient samples. Each experiment was performed as three technical repeats, and the average of all three repeats are shown. Samples from 11 B-ALL patients using PB and 4 from BM are shown. In addition, healthy donors, 4 from PB (N-43, N12, N8 and N10) and 1 from BM (N-21) were analysed. Average of two reference was presented. (A) represents reference genes for each sample and (B) cluster of all samples.

6.3.3 ICC for TEAD4 and SMAD3 in B-ALL samples compared to CD19 B cell

TEAD4 and SMAD3 levels were determined in K562 (positive cell line) ensuring the antibodies are working (**Figure 6.10 A**). Comparison of isotypes antibodies shows that both SMAD3 and TEAD4 are recognised by B-cells. TEAD4 protein was mostly found in nucleus whereas SMAD3 was found in both the nucleus and cytoplasm. The levels for TEAD4 and SMAD3 had significantly higher H-Scores compared to their isotype control antibodies aB-ALL samples (**Figure 6.10B**). Figure 6.10B shows H-scores analysis for SMAD3 and TEAD4 in B-ALL samples compared to CD19 B-cell were available only from 1 donor in 3 technical replicates, H-score values of SMAD3 and TEAD4 were upregulated compared to CD19 B-cell. However, TEAD4 expression in BM is significantly higher than in PB.

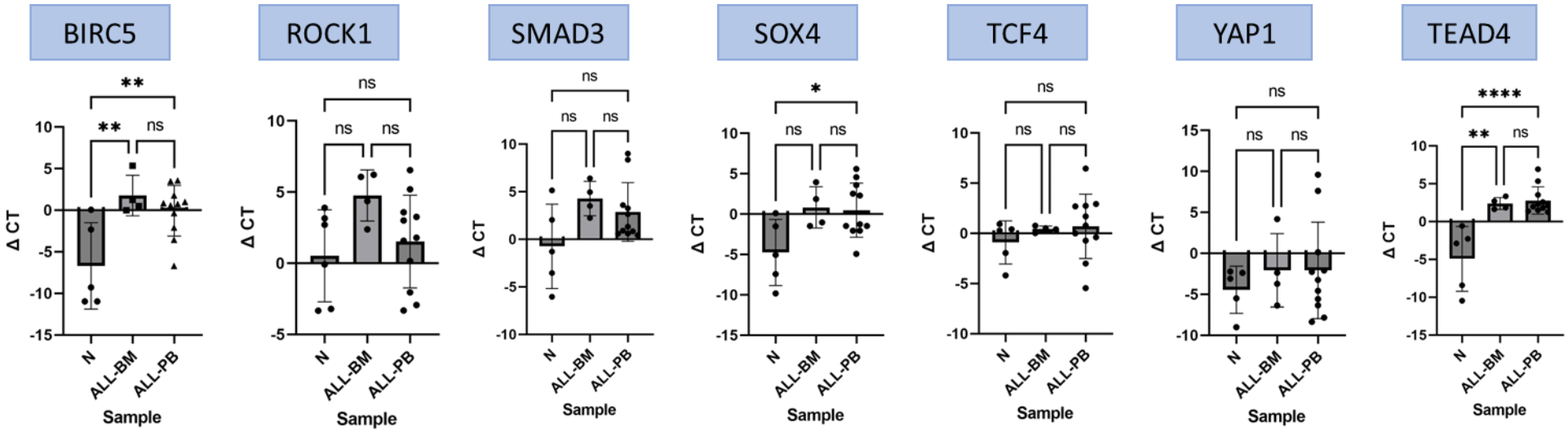


Figure 6.3 Relative expression of BIRC5, ROCK1, SMAD3, SOX4, TCF1, YAP1, and TEAD4 compared to the house keeping gene TBP1

PB samples from B-ALL (11), BM samples from B-ALL (4) and healthy donors (N=5 with 4 samples from PB and 1 from BM). Data is normally disturbed (Kolmogorov-Smirnov and Shapiro-Wilk tests) shown as mean ± SD. Statistical significance was calculated using one ANOVA test (Dunnett test). BIRC5 upregulation in aB-ALL compared to normal donor PB (**). SOX4 was upregulated in aB-ALL PB compared to normal PB (*). TEAD4 was found to be highly upregulated in aB-ALL (***). *p≤0.05, **p≤ 0.01, and ***p≤0.001. Each experiment was performed as three technical replicates and the average is shown.

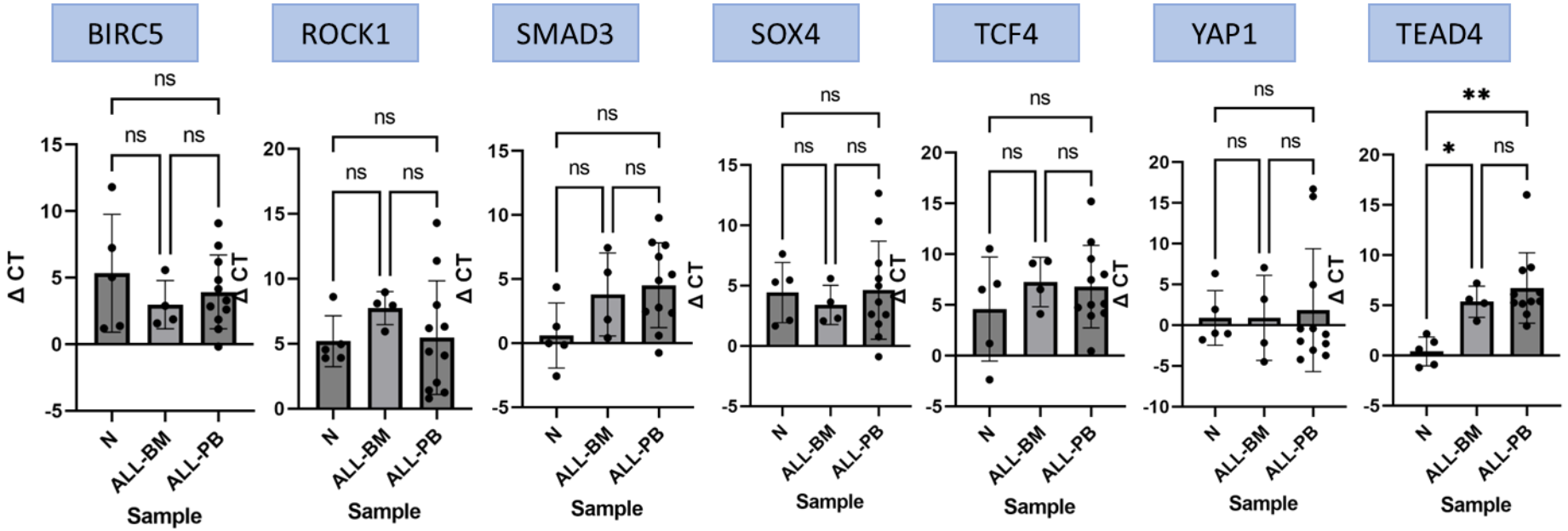


Figure 6.4 Relative expression of BIRC5, ROCK1, SMAD3, SOX4, TCF1, YAP1, and TEAD4 when compared to the PRKG1 housekeeping gene

PB samples from B-ALL (11), BM samples from B-ALL (4) and healthy donors (N=5 with 4 samples from PB and 1 from BM) were analysed. Data is normally disturbed (Kolmogorov-Smirnov and Shapiro-Wilk tests) shown as mean ± SD. Statistical significance was calculated using one ANOVA test (Dunnett test). All genes show no statistical significance between samples from aB-ALL patients and normal donors except for TEAD4 which was highly expressed in aB-ALL PB and BM samples compared to healthy donor (N) samples ($p<0.01$). TEAD4 was found to be highly upregulated in aB-ALL (***). *$p<0.05$, **$p< 0.01$, and ***$p<0.001$. Each data point is representative of three technical replicates and the average is shown.

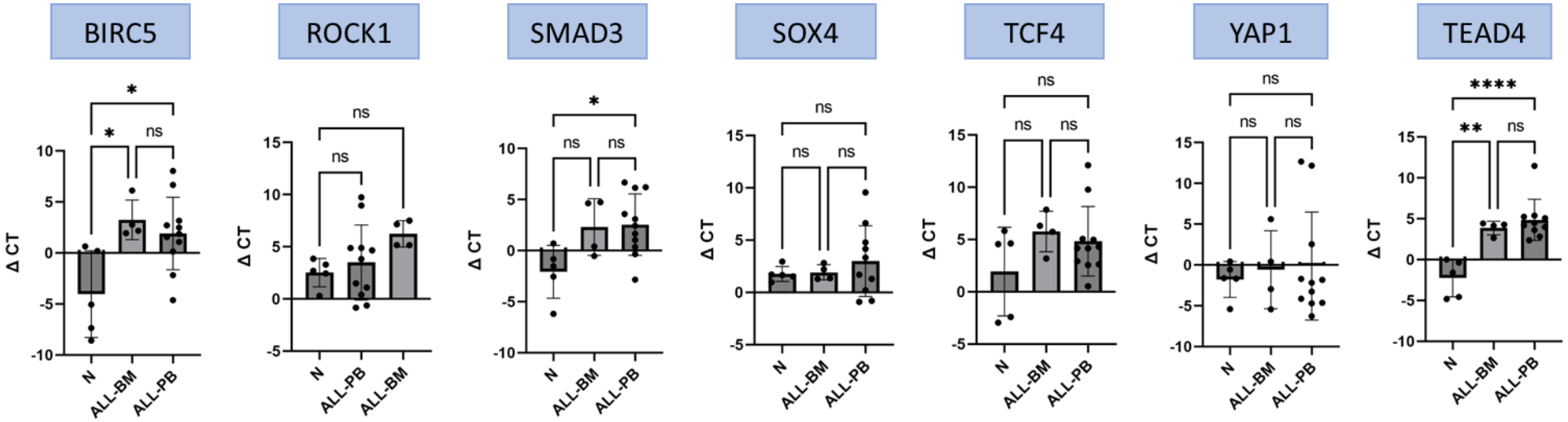
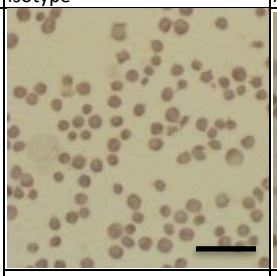
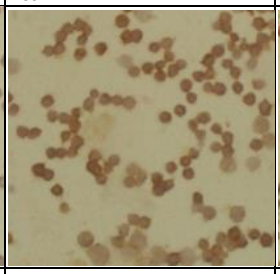
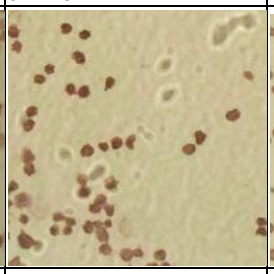
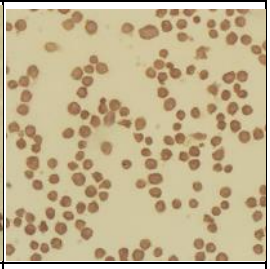
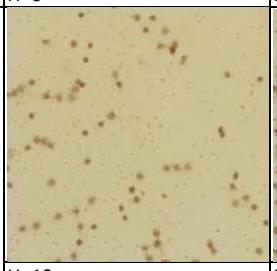
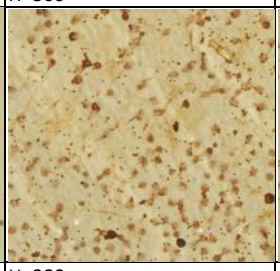
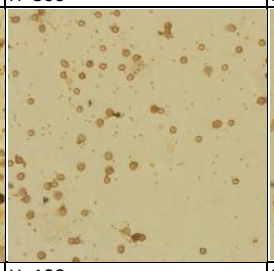
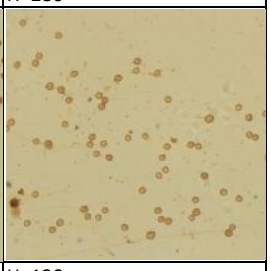
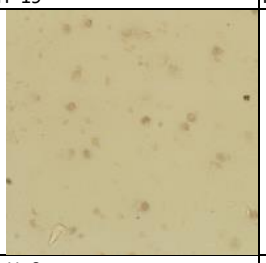
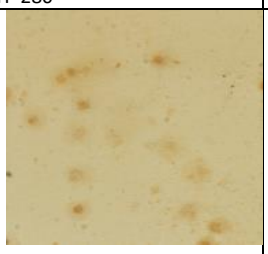
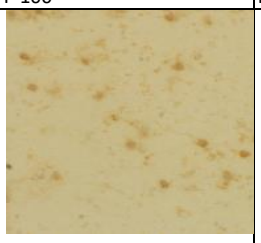
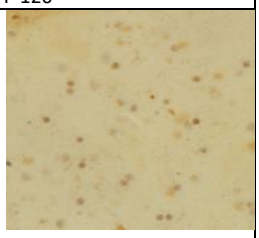
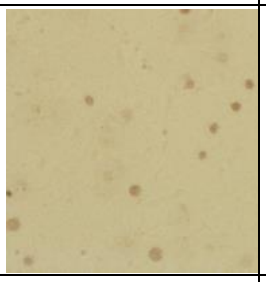
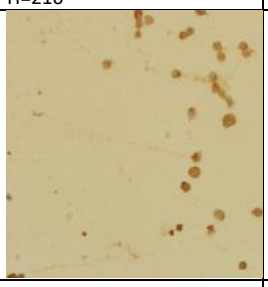
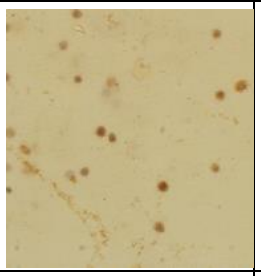
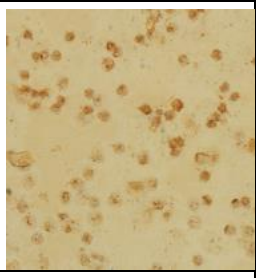


Figure 6.9 Relative expression of BIRC5, ROCK1, SMAD3, SOX4, TCF1, YAP1, and TEAD4 compared to the average of two reference genes

PB samples from B-ALL (11), BM samples from B-ALL (4) and healthy donors (N=5 with 4 samples from PB and 1 from BM) were analysed. Data is normally disturbed (Kolmogorov-Smirnov and Shapiro-Wilk tests) shown as mean ± SD. Statistical significance was calculated using one ANOVA test (Dunnett test). All genes show no statistical significance between samples from aB-ALL patients and normal donors except for BIRC5, SMAD3, and TEAD4 which was highly expressed in aB-ALL PB and BM samples compared to healthy donor (N) samples ($p<0.05$, $p<0.05$ and $p<0.01$ respectively).

A

	Isotype	Actin	SMAD3	TEAD4
K562				
	H=0	H=300	H=300	H=280
CD19 B-cell				
	H=19	H=280	H=100	H=120
ALL020 PB				
	H=9	H=210	H=200	H=190
ALL020 BM				
	H=0	H=305	H=200	H=270

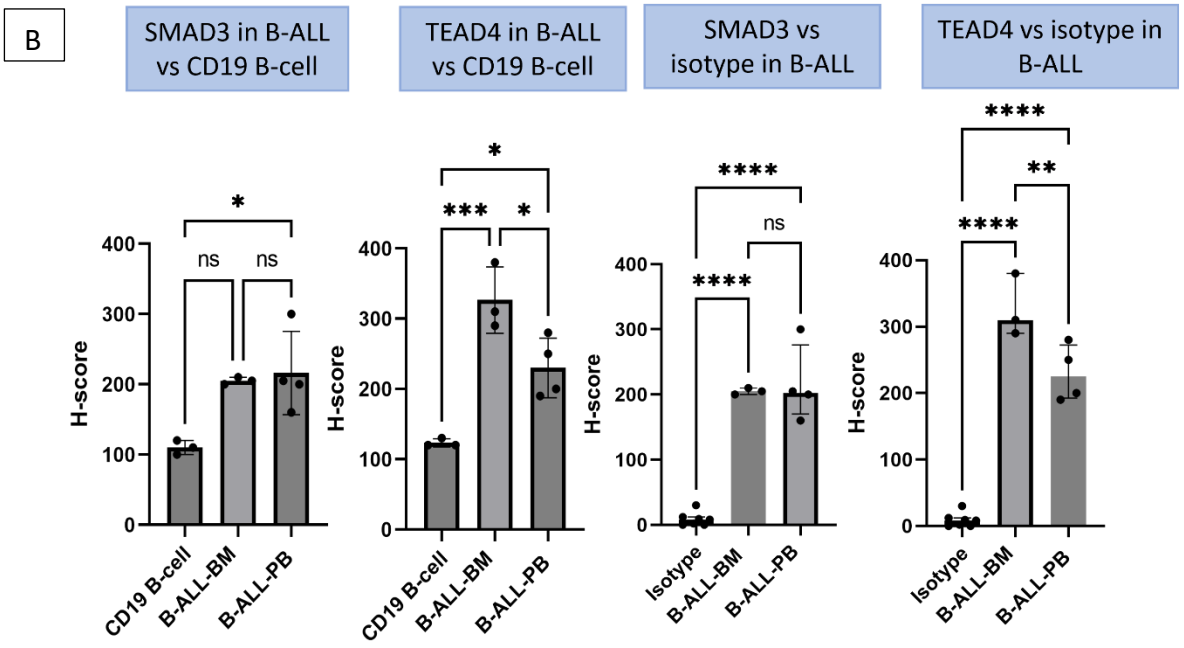


Figure 6.10 Immunolabelling of SMAD3 and TEAD4 in cell lines and patient samples
(A) Cells labelled with isotype control appearing purple/or slightly brown as it was used as negative control and has a low H-score. Actin was used as a positive control appearing mostly brown and receiving a high h-score. SMAD3 and TEAD4 were expressed in all patient samples but at different intensities indicating their varied expression. Immunoreactivity scores were as follows: 0: negative; 1–29: weak; 30–143: moderate (mod) and 144–227: high; >228: very high. The bar size was 20µm (40x magnification). (B) Bar graphs show H-scores analysis for SMAD3 and TEAD4 in B-ALL samples compared to CD19 B-cell (1 sample performing three technical repeats). H-score values were normally disturbed (Kolmogorov-Smirnov and Shapiro-Wilk tests) and Dunnett test was performed, both SMAD3 and TEAD4 upregulated compared to CD19 B-cell. However, TEAD4 expression in aB-ALL BM was significantly different from PB samples. Compared to isotype, SMAD3 and TEAD4 were highly significant ***. *p<0.05, **p< 0.01, and ***p<0.001. Each experiment was performed as a technical triplicate and average shown.

6.4 Discussion

This chapter identified three pathways (TGFβ, Wnt, and Hippo pathways) highly enriched in aB-ALL and focused on the expression of antigens in B-ALL patient samples, compared to healthy donors. The selected antigens, namely SOX4, ROCK1, YAP1, SMAD3, TCF4, and TEAD4, have all been previously linked to cancer and were chosen specifically due to their high values from Cheever model ranking. After thorough analysis, it was found that TEAD4 and SOX4 were upregulated in B-ALL (p-value of 0.001 and 0.05 respectively). There were non-significant trends in expression of ROCK1 and

TCF4 in B-ALL, with more samples are required to verify their expression in B-ALL to have confidence interval >95% in qPCR. Discordance of transcript and protein levels of gene identified in literature (Buccitelli & Selbach, 2020), this may be due to posttranscriptional modifications, differences in mRNA and protein turnover.

SOX4 belongs to the sex-determining region Y-related HMG box family. It is a transcriptional factor that plays a significant role in regulating cell stemness, proliferation and differentiation. Its abnormal expression is associated with cancer and is known to regulate several cancer-promoting signalling pathways such as Wnt, TGFβ, and PI3K (Luo et al., 2022). In cB-ALL, SOX4 is upregulated and linked to poor survival. It is also crucial in activating MAPK and PI3k/Akt in ALL (Ramezani-Rad *et al*., 2013). In T-ALL, SOX4 is highly expressed and associated with the downregulation of miR-204, which acts as a tumour suppressor. Notably, SOX4 is essential in B-cell development, and its absence results in development arrest at the pro-B to pre-cell transition. In SOX4-deficient mice, pro-B cells were not able to proliferate because of the absence of IL-7 production, which led to their arrest at the pre-B cell receptor checkpoint and prevented further differentiation. In addition to affecting the production of IL-7, SOX4 has been found to interact with IL5Rα chain in early B-cell development through the recruitment of syntenin (Ramezani-Rad *et al*., 2013). SOX4 promotes cancer cell survival in Ph+ subtype by directly binding to the promoters of genes involved in PI3K and BCL-xL. Consequently, this leads to an increase in the expression of MAPK8(JNK1), PI3k/Akt, and BCL-xL mRNA levels. In light of this, pharmacologic targeting of the downstream signalling pathway of PI3K/AKT has been identified as an important pathway in human Ph+ ALL. The BCR-ABL1 oncogene activates the JNK pathway, and when MAPk8 (JNK1) is disrupted, it results in defective BCR-ABL1 driven pre-B ALL transformation. In Ph+ leukaemia, Bcl-xL is a vital survival molecule. Stat5-dependent transcriptional activation of Bcl-xL promotes the inhibition of apoptosis. An important discovery is that SOX4 is a critical upstream regulator that enables Bcl-xL, JNK1, and PI3K/AKT signalling in Ph+ ALL. Bcl-xL is an antiapoptotic protein that promotes survival during the early stages of B-cell development. Pro-B cells exhibit high levels of Bcl-xL expression.

It is noteworthy that transgenic mice exhibiting increased expression of Bcl-xL show significant expansions in their B-cell populations. This expansion is attributed to

improved survival during the VH-DJH gene rearrangement process from pro- to pre-B cells. The stage-specific and gradient regulation of Bcl-xL expression by SOX4 aligns with the observed differentiation block in Sox4-deficient mice at the pro-B cell stage. In the absence of SOX4-induced Bcl-xL, pro-B cells are unable to survive the VH-DJH gene rearrangement process at the pre-B cell checkpoint, leading to their inevitable demise, as explained by Ramezani-Rad *et al.* (2013) (Ramezani-Rad *et al*., 2013).

In addition, SOX4 is involved in the negative regulation of p53 and Arf (CDKN2A) expression, which are critical negative regulators of self-renewal. This negative regulation is particularly important in mediating leukaemia colony formation. Interestingly, downregulation of Arf has been found to result in a loss of viability of leukaemic cells following treatment with TKI. SOX4 deletion results in rapid loss of viability and apoptosis in BCR-ABL1–transformed ALL, but not normal pre-B cells due to reduced expression of Bcl-xL (Bcl2l1) and increased levels of Arf/p53 (Ramezani-Rad *et al*., 2013). Moreover, SOX4 expression is a predictor of the outcome of B-ALL, with patients who have lower expression of SOX4 having better outcomes and survival ($p=0.07$) compared to those with high SOX4 expression. This supports that high-risk ALL has high SOX4 expression due to lower levels of SOX4 promoter methylation. SOX4 levels may independently predict short survival ($p=0.04$), from analysing 206 ALL patients (Ramezani-Rad *et al*., 2013). SOX4 has been identified as a crucial upstream regulator of B-ALL survival via enhancing BCL-xL and PI3K/Akt signalling and negative regulation of Arf and p53. SOX4 is also upregulated in AML and is associated with poor survival (Ramezani-Rad *et al*., 2013). In addition, SOX4 is known oncogene and upregulated in various solid cancers, such as prostate, breast, and adenoid cystic carcinoma (ACC) and melanoma, and is linked to malignant transformation and metastasis (Lu *et al*., 2017).

The inhibition of SOX4 through siRNA treatment resulted in a significant decrease in proliferation, invasion, and migration of the ACC cell line (Lu *et al*., 2017). The knockdown of SOX4 protein led to a 51% decrease in cell viability and an increase in apoptosis to 85% (Lu *et al*., 2017). This was confirmed through various methods, including cell morphology, DNA fragmentation, and flow cytometry. Interestingly, co-transfection with a functional SOX4 construct was able to rescue cells from the proapoptotic effects of SOX4 RNAi (Lu *et al*., 2017). Microarray gene expression profiling

revealed that the downregulation of SOX4 influenced the expression of critical genes involved in apoptosis (p53) and cell cycle control (CDKN2A) (Pramoonjago *et al*., 2006).

Taken altogether, SOX4 may be an attractive target in aB-ALL. While transcription factors are generally challenging to inhibit pharmacologically, BCL6 may be an exception due to its reliance on protein-protein interactions with important cofactors (Ramezani-Rad *et al*., 2013). The identification of critical cofactors involved in SOX4-mediated transcriptional activation of BCL2L1 (BCL-xL) and PI3K/AKT, suggests that this oncogenic pathway may become a target of small molecules inhibitors, such as those targeting BCL6.

TCF4 is upregulated in T-ALL and associated with poor prognosis. TCF4 up regulates BIRC5 expression, potentially increasing cell viability (Li et al., 2020b). Additionally, TCF4 is significantly up-regulated HSCs and promotes tumour progression. TCF4 as a target gene of miR-7-5p, which negatively regulates its expression, while ANRIL positively regulates it (Li et al., 2020b). This suggests that the ANRIL/miR-7-5p/TCF4 axis may play a role in the tumourigenesis and progression of T-ALL. However, future investigations of TCF4 in B-ALL are needed.

The core component of the Hippo pathway includes YAP1 and TEAD4. YAP has been found to be overexpressed in CML compared to healthy donors, and suppressing its expression can inhibit cell proliferation and induce cell cycle arrest in the G1 phase and the induction of apoptosis in CML cells (Li *et al*., 2016). YAP can have a significant anti-tumour effect in CML. The findings obtained from the *in vivo* leukemogenesis assay further supported the notion that suppressing YAP can augment the efficacy of imatinib (IM) against CML cells (Li *et al*., 2016). BCR/ABL positive cell lines exhibited elevated levels of YAP protein expression compared to their BCR/ABL negative counterparts, although there was no marked variation in YAP transcription levels between them (Li *et al*., 2016). Treatment with IM, which inhibits BCR/ABL kinase activity, led to a reduction in YAP protein expression.

Compared to normal healthy tissues (excluding stem/progenitor populations), cancer cells express high levels of YAP, making them a potential target for cancer-specific drugs (Chen et al., 2019). However, since YAP primarily functions in the nucleus as a

transcriptional coactivator, their accessibility to administered drugs may be limited, particularly for macromolecules like antibodies (Steinhardt et al., 2008). Therefore, small molecule inhibitors to YAP/TAZ are being sought after. The Porphyrin family, including veterporfin (VP), hematoporphyrin (HP) and protoporphyrin IX (PPIX), have been discovered to inhibit YAP by disrupting its interaction with TEAD *(*Qiao *et al.,* 2018*)*. VP suppresses CML cell proliferation and enhances the efficacy of IM both *in- vitro* and *in-vivo (*Qiao *et al.,* 2018*)*. VP induces a conformational change in YAP and up-regulates the protein level of 14-3-3σ, a chaperone protein that retains YAP/TAZ in the cytoplasm (Wang et al., 2016a). Forskolin can prevent cAMP breakdown and promote YAP phosphorylation, which may be beneficial in treating cancers with YAP oncogenic activity (Wang et al., 2016a). PDE inhibitors could also be used for this purpose. Dobutamine has anti-cancer properties and has been found to cause YAP-Ser127 phosphorylation, which can suppress YAP-dependent gene transcription (Yu et al., 2013b). Latrunculin B and cytochalasin D can limit YAP activation by disrupting the actin cytoskeleton and inhibiting nuclear YAP localization through increased LATS activity (Seo & Kim, 2018). Finally, dihydrexidine has been found to increase YAP phosphorylation and inhibit the hippo signalling pathway (Qiao et al., 2018).

Silencing YAP has been shown to improve the inhibition of survivin, suggesting that simultaneously targeting both molecules may lead to better therapeutic outcomes (Li *et al*., 2016). In CML, c-Myc and survivin are involved in disease progression, with BCR/ABL up-regulating c-Myc. Imatinib (IM) down-regulates c-Myc in CML cell lines, and higher levels of c-Myc at diagnosis correlate with poorer response to IM (Huang *et al*., 2013). Additionally, targeting survivin sensitizes CML cells to IM and other cytotoxic drugs. It was previously shown that genetic and pharmacological inhibition of YAP significantly reduces the expression of c-Myc and survivin (Li *et al*., 2016).

However, YAP may act as a tumour suppressor in MM and is inactivated due to focal deletion in all MM cell lines and 3-15% of MM samples (Maruyama et al., 2018). This deletion targets BIRC1 and BIRC3, regulating the oncogenic pathway NF-κB. YAP also interacts with the RUNX decreasing p73 degradation (Walker *et al*., 2010; Cottini *et al*., 2014). Thus, interacting with the oncogenic modulator keeps YAP1 from p73, granting YAP1 proliferative and anti-apoptotic features. Low YAP1 expression is associated with

poor survival compared to the MM sample with high expression. Furthermore, YAP1 knockdown may lead the normal plasma cell progress to MM (Cottini *et al*., 2014). Rescuing YAP activity may be achieved by synthetic lethality for selective target cancer cells with endogenous DNA damage and low YAP1 levels (Kaelin, 2005). STK4 knockdown via shRNA restores YAP expression enhancing cancer death *in-vivo* and *in-vitro*. Therefore, STK4 may be a therapeutic target for cancer with p53 inactivation and poor prognosis (Kaelin, 2005). YAP1 expression in B-ALL needs to be verified in large sample size to confirm upregulation and functions in B-ALL. Similar to our study, YAP1 expression is found in 33% of ALL without any significant difference between ALL and healthy donors (Machado-Neto et al., 2014).

TEAD4, the second element of the Hippo pathway, is activated by co-regulators such as YAP, TAZ, p160, and VGLL-4. EMT is driven by YAP, and the transcription factor TEAD is essential for cell division. Normal cell growth requires that TEAD-YAP activity be tightly regulated. In quiescent adult normal cells, the Hippo route closely regulates YAP through cell-cell interaction. But during neoplastic transformation, YAP is frequently dysregulated, which results in increased and carcinogenic TEAD-YAP activity. Through the overexpression of target genes like CTGF and CYR61, this aberrant activity promotes cell proliferation and transformation, which aids in the aetiology of a number of human malignancies (Steinhardt et al., 2008).

TEAD4 is upregulated in CLL. Overexpression of TEAD4 in CLL is associated with the Richter transformation (RT), which is a significant progression of CLL that results in the development of an extremely aggressive large B cell lymphoma. This transformation is associated with a poor patient prognosis (Nadeu *et al*., 2022) and activates genes that involved in oxidative phosphorylation via mTOR activation.

Overexpression of TEAD4 has been reported in ADC, GC, and urinary bladder cancer with different mechanisms promoting tumourigenesis and eventually associated with poor survival (Hu et al., 2021). TEAD4 targets pyruvate kinase isozymes M2 (PKM2), a crucial glycolytic regulatory enzyme, confirmed by in-vitro experiments with a luciferase reporter gene assay (Hu et al., 2021). This resulted in increased activity of HIF-1α and subsequent tumour progression.

TEAD4 overexpression was associated with poor prognosis in GC patients, significantly correlating with advanced tumour stage, lymph node metastasis, and poor survival (Giraud *et al*., 2020). TEAD4 overexpression promotes GC cell proliferation by activating the expression of cyclin D1 and c-myc genes through the Hippo signalling pathway. TEAD4 overexpression enhances gastric cancer cell migration and invasion by upregulating matrix metalloproteinase-7 (MMP-7) expression (Giraud *et al*., 2020). TEAD4 expression is correlated to EMT genes, and inhibition of Twist1 suppresses EMT suggesting that TEAD4 may be an upstream positive regulator of Twist1, negatively regulated by FOXO3a (Shen *et al*., 2017).

Hypoxia plays a central role in the leukaemia microenvironment modulating angiogenesis, apoptosis, and energy metabolism (Petit *et al*., 2016). This also induces resistance in ALL cell lines towards multiple chemotherapies by stabilizing HIF-1α increasing antiapoptotic genes BCL1 and MCL1(Petit *et al*., 2016). TEAD4, via PKM2, triggers HIF-1α overexpression, associated with chemotherapy resistance potentially linked to B-ALL. By examining the expression of TEAD4 in B-ALL samples and normal bone marrow in the GSE13204 dataset, TEAD4 exhibited an increase in gene expression. It was only associated with survival (p=0.02) compared to the other three members of the TEAD family. As TEAD4 is upregulated in B-ALL in our findings, further study is required to investigate the associations between TEAD4 overexpression and clinicopathological parameters and prognosis. Additionally, functional studies will be needed to determine if TEAD4 overexpression promotes leukemogenesis.

Upregulation of SMAD3 is found in CML, especially the BCR-ABL subtype (Møller *et al*., 2007). TF-1 myeloid cells expressing BCR-ABL were more effectively arrested by TGFβ than the original TF-1 cell line (Møller *et al*., 2007). Furthermore, TGFβ inhibited the growth of CD34+ cells expressing BCR-ABL from chronic phase CML patients. Interestingly, when a non-proliferating CD34+ CML cell sub-population was treated with the TGFβ kinase inhibitor SB431542, it enhanced cell death mediated by the BCR-ABL inhibitor imatinib. These results suggest that BCR-ABL expression increases myeloid cell sensitivity to TGFβ. This novel cross-regulatory mechanism may play a crucial role in maintaining CML transformed progenitor cells (Møller *et al*., 2007). Therefore,

combining imatinib with TGFβ kinase inhibitors could offer additional therapeutic benefits for CML patients.

SMAD3 binds to Egr3 upregulating its expression, resulting in arrest of HSC proliferation. In a non-irradiated AML model, Egr3 was highly expressed in HSCs within the leukemic bone marrow. Overexpression of Egr3 caused cell cycle arrest and the restoration of HSC defects observed in leukaemia, while knockdown of Egr3 promoted HSC proliferation and improved their engraftment ability (Cheng et al., 2016). Notably, Egr3 knockdown restored the cycling of HSCs in leukemic marrow, indicating its role in suppressing HSCs in leukaemia. Conversely, upregulation of MafF and Hey1 enhanced the function of hematopoietic stem and progenitor cells (HSPCs), suggesting their role in preserving HSCs in leukaemia (Cheng et al., 2016). The Hes1-Cdkn1a axis was identified as a mediator for cell cycle alteration of HSCs in the ALL model. Exogenous activation of Hes1 silenced HSPCs and preserved their function in leukaemia. Additionally, microarray analysis identified other genes, such as Nr4a2 and Egr1, that suppress HSC function in leukaemia (Cheng et al., 2018). A small molecule inhibitor, RepSox, inhibits TGF-β-RI/ALK5 pathway, specially SMAD2/3 phosphorylation. RepSox also increases histone H3S10 phosphorylation and reduced H3S28 phosphorylation, thus resulting in programming of embryonic fibroblasts into iPSCs, in a murine model (Guo et al., 2021). This suggests RepSox induced differentiation of fibroblasts via activating of histone methylation and decreasing of histone acetylation. Overexpression of SMAD3 in B-ALL from our findings may suggest that SMAD3 have a role in B-ALL pathogenicity and further studies may reveal its function and how this may be targeted in B-ALL to improve treatment outcomes.

SOX4, ROCK1, YAP1, SMAD3, TCF4 and TEAD4 are involved in the maintenance of CSC features (Moreno, 2020). SOX4 is involved in EMT regulation, angiogenesis, increasing cell-stemness, pro-survival, and proliferative signalling suggesting its oncogenic function (Pramoonjago *et al*., 2006; Medina *et al*., 2009). TGFBR2, on the other hand, regulates SOX4-mediated PI3K activation and is a direct target of SOX4 activating the downstream PI3K/Akt signalling. Hence, TGFBR2 plays a crucial role in activating non-canonical PI3K signalling (Moreno, 2020). Both SOX4 and TGFβ activate SMAD signalling and induce EMT. ROCK1 is a non-canonical activator of Wnt promoting breast cancer proliferation,

stemness, and migration (Mohammadi-Yeganeh et al., 2016). ROCK1 is overexpressed in AML and associated with poor survival ($p < 0.01$). ROCK1 knockdown of Kasumi-1 and AML-193 cell lines enhances tumour cell apoptosis and significantly inhibits the blast proliferation ($P < 0.05$). It has been suggested that miR340-5p downregulation is the main cause of ROCK1 upregulation (Liu et al., 2019).

Targeting Wnt components have been tested in clinical trials. The tankyrase inhibitor, XAV939, has a mechanism of action targeting interaction between β-catenin and tankyrase disturbing the BM niche and increasing β-catenin degradation (Tanton et al., 2022). A CBP/β-catenin transcription inhibitor, ICG-001, decreased self-renewal of B-ALL, BIRC5 expression and the chemo-resistance (both *in-vitro* and *in-vivo*). Further, 18.3% of relapse cases had CBP mutations, associated with an adverse prognosis (Tanton et al., 2022). Given that Wnt pathway is involved in the HSC maintenance and differentiation, further work is needed to assess the safety of selective inhibition.

The result variation highlights the heterogeneity of aB-ALL and finding an antigen that is expressed in all subtypes but not healthy tissues remain challenging. Therefore, these antigens may be used to stratify patients according to their expression and determine which group may have the high benefit from targeting these TAs either alone or in combination with available treatment options. Biomarkers can be utilized to monitor treatment progress and inform drug and dosage modifications. Well-designed clinical trials must be conducted to gain insights and to construct more effective immune targeting therapies and treatment protocols. Further analysis using an *in-vitro* investigation of the possible interactions among B-ALL signalling and other identified pathways (TGFβ, Wnt, Hippo pathways) are warranted to clarify the role of these pathways landscape as a predictor for targeting and prognosis in different types of aB-ALL along with the known pathways such as FOXO and B-cell signalling (Table 9.17 and Table 9.18). However, cluster analysis using cytogenetic prognosis rather than individual subgroup analysis may provide more rapid results and may be less expensive in the future. Consequently, this analysis may provide in-depth understanding of B-ALL pathogenesis and which targets may have clinical benefits, and others are not with tolerable side effects.

Chapter 7: General discussion

7.1 Introduction

This chapter describes and discusses the findings of the studies of the mixed method designs used in this PhD thesis. The designs include a systematic review, which was carried out to provide a comprehensive and focused view of the diagnostic biomarkers: used in the early detection of NSCLC; followed by the analysis of RNAseq (GSE81089) for the identification of the biomarkers which enable the earlier diagnosis of NSCLC. The second part analysis of aB-ALL sera for the identification of novel antigens which are potential therapeutic targets for treatment intervention.

7.2 Potential role of highly sensitive and specific NSCLC biomarkers

This study has investigated and established the TAs that act as biomarkers for the early detection of NSCLC. Early detection of the disease condition relates to a favourable prognosis and a 5-year survival rate of 70-90% for small-localised tumours (Goldstraw et al, 2016), emphasising the importance of early detection. However, even with recent significant research in NSCLC diagnosis, ~75% of cases are diagnosed at the advanced stage of the disease (Blandin Knight et al, 2017). While CEA and CYFRA21 have been found to be highly expressed during the later stages of NSCLC as non-invasive LC diagnosis, they have shown low sensitivity in the earliest stages of the diseases. Therefore, early detection of LC would need the biomarkers having the combination of with high sensitivity and specificity, especially in individuals with inconclusive high-risk features.

Findings from our systematic review of 79 articles have identified the potential biomarkers that could be used for the early diagnosis of the NSCLC, with a pooled area under the curve of 0.86, suggesting an excellent diagnostic performance (Figure 2.3). Biomarkers with high sensitivities included Ciz1, exoGCC2 and ITGA2B, while those with high specificities were CYFRA21-1, antiHE4 and OPNV, along with miR-15b and miR-27b (Table 2.1). However, these biomarkers have the potential to improve early NSCLC detection. Importantly, Ciz1 has previously been implicated in the diagnosis of neurological tumours (Main, 2021), and its variant-Ciz1b has shown a promising

biomarker for the early detection of NSCLC in this research. This evidence supports other researchers (Sun et al, 2019; Odera etal, 2021; Chen et al, 2022), who have emphasised the importance of this biomarker in the early diagnosis of NSCLC at the early stage for case management. The Ciz1 biomarker is essentially sensitive for NSCLC. Furthermore, while the challenges in the field of biomarkers for LC screening have been emphasised, this thesis highlights the biomarkers with high specificities (CYFRA21-1, antiHE4 and OPNV, along with miR-15b and miR-27b). These biomarkers could be targeted alongside the high sensitive biomarkers in order to improve the diagnosis at the early stage of LC. This is because, to our knowledge, this is the first systematic review that has identified these three biomarkers together in their role of early detection of NSCLC. Our findings also reiterated the importance of GCC2 and ITGA2B in the early detection of NSCLC (Table 2.2). These biomarkers are upregulated in cases of NSCLC and could however play a significant role in the early diagnosis of LC. While there are challenges including time constraints in the analysis methodology for these biomarkers as potential diagnostic tools (Huang et al, 2022), GCC2 has shown a distinguishing feature and exosome potential for the early diagnosis or the detection of LC and it could be validated to distinguish individuals with stage I-II NSCLC. Now, these highly sensitive Ciz1, exoGCC2 and ITGA2B biomarkers could be targeted together in the early detection of LC as the result of the findings of this study. This would primarily benefit early LC detection. More studies in this regard are solicited to fully provide the yardstick for future early diagnosis of NSCLC.

Moreover, CTAs are commonly expressed in cancers including LC and their expression is associated with advanced disease stages and poor prognosis. COLL11A1 was found to be correlated with early NSCLC (Figure 4.3). Finding another dataset containing different controls such as inflammatory lung diseases to examine the changes of these DEGs will be the next stage for results validation.

7.3 The role of biomarkers studies in the diagnosis of cancer

An effective biomarker would lead to significant change in clinical decisions and management. In case of indetermined nodules, many biomarkers do not work may limit clinical applications. The appropriate control selection and study design are important for identifying diagnostic biomarkers (Ost & Gould, 2012; Deppen & Grogan, 2015).

Misdiagnosis may result when nodule characteristics on a CT scan overlap between benign and malignant nodules. Sometimes, biomarkers may not add the diagnostic value above thar of the radiological methods (Deppen & Grogan, 2015). Furthermore, sample types such as plasma, urine or image types and patient characteristic with their history are crucial for biomarker studies. Tumour heterogeneity makes the use of a single biomarker for cancer diagnosis, for clinical decision virtually impossible. However, the use of biomarker panels shows promise in such settings. Genomic and protein techniques have their own challenges for quantification, because the validation of results may be yet to be determined (Cohen et al., 2018). At times, biomarker discovery may not add value to the diagnostic value as it is repetitive of the findings of radiological image. For example, in the case of a 65-year-old heavy smoker patient having 30 mm lung nodule with spiculated margin, blood sample analysis would not assist much and tissue biopsy would be required for diagnosis confirmation (Kerr et al., 2019). Consequently, biomarkers testing requires the representation of a meaningful clinical endpoint and validation in line with the clinical context. Biomarkers validation is required to include the current clinical guidelines in order to evaluate true diagnostic performance. Appropriate statistical tests and assessment of risk classification are crucial for improving biomarker use.

Biomarkers may have the benefit of providing more information about the patient status, not known to the clinician and leading to another relevant decision. Biomarkers screening strategies are therefore required for meaningful clinical point and accuracy (Kerr et al., 2019). Rigorous quality-controlled procedure ensures high sensitivity and specificity, with high throughput screening is not validated against the clinical standard to assist early lung diagnosis. However, as described in the introduction, every candidate biomarker for detection of cancer should be taken through the standard stages of biomarker discovery. According to the EDRN (See Section 1.5.1), there are five phases of biomarker discovery. It is important to pay attention to the pitfalls of these procedures and ensure that the discovered biomarker is a representation of the variations in the early stage of NSCLC. The EDRN develops a reference set which includes a clear clinical application and specimens that represent the disease without bias, and it must mainly match age, sex to improve biomarker discovery.

Moreover, as most of the previous studies focusing on biomarker discovery start with retrospective studies, limited samples availability, and storage may affect the stability of analytes particularly RNA. In addition, it is essential to conduct further validation studies on larger and diverse patient populations to assess the performance of the biomarkers. This helps to determine if the initial findings are true across different cohorts and if they can reliably distinguish between cancer patients and healthy individuals. Furthermore, the next step is to explore the possibility of using multiple biomarkers in order to enhance their sensitivity and specificity. By combining different discovered biomarkers that complement each other, it is possible to achieve higher accuracy in cancer detection. Also, it is essential to refine the cut-off values, with adjustments used to define positive or negative results for the biomarkers. This optimisation process can help strike a balance between sensitivity and specificity, ensuring that true positives are not missed while minimizing false positives. Integration of the biomarker investigation with other diagnostic tools is the next step. Combination of biomarkers with other diagnostic techniques such as imaging (e.g., CT scans) or genetic testing (e.g., gene expression profiling), integrates multiple approaches providing a more comprehensive assessment of cancer presence and improve the overall accuracy (Kerr et al., 2019). Conducting a longitudinal analysis of biomarkers is imperative to evaluate the biomarker ability to detect cancer at different stages or/and monitor disease progression. This approach helps in the assessment of whether the biomarker is effective in early-stage detection or monitoring treatment response. The data obtained so far are managed by engaging the machine learning algorithms. This step involves the utilisation of advanced computational methods like machine learning algorithms to develop predictive models which incorporate various clinical parameters alongside biomarker data. These models can enhance diagnostic accuracy by considering multiple factors simultaneously. The lessons obtained from the data obtained so far would assist with the design of clinical trials. It is important to conduct a prospective clinical trial to evaluate the performance of the identified biomarkers in real-world settings, where their effectiveness can be assessed against the standard diagnostic methods or existing gold standards. The last evaluation procedure is to conduct an external validation. This is where there is collaboration with independent research groups or institutions for external validation of the identified biomarkers. This

helps ensure the reproducibility and generalizability of the findings, increasing confidence in biomarker clinical utility.

7.4 Identifying TAs as targets for B-ALL therapy

Due to the existence of escape mutants, the heterogeneity between and within tumours and inevitable limitations with existing immunotherapy treatments, it is essential to broaden the number, specificity and sensitivity of the TA targets available for clinical use. Finding new TAs has the potential to improve outcomes and decrease the risk of relapse, critical for extending long-term survival.

Characteristically, around 40% of identified clones corresponded to uncharacterised genes. For the known genes (72), these were prioritised by Cheever criteria as target for vaccine (Figure 5.4). Identifying the expression in different cancers were performed as well as normal tissues (Figures (accessing 10.5281/zenodo.8419232)). Characterisation of immunogenicity and identifying epitopes are the next stage for validating TAs for immunotherapy via utilising direct isolation of peptides from the HLA-I complex and *in-silico* prediction of relevant antigens (Backert & Kohlbacher, 2015; Caron et al., 2017). The direct identification of peptides from the HLA-I complex involves techniques such as immunoaffinity purification and extraction of HLA peptides. These methods have been well-established but are highly laborious and time-consuming. Immunoaffinity purification, in particular, has been reported to produce a low yield, which may result in a significant loss of samples (Caron et al., 2017). This bottleneck in the peptide isolation procedure needs to be improved to enhance the efficiency and reliability of identifying TAs. However, the advancements in antigen discovery have provided a better understanding of the intricate interaction between peptides and the HLA-I complex, as well as between HLA-I-restricted peptides and CD8+ T cells (Kanaseki & Torigoe, 2019). In contrast, *in-silico* tools were employed to predict antigens that could be significant for the immune system. By utilizing bioinformatic methods, the structure of genomic and proteomic data to pinpoint potential targets can be analysed. Epitope prediction tools, which integrate information on gene expression disparities between cancerous and healthy tissues with the likelihood of these candidates being displayed on cell surfaces through HLA molecules, can be advantageous for next-generation sequencing (NGS) approaches (Backert & Kohlbacher, 2015).

The identification of TAs is crucial for the development of effective tumour protection and rejection strategies. Currently, whole TA vaccines are developed based on the understanding of the human immune system ability to recognize immunodominant antigens. These are the antigens that are most easily recognized by the immune system and elicit antigen-specific T cell and antibody responses. However, only a small number of patients have achieved tumour regression with these vaccines containing immunodominant antigens (Chiang et al., 2015). One reason for this limited success is that many immunodominant TAs are self-proteins. As self-proteins, they are subjected to both central and peripheral tolerances that dampen immune responses to avoid autoimmunity (Chiang et al., 2015). This ironic situation suggests that breaking self-tolerance is necessary to achieve effective tumour control and suppression.

To address this challenge, highly immunogenic cancer vaccines and robust immunization regimens can be used. These may include T cell checkpoint blockade and/or anti-angiogenesis strategies (Chiang et al., 2015). By using these approaches, self-tolerance can be overcome, leading to improved tumour control (Chiang et al., 2015). The Cancer Genome Atlas (TCGA) data has revealed that individual patients' cancers have a high degree of unique mutations that are not shared among patients with the same type of cancer (Chiang et al., 2015). This finding suggests an attractive alternative approach: personalized autologous whole tumour cancer vaccines containing the unique mutated neo-tumour antigens of each individual patient (e.g., mutanome-based vaccines).

There are several factors that can influence the immune responses to whole tumour antigens. One important factor is the availability of antigen-presenting cells (APCs) that are responsible for presenting the tumour antigens to the immune system (Jiménez-Morales et al., 2021). The numbers of APCs, particularly dendritic cells (DCs), can impact the immune-dominance of certain TAs. Another factor to consider is the presence of cytokines, such as IL-12, which can modulate the immune response to tumour antigens. IL-12 has been shown to enhance the immunodominance of certain antigens (Jiménez-Morales et al., 2021). Furthermore, the immunosuppressive microenvironment created by tumour cells can hinder anti-tumour T cell responses. Tumour cells can produce immunosuppressive cytokines like IL-10 and TGFβ, which can inhibit the immune response. In addition, TGFβ can also inhibit DC differentiation and natural killer (NK) cell

functions (Thomas & Massagué, 2005). To overcome these challenges, it is desirable to develop whole tumour lysate preparations that induce immunogenic cell death in tumour cells and simultaneously inhibits the production of immunosuppressive cytokines (Chiang et al., 2006). This would help in creating a more favourable environment for the immune response to TAs.

B-ALL cells increase resistance to apoptosis (Elden et al., 2018) by overexpressing anti-apoptotic molecules like BCl-2, survivin, or disrupting apoptotic signalling pathways. The significance of death receptor signalling (via FADD and TNFRSF10B) and the potential enhancement of cytotoxicity with SMAC mimetics has been revealed through high-throughput small-molecule and genome-scale CRISPR-Cas9 loss-of-function screens (Dufva et al., 2020). SMAC mimetics are a group of compounds that imitate the function of SMAC, a caspase activator derived from mitochondria. These compounds are specifically designed to block inhibitor of apoptosis proteins (IAPs), which are cellular proteins responsible for preventing cell death. They act as antagonists to IAPs, disrupting their ability to block caspase activation and allowing the apoptotic pathway to proceed. In essence, SMAC mimetics help counterbalance the anti-apoptotic effects of IAPs, leading to increased cell death. As BIRC5, a member of the inhibitor of apoptosis family, has been found upregulated in B-ALL at the diagnostic stage from the Guinn group (Boullosa et al., 2018), we have confirmed here its significant overexpression in B-ALL compared to healthy controls (p-value=0.01), suggesting this could be a target for treatment. BIRC5 knockdown results in the induction of apoptosis of leukaemia cells and increased chemosensitivity *in-vitro*. Phase I trial using antisense oligonucleotide targeting BIRC5 has been examined for relapsed cB-ALL. However, the trial discontinued due to high toxicity and off target effects (Li et al., 2019b).

In addition to immunotherapy, TAs targeting via protein-protein interactions may facilitate an attractive approach in B-ALL treatment. B-ALL originates from antibody-producing cells, which heavily rely on proper protein folding at the endoplasmic reticulum (ER) and efficient degradation of unfolded substrates through the ubiquitin-proteasome system (UPS). Consequently, inhibiting the proteasome and inducing the unfolded protein response (UPR) may help make these cancer types more immunogenic (Kharabi Masouleh et al., 2014). Impaired lymphoid factors make pre-B ALL vulnerable

to ER stress via upregulation of the UPR such as ERN1, HSPA5, PRDM1, and XBP1 (Scheuner et al., 2001). The UPR plays a critical role in the transition from surface Ig-dependent B cells to Ig-secreting plasma cells (Scheuner et al., 2001). The UPR also regulates the pre-B-cell stage when Ig heavy-chain variable region genes are rearranged and expressed (Scheuner et al., 2001). The deletion of UPR leads to programmed cell death in both normal pre-B-cells and B-ALL. The protein XBP1, which is activated by UPR, is positively regulated by STAT5 but negatively regulated by BACH2 and BCL6. In adults with ALL, elevated levels of XBP1 mRNA at the time of diagnosis are associated with a poor prognosis. In laboratory experiments, a small molecule inhibitor that blocks the activation of XBP1 through ERN1 induces cell death specifically in pre-B ALL cells derived from patients. Furthermore, this inhibitor significantly prolongs the survival of mice that received transplants of pre-B ALL cells (Kharabi Masouleh et al., 2014). Using SEREX, we identified eIF2α through the immunoscreening of a testis cDNA library using ALL sera. The activation of protein kinase RNA-like endoplasmic reticulum kinase (PERK) leads to the phosphorylation of eukaryotic initiation factor 2α (eIF2α). This phosphorylation event promotes the translation of activating transcription factor 4 (ATF4) while reducing overall protein synthesis (Fun & Thibault, 2020). ATF4, as a transcription factor, has the ability to increase the expression of CCAAT/enhancer-binding protein homologous protein (CHOP). ATF4 and CHOP together can also upregulate the production of growth arrest and DNA damage-inducible protein-34 (GADD34), which acts to dephosphorylate phospho-eIF2α (Fun & Thibault, 2020). Therefore, prolonged ER stress persisting without restoration of protein synthesis homeostasis, can ultimately result in apoptotic cell death and may be worth for further investigation in B-ALL.

There is growing evidence that malignant cells have the ability to disrupt the normal functioning of the BM haematopoietic niche, leading to the creation of a new microenvironment which supports leukaemia growth (Chiarini et al., 2016). This is facilitated by reprogrammed BM stromal cells and various immune cells, which work together to create self-reinforcing niches that provide leukaemic cells with the necessary signals for their survival, proliferation, and protection against chemotherapy (Chiarini et al., 2016). In our IPA analysis, we found that p38 MAPK is an upstream regulator, p38 MAPK was found to be important for stromal cell support of leukemic cell

proliferation and survival (van den Berk et al., 2014). Given that p38 MAPK is activated by CXCL12, the CXCL12/CXCR4 axis plays a crucial role in the homing of B-ALL cells to the BM, while the involvement of VLA-4 is also found in BM homing of human B-ALL cells.

The PKC cascade is the most dominant for BM migration and homing while considering the signalling pathways activated downstream of CXCR4. The inhibition of p38 MAPK could therefore be a potential therapeutic target for B-ALL progression (van den Berk et al., 2014). The CXCR4 upregulation by leukaemic cells is a strong predictor of the involvement of extra-medullary organ in B-ALL, regardless of the count of peripheral lymphoblast (van den Berk et al., 2014). CXCR4 upregulation in the mature B-ALL, is associated with substantial leukemic infiltration in extra-medullary sites including CNS (Crazzolara et al., 2001). The evaluation of CXCL12/CXCR4 in BM-MSCs from young adults and adolescents with ALL revealed that low CXCR4 levels had an inverse relation to CXCL12 expression compared to healthy donors. In the pre-clinical models of B-ALL, the efficacy of CXCR4 antagonists such as T140 and AMD3100 (Plerixafor) in blocking CXCL12-driven chemotaxis in both cell lines and primary B-ALL cells have been demonstrated (Juarez et al., 2007). While AMD3100 had little effect on the survival of B-ALL cells in co-culture with stromal cells, this significantly impacted B-ALL proliferation due to the co-stimulatory function of CXCL12. Leukaemic cell numbers in peripheral blood and spleen significantly decreased upon prolonged administration of CXCR4 antagonists in mice or B-ALL accompanied by a marked decrease in dissemination to extra-medullary sites (Juarez et al., 2007). Lestaurtinib, an FLT3 inhibitor, combined with AMD3100/G-CSF significantly reduced leukaemia initiating cells (LIC) engraftment in MLL-R ALL (Sison et al., 2013). These results suggest that disrupting leukaemia/BM-MSC signalling is crucial for treating MLL-R ALL.

Our result from SEREX identified many genes related to ubiquitin ligase (CUL1, UBE3C, UBE2D2, Ubiquitin C). Protein ubiquitination mainly promotes degradation by proteasome. It also enables cyclin binding activity (Huang et al., 2015). AHNAK was identified as a novel substrate of UBE3C, which regulates LC cell stemness by ubiquitinating and degrading AHNAK. Mechanistically, AHNAK was found to be a p53 cofactor, necessary for p53-mediated inhibition of the transcription of stemness-related genes by binding to their respective promoter regions. Enhanced UBE3C-mediated

downregulation of AHNAK disrupted the AHNAK-p53 complex, leading to the loss of p53-mediated transcriptional inhibition and ultimately resulting in enhanced cell stemness (Gu et al., 2019). These findings suggest that UBE3C acts as a tumour promoter in NSCLC by maintaining LC cell stemness through the disruption of the AHNAK-p53 complex. By gaining a deeper understanding of the signalling pathways in leukaemia in which these E3 ligases are involved, we can accelerate the development of improved therapeutic strategies for acute leukaemia.

Dysfunction of lymphoid progenitors is one of the hallmarks of B-ALL and a deep understanding of molecular interactions may assist in identifying new targets for B-ALL. Synthetic lethality interaction targets cell stress induced by specific oncogenes or genes associated with the loss of tumour suppressors. For example, Ikaros is known as a crucial regulator of haematopoiesis, and its dysfunction or deletion can lead to the development of leukaemia (Iacobucci & Mullighan, 2017). It controls multiple biological pathways by regulating the expression of numerous genes. Specifically, Ikaros binds to the promoter region of the Igll1 gene which encodes Lambda 5, a component of the pre-BCR that is essential for progression beyond the pre-B cell stage (Faderl et al., 2010; Iacobucci & Mullighan, 2017). Ikaros binds to the promoter region of Rag genes upregulating expression of the RAG1 and RAG2 proteins, which are necessary for B cell differentiation that play critical role in the heavy chain rearrangement of immunoglobulin (Gurel et al., 2008). Casein kinase II (CK2) phosphorylates the Ikaros and impairs its ability to bind to DNA, regulating gene expression. CK2 acts an oncogenic kinase that is often upregulated in ALL. Moreover, Ikaros suppresses genes that promote cell cycle progression and the PI3K pathway, but its tumour suppressor functions are impaired by CK2 phosphorylation (Gurel et al., 2008). Inhibiting CK2 through molecular or pharmacological means can increase Ikaros ability to repress these genes in B-ALL cells. The repression of these genes is facilitated through the restoration of Ikaros transcriptional repressor function, as confirmed by the lack of repression in a group where the function of Ikaros was knocked out (Gurel et al., 2008). The CK2 inhibitors demonstrated efficacy in the PDX mice which were generated from primary B-ALL patient samples, resulting in decreased leukaemia progression and prolonged survival (Gowda et al., 2017). The interaction of Ikaros and HDAC1 with a target gene can either

induce or suppress gene expression through a unique histone modification. The effects of CK2 on Ikaros-mediated chromatin remodelling and transcriptional regulation were studied in high-risk B-ALL cells. It was shown that the inhibition of CK2 restored the epigenetic regulation of cell cycle progression by Ikaros. Similarly, KDM5B, a histone H3K4 demethylase, overexpressed in leukaemia is suppressed by Ikaros, leading to upregulation of H3K4 trimethylation. However, the KDM5B inhibition results in the arrest of cell growth, but the CK2-mediated phosphorylation inhibits the regulation of KDM5B expression by Ikaros. Utilising of CK2 inhibitor to restore Ikaros regulatory functions represents an attractive target for B-ALL having reduction in leukaemia burden and decreased KDM5B expression (Gowda et al., 2017). Ge *et al.* (2016) have identified high expression of IL-7Rα (IL-7R), low expression of SH2B3, and Ikaros dysfunction a subgroup of aB-ALL patients with poor prognosis (Ge et al., 2016; Wang et al., 2016b). IL-7R and SH2B3 are target genes of Ikaros because it regulates their expression through a chromatin remodelling mechanism (Ge et al., 2016). While IL-7R is necessary for lymphoid development, its overexpression can promote oncogenesis. The SH2B3 is a negative regulator of cytokine signalling that plays a crucial role in the HSC and lymphoid progenitor homeostasis. Ge *et al.* also demonstrated that reduction of IL-7R and upregulation of SH2B3 have been observed due to Ikaros overexpression and CK2 inhibition, while CK2 inhibition after Ikaros silencing blocked changes in both gene expression. These findings suggest that CK2 regulates Ikaros ability to activate or repress its target genes, including IL-7R and SH2B3. The study sheds light on the interaction between the IL-7R signalling and CK2-Ikaros axis providing a rationale for using CK2 inhibitors to treat the high-risk B-ALL subtype (Ge et al., 2016). Furthermore, Ikaros deletion is associated with overexpression of CRLF2 in a subset of ALL patients (Ge et al., 2016). However, the study did not include patients with increased CRLF2 expression due to CRLF2 rearrangement in the Hispanic population raising questions about the generalizability of the findings to other populations. The role of Ikaros in suppressing transcription of CRLF2 through chromatin remodelling is intriguing. High levels of CK2 can lead to Ikaros hyperphosphorylation and impaired function. This highlights the importance of the CK2-Ikaros interaction in regulating tumour suppression in high-risk B-ALL. Restoring Ikaros binding to the CRLF2 promoter via CK2 inhibition represses the transcription of the gene, indicating that the CK2-Ikaros interaction regulates tumour

suppression in high-risk B-ALL. The use of CK2-specific inhibitors can restore Ikaros tumour suppressor function and have an anti-leukemic impact is a promising therapeutic approach (Ge et al., 2016). Further research on this signalling axis will provide insight into regulating cellular proliferation in hematopoietic malignancies.

Detailed understanding of how pathways cross-interact is crucial for selected inhibitions. RAG activity is required for a differentiation arrest at the pro-B to pre-B stage and Ig heavy chain arrangement. Downregulation of RAG1, RAG2, TDT are found in B-ALL and is correlated to high expression of NF-κB. Inhibition of NF-κB may result in the increase of RAG-dependent DNA damage in leukaemia but it may deplete Ikaros enhancing leukaemogenesis (Papaemmanuil et al., 2014). Targeting NF-κB should be used with cautious considering Ikoros expression to monitor therapy efficiency. Understanding of co-drivers of pre and pro-B-cells allows identifying the transformation of pre leukaemia cells with inhibited differentiation gene expression to acquired stem cell features, such as HOX, and may identify new targets for B-ALL (Bueno et al., 2012). Our result identified HOX4 that has been recognised by B-ALL patient sera. Upregulation of HOX genes in B-ALL is correlated to acquisition of stem-cell features and arrest of B-cell differentiation. HOX4 acts as a tumour suppressor inducing apoptosis. However, HOX4 is downregulated in TNBC while its expression impairs cancer proliferation via inhibiting AKT/mTOR pathway (Zhang et al., 2021).

TAs may be difficult for direct targeting because they are involved in the maintenance of homeostatic functions. For example, the oncogenic protein c-Myc plays a critical role in survival, proliferation, and drug resistance in B -ALL (Chen et al., 2018a). While direct targeting of Myc has been difficult due to its “undruggable” protein structure, recent studies have shown promise in targeting c-Myc transcription by interfering with chromatin-dependent signal transduction to RNA polymerase through BRD4 inhibition (Chen et al., 2018a). BRD4 belongs to the bromodomain and extra-terminal domain (BET) family and is a critical chromatin regulator that maintains disease progression in acute myeloid leukemia (AML), and suppression of BRD4 with shRNA or JQ1 (a bromodomain inhibitor) has shown anti-leukemic effects *in-vitro* and *in-vivo*. BET inhibition has also been shown to be effective against primary childhood B-ALL by decreasing c-Myc protein stability, suppressing progression at DNA replication forks, and

sensitizing primary B-ALL towards dexamethasone *in-vitro* and *in-vivo* (Ruan et al., 2021).

In our study (Figure 6.1), we presented a novel concept proposing targets for B-ALL patients, based on protoarray, transcriptional, and epigenetic profiling. Dissecting B-ALL pathogenesis may allow to identify the molecular vulnerabilities which may act as targets for the treatment. For the gene enrichment program analysis, the findings showed that Wnt, Hippo, and TGFβ are enriched in B-ALL subtypes (Figure 6.3). Wnt seems to be a determining marker for B-ALL diagnosis and survival as indicated in the SEREX technique. Hippo pathways have been known to restrict cell growth in adults, while modulating cell proliferation, differentiation, and migration in developing organs (Ma et al, 2016; Hsu et al, 2020); Hippo pathway dysregulation results in aberrant cell growth and neoplasia. In this study, however, the result reiterates Hippo ability to be used as a survival instrument in cancer prognosis or treatment approach for a wide range of diseases. Its implication in the survival of B-ALL could be the yardstick in the diagnosis and treatment of cancer and further lead to significant changes in clinical decisions. Furthermore, we suggest the design that inhibitor peptides that disrupt the interaction between TEAD and YAP, which is crucial for cancer progression. One such peptide, super-TDU, was designed based on the interface between TEAD4 and VGLL4 and effectively blocked TEAD-YAP complex formation *in-vitro*. It also produced therapeutic effects in mouse models with high YAP expression (Jiao et al., 2014). Smaller cyclic peptides were also engineered to disrupt the most crucial interface of the TEAD-YAP complex (Zhou et al., 2015). However, peptide-mediated cancer therapies face challenges such as poor stability, low membrane permeability, and susceptibility to proteolytic digestion (Zhou et al., 2015). Strategies such as amino acid substitution, fusion of functional peptides, and conjugation with chemotherapeutic drugs are being adopted to improve their efficacy (Zhou et al., 2015). Similarly, YAP knockdown by shRNA induces apoptosis in HL-60 cells and impedes proliferation. YAP may also be a potential treatment strategy for acute pro-myelocytic leukaemia (APL). However, additional investigations are necessary to determine its clinical suitability (Chen et al., 2017).

The TEAD4-YAP complex regulates various cancer functions, such as cellular growth, migration, and invasiveness. When the LATS1/2 tumour suppressor kinase phosphorylates YAP, this becomes inactive and translocates to the cytoplasm. TEAD4 knockdown leads to increased levels of YAP phosphorylation and decreased YAP in the nuclei, which suggests that TEAD4 expression levels may control DNA binding and transcriptional activities with YAP and the phosphorylation levels of YAP. Furthermore, CDK6 is a downstream target of the TEAD4-YAP complex for cellular senescence. As CDK6 is a critical molecule in the G1 phase of the cell cycle, inactivated TEAD4-YAP complex (TEAD4 knockdown) would result in cell-cycle arrest in the G1 phase (Takeuchi *et al*., 2017). This finding supports shTEAD4 cells had less tumour progression and patients with human oral squamous cell carcinomas (OSCCs) negative for TEAD4. Lastly, the most effective strategy for managing relapse is to prevent it altogether. Therefore, integrating these agents into the initial treatment plan holds potential for enhancing outcomes in patients.

7.5 Repurposing of existing drugs for cancer treatment

Recently, the repurposing of drugs has been investigated for cancer treatment (Zhang et al., 2020). Ivermectin is an example, as this is effective against parasites by increasing permeability of parasite cell membranes, leading to paralysis and death. Interestingly, there have been preliminary findings suggesting its potential as an anti-cancer agent (Armando et al., 2020). In various tumour and cancer cells, Ivermectin has shown its activities against the proliferation, metastasis, and angiogenic activity. These effects may be attributed to its regulation of signalling pathways, particularly through p21-activated kinase 1. In addition, Ivermectin promotes programmed cell death mechanisms such as apoptosis and autophagy. This have been observed to sensitize multidrug-resistant cells to chemotherapeutic agents, exhibiting optimal results when combined with other chemotherapeutics (Tang et al., 2020).

Notably, Jiang et al. (2019) demonstrated that Ivermectin possesses the potency to reverse resistance to anti-cancer activities in tumour cells. It achieves this primarily by reducing the expression of P-glycoprotein (P-gp) through the inhibition of the epidermal growth factor receptor (EGFR). Ivermectin opposes the activities of the ERK/Akt/NF-κB pathway by binding to the extracellular domain of EGFR and opposing its activation, and

it ultimately leads to a significant reduction in transcription of P-gp. Additionally, Ivermectin restricts the activity of YAP1, a transcription activating molecule responsible for the upregulation of genes that are related to cell proliferation and suppression of apoptotic activities (Armando et al., 2020). Moreover, Ivermectin interferes with the function of karyopherin β1 (KPNB1), which encodes the nuclear transport factors. Research by Kodama et al. (2017) demonstrated that Ivermectin inhibits KPNB1 in ovarian cancer cells, resulting in apoptosis and cell cycle arrest. *In-vivo* studies also indicated synergistic anti-cancer effects when Ivermectin was combined with paclitaxel (Kodama et al., 2017). Additionally, Ivermectin exhibits inhibitory effects on the canonical Wnt signalling pathway, specifically targeting the transcriptional factor T-cell factor (TCF) family. This inhibition results in the suppression of colon and LC proliferation (Armando et al., 2020). However, versatility of Ivermectin extends with its ability to modulate genes involved in the epithelial-mesenchymal transition (EMT) and sustaining of a CSC phenotype in a triple-negative breast cancer (TNBC). Consequently, Ivermectin disrupts the self-renewal *in-vitro* and restricts the growth of tumour and metastasis *in-vivo* (Kwon et al., 2015). Overall, Ivermectin exhibits a range of mechanisms for exerting anti-tumour effects. It interacts with multiple targets, including multidrug resistance (MDR) proteins.

Another anti-parasitic drug is pyrvinium pamoate which exhibits a potent anti-cancer property, particularly against CSC. Its mechanism of action involves the suppression of Wnt- and Hedgehog-dependent signalling pathways, as demonstrated by Dattilo et al. (2020). Additionally, the drug inhibits mitochondrial respiration and oncogenic PI3K-dependent signalling (Dattilo et al., 2020). This drug reduces CSC metastasis by impacting lipid anabolism, disrupting the anabolic flux from glucose to cholesterol and fatty acids. Nair *et al.* (2020) assessed the activity of pyrvinium pamoate against human B-ALL cells, including REH and RS4 cells. The metabolic function of these cells was examined by measuring the acidification of extracellular environment and oxygen consumption rates. The overall results demonstrated significant anti-leukaemia effects of pyrvinium pamoate, with IC50 values of 0.17 μM and 1 μM for REH and RS4 cells, respectively. The study also revealed a notable reduction in cell numbers in the S phase (by ~26%) and M phase (by ~50%) compared to the control group, indicating an impact

on the REH cell cycle (Nair et al., 2020). Furthermore, Laudisi et al. (2020) investigated the anti-tumour activity of pyrvinium pamoate using a xenograft model of human pancreatic cancer PANC1 cells. Oral administration of pyrvinium pamoate at doses of 100 µg and 200 µg resulted in reduced tumour growth and inhibited Akt phosphorylation. The study revealed that pyrvinium pamoate at concentrations of 0.1 µM and 0.3 µM suppressed the transcription of critical chaperones, namely GRP78 and GRP94. This inhibition restricted the UPR caused by glucose deprivation. Additionally, the drug exhibited inhibitory effects on other UPR pathways, including XBP-1 and ATF4, induced by glucose starvation (Laudisi et al., 2020).

7.6 Regulation of TAs via miRNA modulation

miRNAs are involved in normal lymphopoiesis and their expression may differ in peripheral blood compared to BM reflecting their function and target genes (Luan et al., 2015). miRNAs dysregulation has been found in B-ALL (Mi et al., 2007). Upregulation of oncomiR is correlated with tumour initiation and progression. For example, miRNA-155 overexpression enhances blast proliferation and consequently ALL development. miRNAs could also be utilised in differentiate B-ALL from T-ALL using a panel of expression of miRNA-148, miRNA-151, and miRNA-424. Five miRNAs including miRNA-23a, miRNA-27a, miRNA-199b, miRNA-221, and miRNA-223 can help discriminate AML from B-ALL (Mi et al., 2007). miRNA-425-5p is downregulated in Ph+ and may differentiate B-ALL groups. hsa-miR-511 was suggested to differentiate B-ALL from normal B-cells with 100% sensitivity and specificity; however, these findings lack validation (Luna-Aguirre et al., 2015).

Targeting miRNAs (Abd-Aziz et al., 2020) has been suggested to reduce drug resistance and improve disease survival. Inhibition of oncomiRs includes using of complementary chemically modified oligonucleotides, such as anti-miR oligonucleotides and antagomir, as well as small molecule inhibitors of miRNAs, a combination of small interfering RNAs (siRNAs), and microRNA sponges. It has been found that a locked-nucleic acid (LNA) antisense (Harada et al., 2012) targeting miR-17 induces apoptosis and reverses the resistance to dexamethasone via increasing BIM levels in c-ALL. In addition to silencing miRNAs, ectopic expression of tumour suppressor miRNAs has been achieved in the preclinical model via introduction synthetic double-stranded RNA mimics via lipid carrier

or viral vector expressing the oligonucleotide (Sbirkov et al., 2022). Restoring expression of miR-101 in T-ALL induces chemotherapy sensitivity and apoptosis through targeting NOTCH1 pathway (Qian et al., 2016). As epigenetic dysregulation is common in B-ALL, it is possible to reactivate silenced miRNAs by targeting them with epigenetic modifiers. For example, HDAC inhibitors that activates DICER suppressing in Adult T cell Leukaemia (ATL) leading to rescuing the expression of miR-31. However, HDAC inhibitors are not specific to miRNAs enhancing other genes transcription silenced in B-ALL (Gazon et al., 2016). miRNAs targeting has a limitation to be translated into the clinic because miRNAs regulate several target genes either belonging to same or different biological processes, leading to off-target effects and adverse consequences (Chen et al., 2015). The small length of oligonucleotides may not be unique to the specific targets as miRNAs can bind to their target mRNA via incomplete sequence pairing and thus they may not be specific (Chen et al., 2015). miRNA drug interaction may result in regulating of unknown transcript regulated by this miRNA, while miRNA/mRNA association is not fully understood (Lal et al., 2009). Bioinformatics tools that predict miRNA/mRNA interaction are not systematically validated *in-vitro* and *in-vivo* models are needed to assess their suitability as therapeutic molecules. Animals models are poorly validated to mimic ALL pathology and interfering with the endogenous miRNAs (Leclercq et al., 2017). Some miRNAs such as let-7 may modify immune response interacting with transcription factors and pathways. Treatment with let-7 containing lentivirus causes tumour inhibition in NSCLC xenograft murine model, after prolonged treatment, however disease relapse may occur, as some oncogenes (such as HMGA2) cause the loss of the let-7 binding site (Chakraborty et al., 2020). Another hurdle is that modification of RNA drugs to increase their specificity and inhibit nuclease degradation may interfere with cytokine productions and severe adverse consequences (Egli & Manoharan, 2019).

7.7 Future direction

B-ALL is a complex disease with various factors influencing the overall prognosis, including cytogenetic abnormalities, MDR, and response to therapy. Simply targeting the malignant cells is not enough to eradicate tumour growth, as the tumour microenvironment within the BM plays a crucial role in treatment outcomes. The BM serves as the tumour microenvironment and consists of a variety of hematopoietic and

non-hematopoietic cells surrounded by extracellular matrix (ECM). Sensitivity to immunotherapy in B-ALL is not solely determined by intrinsic biological factors but also by the diverse interactions between leukaemia cells and the bone marrow microenvironment (Brown, 2022). Leukaemia cells exploit this microenvironment to sustain their proliferation and survival, taking advantage of the tightly controlled signalling pathways and transcriptional factors that regulate normal lymphopoiesis. B-ALL is a heterogeneous disease, and single targets may not be effective for disease elimination. Targeting multiple genes (**Figure 7.1**) involved in maintaining tumour stem cell features as well as tumour microenvironment may be improving overall prognosis and treatment. As we identified three pathways (TGFβ, Wnt, Hippo pathways) enriched in B-ALL, generating an animal model may allow reprogramming of these pathways combined with other known pathways involved in B-ALL tumourigenesis to identify any potential therapeutic interventions. The unique challenge is similarity between the normal haematopoiesis and B-ALL in terms of complexity, including gene expression, differentiation, etc as well as hierarchies and clonal evolution (Brown, 2022). Dissecting and understanding these may be useful for analysis between normal and abnormal processes. It is important to accurately identify patients who may benefit from targeted therapies, and one way to do this is through the characterisation of immunophenotypic leukemic stem cells (LSCs) (Zhang & Oak, 2023). However, relying solely on the expression of CD34+ CD38- may not provide enough information, as this represents only a fraction of LSCs (Hansen et al., 2022). It is also essential to confirm the effectiveness of these therapies by testing their ability to eliminate dormant LSCs through *in-vivo* xenograft experiments. Furthermore, evaluating LSC activity should involve both functional and genetic studies. With advancements in technology, such as the ability to analyse stemness signatures, mutational profiles, and cell origin at a single-cell level, there is substantial potential for revolutionizing future ALL treatments. Global transcriptome and epigenome profiling are a valuable tool to enhance diagnostic testing in clinical settings. By incorporating gene expression profiles, we can potentially improve the classification and risk assessment of B-ALL, which is crucial for personalized medicine. However, due to the complexity and heterogeneity of ALL, understanding the clonal evolution at a single-cell level is still in its early stages.

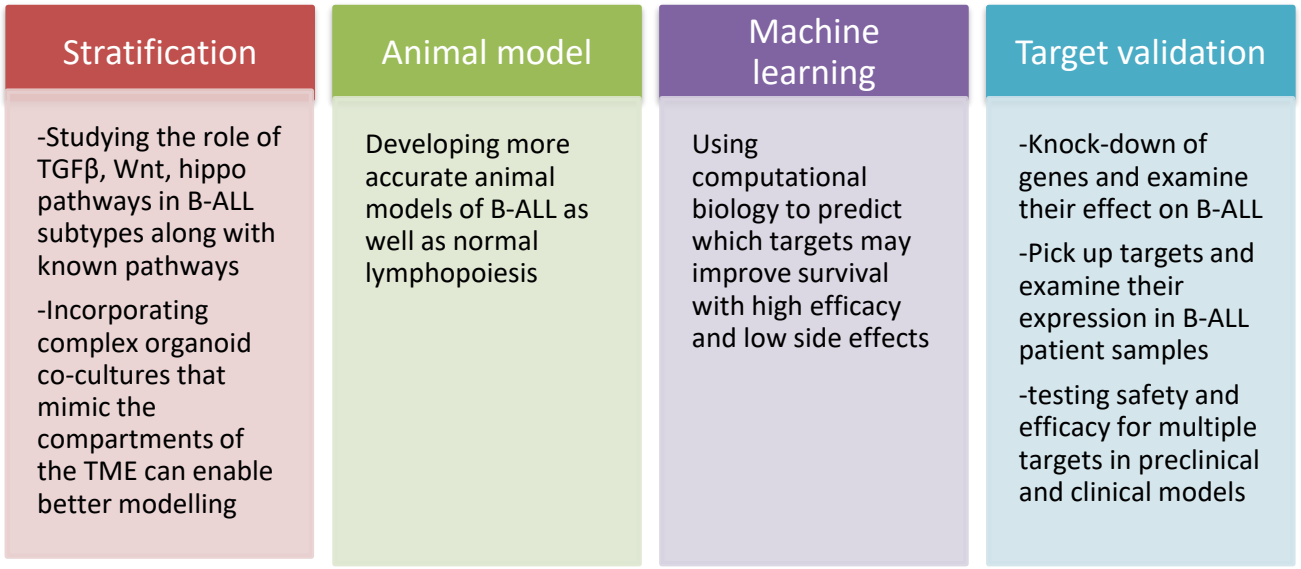


Figure 7.1 Future direction for B-ALL

Tackling tumour heterogeneity of B-ALL by finding specific signatures that are unique to B-ALL subtypes.

Therefore, an integrated approach that combines various sequencing techniques like RNAseq, genome sequencing, exome sequencing, methylome sequencing or targeted resequencing would be ideal for capturing the interactions between different components in ALL. Incorporating complex organoid co-cultures that mimic the fibroblastic, immune, and vascular compartments of the TME can enable better modelling of tumour microenvironmental influences on cell state. The use of 3D cultures of primary tumour tissue is highlighted as a valuable tool for maintaining cellular heterogeneity and observing cell state transitions observed *in vivo*. These 3D cultures also allow for analysis of tumour-level phenotypes that are difficult to model in traditional 2D cultures, such as collective invasion, immune surveillance, and organ colonization (Yuki et al., 2020). The development of more accurate animal models of B-ALL as well as normal lymphopoiesis could help identify genetic aberrations during disease development. *In-vitro* studies may not fully capture the complexities of *in-vivo* tumour biology, and animal models may not always accurately represent human disease (Mian et al., 2021). Therefore, relying solely on experimental evidence from limited models may not provide a comprehensive understanding of target validity or predict clinical outcomes accurately. Furthermore, creating and characterizing more experimental models should indeed be a priority to enhance our understanding of specific categories of cancer targets. However, it is essential to ensure that these models adequately represent human disease and encompass its heterogeneity. Developing relevant preclinical models that mimic human tumours more closely will be crucial for improving translational success. Furthermore, single-cell RNA sequencing (scRNAseq) analyses have revealed multiple distinct cell states within tumours, with unequal contributions to tumour relapse (Iacobucci et al., 2023). The selective targeting and depletion of specific populations, such as stem-like progenitor cells that may seed tumour relapse, is suggested as a more achievable goal than completely eliminating all residual tumour cells (Hahn et al., 2021). Understanding the processes via which persisting cells survive therapeutic stress and regrow may lead to new therapeutic approaches. To improve patient outcomes, it is crucial to target cancer not only in isolation but also as part of a larger ecosystem (Hahn et al., 2021). While our understanding of the composition and activities of the TME is still limited, there have

been notable functional demonstrations showing clinical efficacy with immune- and blood vessel-targeted therapy (Hahn et al., 2021). This raises the potential for further target identification and effective disruption of the tumour ecosystem. While collapsing the complexity of mutational patterns and TME into distinct tumour cell states can aid in assessing potential targets for therapeutic intervention, it is important to critically evaluate the limitations and challenges associated with this approach. Thorough evaluation of potential toxicity, consideration of tumour heterogeneity, and the development of relevant experimental models should all be prioritized to improve the translation of promising targets into effective clinical therapies.

Discovering and validating these TME and non-oncogenes may help develop novel cancer treatment classes. This is important is to identify indicators or characteristics that can be used to group patients who will most likely benefit from these treatments. The need for novel instruments to evaluate multi-parameter gene expression in cancer tissue is one of the primary issues raised (Chari et al., 2019). While immunohistochemical methods or molecular profiling may be sufficient for some targets, others will require more advanced techniques. This may require substantial changes to sample acquisition and testing, which may pose logistical and practical challenges in clinical practice. Despite these challenges, advancements in research and technology have made it feasible to identify potential therapeutic targets beyond traditional oncogenes. However, it is important to note that the effectiveness of these targets still needs to be validated through rigorous testing and clinical trials. Furthermore, there is need to developing robust methods to test potential combinations of therapies targeting both tumour intrinsic (e.g. oncogenes) and extrinsic factors (i.e. TME) (Chari et al., 2019). This approach aims to achieve rational combinations that can provide more effective and durable outcomes for patients. However, scaling up these approaches and identifying novel endpoints beyond cell fitness will be necessary; in future.

Chapter 8: References

Pioglitazone Hydrochloride in Treating Patients With Stage IA-IIIA Non-small Cell Lung Cancer. https://ClinicalTrials.gov/show/NCT01342770.

Abd-Aziz, N., Kamaruzman, N. I. & Poh, C. L. (2020) Development of microRNAs as potential therapeutics against cancer. *Journal of oncology*, 2020.

Aberle, D., Adams, A., Berg, C., Black, W., Clapp, J., Fagerstrom, R., Gareen, I., Gatsonis, C., Marcus, P. & Sicks, J. (2011) Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening New England Journal of Medicine 365 (5): 395-409 DOI 10.1056. *NEJMoa1102873*.

Abou-Zeid, A., Hashad, D., Baess, A., Mosaad, M. & Tayae, E. (2023) HOXA9 gene promotor methylation and copy number variation of SOX2 and HV2 genes in cell free DNA: A potential diagnostic panel for non-small cell lung cancer. *BMC Cancer*, 23(1), 329.

Adams, S., Sahota, S., Mijovic, A., Czepulkowski, B., Padua, R., Mufti, G. & Guinn, B. (2002) Frequent expression of HAGE in presentation chronic myeloid leukaemias. *Leukemia*, 16(11), 2238-2242.

Affaitati, A., Cardone, L., De Cristofaro, T., Carlucci, A., Ginsberg, M. D., Varrone, S., Gottesman, M. E., Avvedimento, E. V. & Feliciello, A. (2003) Essential role of A-kinase anchor protein 121 for cAMP signaling to mitochondria. *Journal of Biological Chemistry*, 278(6), 4286-4294.

Ajona, D., Remirez, A., Sainz, C., Bertolo, C., Gonzalez, A., Varo, N., Lozano, M., Zulueta, J., Mesa-Guzman, M., Martin, A., Perez-Palacios, R., Perez-Gracia, J., Massion, P., Montuenga, L. & Pio, R. (2021) A model based on the quantification of complement C4c, CYFRA 21-1 and CRP exhibits high specificity for the early diagnosis of lung cancer. *Translational Research*, 233, 77-91.

Akbari Moqadam, F., Lange-Turenhout, E. A. M., Ariës, I. M., Pieters, R. & den Boer, M. L. (2013) MiR-125b, miR-100 and miR-99a co-regulate vincristine resistance in childhood acute lymphoblastic leukemia. *Leukemia Research*, 37(10), 1315-1321.

Akhmedov, M., Martinelli, A., Geiger, R. & Kwee, I. (2020) Omics Playground: a comprehensive self-service platform for visualization, analytics and exploration of Big Omics Data. *NAR Genom Bioinform*, 2(1), lqz019.

Al-Ameri, A., Malhotra, P., Thygesen, H., Plant, P. K., Vaidyanathan, S., Karthik, S., Scarsbrook, A. & Callister, M. E. (2015) Risk of malignancy in pulmonary nodules: a validation study of four prediction models. *Lung cancer*, 90(1), 119-120.

Al-Odat, O. S., Guirguis, D. A., Schmalbach, N. K., Yao, G., Budak-Alpdogan, T., Jonnalagadda, S. C. & Pandey, M. K. (2022) Autophagy and Apoptosis: Current Challenges of Treatment and Drug Resistance in Multiple Myeloma. *International Journal of Molecular Sciences*, 24(1), 644.

Alchahin, A. M., Mei, S., Tsea, I., Hirz, T., Kfoury, Y., Dahl, D., Wu, C.-L., Subtelny, A. O., Wu, S. & Scadden, D. T. (2022) A transcriptional metastatic signature predicts survival in clear cell renal cell carcinoma. *Nature communications*, 13(1), 1-15.

Aldoss, I., Forman, S. J. & Pullarkat, V. (2019) Acute lymphoblastic leukemia in the older adult. *Journal of oncology practice*, 15(2), 67-75.

Allegra, A., Pioggia, G., Innao, V., Musolino, C. & Gangemi, S. (2021) New insights into yes-associated protein signaling pathways in hematological malignancies: diagnostic and therapeutic challenges. *Cancers*, 13(8), 1981.

Almutairi, M., Alrubie, T., Alamri, A., Almutairi, B., Alrefaei, A., Arafah, M., Alanazi, M. & Semlali, A. (2022) Cancer-Testis Gene Biomarkers discovered in Colon Cancer Patients. Genes 2022, 13, 807. s Note: MDPI stays neu-tral with regard to jurisdictional claims in ….

Altintas, Z. & Tothill, I. (2013) Biomarkers and biosensors for the early diagnosis of lung cancer. *Sensors and Actuators B-Chemical*, 188, 988-998.

Amoorahim, M., Valipour, E., Hoseinkhani, Z., Mahnam, A., Rezazadeh, D., Ansari, M., Shahlaei, M., Gamizgy, Y. H., Moradi, S. & Mansouri, K. (2020) TSGA10 overexpression inhibits angiogenesis of HUVECs: A HIF-2α biased perspective. *Microvascular research*, 128, 103952.

Anderson, K. S. & LaBaer, J. (2005) The sentinel within: exploiting the immune system for cancer biomarkers. *Journal of proteome research*, 4(4), 1123-1133.

Angi, M., Kamath, V., Yuvarani, S., Meena, J., Sitaram, U., Manipadam, M. T., Nair, S., Ganapule, A., Fouzia, N. A., Abraham, A., Viswabandya, A., Poonkuzhali, B., George, B., Mathews, V., Srivastava, A. & Srivastava, V. M. (2017) The t(8;14)(q24.1;q32) and its variant translocations: A study of 34 cases. *Hematology/Oncology and Stem Cell Therapy*, 10(3), 126-134.

Antonangelo, L., Sales, R., Corá, A., Acencio, M., Teixeira, L. & Vargas, F. (2015) Pleural fluid tumour markers in malignant pleural effusion with inconclusive cytologic results. *Current Oncology*, 22(5), e336.

Ariës, I. M., Bodaar, K., Karim, S. A., Chonghaile, T. N., Hinze, L., Burns, M. A., Pfirrmann, M., Degar, J., Landrigan, J. T., Balbach, S., Peirs, S., Menten, B., Isenhart, R., Stevenson, K. E., Neuberg, D. S., Devidas, M., Loh, M. L., Hunger, S. P., Teachey, D. T., Rabin, K. R., Winter, S. S., Dunsmore, K. P., Wood, B. L., Silverman, L. B., Sallan, S. E., Van Vlierberghe, P., Orkin, S. H., Knoechel, B., Letai, A. G. & Gutierrez, A. (2018) PRC2 loss induces chemoresistance by repressing apoptosis in T cell acute lymphoblastic leukemia. *J Exp Med*, 215(12), 3094-3114.

Ariës, I. M., Jerchel, I. S., van den Dungen, R. E. S. R., van den Berk, L. C. J., Boer, J. M., Horstmann, M. A., Escherich, G., Pieters, R. & den Boer, M. L. (2014) EMP1, a novel poor prognostic factor in pediatric leukemia regulates prednisolone resistance, cell proliferation, migration and adhesion. *Leukemia*, 28(9), 1828-1837.

Atanackovic, D., Hildebrandt, Y., Jadczak, A., Cao, Y., Luetkens, T., Meyer, S., Kobold, S., Bartels, K., Pabst, C. & Lajmi, N. (2010) Cancer-testis antigens MAGE-C1/CT7 and MAGE-A3 promote the survival of multiple myeloma cells. *Haematologica*, 95(5), 785.

Ayan, A. K., Erdemci, B., Orsal, E., Bayraktutan, Z., Akpinar, E., Topcu, A., Turkeli, M. & Seven, B. (2016) Is there any correlation between levels of serum ostepontin, CEA, and FDG uptake in lung cancer patients with bone metastasis? *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)*, 35(2), 102-106.

Azzolin, L., Zanconato, F., Bresolin, S., Forcato, M., Basso, G., Bicciato, S., Cordenonsi, M. & Piccolo, S. (2012) Role of TAZ as mediator of Wnt signaling. *Cell*, 151(7), 1443-1456.

Baba, Y., Pelayo, R. & Kincade, P. W. (2004) Relationships between hematopoietic stem cells and lymphocyte progenitors. *TRENDS in Immunology*, 25(12), 645-649.

Backert, L. & Kohlbacher, O. (2015) Immunoinformatics and epitope prediction in the age of genomic medicine. *Genome medicine*, 7, 1-12.

Baker, S. G. (2000) Identifying combinations of cancer markers for further study as triggers of early intervention. *Biometrics*, 56(4), 1082-1087.

Balata, H., Evison, M., Sharman, A., Crosbie, P. & Booton, R. (2019) CT screening for lung cancer: Are we ready to implement in Europe? *Lung Cancer*, 134, 25-33.

Baldassarre, M., Razinia, Z., Burande, C. F., Lamsoul, I., Lutz, P. G. & Calderwood, D. A. (2009) Filamins regulate cell spreading and initiation of cell migration. *PloS one*, 4(11), e7830.

Baldwin, D. R. & Callister, M. E. (2015) The British Thoracic Society guidelines on the investigation and management of pulmonary nodules. *Thorax*, 70(8), 794-798.

Balekian, A. A., Silvestri, G. A., Simkovich, S. M., Mestaz, P. J., Sanders, G. D., Daniel, J., Porcel, J. & Gould, M. K. (2013) Accuracy of clinicians and models for estimating the probability that a pulmonary nodule is malignant. *Annals of the American Thoracic Society*, 10(6), 629-635.

Banerjee-Basu, S. & Baxevanis, A. D. (2001) Molecular evolution of the homeodomain family of transcription factors. *Nucleic acids research*, 29(15), 3258-3269.

Bannert, N., Hofmann, H., Block, A. & Hohn, O. (2018) HERVs new role in cancer: from accused perpetrators to cheerful protectors. *Frontiers in microbiology*, 178.

Bänziger, C., Soldini, D., Schütt, C., Zipperlen, P., Hausmann, G. & Basler, K. (2006) Wntless, a conserved membrane protein dedicated to the secretion of Wnt proteins from signaling cells. *Cell*, 125(3), 509-522.

Barbosa, C., Peixeiro, I. & Romão, L. (2013) Gene expression regulation by upstream open reading frames and human disease. *PLoS genetics*, 9(8), e1003529.

Barlési, F., Gimenez, C., Torre, J.-P., Doddoli, C., Mancini, J., Greillier, L., Roux, F. & Kleisbauer, J.-P. (2004) Prognostic value of combination of Cyfra 21-1, CEA and NSE in patients with advanced non-small cell lung cancer. *Respiratory medicine*, 98(4), 357-362.

Becker, N., Motsch, E., Trotter, A., Heussel, C. P., Dienemann, H., Schnabel, P. A., Kauczor, H. U., Maldonado, S. G., Miller, A. B. & Kaaks, R. (2020) Lung cancer mortality reduction by LDCT screening—Results from the randomized German LUSI trial. *International Journal of Cancer*, 146(6), 1503-1513.

Behrendt, L., Hoischen, C. & Kaether, C. (2021) Disease-causing mutated ATLASTIN 3 is excluded from distal axons and reduces axonal autophagy. *Neurobiology of Disease*, 155, 105400.

Belinsky, S. A. (2004) Gene-promoter hypermethylation as a biomarker in lung cancer. *Nature Reviews Cancer*, 4(9), 707-717.

Berge, G., Pettersen, S., Grotterød, I., Bettum, I. J., Boye, K. & Mælandsmo, G. M. (2011) Osteopontin—An important downstream effector of S100A4 - mediated invasion and metastasis. *International journal of cancer*, 129(4), 780-790.

Bhushan, N. & Rai, K. (2007) *Strategic decision making: applying the analytic hierarchy process*.Springer Science & Business Media.

Bhutani, M., Pathak, A. K., Tang, H., Fan, Y. H., Liu, D. D., Lee, J. J., Kurie, J., Morice, R. C., Hong, W. K. & Mao, L. (2011) Frequent expression of MAGE1 tumor antigens in bronchial epithelium of smokers without lung cancer. *Exp Ther Med*, 2(1), 137-142.

Biaoxue, R., Hua, L., Wenlong, G. & Shuanying, Y. (2016) Increased serum amyloid A as potential diagnostic marker for lung cancer: a meta-analysis based on nine studies. *BMC cancer*, 16(1), 836.

Biesterfeld, S., Veuskens, U., Schmitz, F., Amo-Takyi, B. & Böcking, A. (1996) Interobserver reproducibility of immunocytochemical estrogen-and progesterone receptor status assessment in breast cancer. *Anticancer research*, 16(5A), 2497-2500.

Bigbee, W. L., Gopalakrishnan, V., Weissfeld, J. L., Wilson, D. O., Dacic, S., Lokshin, A. E. & Siegfried, J. M. (2012) A multiplexed serum biomarker immunoassay panel discriminates clinical lung cancer patients from high-risk individuals found to be cancer-free by CT screening. *Journal of thoracic oncology*, 7(4), 698-708.

Bingle, L., Singleton, V. & Bingle, C. D. (2002) The putative ovarian tumour marker gene HE4 (WFDC2), is expressed in normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms. *Oncogene*, 21(17), 2768-2773.

Bishnoi, S., Pittman, K., Colbeck, M., Townsend, A., Hardingham, J., Sukumaran, S., Hooper, B., Tuck, M., Roder, D. & Luke, C. (2011) Small Cell Lung Cancer: Patterns of care and their influence on survival – 25 years experience of a single Australian oncology unit. *Asia - Pacific Journal of Clinical Oncology*, 7(3), 252-257.

Boeri, M., Verri, C., Conte, D., Roz, L., Modena, P., Facchinetti, F., Calabrò, E., Croce, C. M., Pastorino, U. & Sozzi, G. (2011) MicroRNA signatures in tissues and plasma predict development and prognosis of computed tomography detected lung cancer. *Proceedings of the National Academy of Sciences*, 108(9), 3713-3718.

Boncheva, V. B. (2013) The identification of tumour antigens recognized by patients with Duke's B (Stage II) reactive colorectal cancers using SEREX.

Boullosa, L. F., Savaliya, P., Bonney, S., Orchard, L., Wickenden, H., Lee, C., Smits, E., Banham, A. H., Mills, K. I. & Orchard, K. (2018) Identification of survivin as a promising target for the immunotherapy of adult B-cell acute lymphoblastic leukemia. *Oncotarget*, 9(3), 3853.

Boullosa, L. F., Savaliya, P., Bonney, S., Orchard, L., Wickenden, H., Lee, C., Smits, E., Banham, A. H., Mills, K. I., Orchard, K. & Guinn, B.-A. (2017) Identification of survivin as a promising target for the immunotherapy of adult B-cell acute lymphoblastic leukemia. *Oncotarget*, 9(3), 3853-3866.

Boyle, P., Chapman, C., Holdenrieder, S., Murray, A., Robertson, C., Wood, W., Maddison, P., Healey, G., Fairley, G. & Barnes, A. (2011) Clinical validation of an autoantibody test for lung cancer. *Annals of Oncology*, 22(2), 383-389.

Bray, F., Jemal, A., Grey, N., Ferlay, J. & Forman, D. (2012) Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. *The lancet oncology*, 13(8), 790-801.

Bretthauer, M., Kaminski, M. F., Hassan, C., Kalager, M., Holme, Ø., Hoff, G., Løberg, M., Regula, J., Castells, A. & Adami, H.-O. (2017) America, we are confused: the updated US Preventive Services Task Force recommendation on colorectal cancer screening. American College of Physicians.

Brown, G. (2022) Lessons to cancer from studies of leukemia and hematopoiesis. *Frontiers in cell and developmental biology*, 10, 993915.

Buccitelli, C. & Selbach, M. (2020) mRNAs, proteins and the emerging principles of gene expression control. *Nature Reviews Genetics*, 21(10), 630-644.

Buchner, M., Park, E., Geng, H., Klemm, L., Flach, J., Passegué, E., Schjerven, H., Melnick, A., Paietta, E., Kopanja, D., Raychaudhuri, P. & Müschen, M. (2015) Identification of FOXM1 as a therapeutic target in B-cell lineage acute lymphoblastic leukaemia. *Nature Communications*, 6(1), 6471.

Bueno, C., Montes, R., Melen, G. J., Ramos-Mejia, V., Real, P. J., Ayllón, V., Sanchez, L., Ligero, G., Gutierrez-Aranda, I. & Fernández, A. F. (2012) A human ESC model for MLL-AF4 leukemic fusion gene reveals an impaired early hematopoietic-endothelial specification. *Cell research*, 22(6), 986-1002.

Burotto, M., Thomas, A., Subramaniam, D., Giaccone, G. & Rajan, A. (2014) Biomarkers in Early-Stage Non–Small-Cell Lung Cancer: Current Concepts and Future Directions. *Journal of Thoracic Oncology*, 9(11), 1609-1617.

Burt, J. R., Iribarren, C., Fair, J. M., Norton, L. C., Mahbouba, M., Rubin, G. D., Hlatky, M. A., Go, A. S. & Fortmann, S. P. (2008) Incidental findings on cardiac multidetector row computed tomography among healthy older adults: prevalence and clinical correlates. *Archives of internal medicine*, 168(7), 756-761.

Caballero, O. L., Cohen, T., Gurung, S., Chua, R., Lee, P., Chen, Y.-T., Jat, P. & Simpson, A. J. (2013) Effects of CT-Xp gene knock down in melanoma cell lines. *Oncotarget*, 4(4), 531.

Callister, M., Baldwin, D., Akram, A., Barnard, S., Cane, P., Draffan, J., Franks, K., Gleeson, F., Graham, R. & Malhotra, P. (2015) British Thoracic Society guidelines for the investigation and management of pulmonary nodules: accredited by NICE. *Thorax*, 70(Suppl 2), ii1-ii54.

Campagnolo, C., Meyers, K. J., Ryan, T., Atkinson, R. C., Chen, Y.-T., Scanlan, M. J., Ritter, G., Old, L. J. & Batt, C. A. (2004) Real-Time, label-free monitoring of tumor antigen and serum antibody interactions. *Journal of biochemical and biophysical methods*, 61(3), 283-298.

Cancer Research UK (2020) *Leukaemia incidence statistics* Available online: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia/incidence#heading-One [Accessed 23/04/2020].

Capo-chichi, C. D., Agbangbatin, C., Aguida, B., Houngue, S. & Sanni, A. (2022) The Inverse Correlation of Nuclear Protein Lamin A and Ribosomal Proteins L28 Determines Molecular Initiation of Cervical Carcinogenesis in the Context of HR-HPV Infection. *American Journal of Epidemiology*, 10(1), 24-30.

Cappell, K. M., Sinnott, R., Taus, P., Maxfield, K., Scarbrough, M. & Whitehurst, A. W. (2012) Multiple cancer testis antigens function to support tumor cell mitotic fidelity. *Molecular and cellular biology*, 32(20), 4131-4140.

Caron, E., Aebersold, R., Banaei-Esfahani, A., Chong, C. & Bassani-Sternberg, M. (2017) A case for a human immuno-peptidome project consortium. *Immunity*, 47(2), 203-208.

Carozzi, F. M., Bisanzi, S., Carrozzi, L., Falaschi, F., Lopes Pegna, A., Mascalchi, M., Picozzi, G., Peluso, M., Sani, C., Greco, L., Ocello, C., Paci, E. & Group, t. I. W. (2017) Multimodal lung cancer screening using the ITALUNG biomarker panel and low dose computed tomography. Results of the ITALUNG biomarker study. *International Journal of Cancer*, 141(1), 94-101.

Casey, M., Odhiambo, L., Aggarwal, N., Shoukier, M., Garner, J., Islam, K. & Cortes, J. E. (2022) Are pivotal clinical trials for drugs approved for leukemias and multiple myeloma representative of the population at risk? *Journal of Clinical Oncology*, 40(32), 3719-3729.

Cassaro, A., Grillo, G., Notaro, M., Gliozzo, J., Esposito, I., Reda, G., Trojani, A., Valentini, G., Di Camillo, B., Cairoli, R. & Beghini, A. (2021) FZD6 triggers Wnt–signalling driven by WNT10BIVS1 expression and highlights new targets in T-cell acute lymphoblastic leukemia. *Hematological Oncology*, 39(3), 364-379.

Castle, J. C., Kreiter, S., Diekmann, J., Löwer, M., Van de Roemer, N., de Graaf, J., Selmi, A., Diken, M., Boegel, S. & Paret, C. (2012) Exploiting the mutanome for tumor vaccination. *Cancer research*, 72(5), 1081-1091.

Cavé, H., Caye, A., Strullu, M., Aladjidi, N., Vignal, C., Ferster, A., Méchinaud, F., Domenech, C., Pierri, F., Contet, A., Cacheux, V., Irving, J., Kratz, C., Clavel, J. & Verloes, A. (2016) Acute lymphoblastic leukemia in the context of RASopathies. *European Journal of Medical Genetics*, 59(3), 173-178.

Cazzoli, R., Buttitta, F., Di Nicola, M., Malatesta, S., Marchetti, A., Rom, W. N. & Pass, H. I. (2013) microRNAs derived from circulating exosomes as noninvasive biomarkers for screening and diagnosing lung cancer. *Journal of thoracic oncology*, 8(9), 1156-1162.

Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T. & Pfanner, N. (2009) Importing mitochondrial proteins: machineries and mechanisms. *Cell*, 138(4), 628-644.

Chakraborty, C., Sharma, A. R., Sharma, G. & Lee, S.-S. (2020) The interplay among miRNAs, major cytokines, and cancer-related inflammation. *Molecular Therapy-Nucleic Acids*, 20, 606-620.

Chambost, H., van Baren, N., Brasseur, F. & Olive, D. (2001) MAGE-A genes are not expressed in human leukemias. *Leukemia*, 15(11), 1769-1771.

Chang, Y.-X., Lin, Y.-F., Chen, C.-L., Huang, M.-S., Hsiao, M. & Liang, P.-H. (2020) Chaperonin-containing TCP-1 promotes cancer chemoresistance and metastasis through the AKT-GSK3β-β-catenin and XIAP-survivin pathways. *Cancers*, 12(12), 3865.

Chari, A., Vogl, D. T., Gavriatopoulou, M., Nooka, A. K., Yee, A. J., Huff, C. A., Moreau, P., Dingli, D., Cole, C. & Lonial, S. (2019) Oral selinexor–dexamethasone for triple-class refractory multiple myeloma. *New England Journal of Medicine*, 381(8), 727-738.

Cheever, M. A., Allison, J. P., Ferris, A. S., Finn, O. J., Hastings, B. M., Hecht, T. T., Mellman, I., Prindiville, S. A., Viner, J. L., Weiner, L. M. & Matrisian, L. M. (2009) The prioritization of cancer antigens: a national cancer institute pilot project for the acceleration of translational research. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 15(17), 5323-5337.

Chen, C., Huang, X., Yin, W., Peng, M., Wu, F., Wu, X., Tang, J., Chen, M., Wang, X. & Hulbert, A. (2020a) Ultrasensitive DNA hypermethylation detection using plasma for early detection of NSCLC: a study in Chinese patients with very small nodules. *Clinical Epigenetics*, 12, 1-11.

Chen, H., Liu, H. & Qing, G. (2018a) Targeting oncogenic Myc as a strategy for cancer treatment. *Signal transduction and targeted therapy*, 3(1), 5.

Chen, L.-Y., Kang, L.-Q., Zhou, H.-X., Gao, H.-Q., Zhu, X.-F., Xu, N., Yu, L., Wu, D.-P., Xue, S.-L. & Sun, A.-N. (2020b) Successful application of anti-CD19 CAR-T therapy with IL-6 knocking down to patients with central nervous system B-cell acute lymphocytic leukemia. *Translational Oncology*, 13(11), 100838.

Chen, L., Han, S., Li, Y., Zheng, Y. & Zhang, Q. (2022a) SEZ6L2, regulated by USF1, accelerates the growth and metastasis of breast cancer. *Exp Cell Res*, 417(1), 113194.

Chen, L., Peng, M., Li, N., Song, Q., Yao, Y., Xu, B., Liu, H. & Ruan, P. (2018b) Combined use of EpCAM and FRα enables the high-efficiency capture of circulating tumor cells in non-small cell lung cancer. *Scientific Reports*, 8(1), 1188.

Chen, M., Wang, J., Yao, S. F., Zhao, Y., Liu, L., Li, L. W., Xu, T., Gan, L. G., Xiao, C. L., Shan, Z. L., Zhong, L. & Liu, B. Z. (2017) Effect of YAP Inhibition on Human Leukemia HL-60 Cells. *Int J Med Sci*, 14(9), 902-910.

Chen, R., Pan, S., Lai, K., Lai, L. A., Crispin, D. A., Bronner, M. P. & Brentnall, T. A. (2014a) Up-regulation of mitochondrial chaperone TRAP1 in ulcerative colitis associated colorectal cancer. *World J Gastroenterol*, 20(45), 17037-48.

Chen, S.-S., Li, K., Wu, J., Peng, Z.-Y., Wang, Z.-D., Wang, J.-C., Xu, C.-W., Zhu, C.-l., Li, B.-C., Ren, H., Tang, S.-C. & Sun, X. (2021a) Stem signatures associated antibodies yield early diagnosis and precise prognosis predication of patients with non-small cell lung cancer. *Journal of Cancer Research and Clinical Oncology*, 147(1), 223-233.

Chen, X., Liu, Y., Wang, Y., Wang, C., Chen, X., Xiong, Y., Liu, L., Yuan, X., Tang, H., Shu, C., Zhang, J., Guo, A. M., Chen, H. & Yang, J. (2022b) CYP4F2-Catalyzed Metabolism of Arachidonic Acid Promotes Stromal Cell-Mediated Immunosuppression in Non–Small Cell Lung Cancer. *Cancer Research*, 82(21), 4016-4030.

Chen, Y.-A., Lu, C.-Y., Cheng, T.-Y., Pan, S.-H., Chen, H.-F. & Chang, N.-S. (2019) WW domain-containing proteins YAP and TAZ in the hippo pathway as key regulators in stemness maintenance, tissue homeostasis, and tumorigenesis. *Frontiers in Oncology*, 9, 60.

Chen, Y.-T., Venditti, C. A., Theiler, G., Stevenson, B. J., Iseli, C., Gure, A. O., Jongeneel, C. V., Old, L. J. & G Simpson, A. J. (2005) Identification of CT46/HORMAD1, an immunogenic cancer/testis antigen encoding a putative meiosis-related protein. *Cancer immunity*, 5(1).

Chen, Y.-X., Wang, C.-J., Xiao, D.-S., He, B.-M., Li, M., Yi, X.-P., Zhang, W., Yin, J.-Y. & Liu, Z.-Q. (2021b) eIF3a R803K mutation mediates chemotherapy resistance by inducing cellular senescence in small cell lung cancer. *Pharmacological Research*, 174, 105934.

Chen, Y., Hong, C., Qu, J., Chen, J. & Qin, Z. (2022c) Knockdown of lncRNA PCAT6 suppresses the growth of non-small cell lung cancer cells by inhibiting macrophages M2 polarization via miR-326/KLF1 axis. *Bioengineered*, 13(5), 12834-12846.

Chen, Y., Zhao, H., Tan, Z., Zhang, C. & Fu, X. (2015) Bottleneck limitations for microRNA-based therapeutics from bench to the bedside. *Die Pharmazie-An International Journal of Pharmaceutical Sciences*, 70(3), 147-154.

Chen, Z., Fillmore, C. M., Hammerman, P. S., Kim, C. F. & Wong, K.-K. (2014b) Non-small-cell lung cancers: a heterogeneous set of diseases. *Nature Reviews Cancer*, 14(8), 535-546.

Chen, Z., Shi, C., Gao, S., Song, D. & Feng, Y. (2018c) Impact of protamine I on colon cancer proliferation, invasion, migration, diagnosis and prognosis. *Biol Chem*, 399(3), 265-275.

Cheng, H., Liu, Y., Jia, Q., Ma, S., Yuan, W., Jia, H. & Cheng, T. (2016) Novel regulators in hematopoietic stem cells can be revealed by a functional approach under leukemic condition. *Leukemia*, 30(10), 2074-2077.

Cheng, H., Sun, G. & Cheng, T. (2018) Hematopoiesis and microenvironment in hematological malignancies. *Cell Regeneration*, 7(1), 22-26.

Chessells, J., Harrison, G., Richards, S., Bailey, C., Hill, F., Gibson, B. & Hann, I. (2001) Down's syndrome and acute lymphoblastic leukaemia: clinical features and response to treatment. *Archives of disease in childhood*, 85(4), 321-325.

Chiang, C. L.-L., Coukos, G. & Kandalaft, L. E. (2015) Whole Tumor Antigen Vaccines: Where Are We? *Vaccines*, 3(2), 344-372.

Chiang, C. L.-L., Ledermann, J. A., Rad, A. N., Katz, D. R. & Chain, B. M. (2006) Hypochlorous acid enhances immunogenicity and uptake of allogeneic ovarian tumor cells by dendritic cells to cross-prime tumor-specific T cells. *Cancer Immunology, Immunotherapy*, 55, 1384-1395.

Chiarini, F., Lonetti, A., Evangelisti, C., Buontempo, F., Orsini, E., Evangelisti, C., Cappellini, A., Neri, L. M., McCubrey, J. A. & Martelli, A. M. (2016) Advances in understanding the acute lymphoblastic leukemia bone marrow microenvironment: From biology to therapeutic targeting. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1863(3), 449-463.

Chiarini, F., Paganelli, F., Martelli, A. M. & Evangelisti, C. (2020) The Role Played by Wnt/β-Catenin Signaling Pathway in Acute Lymphoblastic Leukemia. *Int J Mol Sci*, 21(3).

Chikwe, J., Cooke, D. & Weiss, A. (2013) *Cardiothoracic surgery*.OUP Oxford.

Chilton, L., Buck, G., Harrison, C., Ketterling, R., Rowe, J., Tallman, M., Goldstone, A., Fielding, A. & Moorman, A. (2014) High hyperdiploidy among adolescents and adults with acute lymphoblastic leukaemia (ALL): cytogenetic features, clinical characteristics and outcome. *Leukemia*, 28(7), 1511-1518.

Chiou, S.-S., Wang, L.-T., Huang, S.-B., Chai, C.-Y., Wang, S.-N., Liao, Y.-M., Lin, P.-C., Liu, K.-Y. & Hsu, S.-H. (2014) Wntless (GPR177) expression correlates with poor prognosis in B-cell precursor acute lymphoblastic leukemia via Wnt signaling. *Carcinogenesis*, 35(10), 2357-2364.

Chiu, D. S., Oram, J. F., LeBoeuf, R. C., Alpers, C. E. & O'Brien, K. D. (1997) High-density lipoprotein-binding protein (HBP)/vigilin is expressed in human atherosclerotic lesions and colocalizes with apolipoprotein E. *Arteriosclerosis, thrombosis, and vascular biology*, 17(11), 2350-2358.

Cho, S., Yang, H., Kim, K. & Jheon, S. (2013) Pathology and Prognosis of Persistent Stable Pure Ground-Glass Opacity Nodules After Surgical Resection. *The Annals of Thoracic Surgery*, 96(4), 1190-1195.

Chorzalska, A., Kim, J. F., Roder, K., Tepper, A., Ahsan, N., Rao, R. S. P., Olszewski, A. J., Yu, X., Terentyev, D. & Morgan, J. (2017) Long-term exposure to imatinib mesylate downregulates hippo pathway and activates YAP in a model of chronic myelogenous leukemia. *Stem cells and development*, 26(9), 656-677.

Coakley, M. & Popat, S. (2020) Management of lung cancer. *Medicine*, 48(4), 273-278.

Cohen, J., Korevaar, D., Altman, D., Bruns, D., Gatsonis, C., Hooft, L., Irwig, L., Levine, D., Reitsma, J. & de Vet, H. (2016) STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016 Nov 14; 6 (11): e012799. doi: 10.1136/bmjopen-2016-012799.

Cohen, J. D., Li, L., Wang, Y., Thoburn, C., Afsari, B., Danilova, L., Douville, C., Javed, A. A., Wong, F., Mattox, A., Hruban, R. H., Wolfgang, C. L., Goggins, M. G., Dal Molin, M., Wang, T. L., Roden, R., Klein, A. P., Ptak, J., Dobbyn, L., Schaefer, J., Silliman, N., Popoli, M., Vogelstein, J. T., Browne, J. D., Schoen, R. E., Brand, R. E., Tie, J., Gibbs, P., Wong, H. L., Mansfield, A. S., Jen, J., Hanash, S. M., Falconi, M., Allen, P. J., Zhou, S., Bettegowda, C., Diaz, L. A., Jr., Tomasetti, C., Kinzler, K. W., Vogelstein, B., Lennon, A. M. & Papadopoulos, N. (2018) Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*, 359(6378), 926-930.

Collin, J. F., Wells, J. W., Czepulkowski, B., Lyne, L., Duriez, P. J., Banham, A. H., Mufti, G. J. & Guinn, B. A. (2015) A novel zinc finger gene, ZNF465, is inappropriately expressed in acute myeloid leukaemia cells. *Genes Chromosomes Cancer*, 54(5), 288-302.

Comeaux, E. Q. & Mullighan, C. G. (2017) TP53 mutations in hypodiploid acute lymphoblastic leukemia. *Cold Spring Harbor perspectives in medicine*, 7(3), a026286.

Connerty, P., Moles, E., de Bock, C. E., Jayatilleke, N., Smith, J. L., Meshinchi, S., Mayoh, C., Kavallaris, M. & Lock, R. B. (2021) Development of siRNA-Loaded Lipid Nanoparticles Targeting Long Non-Coding RNA LINC01257 as a Novel and Safe Therapeutic Approach for t(8;21) Pediatric Acute Myeloid Leukemia. *Pharmaceutics*, 13(10), 1681.

Cottini, F., Hideshima, T., Xu, C., Sattler, M., Dori, M., Agnelli, L., Ten Hacken, E., Bertilaccio, M. T., Antonini, E. & Neri, A. (2014) Rescue of Hippo coactivator YAP1 triggers DNA damage–induced apoptosis in hematological cancers. *Nature medicine*, 20(6), 599-606.

Coulie, P. G., Lehmann, F., Lethe, B., Herman, J., Lurquin, C., Andrawiss, M. & Boon, T. (1995) A mutated intron sequence codes for an antigenic peptide recognized by cytolytic T lymphocytes on a human melanoma. *Proceedings of the National Academy of Sciences*, 92(17), 7976-7980.

Counts, S. J. & Kim, A. W. (2017) Diagnostic imaging and newer modalities for thoracic diseases: PET/computed tomographic imaging and endobronchial ultrasound for staging and its implication for lung cancer. *Surgical Clinics*, 97(4), 733-750.

Coyaud, E., Struski, S., Prade, N., Familiades, J., Eichner, R., Quelen, C., Bousquet, M., Mugneret, F., Talmant, P. & Pages, M.-P. (2010) Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenetique Hematologique study. *Blood, The Journal of the American Society of Hematology*, 115(15), 3089-3097.

Crazzolara, R., Kreczy, A., Mann, G., Heitger, A., Eibl, G., Fink, F. M., Möhle, R. & Meister, B. (2001) High expression of the chemokine receptor CXCR4 predicts extramedullary organ infiltration in childhood acute lymphoblastic leukaemia. *British journal of haematology*, 115(3), 545-553.

Czabotar, P. E., Lessene, G., Strasser, A. & Adams, J. M. (2014) Control of apoptosis by the BCL-2 protein family: implications for physiology and therapy. *Nature reviews Molecular cell biology*, 15(1), 49-63.

D'Ambrosi, S., Giannoukakos, S., Antunes-Ferreira, M., Pedraz-Valdunciel, C., Bracht, J. W. P., Potie, N., Gimenez-Capitan, A., Hackenberg, M., Fernandez Hilario, A., Molina-Vila, M. A., Rosell, R., Würdinger, T. & Koppers-Lalic, D. (2023) Combinatorial Blood Platelets-Derived circRNA and mRNA Signature for Early-Stage Lung Cancer Detection. *Int J Mol Sci*, 24(5).

D'Arcy, P., Maruwge, W., Wolahan, B., Ma, L. & Brodin, B. (2014) Oncogenic functions of the cancer-testis antigen SSX on the proliferation, survival, and signaling pathways of cancer cells. *PloS one*, 9(4), e95136.

Dai, T., Ye, L., Deng, M., Lin, G., Liu, R., Yu, H., Liu, W., Yang, Y. & Wang, G. (2022) Upregulation of TMCO3 Promoting Tumor Progression and Contributing to the Poor Prognosis of Hepatocellular Carcinoma. *Journal of Clinical and Translational Hepatology*, 10(5), 913-924.

Daly, S., Rinewalt, D., Fhied, C., Basu, S., Mahon, B., Liptay, M. J., Hong, E., Chmielewski, G., Yoder, M. A., Shah, P. N., Edell, E. S., Maldonado, F., Bungum, A. O. & Borgia, J. A. (2013) Development and Validation of a Plasma Biomarker Panel for Discerning Clinical Significance of Indeterminate Pulmonary Nodules. *Journal of Thoracic Oncology*, 8(1), 31-36.

Dandekar, S., Romanos‐Sirakis, E., Pais, F., Bhatla, T., Jones, C., Bourgeois, W., Hunger, S. P., Raetz, E. A., Hermiston, M. L. & Dasgupta, R. (2014) Wnt inhibition leads to improved chemosensitivity in paediatric acute lymphoblastic leukaemia. *British journal of haematology*, 167(1), 87-99.

Dander, E., Fallati, A., Gulić, T., Pagni, F., Gaspari, S., Silvestri, D., Cricrì, G., Bedini, G., Portale, F., Buracchi, C., Starace, R., Pasqualini, F., D'Angiò, M., Brizzolara, L., Maglia, O., Mantovani, A., Garlanda, C., Valsecchi, M. G., Locatelli, F., Biondi, A., Bottazzi, B., Allavena, P. & D'Amico, G. (2021) Monocyte–macrophage polarization and recruitment pathways in the tumour microenvironment of B-cell acute lymphoblastic leukaemia. *British Journal of Haematology*, 193(6), 1157-1171.

de Koning, H. J., van der Aalst, C. M., de Jong, P. A., Scholten, E. T., Nackaerts, K., Heuvelmans, M. A., Lammers, J.-W. J., Weenink, C., Yousaf-Khan, U. & Horeweg, N. (2020) Reduced lung-cancer mortality with volume CT screening in a randomized trial. *New England Journal of Medicine*, 382(6), 503-513.

De Paula, A. M., Franques, J., Fernandez, C., Monnier, N., Lunardi, J., Pellissier, J.-F., Figarella-Branger, D. & Pouget, J. (2009) A TPM3 mutation causing cap myopathy. *Neuromuscular Disorders*, 19(10), 685-688.

DeKelver, R. C., Lewin, B., Lam, K., Komeno, Y., Yan, M., Rundle, C., Lo, M.-C. & Zhang, D.-E. (2013) Cooperation between RUNX1-ETO9a and novel transcriptional partner KLF6 in upregulation of Alox5 in acute myeloid leukemia. *PLoS genetics*, 9(10), e1003765.

del Valle-Pérez, B., Martínez, V. G., Lacasa-Salavert, C., Figueras, A., Shapiro, S. S., Takafuta, T., Casanovas, O., Capellà, G., Ventura, F. & Viñals, F. (2010) Filamin B plays a key role in vascular endothelial growth factor-induced endothelial cell motility through its interaction with Rac-1 and Vav-2. *Journal of Biological Chemistry*, 285(14), 10748-10760.

Deng, Z., Hasegawa, M., Aoki, K., Matayoshi, S., Kiyuna, A., Yamashita, Y., Uehara, T., Agena, S., Maeda, H. & Xie, M. (2014) A comprehensive evaluation of human papillomavirus positive status and p16INK4a overexpression as a prognostic biomarker in head and neck squamous cell carcinoma. *International journal of oncology*, 45(1), 67-76.

Depil, S., Duchateau, P., Grupp, S., Mufti, G. & Poirot, L. (2020) 'Off-the-shelf'allogeneic CAR T cells: development and challenges. *Nature reviews Drug discovery*, 19(3), 185-199.

Deppen, S. A., Blume, J. D., Aldrich, M. C., Fletcher, S. A., Massion, P. P., Walker, R. C., Chen, H. C., Speroff, T., Degesys, C. A. & Pinkerman, R. (2014) Predicting lung cancer prior to surgical resection in patients with lung nodules. *Journal of Thoracic Oncology*, 9(10), 1477-1484.

Deppen, S. A. & Grogan, E. L. (2015) Using Clinical Risk Models for Lung Nodule Classification. *Semin Thorac Cardiovasc Surg*, 27(1), 30-5.

Derakhshan, Z., Khamisipour, G., Soleimani, F. H. & Motamed, N. (2022) Serum MicroRNAs: -28-3p, -31-5p, -378a-3p, and -382-5p as novel potential biomarkers in acute lymphoblastic leukemia. *Gene Reports*, 27, 101582.

Desterke, C., Hugues, P., Hwang, J. W., Bennaceur-Griscelli, A. & Turhan, A. G. (2020) Embryonic Program Activated during Blast Crisis of Chronic Myelogenous Leukemia (CML) Implicates a TCF7L2 and MYC Cooperative Chromatin Binding. *Int J Mol Sci*, 21(11).

Dharaneeswaran, H., Abid, M. R., Yuan, L., Dupuis, D., Beeler, D., Spokes, K. C., Janes, L., Sciuto, T., Kang, P. M. & Jaminet, S.-C. S. (2014) FOXO1-mediated activation of Akt plays a critical role in vascular homeostasis. *Circulation research*, 115(2), 238-251.

Diederich, S., Wormanns, D., Lenzen MS, H., Semik, M., Thomas, M. & Peters, P. E. (2000) Screening for asymptomatic early bronchogenic carcinoma with low dose CT of the chest. *Cancer*, 89(S11), 2483-2484.

Diederich, S., Wormanns, D., Semik, M., Thomas, M., Lenzen, H., Roos, N. & Heindel, W. (2002) Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers. *Radiology*, 222(3), 773-781.

Dietel, M., Arps, H., Klapdor, R., Müller-Hagen, S., Sieck, M. & Hoffmann, L. (1986) Antigen detection by the monoclonal antibodies CA 19-9 and CA 125 in normal and tumor tissue and patients' sera. *Journal of cancer research and clinical oncology*, 111(3), 257-265.

Dittadi, R. & Gion, M. (2013) Re: biological variation of neuroendocrine tumor markers chromogranin A and neuron-specific enolase. *Clinical Biochemistry*, 12(46), 1145.

Djebbi, K., Xing, J., Weng, T., Bahri, M., Elaguech, M. A., Du, C., Shi, B., Hu, L., He, S. & Liao, P. (2022) Highly sensitive fluorescence multiplexed miRNAs biosensors for accurate clinically diagnosis lung cancer disease using LNA-modified DNA probe and DSN enzyme. *Analytica Chimica Acta*, 1208, 339778.

Djureinovic, D., Hallström, B. M., Horie, M., Mattsson, J. S. M., La Fleur, L., Fagerberg, L., Brunnström, H., Lindskog, C., Madjar, K., Rahnenführer, J., Ekman, S., Ståhle, E., Koyi, H., Brandén, E., Edlund, K., Hengstler, J. G., Lambe, M., Saito, A., Botling, J., Pontén, F., Uhlén, M. & Micke, P. (2016) Profiling cancer testis antigens in non-small-cell lung cancer. *JCI Insight*, 1(10), e86837.

Dong, L., Lin, F., Wu, W., Huang, W. & Cai, Z. (2016) Transcriptional cofactor Mask2 is required for YAP-induced cell growth and migration in bladder cancer cell. *J Cancer*, 7(14), 2132-2138.

Dong, X., Chang, M., Song, X., Ding, S., Xie, L. & Song, X. (2021a) Plasma miR-1247-5p, miR-301b-3p and miR-105-5p as potential biomarkers for early diagnosis of non-small cell lung cancer. *Thoracic Cancer*, 12(4), 539-548.

Dong, X., Song, X., Ding, S., Yu, M., Shang, X., Wang, K., Chang, M., Xie, L. & Song, X. (2021b) Tumor-educated platelet SNORD55 as a potential biomarker for the early diagnosis of non-small cell lung cancer. *Thoracic Cancer*, 12(5), 659-666.

Doseeva, V., Colpitts, T., Gao, G., Woodcock, J. & Knezevic, V. (2015) Performance of a multiplexed dual analyte immunoassay for the early detection of non-small cell lung cancer. *Journal of translational medicine*, 13, 55-55.

Dou, Y., Zhu, Y., Ai, J., Chen, H., Liu, H., Borgia, J. A., Li, X., Yang, F., Jiang, B. & Wang, J. (2018) Plasma small ncRNA pair panels as novel biomarkers for early-stage lung adenocarcinoma screening. *BMC genomics*, 19(1), 1-10.

Doyle, J. M., Gao, J., Wang, J., Yang, M. & Potts, P. R. (2010) MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. *Molecular cell*, 39(6), 963-974.

Drapkin, R., Von Horsten, H. H., Lin, Y., Mok, S. C., Crum, C. P., Welch, W. R. & Hecht, J. L. (2005) Human epididymis protein 4 (HE4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. *Cancer research*, 65(6), 2162-2169.

Du, Q., Yu, R., Wang, H., Yan, D., Yuan, Q., Ma, Y., Slamon, D., Hou, D., Wang, H. & Wang, Q. (2018) Significance of tumor‐associated autoantibodies in the early diagnosis of lung cancer. *The Clinical Respiratory Journal*, 12(6), 2020-2028.

Du, X. L. & Chen, Q. (2013) Recent Advancements of Bortezomib in Acute Lymphocytic Leukemia Treatment. *Acta Haematologica*, 129(4), 207-214.

Du, Z., Fei, T., Verhaak, R. G., Su, Z., Zhang, Y., Brown, M., Chen, Y. & Liu, X. S. (2013) Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nature structural & molecular biology*, 20(7), 908-913.

Duan, X., Qiao, S., Li, D., Li, S., Zheng, Z., Wang, Q. & Zhu, X. (2021) Circulating miRNAs in Serum as Biomarkers for Early Diagnosis of Non-small Cell Lung Cancer. *Frontiers in Genetics*, 12.

Dufva, O., Koski, J., Maliniemi, P., Ianevski, A., Klievink, J., Leitner, J., Pölönen, P., Hohtari, H., Saeed, K. & Hannunen, T. (2020) Integrated drug profiling and CRISPR screening identify essential pathways for CAR T-cell cytotoxicity. *Blood*, 135(9), 597-609.

Dunn, B. K., Jegalian, K. & Greenwald, P. (2010) Biomarkers for early detection and as surrogate endpoints in cancer prevention trials: issues and opportunities. *Clinical Cancer Prevention*, 21-47.

Ebrahimi, E., Shabestari, R. M., Bashash, D. & Safa, M. (2022) Synergistic apoptotic effect of Mcl-1 inhibition and doxorubicin on B-cell precursor acute lymphoblastic leukemia cells. *Molecular Biology Reports*, 49(3), 2025-2036.

Edelman, M. J., Hodgson, L., Rosenblatt, P. Y., Christenson, R. H., Vokes, E. E., Wang, X. & Kratzke, R. (2012) CYFRA 21-1 as a prognostic and predictive marker in advanced non-small-cell lung cancer in a prospective trial: CALGB 150304. *J Thorac Oncol*, 7(4), 649-54.

Edey, A. J. & Hansell, D. M. (2009) Incidentally detected small pulmonary nodules on CT. *Clinical Radiology*, 64(9), 872-884.

Egeblad, M. & Werb, Z. (2002) New functions for the matrix metalloproteinases in cancer progression. *Nature reviews cancer*, 2(3), 161-174.

Egland, K. A., Kumar, V., Duray, P. & Pastan, I. (2002) Characterization of overlapping XAGE-1 transcripts encoding a cancer testis antigen expressed in lung, breast, and other types of cancers. *Molecular Cancer Therapeutics*, 1(7), 441-450.

Egli, M. & Manoharan, M. (2019) Re-engineering RNA molecules into therapeutic agents. *Accounts of Chemical Research*, 52(4), 1036-1047.

Elden, S. M. K., Azzam, A., Elbassal, F., El-Hawy, M. A. & Saleh, N. Y. (2018) Evaluation of survivin gene expression as a prognostic biomarker in pediatric B-acute lymphoblastic leukemia. *Menoufia Medical Journal*, 31(3), 952.

Elsallab, M., Ellithi, M., Hempel, S., Abdel-Azim, H. & Abou-el-Enein, M. (2023) Long-term response to autologous anti-CD19 chimeric antigen receptor T cells in relapsed or refractory B cell acute lymphoblastic leukemia: a systematic review and meta-analysis. *Cancer Gene Therapy*, 1-10.

Erasmus, J. J., Connolly, J. E., McAdams, H. P. & Roggli, V. L. (2000) Solitary pulmonary nodules: Part I. Morphologic evaluation for differentiation of benign and malignant lesions. *Radiographics*, 20(1), 43-58.

Escudero, J. M., Auge, J. M., Filella, X., Torne, A., Pahisa, J. & Molina, R. (2011) Comparison of serum human epididymis protein 4 with cancer antigen 125 as a tumor marker in patients with malignant and nonmalignant diseases. *Clinical Chemistry*, 57(11), 1534-1544.

Etzioni, R., Pepe, M., Longton, G., Hu, C. & Goodman, G. (1999) Incorporating the time dimension in receiver operating characteristic curves: a case study of prostate cancer. *Medical Decision Making*, 19(3), 242-251.

Ezzatifar, F., Rafiei, A., Valadan, R., Asgarian-Omran, H. & Jeddi-Tehrani, M. (2022) Detection of Novel Autoantibodies to Nucleolin's RNA-binding Domains as a Serum Tumor Biomarker Through ELISA. *Iran J Allergy Asthma Immunol*, 21(6), 616-625.

Faderl, S., Jeha, S. & Kantarjian, H. M. (2003) The biology and therapy of adult acute lymphoblastic leukemia. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 98(7), 1337-1354.

Faderl, S., O'Brien, S., Pui, C. H., Stock, W., Wetzler, M., Hoelzer, D. & Kantarjian, H. M. (2010) Adult acute lymphoblastic leukemia: concepts and strategies. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 116(5), 1165-1176.

Fahrmann, J. F., Marsh, T., Irajizad, E., Patel, N., Murage, E., Vykoukal, J., Dennison, J. B., Do, K. A., Ostrin, E., Spitz, M. R., Lam, S., Shete, S., Meza, R., Tammemägi, M. C., Feng, Z. & Hanash, S. M. (2022) Blood-Based Biomarker Panel for Personalized Lung Cancer Risk Assessment. *J Clin Oncol*, 40(8), 876-883.

Falà, F., Blalock, W. L., Tazzari, P. L., Cappellini, A., Chiarini, F., Martinelli, G., Tafuri, A., McCubrey, J. A., Cocco, L. & Martelli, A. M. (2008) Proapoptotic activity and chemosensitizing effect of the novel Akt inhibitor (2S)-1-(1H-Indol-3-yl)-3-[5-(3-methyl-2H-indazol-5-yl) pyridin-3-yl] oxypropan2-amine (A443654) in T-cell acute lymphoblastic leukemia. *Molecular pharmacology*, 74(3), 884-895.

Fan, L., Sha, J., Teng, J., Li, D., Wang, C., Xia, Q., Chen, H., Su, B. & Qi, H. (2018) Evaluation of serum paired microRNA ratios for differential diagnosis of non-small cell lung cancer and benign pulmonary diseases. *Molecular Diagnosis & Therapy*, 22(4), 493-502.

Fanipakdel, A., Seilanian Toussi, M., Rezazadeh, F., Mohamadian Roshan, N. & Javadinia, S. A. (2019) Overexpression of cancer-testis antigen melanoma-associated antigen A1 in lung cancer: A novel biomarker for prognosis, and a possible target for immunotherapy. *Journal of Cellular Physiology*, 234(7), 12080-12086.

Farlow, E., Vercillo, M., Coon, J., Basu, S., Kim, A., Faber, L., Warren, W., Bonomi, P., Liptay, M. & Borgia, J. A. (2010a) A multi-analyte serum test for the detection of non-small cell lung cancer. *British journal of cancer*, 103(8), 1221-1228.

Farlow, E. C., Patel, K., Basu, S., Lee, B. S., Kim, A. W., Coon, J. S., Faber, L. P., Bonomi, P., Liptay, M. J. & Borgia, J. A. (2010b) Development of a multiplexed tumor-associated autoantibody-based blood test for the detection of non-small cell lung cancer. *Clin Cancer Res*, 16(13), 3452-62.

Felix, L., Serra-Tosio, G., Lantuejoul, S., Timsit, J.-F., Moro-Sibilot, D., Brambilla, C. & Ferretti, G. (2011) CT characteristics of resolving ground-glass opacities in a lung cancer screening programme. *European journal of radiology*, 77(3), 410-416.

Field, J. K., Duffy, S., Baldwin, D. R., Whynes, D., Devaraj, A., Brain, K. E., Eisen, T., Gosney, J., Green, B. & Holemans, J. (2016) UK Lung Cancer RCT Pilot Screening Trial: baseline findings from the screening arm provide evidence for the potential implementation of lung cancer screening. *Thorax*, 71(2), 161-170.

Fleming, T. R. (2005) Surrogate endpoints and FDA's accelerated approval process. *Health affairs*, 24(1), 67-78.

Fleming, T. R. & DeMets, D. L. (1996) Surrogate end points in clinical trials: are we being misled? *Annals of internal medicine*, 125(7), 605-613.

Fong, K. M., Sekido, Y., Gazdar, A. F. & Minna, J. D. (2003) Lung cancer - 9: Molecular biology of lung cancer: clinical implications. *Thorax*, 58(10), 892-900.

Fratta, E., Coral, S., Covre, A., Parisi, G., Colizzi, F., Danielli, R., Nicolay, H. J. M., Sigalotti, L. & Maio, M. (2011) The biology of cancer testis antigens: putative function, regulation and therapeutic potential. *Molecular oncology*, 5(2), 164-182.

Fun, X. H. & Thibault, G. (2020) Lipid bilayer stress and proteotoxic stress-induced unfolded protein response deploy divergent transcriptional and non-transcriptional programmes. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1865(1), 158449.

Gao, H.-Y., Luo, X.-G., Chen, X. & Wang, J.-H. (2015a) Identification of key genes affecting disease free survival time of pediatric acute lymphoblastic leukemia based on bioinformatic analysis. *Blood Cells, Molecules, and Diseases*, 54(1), 38-43.

Gao, L., Xie, E., Yu, T., Chen, D., Zhang, L., Zhang, B., Wang, F., Xu, J., Huang, P. & Liu, X. (2015b) Methylated APC and RASSF1A in multiple specimens contribute to the differential diagnosis of patients with undetermined solitary pulmonary nodules. *Journal of thoracic disease*, 7(3), 422.

Garcia-Gonzalo, F. R., Corbit, K. C., Sirerol-Piquer, M. S., Ramaswami, G., Otto, E. A., Noriega, T. R., Seol, A. D., Robinson, J. F., Bennett, C. L. & Josifova, D. J. (2011) A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition. *Nature genetics*, 43(8), 776-784.

Garcillán, B., Figgett, W. A., Infantino, S., Lim, E. X. & Mackay, F. (2018) Molecular control of B-cell homeostasis in health and malignancy. *Immunology & Cell Biology*, 96(5), 453-462.

Gardner, R., Wu, D., Cherian, S., Fang, M., Hanafi, L.-A., Finney, O., Smithers, H., Jensen, M. C., Riddell, S. R., Maloney, D. G. & Turtle, C. J. (2016) Acquisition of a CD19-negative myeloid phenotype allows immune escape of MLL-rearranged B-ALL from CD19 CAR-T-cell therapy. *Blood*, 127(20), 2406-2410.

Gasparri, R., Noberini, R., Cuomo, A., Yadav, A., Tricarico, D., Salvetto, C., Maisonneuve, P., Caminiti, V., Sedda, G., Sabalic, A., Bonaldi, T. & Spaggiari, L. (2023) Serum proteomics profiling identifies a preliminary signature for the diagnosis of early-stage lung cancer. *PROTEOMICS – Clinical Applications*, 17(2), 2200093.

Gaundar, S. S., Bradstock, K. F. & Bendall, L. J. (2009) p38MAPK inhibitors attenuate cytokine production by bone marrow stromal cells and reduce stroma-mediated proliferation of acute lymphoblastic leukemia cells. *Cell Cycle*, 8(18), 2977-2985.

Gauthier, J. & Turtle, C. J. (2021) Chimeric Antigen Receptor T-Cell Therapy for B-Cell Acute Lymphoblastic Leukemia: Current Landscape in 2021. *The Cancer Journal*, 27(2), 98-106.

Gazon, H., Belrose, G., Terol, M., Meniane, J.-C., Mesnard, J.-M., Cesaire, R. & Peloponese Jr, J.-M. (2016) Impaired expression of DICER and some microRNAs in HBZ expressing cells from acute adult T-cell leukemia patients. *Oncotarget*, 7(21), 30258.

Ge, Z., Gu, Y., Xiao, L., Han, Q., Li, J., Chen, B., Yu, J., Kawasawa, Y. I., Payne, K. J., Dovat, S. & Song, C. (2016) Co-existence of IL7R high and SH2B3 low expression distinguishes a novel high-risk acute lymphoblastic leukemia with Ikaros dysfunction. *Oncotarget*, 7(29), 46014-46027.

Geng, H., Brennan, S., Milne, T. A., Chen, W.-Y., Li, Y., Hurtz, C., Kweon, S.-M., Zickl, L., Shojaee, S. & Neuberg, D. (2012) Integrative epigenomic analysis identifies biomarkers and therapeutic targets in adult B-acute lymphoblastic leukemia. *Cancer discovery*, 2(11), 1004-1023.

Gholami, M., Mirfakhraie, R., Movafagh, A., Jalaeekhoo, H., Kalahroodi, R., Zare-Abdollahi, D. & Zare-Karizi, S. (2014) The expression analysis of LATS2 gene in de novo AML patients. *Medical Oncology*, 31(5), 1-4.

Giard, D. J., Aaronson, S. A., Todaro, G. J., Arnstein, P., Kersey, J. H., Dosik, H. & Parks, W. P. (1973) In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *Journal of the National Cancer Institute*, 51(5), 1417-1423.

Gibbons, D. L., Byers, L. A. & Kurie, J. M. (2014) Smoking, p53 mutation, and lung cancer. *Molecular cancer research : MCR*, 12(1), 3-13.

Giraud, J., Molina‐Castro, S., Seeneevassen, L., Sifré, E., Izotte, J., Tiffon, C., Staedel, C., Boeuf, H., Fernandez, S. & Barthelemy, P. (2020) Verteporfin targeting YAP1/TAZ‐TEAD transcriptional activity inhibits the tumorigenic properties of gastric cancer stem cells. *International Journal of Cancer*, 146(8), 2255-2267.

Gjerstorff, M. F., Andersen, M. H. & Ditzel, H. J. (2015) Oncogenic cancer/testis antigens: prime candidates for immunotherapy. *Oncotarget*, 6(18), 15772.

Gjerstorff, M. F., Harkness, L., Kassem, M., Frandsen, U., Nielsen, O., Lutterodt, M., Møllgård, K. & Ditzel, H. J. (2008) Distinct GAGE and MAGE-A expression during early human development indicate specific roles in lineage differentiation. *Human reproduction*, 23(10), 2194-2201.

Goebel, C., Louden, C. L., McKenna, R., Onugha, O., Wachtel, A. & Long, T. (2019) Diagnosis of non-small cell lung cancer for early stage asymptomatic patients. *Cancer Genomics & Proteomics*, 16(4), 229-244.

Goldstone, A. H., Richards, S. M., Lazarus, H. M., Tallman, M. S., Buck, G., Fielding, A. K., Burnett, A. K., Chopra, R., Wiernik, P. H. & Foroni, L. (2008) In adults with standard-risk acute lymphoblastic leukemia, the greatest benefit is achieved from a matched sibling allogeneic transplantation in first complete remission, and an autologous transplantation is less effective than conventional consolidation/maintenance chemotherapy in all patients: final results of the International ALL Trial (MRC UKALL XII/ECOG E2993). *Blood, The Journal of the American Society of Hematology*, 111(4), 1827-1833.

González-Sancho, J. M., Larriba, M. J. & Muñoz, A. (2020) Wnt and Vitamin D at the Crossroads in Solid Cancer. *Cancers (Basel)*, 12(11).

Gordeeva, O. (2018) Cancer-testis antigens: Unique cancer stem cell biomarkers and targets for cancer therapy, *Seminars in cancer biology*. Elsevier.

Gould, M. K., Donington, J., Lynch, W. R., Mazzone, P. J., Midthun, D. E., Naidich, D. P. & Wiener, R. S. (2013) Evaluation of individuals with pulmonary nodules: When is it lung cancer?: Diagnosis and management of lung cancer: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*, 143(5), e93S-e120S.

Gowda, C., Song, C., Kapadia, M., Payne, J. L., Hu, T., Ding, Y. & Dovat, S. (2017) Regulation of cellular proliferation in acute lymphoblastic leukemia by Casein Kinase II (CK2) and Ikaros. *Adv Biol Regul*, 63, 71-80.

Green, T. M., de Stricker, K. & Møller, M. B. (2009) Validation of putative reference genes for normalization of Q-RT-PCR data from paraffin-embedded lymphoid tissue. *Diagnostic Molecular Pathology*, 18(4), 243-249.

Greer, P. (2002) Closing in on the biological functions of Fps/Fes and Fer. *Nature Reviews Molecular Cell Biology*, 3(4), 278-289.

Greiner, J., Ringhoffer, M., Taniguchi, M., Li, L., Schmitt, A., Shiku, H., Döhner, H. & Schmitt, M. (2004) mRNA expression of leukemia‐associated antigens in patients with acute myeloid leukemia for the development of specific immunotherapies. *International journal of cancer*, 108(5), 704-711.

Greve, K. B., Lindgreen, J. N., Terp, M. G., Pedersen, C. B., Schmidt, S., Mollenhauer, J., Kristensen, S. B., Andersen, R. S., Relster, M. M. & Ditzel, H. J. (2015) Ectopic expression of cancer/testis antigen SSX2 induces DNA damage and promotes genomic instability. *Molecular oncology*, 9(2), 437-449.

Grewal, R. G. & Austin, J. (1994) CT demonstration of calcification in carcinoma of the lung. *Journal of computer assisted tomography*, 18(6), 867-871.

Gridelli, C., Rossi, A., Carbone, D. P., Guarize, J., Karachaliou, N., Mok, T., Petrella, F., Spaggiari, L. & Rosell, R. (2015) Non-small-cell lung cancer. *Nature reviews Disease primers*, 1(1), 1-16.

Grunnet, M. & Sorensen, J. (2012) Carcinoembryonic antigen (CEA) as tumor marker in lung cancer. *Lung cancer*, 76(2), 138-143.

Gu, J., Mao, W., Ren, W., Xu, F., Zhu, Q., Lu, C., Lin, Z., Zhang, Z., Chu, Y., Liu, R. & Ge, D. (2019) Ubiquitin-protein ligase E3C maintains non-small-cell lung cancer stemness by targeting AHNAK-p53 complex. *Cancer Letters*, 443, 125-134.

Guillamot, M., Cimmino, L. & Aifantis, I. (2016) The Impact of DNA Methylation in Hematopoietic Malignancies. *Trends in Cancer*, 2(2), 70-83.

Guimaraes, M. D., Marchiori, E., Hochhegger, B., Chojniak, R. & Gross, J. L. (2014) CT-guided biopsy of lung lesions: defining the best needle option for a specific diagnosis. *Clinics*, 69(5), 335-340.

Guinn, B.-a., Bland, E. A., Lodi, U., Liggins, A. P., Tobal, K., Petters, S., Wells, J. W., Banham, A. H. & Mufti, G. J. (2005) Humoral detection of leukaemia-associated antigens in presentation acute myeloid leukaemia. *Biochemical and biophysical research communications*, 335(4), 1293-1304.

Guinn, B.-a., Collin, J. F., Li, G., Rees, R. C. & Mufti, G. J. (2002) Optimised SEREX technique for the identification of leukaemia-associated antigens. *Journal of Immunological Methods*, 264(1), 207-214.

Gumireddy, K., Li, A., Chang, D. H., Liu, Q., Kossenkov, A. V., Yan, J., Korst, R. J., Nam, B. T., Xu, H., Zhang, L., Ganepola, G. A., Showe, L. C. & Huang, Q. (2015) AKAP4 is a circulating biomarker for non-small cell lung cancer. *Oncotarget*, 6(19), 17637-47.

Guo, Y., Zhu, H., Li, X., Ma, C., Li, Y., Sun, T., Wang, Y., Wang, C., Guan, W. & Liu, C. (2021) RepSox effectively promotes the induced differentiation of sheep fibroblasts into adipocytes via the inhibition of the TGF-β1/Smad pathway. *Int J Mol Med*, 48(2), 148.

Gure, A. O., Chua, R., Williamson, B., Gonen, M., Ferrera, C. A., Gnjatic, S., Ritter, G., Simpson, A. J., Chen, Y.-T. & Old, L. J. (2005) Cancer-testis genes are coordinately expressed and are markers of poor outcome in non–small cell lung cancer. *Clinical Cancer Research*, 11(22), 8055-8062.

Gurel, Z., Ronni, T., Ho, S., Kuchar, J., Payne, K. J., Turk, C. W. & Dovat, S. (2008) Recruitment of ikaros to pericentromeric heterochromatin is regulated by phosphorylation. *Journal of Biological Chemistry*, 283(13), 8291-8300.

Gurzov, E. N., Stanley, W. J., Brodnicki, T. C. & Thomas, H. E. (2015) Protein tyrosine phosphatases: molecular switches in metabolism and diabetes. *Trends in Endocrinology & Metabolism*, 26(1), 30-39.

Guyatt, G. H., Tugwell, P. X., Feeny, D. H., Haynes, R. B. & Drummond, M. (1986) A framework for clinical evaluation of diagnostic technologies. *CMAJ: Canadian Medical Association Journal*, 134(6), 587.

Hahn, W. C., Bader, J. S., Braun, T. P., Califano, A., Clemons, P. A., Druker, B. J., Ewald, A. J., Fu, H., Jagu, S., Kemp, C. J., Kim, W., Kuo, C. J., McManus, M. T., B. Mills, G., Mo, X., Sahni, N., Schreiber, S. L., Talamas, J. A., Tamayo, P., Tyner, J. W., Wagner, B. K., Weiss, W. A., Gerhard, D. S., Dancik, V., Gill, S., Hua, B., Sharifnia, T., Viswanathan, V., Zou, Y., Dela Cruz, F., Kung, A., Stockwell, B., Boehm, J., Dempster, J., Manguso, R., Vazquez, F., Cooper, L. A. D., Du, Y., Ivanov, A., Lonial, S., Moreno, C. S., Niu, Q., Owonikoko, T., Ramalingam, S., Reyna, M., Zhou, W., Grandori, C., Shmulevich, I., Swisher, E., Cai, J., Chan, I. S., Dunworth, M., Ge, Y., Georgess, D., Grasset, E. M., Henriet, E., Knútsdóttir, H., Lerner, M. G., Padmanaban, V., Perrone, M. C., Suhail, Y., Tsehay, Y., Warrier, M., Morrow, Q., Nechiporuk, T., Long, N., Saultz, J., Kaempf, A., Minnier, J., Tognon, C. E., Kurtz, S. E., Agarwal, A., Brown, J., Watanabe-Smith, K., Vu, T. Q., Jacob, T., Yan, Y., Robinson, B., Lind, E. F., Kosaka, Y., Demir, E., Estabrook, J., Grzadkowski, M., Nikolova, O., Chen, K., Deneen, B., Liang, H., Bassik, M. C., Bhattacharya, A., Brennan, K., Curtis, C., Gevaert, O., Ji, H. P., Karlsson, K. A. J., Karagyozova, K., Lo, Y.-H., Liu, K., Nakano, M., Sathe, A., Smith, A. R., et al (2021) An expanded universe of cancer targets. *Cell*, 184(5), 1142-1155.

Hajian-Tilaki, K. (2014) Sample size estimation in diagnostic test studies of biomedical informatics. *Journal of Biomedical Informatics*, 48, 193-204.

Haldrup, J., Strand, S. H., Cieza-Borrella, C., Jakobsson, M. E., Riedel, M., Norgaard, M., Hedensted, S., Dagnaes-Hansen, F., Ulhoi, B. P., Eeles, R., Borre, M., Olsen, J. V., Thomsen, M., Kote-Jarai, Z. & Sorensen, K. D. (2021) FRMD6 has tumor suppressor functions in prostate cancer. *Oncogene*, 40(4), 763-776.

Hammarström, S. (1999) The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. *Seminars in Cancer Biology*, 9(2), 67-81.

Han, J., Jin, R., Zhang, M., Guo, Q. & Zhou, F. (2017) Ikaros 6 protects acute lymphoblastic leukemia cells against daunorubicin-induced apoptosis by activating the Akt-FoxO1 pathway. *Journal of Leukocyte Biology*, 101(3), 675-681.

Han, S. P., Tang, Y. H. & Smith, R. (2010) Functional diversity of the hnRNPs: past, present and perspectives. *Biochemical Journal*, 430(3), 379-392.

Han, X., Garcia-Manero, G., McDonnell, T. J., Lozano, G., Medeiros, L. J., Xiao, L., Rosner, G., Nguyen, M., Fernandez, M., Valentin-Vega, Y. A., Barboza, J., Jones, D. M., Rassidakis, G. Z., Kantarjian, H. M. & Bueso-Ramos, C. E. (2007) HDM4 (HDMX) is widely expressed in adult pre-B acute lymphoblastic leukemia and is a potential therapeutic target. *Modern Pathology*, 20(1), 54-62.

Han, Y. (2019) Analysis of the role of the Hippo pathway in cancer. *Journal of Translational Medicine*, 17(1), 116.

Hanagiri, T., Sugaya, M., Takenaka, M., Oka, S., Baba, T., Shigematsu, Y., Nagata, Y., Shimokawa, H., Uramoto, H. & Takenoyama, M. (2011) Preoperative CYFRA 21-1 and CEA as prognostic factors in patients with stage I non-small cell lung cancer. *Lung cancer*, 74(1), 112-117.

Hansell, D. M., Bankier, A. A., MacMahon, H., McLoud, T. C., Muller, N. L. & Remy, J. (2008) Fleischner Society: glossary of terms for thoracic imaging. *Radiology*, 246(3), 697-722.

Hansen, Q., Bachas, C., Smit, L. & Cloos, J. (2022) Characteristics of leukemic stem cells in acute leukemia and potential targeted therapies for their specific eradication. *Cancer Drug Resistance*, 5(2), 344.

Hansson, A., Bloor, B. K., Sarang, Z., Haig, Y., Morgan, P. R., Stark, H. J., Fusenig, N. E., Ekstrand, J. & Grafström, R. C. (2003) Analysis of proliferation, apoptosis and keratin expression in cultured normal and immortalized human buccal keratinocytes. *European journal of oral sciences*, 111(1), 34-41.

Haque, S. & Vaiselbuh, S. R. (2022) Exosomal DNMT1 mRNA transcript is elevated in acute lymphoblastic leukemia which might reprograms leukemia progression. *Cancer Genet*, 260-261, 57-64.

Harada, M., Pokrovskaja-Tamm, K., Söderhäll, S., Heyman, M., Grander, D. & Corcoran, M. (2012) Involvement of miR17 pathway in glucocorticoid-induced cell death in pediatric acute lymphoblastic leukemia. *Leukemia & lymphoma*, 53(10), 2041-2050.

Hartmann, E. M., Campo, E., Wright, G., Lenz, G., Salaverria, I., Jares, P., Xiao, W., Braziel, R. M., Rimsza, L. M. & Chan, W.-C. (2010) Pathway discovery in mantle cell lymphoma by integrated analysis of high-resolution gene expression and copy number profiling. *Blood, The Journal of the American Society of Hematology*, 116(6), 953-961.

Hatzis, P., van der Flier, L. G., van Driel, M. A., Guryev, V., Nielsen, F., Denissov, S., Nijman, I. c. J., Koster, J., Santo, E. E. & Welboren, W. (2008) Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. *Molecular and cellular biology*, 28(8), 2732-2744.

Hauer, J., Mullighan, C., Morillon, E., Wang, G., Bruneau, J., Brousse, N., Lelorc'h, M., Romana, S., Boudil, A., Tiedau, D., Kracker, S., Bushmann, F. D., Borkhardt, A., Fischer, A., Hacein-Bey-Abina, S. & Cavazzana-Calvo, M. (2011) Loss of p19Arf in a Rag1−/−

B-cell precursor population initiates acute B-lymphoblastic leukemia. *Blood*, 118(3), 544-553.

He, Q., Liu, Z., Liu, Z., Lai, Y., Zhou, X. & Weng, J. (2019) TCR-like antibodies in cancer immunotherapy. *Journal of hematology & oncology*, 12, 1-13.

He, S., Huang, Y., Wang, Y., Tang, J., Song, Y., Yu, X., Ma, J., Wang, S., Yin, H. & Li, Q. (2015) Histamine-releasing factor/translationally controlled tumor protein plays a role in induced cell adhesion, apoptosis resistance and chemoresistance in non-Hodgkin lymphomas. *Leukemia & Lymphoma*, 56(7), 2153-2161.

He, S., Li, Y., Shi, X., Wang, L., Cai, D., Zhou, J. & Yu, L. (2023) DNA methylation landscape reveals LIN7A as a decitabine-responsive marker in patients with t (8; 21) acute myeloid leukemia. *Clinical Epigenetics*, 15(1), 37.

He, Y. D. (2006) Genomic approach to biomarker identification and its recent applications. *Cancer Biomarkers*, 2(3-4), 103-133.

Hendriks, W. J. & Böhmer, F.-D. (2016) Non-transmembrane PTPs in cancer. *Protein Tyrosine Phosphatases in Cancer*, 47-113.

Hendriks, W. J. & Pulido, R. (2013) Protein tyrosine phosphatase variants in human hereditary disorders and disease susceptibilities. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1832(10), 1673-1696.

Hennessey, P. T., Sanford, T., Choudhary, A., Mydlarz, W. W., Brown, D., Adai, A. T., Ochs, M. F., Ahrendt, S. A., Mambo, E. & Califano, J. A. (2012) Serum microRNA biomarkers for detection of non-small cell lung cancer. *PloS one*, 7(2), e32307.

Henschke, C. I., McCauley, D. I., Yankelevitz, D. F., Naidich, D. P., McGuinness, G., Miettinen, O. S., Libby, D. M., Pasmantier, M. W., Koizumi, J. & Altorki, N. K. (1999) Early Lung Cancer Action Project: overall design and findings from baseline screening. *The Lancet*, 354(9173), 99-105.

Henschke, C. I., Yankelevitz, D. F., Naidich, D. P., McCauley, D. I., McGuinness, G., Libby, D. M., Smith, J. P., Pasmantier, M. W. & Miettinen, O. S. (2004) CT screening for lung cancer: suspiciousness of nodules according to size on baseline scans. *Radiology*, 231(1), 164-168.

Herlyn, M., Sears, H. F., Steplewski, Z. & Koprowski, H. (1982) Monoclonal antibody detection of a circulating tumor-associated antigen. I. Presence of antigen in sera of patients with colorectal, gastric, and pancreatic carcinoma. *Journal of clinical immunology*, 2(2), 135-140.

Higgins, G., Roper, K. M., Watson, I. J., Blackhall, F. H., Rom, W. N., Pass, H. I., Ainscough, J. F. & Coverley, D. (2012) Variant Ciz1 is a circulating biomarker for early-stage lung cancer. *Proceedings of the National Academy of Sciences*, 109(45), E3128-E3135.

Hill, V. K., Dunwell, T. L., Catchpoole, D., Krex, D., Brini, A. T., Griffiths, M., Craddock, C., Maher, E. R. & Latif, F. (2011) Frequent epigenetic inactivation of KIBRA,

an upstream member of the Salvador/Warts/Hippo (SWH) tumor suppressor network, is associated with specific genetic event in B-cell acute lymphocytic leukemia. *Epigenetics*, 6(3), 326-32.

Hinze, L., Pfirrmann, M., Karim, S., Degar, J., McGuckin, C., Vinjamur, D., Joshua, S., Stevenson, K. E., Neuberg, D. S., Bauer, D. E., Wagner, F., Stegmaier, K. & Gutierrez, A. (2018) Synthetic Lethality of Wnt Pathway Activation and Asparaginase in Drug-Resistant Acute Leukemias. *Blood*, 132, 891.

Hirokawa, N. & Tanaka, Y. (2015) Kinesin superfamily proteins (KIFs): Various functions and their relevance for important phenomena in life and diseases. *Experimental cell research*, 334(1), 16-25.

Ho, K.-H., Chen, P.-H., Shih, C.-M., Lee, Y.-T., Cheng, C.-H., Liu, A.-J., Lee, C.-C. & Chen, K.-C. (2022) miR-4286 is Involved in Connections Between IGF-1 and TGF-β Signaling for the Mesenchymal Transition and Invasion by Glioblastomas. *Cellular and Molecular Neurobiology*, 42(3), 791-806.

Hoang, N. M. & Rui, L. (2020) DNA methyltransferases in hematological malignancies. *J Genet Genomics*, 47(7), 361-372.

Höffken, V., Hermann, A., Pavenstädt, H. & Kremerskothen, J. (2021) Wwc proteins: Important regulators of hippo signaling in cancer. *Cancers*, 13(2), 306.

Hollingsworth, R. E. & Jansen, K. (2019) Turning the corner on therapeutic cancer vaccines. *npj Vaccines*, 4(1), 7.

Holmfeldt, L., Wei, L., Diaz-Flores, E., Walsh, M., Zhang, J., Ding, L., Payne-Turner, D., Churchman, M., Andersson, A. & Chen, S.-C. (2013) The genomic landscape of hypodiploid acute lymphoblastic leukemia. *Nature genetics*, 45(3), 242-252.

Honda, O., Tsubamoto, M., Inoue, A., Johkoh, T., Tomiyama, N., Hamada, S., Mihara, N., Sumikawa, H., Natsag, J. & Nakamura, H. (2007) Pulmonary cavitary nodules on computed tomography: differentiation of malignancy and benignancy. *Journal of computer assisted tomography*, 31(6), 943-949.

Hopp, K., Heyer, C. M., Hommerding, C. J., Henke, S. A., Sundsbak, J. L., Patel, S., Patel, P., Consugar, M. B., Czarnecki, P. G., Gliem, T. J., Torres, V. E., Rossetti, S. & Harris, P. C. (2011) B9D1 is revealed as a novel Meckel syndrome (MKS) gene by targeted exon-enriched next-generation sequencing and deletion analysis. *Hum Mol Genet*, 20(13), 2524-34.

Horeweg, N., van Rosmalen, J., Heuvelmans, M. A., van der Aalst, C. M., Vliegenthart, R., Scholten, E. T., ten Haaf, K., Nackaerts, K., Lammers, J.-W. J. & Weenink, C. (2014) Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. *The Lancet Oncology*, 15(12), 1332-1341.

Hoseinkhani, Z., Rastegari-Pouyani, M., Oubari, F., Mozafari, H., Rahimzadeh, A. B., Maleki, A., Amini, S. & Mansouri, K. (2019) Contribution and prognostic value of

TSGA10 gene expression in patients with acute myeloid leukemia (AML). *Pathology-Research and Practice*, 215(3), 506-511.

Hounguè, S., Aguida, B. & Capo-chichi, C. D. (2019) Dynamism of Ribosomal Protein RPL28 as Biomarker for Monitoring Epigenetic-initiated-breast-cancer Therapy. *American Journal of Biomedical Research*, 7(1), 21-26.

Hsiao, Y.-J., Su, K.-Y., Hsu, Y.-C., Chang, G.-C., Chen, J.-S., Chen, H.-Y., Hong, Q.-S., Hsu, S.-C., Kang, P.-H. & Hsu, C.-Y. (2016) SPANXA suppresses EMT by inhibiting c-JUN/SNAI2 signaling in lung adenocarcinoma. *Oncotarget*, 7(28), 44417.

Hsu, W.-H., Huang, C.-S., Hsu, H.-S., Huang, W.-J., Lee, H.-C., Huang, B.-S. & Huang, M.-H. (2007) Preoperative serum carcinoembryonic antigen level is a prognostic factor in women with early non-small-cell lung cancer. *The Annals of thoracic surgery*, 83(2), 419-424.

Hu, C., Yu, M., Li, C., Wang, Y., Li, X., Ulrich, B., Su, R., Dong, L., Weng, H. & Huang, H. (2020a) miR-550-1 functions as a tumor suppressor in acute myeloid leukemia via the hippo signaling pathway. *International journal of biological sciences*, 16(15), 2853.

Hu, C., Zhou, H., Liu, Y., Huang, J., Liu, W., Zhang, Q., Tang, Q., Sheng, F., Li, G. & Zhang, R. (2019) ROCK1 promotes migration and invasion of non-small-cell lung cancer cells through the PTEN/PI3K/FAK pathway. *International journal of oncology*, 55(4), 833-844.

Hu, K., Liu, X., Li, Y., Li, Q., Xu, Y., Zeng, W., Zhong, G. & Yu, C. (2020b) Exosomes Mediated Transfer of Circ_UBE2D2 Enhances the Resistance of Breast Cancer to Tamoxifen by Binding to MiR-200a-3p. *Med Sci Monit*, 26, e922253.

Hu, L., Bai, Z., Ma, X., Bai, N. & Zhang, Z. (2018) MRFAP1 plays a protective role in neddylation inhibitor MLN4924-mediated gastric cancer cell death. *Eur Rev Med Pharmacol Sci*, 22(23), 8273-8280.

Hu, Y., Mu, H. & Deng, Z. (2021) The transcription factor TEAD4 enhances lung adenocarcinoma progression through enhancing PKM2 mediated glycolysis. *Cell Biology International*, 45(10), 2063-2073.

Hu, Y., Xing, J., Chen, L., Zheng, Y. & Zhou, Z. (2015) RGS22 inhibits pancreatic adenocarcinoma cell migration through the G12/13 α subunit/F-actin pathway. *Oncology Reports*, 34(5), 2507-2514.

Hua, Y., Dai, C., He, Q., Cai, X. & Li, M. (2022) Autoantibody panel on small extracellular vesicles for the early detection of lung cancer. *Clinical Immunology*, 245, 109175.

Huang, H., Wang, W., Lin, T., Zhang, Q., Zhao, X., Lian, H. & Guo, H. (2016) Comparison of the complications of traditional 12 cores transrectal prostate biopsy with image fusion guided transperineal prostate biopsy. *BMC urology*, 16(1), 68.

Huang, J.-M., Nagatomo, I., Suzuki, E., Mizuno, T., Kumagai, T., Berezov, A., Zhang, H., Karlan, B., Greene, M. I. & Wang, Q. (2013) YAP modifies cancer cell sensitivity to

EGFR and survivin inhibitors and is negatively regulated by the non-receptor type protein tyrosine phosphatase 14. *Oncogene*, 32(17), 2220-2229.

Huang, L. Z., Li, Y. J., Xie, X. F., Zhang, J. J., Cheng, C. Y., Yamashiro, K., Chen, L. J., Ma, X. Y., Cheung, C. M., Wang, Y. S., Zhang, C. F., Bai, Y. J., Hou, J., Chen, X. L., Qi, Y., Li, S. S., Sun, Y. Y., Mei, J. P., Cheng, Y., Yu, W. Z., Hu, X. B., Zhuang, F. F., Fan, L., Lu, Y., Sun, X. H., Zhu, X. J., Shen, D. F., Chan, C. C., Zhao, M. W., Yoshimura, N., Pang, C. P., Wong, T. Y., Khor, C. C., Zhang, K., Zhou, P. & Li, X. X. (2015) Whole-exome sequencing implicates UBE3D in age-related macular degeneration in East Asian populations. *Nat Commun*, 6, 6687.

Huang, S., Huang, Z., Chen, P. & Feng, C. (2020) Aberrant Chloride Intracellular Channel 4 Expression Is Associated With Adverse Outcome in Cytogenetically Normal Acute Myeloid Leukemia. *Front Oncol*, 10, 1648.

Huang, X., Ding, L., Bennewith, K. L., Tong, R. T., Welford, S. M., Ang, K. K., Story, M., Le, Q.-T. & Giaccia, A. J. (2009) Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation. *Molecular cell*, 35(6), 856-867.

Huang, Z. & Liu, F. (2014) Diagnostic value of serum carbohydrate antigen 19-9 in pancreatic cancer: a meta-analysis. *Tumor Biology*, 35(8), 7459-7465.

Huo, Y., Guo, Z., Gao, X., Liu, Z., Zhang, R. & Qin, X. (2020) Case study of an autoantibody panel for early detection of lung cancer and ground-glass nodules. *Journal of cancer research and clinical oncology*, 146(12), 3349-3357.

Hupalowska, A., Pyrzynska, B. & Miaczynska, M. (2012) APPL1 regulates basal NF-κB activity by stabilizing NIK. *Journal of cell science*, 125(17), 4090-4102.

Iacobucci, I. & Mullighan, C. G. (2017) Genetic basis of acute lymphoblastic leukemia. *Journal of Clinical Oncology*, 35(9), 975.

Iacobucci, I., Witkowski, M. T. & Mullighan, C. G. (2023) Single-cell analysis of acute lymphoblastic and lineage-ambiguous leukemia: approaches and molecular insights. *Blood, The Journal of the American Society of Hematology*, 141(4), 356-368.

Infante, M., Sestini, S., Galeone, C., Marchianò, A., Lutman, F. R., Angeli, E., Calareso, G., Pelosi, G., Sozzi, G. & Silva, M. (2017) Lung cancer screening with low-dose spiral computed tomography: evidence from a pooled analysis of two Italian randomized trials. *European journal of cancer prevention: the official journal of the European Cancer Prevention Organisation (ECP)*, 26(4), 324.

Isgrò, M. A., Bottoni, P. & Scatena, R. (2015) Neuron-specific enolase as a biomarker: biochemical and clinical aspects, *Advances in Cancer Biomarkers*Springer, 125-143.

Jabbour, E. J., Faderl, S. & Kantarjian, H. M. (2005) Adult acute lymphoblastic leukemia, *Mayo Clinic Proceedings*. Elsevier.

Jacobson, S., Tedder, M. & Eggert, J. (2016) Adult Acute Lymphoblastic Leukemia: A Genetic Overview and Application to Clinical Practice. *Clinical journal of oncology nursing*, 20(6).

Jäger, E., Maeurer, M., Höhn, H., Karbach, J., Jäger, D., Zidianakis, Z., Bakhshandeh‐Bath, A., Orth, J., Neukirch, C. & Necker, A. (2000) Clonal expansion of Melan A‐specific cytotoxic T lymphocytes in a melanoma patient responding to continued immunization with melanoma‐associated peptides. *International journal of cancer*, 86(4), 538-547.

Jariwala, N., Mendoza, R. G., Garcia, D., Lai, Z., Subler, M. A., Windle, J. J., Mukhopadhyay, N. D., Fisher, P. B., Chen, Y. & Sarkar, D. (2019) Posttranscriptional Inhibition of Protein Tyrosine Phosphatase Nonreceptor Type 23 by Staphylococcal Nuclease and Tudor Domain Containing 1: Implications for Hepatocellular Carcinoma. *Hepatology Communications*, 3(9), 1258-1270.

Jemal, A., Center, M. M., DeSantis, C. & Ward, E. M. (2010) Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiology and Prevention Biomarkers*, 19(8), 1893-1907.

Jeong, H., Choi, B. H., Park, J., Jung, J.-H., Shin, H., Kang, K.-W., Quan, Y. H., Yu, J., Park, J.-H., Park, Y., Choi, Y., Kim, H. K. & Hong, S. (2021) GCC2 as a New Early Diagnostic Biomarker for Non-Small Cell Lung Cancer. *Cancers*, 13(21), 5482.

Jett, J. R. & Midthun, D. E. (2008) Commentary: CT screening for lung cancer--caveat emptor. *The Oncologist*, 13(4), 439.

Jia, D., Liu, Z., Deng, N., Tan, T. Z., Huang, R. Y.-J., Taylor-Harding, B., Cheon, D.-J., Lawrenson, K., Wiedemeyer, W. R. & Walts, A. E. (2016) A COL11A1-correlated pan-cancer gene signature of activated fibroblasts for the prioritization of therapeutic targets. *Cancer letters*, 382(2), 203-214.

Jiang, D., Zhang, X., Liu, M., Wang, Y., Wang, T., Pei, L., Wang, P., Ye, H., Shi, J., Song, C., Wang, K., Wang, X., Dai, L. & Zhang, J. (2021) Discovering Panel of Autoantibodies for Early Detection of Lung Cancer Based on Focused Protein Array. *Frontiers in immunology*, 12, 658922.

Jiang, F., Todd, N. W., Qiu, Q., Liu, Z., Katz, R. L. & Stass, S. A. (2009) Combined genetic analysis of sputum and computed tomography for noninvasive diagnosis of non-small-cell lung cancer. *Lung cancer*, 66(1), 58-63.

Jiang, S., Jiang, T., Huang, H., Chen, X., Li, L., Wang, Z., Fei, J., Liu, C., Liu, Z. & Cheng, Y. (2022a) CHMFL-BMX-078, a BMX inhibitor, overcomes the resistance of melanoma to vemurafenib via inhibiting AKT pathway. *Chemico-Biological Interactions*, 351, 109747.

Jiang, T., Shi, T., Zhang, H., Hu, J., Song, Y., Wei, J., Ren, S. & Zhou, C. (2019) Tumor neoantigens: from basic research to clinical applications. *Journal of hematology & oncology*, 12(1), 93-93.

Jiang, Y. F., Wei, S. N., Geng, N., Qin, W. W., He, X., Wang, X. H., Qi, Y. P., Song, S. & Wang, P. (2022b) Evaluation of circulating small extracellular vesicle-derived

miRNAs as diagnostic biomarkers for differentiating between different pathological types of early lung cancer. *Sci Rep*, 12(1), 17201.

Jiao, S., Wang, H., Shi, Z., Dong, A., Zhang, W., Song, X., He, F., Wang, Y., Zhang, Z. & Wang, W. (2014) A peptide mimicking VGLL4 function acts as a YAP antagonist therapy against gastric cancer. *Cancer cell*, 25(2), 166-180.

Jiménez-Morales, S., Aranda-Uribe, I. S., Pérez-Amado, C. J., Ramírez-Bello, J. & Hidalgo-Miranda, A. (2021) Mechanisms of immunosuppressive tumor evasion: focus on acute lymphoblastic leukemia. *Frontiers in immunology*, 12, 737340.

Jimenez-Velasco, A., Roman-Gomez, J., Agirre, X., Barrios, M., Navarro, G., Vázquez, I., Prosper, F., Torres, A. & Heiniger, A. (2005) Downregulation of the large tumor suppressor 2 (LATS2/KPM) gene is associated with poor prognosis in acute lymphoblastic leukemia. *Leukemia*, 19(12), 2347-2350.

Jin, Q., Cheng, M., Xia, X., Han, Y., Zhang, J., Cao, P. & Zhou, G. (2021) Down-regulation of MYH10 driven by chromosome 17p13.1 deletion promotes hepatocellular carcinoma metastasis through activation of the EGFR pathway. *J Cell Mol Med*, 25(24), 11142-11156.

Jin, Y., Zhou, J., Xu, F., Jin, B., Cui, L., Wang, Y., Du, X., Li, J., Li, P., Ren, R. & Pan, J. (2016) Targeting methyltransferase PRMT5 eliminates leukemia stem cells in chronic myelogenous leukemia. *The Journal of clinical investigation*, 126(10), 3961-3980.

John, T., Starmans, M. H., Chen, Y. T., Russell, P. A., Barnett, S. A., White, S. C., Mitchell, P. L., Walkiewicz, M., Azad, A., Lambin, P., Tsao, M. S., Deb, S., Altorki, N., Wright, G., Knight, S., Boutros, P. C. & Cebon, J. S. (2013) The role of Cancer-Testis antigens as predictive and prognostic markers in non-small cell lung cancer. *PLoS One*, 8(7), e67876.

Johnson, B. E. (1998) Second lung cancers in patients after treatment for an initial lung cancer. *J Natl Cancer Inst*, 90(18), 1335-45.

Jordaens, S., Cooksey, L., Freire Boullosa, L., Van Tendeloo, V., Smits, E., Mills, K. I., Orchard, K. H. & Guinn, B. A. (2020) New targets for therapy: antigen identification in adults with B-cell acute lymphoblastic leukaemia. *Cancer Immunol Immunother*, 69(5), 867-877.

Jordan, M. (2002) Mechanism of action of antitumor drugs that interact with microtubules and tubulin. *Current Medicinal Chemistry-Anti-Cancer Agents*, 2(1), 1-17.

Joseph, S., Harrington, R., Walter, D., Goldberg, J. D., Li, X., Beck, A., Litton, T., Hirsch, N., Blasberg, J., Slomiany, M., Rom, W., Pass, H. & Donington, J. (2012) Plasma osteopontin velocity differentiates lung cancers from controls in a CT screening population. *Cancer biomarkers : section A of Disease markers*, 12(4), 177-184.

Juarez, J., Dela Pena, A., Baraz, R., Hewson, J., Khoo, M., Cisterne, A., Fricker, S., Fujii, N., Bradstock, K. & Bendall, L. (2007) CXCR4 antagonists mobilize childhood acute lymphoblastic leukemia cells into the peripheral blood and inhibit engraftment. *Leukemia*, 21(6), 1249-1257.

Julian, L. & Olson, M. F. (2014) Rho-associated coiled-coil containing kinases (ROCK) structure, regulation, and functions. *Small GTPases*, 5(2), e29846.

Jung, C. H., Kim, E. M., Song, J. Y., Park, J. K. & Um, H. D. (2019) Mitochondrial superoxide dismutase 2 mediates γ-irradiation-induced cancer cell invasion. *Exp Mol Med*, 51(2), 1-10.

Jung, Y. J., Katilius, E., Ostroff, R. M., Kim, Y., Seok, M., Lee, S., Jang, S., Kim, W. S. & Choi, C.-M. (2017) Development of a protein biomarker panel to detect non–small-cell lung cancer in Korea. *Clinical Lung Cancer*, 18(2), e99-e107.

Kaelin, W. G. (2005) The concept of synthetic lethality in the context of anticancer therapy. *Nature reviews cancer*, 5(9), 689-698.

Kagohashi, K., Satoh, H., Ishikawa, H., Ohtsuka, M. & Sekizawa, K. (2008) A re-evaluation of squamous cell carcinoma antigen (SCC) as a serum marker for non-small cell lung cancer. *Medical Oncology*, 25(2), 187-189.

Kakinuma, R., Muramatsu, Y., Kusumoto, M., Tsuchida, T., Tsuta, K., Maeshima, A. M., Asamura, H. & Moriyama, N. (2015) Solitary pure ground-glass nodules 5 mm or smaller: frequency of growth. *Radiology*, 276(3), 873-882.

Kakinuma, R., Noguchi, M., Ashizawa, K., Kuriyama, K., Maeshima, A. M., Koizumi, N., Kondo, T., Matsuguma, H., Nitta, N., Ohmatsu, H., Okami, J., Suehisa, H., Yamaji, T., Kodama, K., Mori, K., Yamada, K., Matsuno, Y., Murayama, S. & Murata, K. (2016) Natural History of Pulmonary Subsolid Nodules: A Prospective Multicenter Study. *J Thorac Oncol*, 11(7), 1012-28.

Kalanjeri, S. & Gildea, T. R. (2016) Electromagnetic navigational bronchoscopy for peripheral pulmonary nodules. *Thoracic Surgery Clinics*, 26(2), 203-213.

Kalejs, M., Ivanov, A., Plakhins, G., Cragg, M. S., Emzinsh, D., Illidge, T. M. & Erenpreisa, J. (2006) Upregulation of meiosis-specific genes in lymphoma cell lines following genotoxic insult and induction of mitotic catastrophe. *BMC cancer*, 6(1), 1-10.

Kammer, M. N. & Massion, P. P. (2020) Noninvasive biomarkers for lung cancer diagnosis, where do we stand? *Journal of Thoracic Disease*, 12(6), 3317-3330.

Kanaseki, T. & Torigoe, T. (2019) Proteogenomics: Advances in cancer antigen research. *Immunological medicine*, 42(2), 65-70.

Kang, B. H., Plescia, J., Dohi, T., Rosa, J., Doxsey, S. J. & Altieri, D. C. (2007) Regulation of tumor cell mitochondrial homeostasis by an organelle-specific Hsp90 chaperone network. *Cell*, 131(2), 257-270.

Kang, S., Wang, L., Xu, L., Wang, R., Kang, Q., Gao, X. & Yu, L. (2022) Decitabine enhances targeting of AML cells by NY-ESO-1-specific TCR-T cells and promotes the maintenance of effector function and the memory phenotype. *Oncogene*, 41(42), 4696-4708.

Karaca Atabay, E., Mecca, C., Wang, Q., Ambrogio, C., Mota, I., Prokoph, N., Mura, G., Martinengo, C., Patrucco, E., Leonardi, G., Hossa, J., Pich, A., Mologni, L., Gambacorti-Passerini, C., Brugières, L., Geoerger, B., Turner, S. D., Voena, C., Cheong, T.-C. & Chiarle, R. (2022) Tyrosine phosphatases regulate resistance to ALK inhibitors in ALK+ anaplastic large cell lymphoma. *Blood*, 139(5), 717-731.

Karimi, S., Mohammadi, F., Porabdollah, M., Mohajerani, S. A., Khodadad, K. & Nadji, S. A. (2012) Characterization of melanoma-associated antigen-a genes family differential expression in non-small-cell lung cancers. *Clinical lung cancer*, 13(3), 214-219.

Karlsson, S., Lindqvist, A., Fransson, M., Paul-Wetterberg, G., Nilsson, B., Essand, M., Nilsson, K., Frisk, P., Jernberg-Wiklund, H. & Loskog, S. (2013) Combining CAR T cells and the Bcl-2 family apoptosis inhibitor ABT-737 for treating B-cell malignancy. *Cancer gene therapy*, 20(7), 386-393.

Karol, S. E., Mattano Jr, L. A., Yang, W., Maloney, K. W., Smith, C., Liu, C., Ramsey, L. B., Fernandez, C. A., Chang, T. Y. & Neale, G. (2016) Genetic risk factors for the development of osteonecrosis in children under age 10 treated for acute lymphoblastic leukemia. *Blood, The Journal of the American Society of Hematology*, 127(5), 558-564.

Kassiotis, G. & Stoye, J. P. (2016) Immune responses to endogenous retroelements: taking the bad with the good. *Nature Reviews Immunology*, 16(4), 207-219.

Katase, N., Nagano, K. & Fujita, S. (2020) DKK3 expression and function in head and neck squamous cell carcinoma and other cancers. *Journal of Oral Biosciences*, 62(1), 9-15.

Kauczor, H.-U., Bonomo, L., Gaga, M., Nackaerts, K., Peled, N., Prokop, M., Remy-Jardin, M., Von Stackelberg, O., Sculier, J.-P. & Radiology, E. S. o. (2015) ESR/ERS white paper on lung cancer screening. *European radiology*, 25(9), 2519-2531.

Kaur, G., Nijhawan, R., Gupta, N., Singh, N. & Rajwanshi, A. (2017) Pleural fluid cytology samples in cases of suspected lung cancer: An experience from a tertiary care centre. *Diagnostic cytopathology*, 45(3), 195-201.

Kavran, J. M. & Leahy, D. J. (2014) Coupling antibody to cyanogen bromide-activated sepharose. *Methods Enzymol*, 541, 27-34.

Kawai, T., Suzuki, M., Kase, K. & Ozeki, Y. (1993) Expression of carbohydrate antigens in human pulmonary adenocarcinoma. *Cancer*, 72(5), 1581-1587.

Keeney, S., Giroux, C. N. & Kleckner, N. (1997) Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell*, 88(3), 375-384.

Kerr, D. L., Haderk, F. & Bivona, T. G. (2021) Allosteric SHP2 inhibitors in cancer: Targeting the intersection of RAS, resistance, and the immune microenvironment. *Current opinion in chemical biology*, 62, 1-12.

Kerr, K. F., Brown, M. D., Marsh, T. L. & Janes, H. (2019) Assessing the Clinical Impact of Risk Models for Opting Out of Treatment. *Med Decis Making*, 39(2), 86-90.

Khan, G. N., Orchard, K. & Guinn, B. A. (2019) Antigenic Targets for the Immunotherapy of Acute Myeloid Leukaemia. *J Clin Med*, 8(2).

Khan, N. I., Bradstock, K. F. & Bendall, L. J. (2007) Activation of Wnt/β‐catenin pathway mediates growth and survival in B‐cell progenitor acute lymphoblastic leukaemia. *British journal of haematology*, 138(3), 338-348.

Khan, S., Ricciardelli, C. & Yool, A. J. (2021) Targeting aquaporins in novel therapies for male and female breast and reproductive cancers. *Cells*, 10(2), 215.

Kharabi Masouleh, B., Geng, H., Hurtz, C., Chan, L. N., Logan, A. C., Chang, M. S., Huang, C., Swaminathan, S., Sun, H. & Paietta, E. (2014) Mechanistic rationale for targeting the unfolded protein response in pre-B acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences*, 111(21), E2219-E2228.

Kikuchi, E., Yamazaki, K., Nakayama, E., Sato, S., Uenaka, A., Yamada, N., Oizumi, S., Dosaka-Akita, H. & Nishimura, M. (2008) Prolonged survival of patients with lung adenocarcinoma expressing XAGE-1b and HLA class I antigens. *Cancer Immun*, 8, 13.

Kim, E., Hurtz, C., Koehrer, S., Wang, Z., Balasubramanian, S., Chang, B. Y., Müschen, M., Davis, R. E. & Burger, J. A. (2017) Ibrutinib inhibits pre-BCR+ B-cell acute lymphoblastic leukemia progression by targeting BTK and BLK. *Blood, The Journal of the American Society of Hematology*, 129(9), 1155-1165.

Kim, J., Kim, Y., Choi, H., Kwon, A., Jekarl, D. W., Lee, S., Jang, W., Chae, H., Kim, J. R., Kim, J. M. & Kim, M. (2018) Ubiquitin C decrement plays a pivotal role in replicative senescence of bone marrow mesenchymal stromal cells. *Cell Death Dis*, 9(2), 139.

Kim, M., Yim, S. H., Cho, N. S., Kang, S. H., Ko, D. H., Oh, B., Kim, T. Y., Min, H. J., She, C. J., Kang, H. J., Shin, H. Y., Ahn, H. S., Yoon, S. S., Kim, B. K., Shin, H. R., Han, K. S., Cho, H. I. & Lee, D. S. (2009a) Homozygous deletion of CDKN2A (p16, p14) and CDKN2B (p15) genes is a poor prognostic factor in adult but not in childhood B-lineage acute lymphoblastic leukemia: a comparative deletion and hypermethylation study. *Cancer Genet Cytogenet*, 195(1), 59-65.

Kim, Y., Park, H. & Jeoung, D. (2009b) CAGE, a cancer/testis antigen, induces c-FLIP L and snail to enhance cell motility and increase resistance to an anti-cancer drug. *Biotechnology letters*, 31, 945-952.

Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-Fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. & Sugano, S. (2006) Diversification of transcriptional modulation: large-scale identification and characterization of putative alternative promoters of human genes. *Genome Res*, 16(1), 55-65.

Kohlmann, A., Kipps, T. J., Rassenti, L. Z., Downing, J. R., Shurtleff, S. A., Mills, K. I., Gilkes, A. F., Hofmann, W.-K., Basso, G., Dell'Orto, M. C., Foà, R., Chiaretti, S., De

Vos, J., Rauhut, S., Papenhausen, P. R., Hernández, J. M., Lumbreras, E., Yeoh, A. E., Koay, E. S., Li, R., Liu, W.-m., Williams, P. M., Wieczorek, L. & Haferlach, T. (2008) An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics: the Microarray Innovations in LEukemia study prephase. *British Journal of Haematology*, 142(5), 802-807.

Köhrer, S., Havranek, O., Seyfried, F., Hurtz, C., Coffey, G. P., Kim, E., Ten Hacken, E., Jäger, U., Vanura, K., O'Brien, S., Thomas, D. A., Kantarjian, H., Ghosh, D., Wang, Z., Zhang, M., Ma, W., Jumaa, H., Debatin, K. M., Müschen, M., Meyer, L. H., Davis, R. E. & Burger, J. A. (2016) Pre-BCR signaling in precursor B-cell acute lymphoblastic leukemia regulates PI3K/AKT, FOXO1 and MYC, and can be targeted by SYK inhibition. *Leukemia*, 30(6), 1246-1254.

Kondo, K., Takahashi, Y., Hirose, Y., Nagao, T., Tsuyuguchi, M., Hashimoto, M., Ochiai, A., Monden, Y. & Tangoku, A. (2006) The reduced expression and aberrant methylation of p16INK4a in chromate workers with lung cancer. *Lung Cancer*, 53(3), 295-302.

Konyukh, M., Delorme, R., Chaste, P., Leblond, C., Lemière, N., Nygren, G., Anckarsäter, H., Rastam, M., Ståhlberg, O., Amsellem, F., Gillberg, I. C., Mouren-Simeoni, M. C., Herbrecht, E., Fauchereau, F., Toro, R., Gillberg, C., Leboyer, M. & Bourgeron, T. (2011) Variations of the candidate SEZ6L2 gene on Chromosome 16p11.2 in patients with autism spectrum disorders and in human populations. *PLoS One*, 6(3), e17289.

Koscielny, S. (2010) Why most gene expression signatures of tumors have not been useful in the clinic. *Sci Transl Med*, 2(14), 14ps2.

Kremer, R., Best, L., Savulescu, D., Gavish, M. & Nagler, R. (2010) Pleural fluid analysis of lung cancer vs benign inflammatory disease patients. *British journal of cancer*, 102(7), 1180-1184.

Krendel, M. & Mooseker, M. S. (2005) Myosins: tails (and heads) of functional diversity. *Physiology*, 20(4), 239-251.

Krishnadas, D. K., Bai, F. & Lucas, K. G. (2013) Cancer testis antigen and immunotherapy. *ImmunoTargets and therapy*, 2, 11.

Kühnl, A., Gökbuget, N., Kaiser, M., Schlee, C., Stroux, A., Burmeister, T., Mochmann, L. H., Hoelzer, D., Hofmann, W.-K. & Thiel, E. (2011) Overexpression of LEF1 predicts unfavorable outcome in adult patients with B-precursor acute lymphoblastic leukemia. *Blood, The Journal of the American Society of Hematology*, 118(24), 6362-6367.

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W. & Ma'ayan, A. (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*, 44(W1), W90-7.

Kulpa, J., Wojcik, E., Reinfuss, M. & Kołodziejski, L. (2002) Carcinoembryonic antigen, squamous cell carcinoma antigen, CYFRA 21-1, and neuron-specific enolase in squamous cell lung cancer patients. *Clinical chemistry*, 48(11), 1931-1937.

Kupert, E., Anderson, M., Liu, Y., Succop, P., Levin, L., Wang, J., Wikenheiser-brokamp, K., Chen, P., Pinney, S. M. & Macdonald, T. (2011a) Plasma secretory phospholipase A2-IIa as a potential biomarker for lung cancer in patients with solitary pulmonary nodules. *Bmc Cancer*, 11(1), 513.

Kupert, E., Anderson, M., Liu, Y., Succop, P., Levin, L., Wang, J., Wikenheiser-brokamp, K., Chen, P., Pinney, S. M. & Macdonald, T. (2011b) Plasma secretory phospholipase A2-IIa as a potential biomarker for lung cancer in patients with solitary pulmonary nodules. *Bmc Cancer*, 11(1), 1-10.

Kwon, W.-A., Joung, J. Y., Lim, J., Oh, C.-M., Jung, K.-W., Kim, S. H., Seo, H. K., Park, W. S., Chung, J. & Lee, K. H. (2018) Risk of second primary Cancer among bladder Cancer patients: a population-based cohort study in Korea. *BMC cancer*, 18(1), 617.

Labriet, A., Lévesque, É., Cecchin, E., De Mattia, E., Villeneuve, L., Rouleau, M., Jonker, D., Couture, F., Simonyan, D., Allain, E. P., Buonadonna, A., D'Andrea, M., Toffoli, G. & Guillemette, C. (2019) Germline variability and tumor expression level of ribosomal protein gene RPL28 are associated with survival of metastatic colorectal cancer patients. *Sci Rep*, 9(1), 13008.

Lai, T.-C., Fang, C.-Y., Jan, Y.-H., Hsieh, H.-L., Yang, Y.-F., Liu, C.-Y., Chang, P. M.-H. & Hsiao, M. (2020) Kinase shRNA screening reveals that TAOK3 enhances microtubule-targeted drug resistance of breast cancer cells via the NF-κB signaling pathway. *Cell Communication and Signaling*, 18(1), 1-14.

Lai, Y., Wang, Y., Wu, Y., Wu, M., Xing, S., Xie, Y., Chen, S., Li, X., Zhang, A., He, Y., Li, H., Dai, S., Wang, J., Lin, S., Bai, Y., Du, H. & Liu, W. (2022) Identification and Validation of Serum CST1 as a Diagnostic Marker for Differentiating Early-Stage Non-Small Cell Lung Cancer from Pulmonary Benign Nodules. *Cancer Control*, 29, 10732748221104661.

Lal, A., Navarro, F., Maher, C. A., Maliszewski, L. E., Yan, N., O'Day, E., Chowdhury, D., Dykxhoorn, D. M., Tsai, P. & Hofmann, O. (2009) miR-24 Inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to "seedless" 3′ UTR microRNA recognition elements. *Molecular cell*, 35(5), 610-625.

Lamers, C. H., Sleijfer, S., Van Steenbergen, S., Van Elzakker, P., Van Krimpen, B., Groot, C., Vulto, A., Den Bakker, M., Oosterwijk, E. & Debets, R. (2013) Treatment of metastatic renal cell carcinoma with CAIX CAR-engineered T cells: clinical evaluation and management of on-target toxicity. *Molecular therapy*, 21(4), 904-912.

Lamprecht, B., Porsch, P., Wegleitner, B., Strasser, G., Kaiser, B. & Studnicka, M. (2012) Electromagnetic navigation bronchoscopy (ENB): Increasing diagnostic yield. *Respiratory medicine*, 106(5), 710-715.

Larance, M., Kirkwood, K. J., Xirodimas, D. P., Lundberg, E., Uhlen, M. & Lamond, A. I. (2012) Characterization of MRFAP1 turnover and interactions downstream of the NEDD8 pathway. *Molecular & Cellular Proteomics*, 11(3).

Lastwika, K. J., Kargl, J., Zhang, Y., Zhu, X., Lo, E., Shelley, D., Ladd, J. J., Wu, W., Kinahan, P. & Pipavath, S. N. (2019) Tumor-derived autoantibodies identify malignant pulmonary nodules. *American Journal of Respiratory and Critical Care Medicine*, 199(10), 1257-1266.

Laupitz, R., Hecht, S., Amslinger, S., Zepeck, F., Kaiser, J., Richter, G., Schramek, N., Steinbacher, S., Huber, R. & Arigoni, D. (2004) Biochemical characterization of Bacillus subtilis type II isopentenyl diphosphate isomerase, and phylogenetic distribution of isoprenoid biosynthesis pathways. *European journal of biochemistry*, 271(13), 2658-2669.

Leclercq, M., Diallo, A. B. & Blanchette, M. (2017) Prediction of human miRNA target genes using computationally reconstructed ancestral mammalian sequences. *Nucleic acids research*, 45(2), 556-566.

Lee, E. J., Jo, M., Rho, S. B., Park, K., Yoo, Y. N., Park, J., Chae, M., Zhang, W. & Lee, J. H. (2009) Dkk3, downregulated in cervical cancer, functions as a negative regulator of β‐catenin. *International journal of cancer*, 124(2), 287-297.

Lee, E. K., Song, K. A., Chae, J. H., Kim, K. M., Kim, S. H. & Kang, M. S. (2015) GAGE12 mediates human gastric carcinoma growth and metastasis. *International journal of cancer*, 136(10), 2284-2292.

Lee, J. Y., Park, S., Han, A.-R., Hwang, H.-S. & Kim, H.-J. (2022) Therapeutic potential of FLT4-targeting peptide in acute myeloid leukemia.

Lee, Y., Yu, Y., Gunawardena, H., Xie, L. & Chen, X. (2012) BCLAF1 is a radiation-induced H2AX-interacting partner involved in γH2AX-mediated regulation of apoptosis and DNA repair. *Cell death & disease*, 3(7), e359-e359.

Lennard, A., Gaston, K. & Fried, M. (1994) The Surf-1 and Surf-2 genes and their essential bidirectional promoter elements are conserved between mouse and human. *DNA Cell Biol*, 13(11), 1117-26.

Lepisto, A. J., Moser, A. J., Zeh, H., Lee, K., Bartlett, D., McKolanis, J. R., Geller, B. A., Schmotzer, A., Potter, D. P. & Whiteside, T. (2008) A phase I/II study of a MUC1 peptide pulsed autologous dendritic cell vaccine as adjuvant therapy in patients with resected pancreatic and biliary tumors. *Cancer therapy*, 6(B), 955.

Leung, M., Freidin, M. B., Freydina, D. V., Von Crease, C., De Sousa, P., Barbosa, M. T., Nicholson, A. G. & Lim, E. (2020) Blood‐based circulating tumor DNA mutations as a diagnostic and prognostic biomarker for lung cancer. *Cancer*, 126(8), 1804-1809.

Leuraud, K., Schnelzer, M., Tomasek, L., Hunter, N., Timarche, M., Grosche, B., Kreuzer, M. & Laurier, D. (2011) Radon, smoking and lung cancer risk: results of a joint analysis of three European case-control studies among uranium miners. *Radiation research*, 176(3), 375-387.

Li, C., Lv, Y., Shao, C., Chen, C., Zhang, T., Wei, Y., Fan, H., Lv, T., Liu, H. & Song, Y. (2019a) Tumor‐derived exosomal lncRNA GAS5 as a biomarker for early‐stage

non‐small‐cell lung cancer diagnosis. *Journal of cellular physiology*, 234(11), 20721-20727.

Li, F., Aljahdali, I. & Ling, X. (2019b) Cancer therapeutics using survivin BIRC5 as a target: what can we do after over two decades of study? *Journal of Experimental & Clinical Cancer Research*, 38(1), 368.

Li, G., Gao, L., Zhao, J., Liu, D., Li, H. & Hu, M. (2020a) LncRNA ANRIL/miR-7-5p/TCF4 axis contributes to the progression of T cell acute lymphoblastic leukemia. *Cancer cell international*, 20, 335-335.

Li, G., Gao, L., Zhao, J., Liu, D., Li, H. & Hu, M. (2020b) LncRNA ANRIL/miR-7-5p/TCF4 axis contributes to the progression of T cell acute lymphoblastic leukemia. *Cancer Cell International*, 20, 1-12.

Li, G., Zhang, H., Zhang, L., Liu, H., Li, S., Wang, Y. & Deng, X. (2022) Serum Markers CA125, CA153, and CEA along with Inflammatory Cytokines in the Early Detection of Lung Cancer in High-Risk Populations. *Biomed Res Int*, 2022, 1394042.

Li, H., Huang, Z., Gao, M., Huang, N., Luo, Z., Shen, H., Wang, X., Wang, T., Hu, J. & Feng, W. (2016) Inhibition of YAP suppresses CML cell proliferation and enhances efficacy of imatinib in vitro and in vivo. *Journal of Experimental & Clinical Cancer Research*, 35(1), 1-11.

Li, J., Li, J., Hao, H., Lu, F., Wang, J., Ma, M., Jia, B., Zhuo, M., Wang, J., Chi, Y., Zhai, X., Wang, Y., Wu, M., An, T., Zhao, J., Yang, F. & Wang, Z. (2023a) Secreted proteins MDK, WFDC2, and CXCL14 as candidate biomarkers for early diagnosis of lung adenocarcinoma. *BMC Cancer*, 23(1), 110.

Li, L., Wang, Y., Wang, Q., Qu, J., Wei, X., Xu, J., Wang, Y., Suo, F. & Zhang, Y. (2019c) High developmental pluripotency-associated 4 expression promotes cell proliferation and glycolysis, and predicts poor prognosis in non-small-cell lung cancer. *Molecular Medicine Reports*, 20(1), 445-454.

Li, P. & Noegel, A. A. (2015) Inner nuclear envelope protein SUN1 plays a prominent role in mammalian mRNA export. *Nucleic Acids Res*, 43(20), 9874-88.

Li, T., Kung, H. J., Mack, P. C. & Gandara, D. R. (2013a) Genotyping and genomic profiling of non-small-cell lung cancer: Implications for current and future therapies. *Journal of Clinical Oncology*, 31(8), 1039-1049.

Li, X.-F., Ren, P., Shen, W.-Z., Jin, X. & Zhang, J. (2020c) The expression, modulation and use of cancer-testis antigens as potential biomarkers for cancer immunotherapy. *American Journal of Translational Research*, 12(11), 7002.

Li, Y., He, W., Gao, X., Lu, X., Xie, F., Um, S.-W., Kang, M.-W., Yang, H., Shang, Y. & Wang, Z. (2023b) Cullin7 induces docetaxel resistance by regulating the protein level of the antiapoptotic protein Survivin in lung adenocarcinoma cells. *Journal of Thoracic Disease*, 15(9).

Li, Z., Qiu, H. Y., Jiao, Y., Cen, J. N., Fu, C. M., Hu, S. Y., Zhu, M. Q., Wu, D. P. & Qi, X. F. (2013b) Growth and differentiation effects of Homer3 on a leukemia cell line. *Asian Pac J Cancer Prev*, 14(4), 2525-8.

Liang, H., Chen, Q., Coles, A. H., Anderson, S. J., Pihan, G., Bradley, A., Gerstein, R., Jurecic, R. & Jones, S. N. (2003) Wnt5a inhibits B cell proliferation and functions as a tumor suppressor in hematopoietic tissue. *Cancer cell*, 4(5), 349-360.

Liang, K., Liao, L., Liu, Q., Ouyang, Q., Jia, L. & Wu, G. (2021) microRNA-377-3p inhibits osteosarcoma progression by targeting CUL1 and regulating Wnt/β-catenin signaling pathway. *Clinical and Translational Oncology*, 23(11), 2350-2357.

Lilleyman, J., Hann, I., Stevens, R., Eden, O. & Richards, S. (1986) French American British (FAB) morphological classification of childhood lymphoblastic leukaemia and its clinical importance. *Journal of clinical pathology*, 39(9), 998-1002.

Lin, D., Feuer, E., Etzioni, R. & Wax, Y. (1997) Estimating medical costs from incomplete follow-up data. *Biometrics*, 419-434.

Lin, K. H., Rutter, J. C., Xie, A., Killarney, S. T., Vaganay, C., Benaksas, C., Ling, F., Sodaro, G., Meslin, P.-A., Bassil, C. F., Fenouille, N., Hoj, J., Washart, R., Ang, H. X., Cerda-Smith, C., Chaintreuil, P., Jacquel, A., Auberger, P., Forget, A., Itzykson, R., Lu, M., Lin, J., Pierobon, M., Sheng, Z., Li, X., Chilkoti, A., Owzar, K., Rizzieri, D. A., Pardee, T. S., Benajiba, L., Petricoin, E., Puissant, A. & Wood, K. C. (2022) P2RY2-AKT activation is a therapeutically actionable consequence of XPO1 inhibition in acute myeloid leukemia. *Nature Cancer*, 3(7), 837-851.

Lin, Y.-L., Liau, J.-Y., Yu, S.-C., Ou, D.-L., Lin, L.-I., Tseng, L.-H., Chang, Y.-L., Yeh, K.-H. & Cheng, A.-L. (2012) KRAS mutation is a predictor of oxaliplatin sensitivity in colon cancer cells. *PloS one*, 7(11), e50701.

Lin, Y., Leng, Q., Jiang, Z., Guarnera, M. A., Zhou, Y., Chen, X., Wang, H., Zhou, W., Cai, L. & Fang, H. (2017) A classifier integrating plasma biomarkers and radiological characteristics for distinguishing malignant from benign pulmonary nodules. *International journal of cancer*, 141(6), 1240-1248.

Liu-Dumlao, T., Kantarjian, H., Thomas, D. A., O'Brien, S. & Ravandi, F. (2012) Philadelphia-positive acute lymphoblastic leukemia: current treatment options. *Current oncology reports*, 14(5), 387-394.

Liu, F., Zhang, X., Weisberg, E., Chen, S., Hur, W., Wu, H., Zhao, Z., Wang, W., Mao, M. & Cai, C. (2013) Discovery of a selective irreversible BMX inhibitor for prostate cancer. *ACS Chemical Biology*, 8(7), 1423-1428.

Liu, H., Heller-Trulli, D. & Moore, C. L. (2022a) Targeting the mRNA endonuclease CPSF73 inhibits breast cancer cell migration, invasion, and self-renewal. *iScience*, 25(8), 104804.

Liu, L., Li, S. W., Yuan, W., Tang, J. & Sang, Y. (2021a) Downregulation of SUN2 promotes metastasis of colon cancer by activating BDNF/TrkB signalling by interacting with SIRT1. *J Pathol*, 254(5), 531-542.

Liu, S., Zhang, X., Jiang, Q. & Liang, T. (2020) Detection of circulating natural antibodies against CD25, MUC1, and VEGFR1 for early diagnosis of non‐small cell lung cancer. *FEBS Open bio*, 10(7), 1288-1294.

Liu, S. X., Xiao, H. R., Wang, G. B., Chen, X. W., Li, C. G., Mai, H. R., Yuan, X. L., Liu, G. S. & Wen, F. Q. (2018) Preliminary investigation on the abnormal mechanism of CD4+FOXP3+CD25high regulatory T cells in pediatric B-cell acute lymphoblastic leukemia. *Exp Ther Med*, 16(2), 1433-1441.

Liu, Y.-F., Wang, B.-Y., Zhang, W.-N., Huang, J.-Y., Li, B.-S., Zhang, M., Jiang, L., Li, J.-F., Wang, M.-J. & Dai, Y.-J. (2016) Genomic profiling of adult and pediatric B-cell acute lymphoblastic leukemia. *EBioMedicine*, 8, 173-183.

Liu, Y., Cheng, H., Cheng, C., Zheng, F., Zhao, Z., Chen, Q., Zeng, W., Zhang, P., Huang, C., Jiang, W., Liu, X. & Liu, G. (2022b) ZNF191 alters DNA methylation and activates the PI3K-AKT pathway in hepatoma cells via transcriptional regulation of DNMT1. *Cancer Med*, 11(5), 1269-1280.

Liu, Y., Gao, X. & Tian, X. (2019) High expression of long intergenic non-coding RNA LINC00662 contributes to malignant growth of acute myeloid leukemia cells by upregulating ROCK1 via sponging microRNA-340-5p. *European Journal of Pharmacology*, 859, 172535.

Liu, Y., Zhang, J., Du, Z., Huang, J., Cheng, Y., Yi, W., Li, T., Yang, J. & Chen, C. (2023) Comprehensive analysis of PTPN family expression and prognosis in acute myeloid leukemia. *Frontiers in Genetics*, 13, 1087938.

Liu, Y. J., Hounye, A. H., Wang, Z., Liu, X., Yi, J. & Qi, M. (2021b) Identification and validation of three autophagy-related long noncoding RNAs as prognostic signature in cholangiocarcinoma. *Frontiers in Oncology*, 11, 780601.

Lo, W.-K., Biswas, S. K., Brako, L., Shiels, A., Gu, S. & Jiang, J. X. (2014) Aquaporin-0 targets interlocking domains to control the integrity and transparency of the eye lens. *Investigative ophthalmology & visual science*, 55(3), 1202-1212.

Lokhandwala, T., Bittoni, M. A., Dann, R. A., D'Souza, A. O., Johnson, M., Nagy, R. J., Lanman, R. B., Merritt, R. E. & Carbone, D. P. (2017) Costs of diagnostic assessment for lung cancer: a medicare claims analysis. *Clinical lung cancer*, 18(1), e27-e34.

Lossos, I. S., Czerwinski, D. K., Wechser, M. A. & Levy, R. (2003) Optimization of quantitative real-time RT-PCR parameters for the study of lymphoid malignancies. *Leukemia*, 17(4), 789-795.

Loverdos, K., Fotiadis, A., Kontogianni, C., Iliopoulou, M. & Gaga, M. (2019) Lung nodules: A comprehensive review on current approach and management. *Annals of thoracic medicine*, 14(4), 226-238.

Low, H. Z., Reuter, S., Topperwien, M., Dankenbrink, N., Peest, D., Kabalak, G., Stripecke, R., Schmidt, R. E., Matthias, T. & Witte, T. (2013) Association of the LILRA3

deletion with B-NHL and functional characterization of the immunostimulatory molecule. *PLoS One*, 8(12), e81360.

Lowe, F. J., Shen, W., Zu, J., Li, J., Wang, H., Zhang, X. & Zhong, L. (2014) A novel autoantibody test for the detection of pre-neoplastic lung lesions. *Mol Cancer*, 13, 78.

Lozzio, C. B. & Lozzio, B. B. (1975) Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome.

Lu, D., Zhao, Y., Tawatao, R., Cottam, H. B., Sen, M., Leoni, L. M., Kipps, T. J., Corr, M. & Carson, D. A. (2004) Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A*, 101(9), 3118-23.

Lu, J. W., Hsieh, M. S., Hou, H. A., Chen, C. Y., Tien, H. F. & Lin, L. I. (2017) Overexpression of SOX4 correlates with poor prognosis of acute myeloid leukemia and is leukemogenic in zebrafish. *Blood Cancer Journal*, 7(8), e593-e593.

Luan, C., Yang, Z. & Chen, B. (2015) The functional role of microRNA in acute lymphoblastic leukemia: relevance for diagnosis, differential diagnosis, prognosis, and therapy. *OncoTargets and therapy*, 2903-2914.

Lubin, J. H., Moore, L. E., Fraumeni, J. F., Jr. & Cantor, K. P. (2008) Respiratory cancer and inhaled inorganic arsenic in copper smelters workers: a linear relationship with cumulative exposure that increases with concentration. *Environ Health Perspect*, 116(12), 1661-5.

Ludwig, J. A. & Weinstein, J. N. (2005) Biomarkers in cancer staging, prognosis and treatment selection. *Nature Reviews Cancer*, 5(11), 845-856.

Lüftl, M., Schuler, G. & Jungbluth, A. (2004) Melanoma or not? Cancer testis antigens may help. *British Journal of Dermatology*, 151(6), 1213-1218.

Luna-Aguirre, C. M., de la Luz Martinez-Fierro, M., Mar-Aguilar, F., Garza-Veloz, I., Treviño-Alvarado, V., Rojas-Martinez, A., Jaime-Perez, J. C., Malagon-Santiago, G. I., Gutierrez-Aguirre, C. H., Gonzalez-Llano, O., Salazar-Riojas, R., Hidalgo-Miranda, A., Martinez-Rodriguez, H. G., Gomez-Almaguer, D. & Ortiz-Lopez, R. (2015) Circulating microRNA expression profile in B-cell acute lymphoblastic leukemia. *Cancer Biomark*, 15(3), 299-310.

Luo, X., Ji, X., Xie, M., Zhang, T., Wang, Y., Sun, M., Huang, W. & Xia, L. (2022) Advance of SOX transcription factors in hepatocellular carcinoma: From role, tumor immune relevance to targeted therapy. *Cancers*, 14(5), 1165.

Luo, Y., Yu, L., Yu, T., Jiang, F., Cai, X., Zhao, Y., Pan, S. & Luo, C. (2015) The association of DNA methyltransferase 1 gene polymorphisms with susceptibility to childhood acute lymphoblastic leukemia. *Biomedicine & Pharmacotherapy*, 73, 35-39.

Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G. & Haluska, F. G. (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non–

small-cell lung cancer to gefitinib. *New England Journal of Medicine*, 350(21), 2129-2139.

Ma, J., Guarnera, M. A., Zhou, W., Fang, H. & Jiang, F. (2017) A prediction model based on biomarkers and clinical characteristics for detection of lung cancer in pulmonary nodules. *Translational oncology*, 10(1), 40-45.

Ma, Y. C., Tian, P. F., Chen, Z. P., Yue, D. S., Liu, C. C., Li, C. G., Chen, C., Zhang, H., Liu, H. L. & Zhang, Z. F. (2021) Urinary malate dehydrogenase 2 is a new biomarker for early detection of non‐small‐cell lung cancer. *Cancer Science*, 112(6), 2349-2360.

Maamari, D., El-Khoury, H., Saifi, O., Muwakkit, S. A. & Zgheib, N. K. (2020) Implementation of Pharmacogenetics to Individualize Treatment Regimens for Children with Acute Lymphoblastic Leukemia. *Pharmacogenomics and Personalized Medicine*, 13, 295-317.

Mäbert, K., Cojoc, M., Peitzsch, C., Kurth, I., Souchelnytskyi, S. & Dubrovska, A. (2014) Cancer biomarker discovery: current status and future perspectives. *International journal of radiation biology*, 90(8), 659-677.

Machado-Neto, J. A., de Melo Campos, P., Olalla Saad, S. T. & Traina, F. (2014) YAP1 expression in myelodysplastic syndromes and acute leukemias. *Leukemia & Lymphoma*, 55(10), 2413-2415.

MacMahon, H., Naidich, D. P., Goo, J. M., Lee, K. S., Leung, A. N., Mayo, J. R., Mehta, A. C., Ohno, Y., Powell, C. A. & Prokop, M. (2017) Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology*, 284(1), 228-243.

Macuks, R., Baidekalna, I. & Donina, S. (2012) An ovarian cancer malignancy risk index composed of HE4, CA125, ultrasonographic score, and menopausal status: use in differentiation of ovarian cancers and benign lesions. *Tumor Biology*, 33(5), 1811-1817.

Magnani, J. L. (2004) The discovery, biology, and drug development of sialyl Lea and sialyl Lex. *Archives of biochemistry and biophysics*, 426(2), 122-131.

Magnani, J. L., Steplewski, Z., Koprowski, H. & Ginsburg, V. (1983) Identification of the gastrointestinal and pancreatic cancer-associated antigen detected by monoclonal antibody 19-9 in the sera of patients as a mucin. *Cancer research*, 43(11), 5489-5492.

Malard, F. & Mohty, M. (2020) Acute lymphoblastic leukaemia. *The Lancet*, 395(10230), 1146-1162.

Maldonado, S. G., Delorme, S., Hüsing, A., Motsch, E., Kauczor, H.-U., Heussel, C.-P. & Kaaks, R. (2020) Evaluation of prediction models for identifying malignancy in pulmonary nodules detected via low-dose computed tomography. *JAMA Network Open*, 3(2), e1921221-e1921221.

Mali, R. S., Kapur, S. & Kapur, R. (2014) Role of Rho kinases in abnormal and normal hematopoiesis. *Current opinion in hematology*, 21(4), 271.

Malik, M., Jividen, K., Padmakumar, V., Cataisson, C., Li, L., Lee, J., Howard, O. Z. & Yuspa, S. H. (2012) Inducible NOS-induced chloride intracellular channel 4 (CLIC4) nuclear translocation regulates macrophage deactivation. *Proceedings of the National Academy of Sciences*, 109(16), 6130-6135.

Mangino, M., Hwang, S.-J., Spector, T. D., Hunt, S. C., Kimura, M., Fitzpatrick, A. L., Christiansen, L., Petersen, I., Elbers, C. C. & Harris, T. (2012) Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. *Human molecular genetics*, 21(24), 5385-5394.

Manteghi, S., Gingras, M. C., Kharitidi, D., Galarneau, L., Marques, M., Yan, M., Cencic, R., Robert, F., Paquet, M., Witcher, M., Pelletier, J. & Pause, A. (2016) Haploinsufficiency of the ESCRT Component HD-PTP Predisposes to Cancer. *Cell Rep*, 15(9), 1893-900.

Marcus, P. M., Bergstralh, E. J., Zweig, M. H., Harris, A., Offord, K. P. & Fontana, R. S. (2006) Extended lung cancer incidence follow-up in the Mayo Lung Project and overdiagnosis. *Journal of the National Cancer Institute*, 98(11), 748-756.

Marin, J. J., Perez-Silva, L., Macias, R. I., Asensio, M., Peleteiro-Vigil, A., Sanchez-Martin, A., Cives-Losada, C., Sanchon-Sanchez, P., Sanchez De Blas, B. & Herraez, E. (2020) Molecular bases of mechanisms accounting for drug resistance in gastric adenocarcinoma. *Cancers*, 12(8), 2116.

Marinovic, A. C., Mitch, W. E. & Price, S. R. (2000) Tools for evaluating ubiquitin (UbC) gene expression: characterization of the rat UbC promoter and use of an unique 3' mRNA sequence. *Biochem Biophys Res Commun*, 274(2), 537-41.

Martin, V., Agirre, X., Jiménez‐Velasco, A., José‐Eneriz, E. S., Cordeu, L., Gárate, L., Vilas‐Zornoza, A., Castillejo, J. A., Heiniger, A. & Prósper, F. (2008) Methylation status of Wnt signaling pathway genes affects the clinical outcome of Philadelphia‐positive acute lymphoblastic leukemia. *Cancer science*, 99(9), 1865-1868.

Martinez, A., Olarte, I., Mergold, M., Gutierrez, M., Rozen, E., Collazo, J., Amancio-Chassin, O., Ordóñez, R., Montesinos, J. & Mayani, H. (2007) mRNA expression of MAGE-A3 gene in leukemia cells. *Leukemia research*, 31(1), 33-37.

Maruyama, J., Inami, K., Michishita, F., Jiang, X., Iwasa, H., Nakagawa, K., Ishigami-Yuasa, M., Kagechika, H., Miyamura, N. & Hirayama, J. (2018) Novel YAP1 activator, identified by transcription-based functional screen, limits multiple myeloma growth. *Molecular Cancer Research*, 16(2), 197-211.

Matsuda, A., Suzuki, Y., Honda, G., Muramatsu, S., Matsuzaki, O., Nagano, Y., Shimotohno, K., Harada, T., Nishida, E. & Hayashi, H. (2003) Large-scale identification and characterization of human genes that activate NF-κB and MAPK signaling pathways. *Oncogene*, 22(21), 3307-3318.

Matsuoka, S., Kurihara, Y., Yagihashi, K., Niimi, H. & Nakajima, Y. (2005) Peripheral solitary pulmonary nodule: CT findings in patients with pulmonary emphysema. *Radiology*, 235(1), 266-273.

Mayeux, R. (2004) Biomarkers: potential uses and limitations. *NeuroRx*, 1(2), 182-188.

Mazieres, J., You, L., He, B., Xu, Z., Lee, A. Y., Mikami, I., McCormick, F. & Jablons, D. M. (2005) Inhibition of Wnt16 in human acute lymphoblastoid leukemia cells containing the t (1; 19) translocation induces apoptosis. *Oncogene*, 24(34), 5396-5400.

Mazzone, P. J., Sears, C. R., Arenberg, D. A., Gaga, M., Gould, M. K., Massion, P. P., Nair, V. S., Powell, C. A., Silvestri, G. A., Vachani, A., Wiener, R. S. & Oncology, A. T. S. A. o. T. (2017) Evaluating Molecular Biomarkers for the Early Detection of Lung Cancer: When Is a Biomarker Ready for Clinical Use? An Official American Thoracic Society Policy Statement. *American journal of respiratory and critical care medicine*, 196(7), e15-e29.

McWilliams, A., Shaipanich, T. & Lam, S. (2013a) Fluorescence and navigational bronchoscopy. *Thoracic surgery clinics*, 23(2), 153-161.

McWilliams, A., Tammemagi, M. C., Mayo, J. R., Roberts, H., Liu, G., Soghrati, K., Yasufuku, K., Martel, S., Laberge, F. & Gingras, M. (2013b) Probability of cancer in pulmonary nodules detected on first screening CT. *New England Journal of Medicine*, 369(10), 910-919.

McWilliams, A., Tammemagi, M. C., Mayo, J. R., Roberts, H., Liu, G., Soghrati, K., Yasufuku, K., Martel, S., Laberge, F., Gingras, M., Atkar-Khattra, S., Berg, C. D., Evans, K., Finley, R., Yee, J., English, J., Nasute, P., Goffin, J., Puksa, S., Stewart, L., Tsai, S., Johnston, M. R., Manos, D., Nicholas, G., Goss, G. D., Seely, J. M., Amjadi, K., Tremblay, A., Burrowes, P., MacEachern, P., Bhatia, R., Tsao, M.-S. & Lam, S. (2013c) Probability of cancer in pulmonary nodules detected on first screening CT. *The New England journal of medicine*, 369(10), 910-919.

Mecklenburg, I., Sienel, W., Schmid, S., Passlick, B. & Kufer, P. (2017) A Threshold of Systemic MAGE-A Gene Expression Predicting Survival in Resected Non–Small Cell Lung Cancer. *Clinical Cancer Research*, 23(5), 1213-1219.

Medina, P. P., Castillo, S. D., Blanco, S., Sanz-Garcia, M., Largo, C., Alvarez, S., Yokota, J., Gonzalez-Neira, A., Benitez, J. & Clevers, H. C. (2009) The SRY-HMG box gene, SOX4, is a target of gene amplification at chromosome 6p in lung cancer. *Human molecular genetics*, 18(7), 1343-1352.

Mehan, M. R., Williams, S. A., Siegfried, J. M., Bigbee, W. L., Weissfeld, J. L., Wilson, D. O., Pass, H. I., Rom, W. N., Muley, T., Meister, M., Franklin, W., Miller, Y. E., Brody, E. N. & Ostroff, R. M. (2014) Validation of a blood protein signature for non-small cell lung cancer. *Clinical proteomics*, 11(1), 32-32.

Meissner, C., Lorenz, H., Weihofen, A., Selkoe, D. J. & Lemberg, M. K. (2011) The mitochondrial intramembrane protease PARL cleaves human Pink1 to regulate Pink1 trafficking. *Journal of neurochemistry*, 117(5), 856-867.

Meklat, F., Zhang, Y., Shahriar, M., Ahmed, S. U., Li, W., Voukkalis, N., Wang, Z., Zhang, J., Mastulov, S., Jewell, A., Giannakouros, T. & Lim, S. H. (2009) Identification of protamine 1 as a novel cancer-testis antigen in early chronic lymphocytic leukaemia. *British Journal of Haematology*, 144(5), 660-666.

Mellberg, S., Dimberg, A., Bahram, F., Hayashi, M., Rennel, E., Ameur, A., Westholm, J. O., Larsson, E., Lindahl, P., Cross, M. J. & Claesson-Welsh, L. (2009) Transcriptional profiling reveals a critical role for tyrosine phosphatase VE-PTP in regulation of VEGFR2 activity and endothelial cell morphogenesis. *The FASEB Journal*, 23(5), 1490-1502.

Memarian, A., Vosough, P., Asgarian-Omran, H., Tabrizi, M., Shabani, M. & Shokri, F. (2012) Differential WNT gene expression in various subtypes of acute lymphoblastic leukemia. *Iran J Immunol*, 9(1), 61-71.

Meng, F., Yu, W., Niu, M., Tian, X., Miao, Y., Li, X., Zhou, Y., Ma, L., Zhang, X., Qian, K., Yu, Y., Wang, J. & Huang, L. (2023) Ratiometric electrochemical OR gate assay for NSCLC-derived exosomes. *J Nanobiotechnology*, 21(1), 104.

Mertens, A. C., Wen, W., Davies, S. M., Steinbuch, M., Buckley, J. D., Potter, J. D. & Robison, L. L. (1998) Congenital abnormalities in children with acute leukemia: a report from the Children's Cancer Group. *The Journal of pediatrics*, 133(5), 617-623.

Metzger, M. B., Hristova, V. A. & Weissman, A. M. (2012) HECT and RING finger families of E3 ubiquitin ligases at a glance. *Journal of cell science*, 125(3), 531-537.

Mezheyeuski, A., Bergsland, C. H., Backman, M., Djureinovic, D., Sjöblom, T., Bruun, J. & Micke, P. (2018) Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *The Journal of Pathology*, 244(4), 421-431.

Mi, S., Lu, J., Sun, M., Li, Z., Zhang, H., Neilly, M. B., Wang, Y., Qian, Z., Jin, J. & Zhang, Y. (2007) MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia. *Proceedings of the National Academy of Sciences*, 104(50), 19971-19976.

Mian, S. A., Anjos-Afonso, F. & Bonnet, D. (2021) Advances in human immune system mouse models for studying human hematopoiesis and cancer immunotherapy. *Frontiers in immunology*, 11, 619236.

Milewicz, D. M. & Kwartler, C. S. (2012) Chapter 97 - Genetic Variants in Smooth Muscle Contraction and Adhesion Genes Cause Thoracic Aortic Aneurysms and Dissections and Other Vascular Diseases, in Hill, J. A. & Olson, E. N. (eds), *Muscle*. Boston/Waltham: Academic Press, 1291-1300.

Minna, J. D., Roth, J. A. & Gazdar, A. F. (2002) Focus on lung cancer. *Cancer cell*, 1(1), 49-52.

Mitsudomi, T. (2014) Molecular epidemiology of lung cancer and geographic variations with special reference to EGFR mutations. *Translational lung cancer research*, 3(4), 205.

Mohammadi-Yeganeh, S., Paryan, M., Arefian, E., Vasei, M., Ghanbarian, H., Mahdian, R., Karimipoor, M. & Soleimani, M. (2016) MicroRNA-340 inhibits the migration, invasion, and metastasis of breast cancer cells by targeting Wnt pathway. *Tumor Biology*, 37(7), 8993-9000.

Mohammadi, S. M., Mohammad Nejad, D. & Nozad Charoudeh, H. (2017) Genetic alterations in B-acute lymphoblastic leukemia. *Acta Haematologica Polonica*, 48(1), 10-17.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P. & Stewart, L. A. (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*, 4(1), 1.

Mohseni, M., Uludag, H. & Brandwein, J. M. (2018) Advances in biology of acute lymphoblastic leukemia (ALL) and therapeutic implications. *American journal of blood research*, 8(4), 29-56.

Mok, S. C., Kwok, T. T., Berkowitz, R. S., Barrett, A. J. & Tsui, F. W. (1995) Overexpression of the protein tyrosine phosphatase, nonreceptor type 6 (PTPN6), in human epithelial ovarian cancer. *Gynecologic oncology*, 57(3), 299-303.

Molina, O., Bataller, A., Thampi, N., Ribera, J., Granada, I., Velasco, P., Fuster, J. L. & Menendez, P. (2021) Near-haploidy and low-hypodiploidy in B-cell acute lymphoblastic leukemia: when less is too much. *Cancers*, 14(1), 32.

Møller, G. M. O., Frost, V., Melo, J. V. & Chantry, A. (2007) Upregulation of the TGFβ signalling pathway by Bcr-Abl: Implications for haemopoietic cell growth and chronic myeloid leukaemia. *FEBS Letters*, 581(7), 1329-1334.

Monte, M., Simonatto, M., Peche, L. Y., Bublik, D. R., Gobessi, S., Pierotti, M. A., Rodolfo, M. & Schneider, C. (2006) MAGE-A tumor antigens target p53 transactivation function through histone deacetylase recruitment and confer resistance to chemotherapeutic agents. *Proceedings of the National Academy of Sciences*, 103(30), 11160-11165.

Moore, L. D., Le, T. & Fan, G. (2013) DNA methylation and its basic function. *Neuropsychopharmacology*, 38(1), 23-38.

Moore, R. G., Miller, M. C., Disilvestro, P., Landrum, L. M., Gajewski, W., Ball, J. J. & Skates, S. J. (2011) Evaluation of the diagnostic accuracy of the risk of ovarian malignancy algorithm in women with a pelvic mass. *Obstetrics and gynecology*, 118(2 Pt 1), 280.

Moorman, A. V., Harrison, C. J., Buck, G. A., Richards, S. M., Secker-Walker, L. M., Martineau, M., Vance, G. H., Cherry, A. M., Higgins, R. R. & Fielding, A. K. (2007) Karyotype is an independent prognostic factor in adult acute lymphoblastic leukemia (ALL): analysis of cytogenetic data from patients treated on the Medical Research Council (MRC) UKALLXII/Eastern Cooperative Oncology Group (ECOG) 2993 trial. *Blood*, 109(8), 3189-3197.

Moreno, C. S. (2020) SOX4: The unappreciated oncogene. *Seminars in Cancer Biology*, 67, 57-64.

Morgan, R. A., Chinnasamy, N., Abate-Daga, D. D., Gros, A., Robbins, P. F., Zheng, Z., Feldman, S. A., Yang, J. C., Sherry, R. M. & Phan, G. Q. (2013) Cancer regression and neurologic toxicity following anti-MAGE-A3 TCR gene therapy. *Journal of immunotherapy (Hagerstown, Md.: 1997)*, 36(2), 133.

Morris, L. G., Sikora, A. G., Patel, S. G., Hayes, R. B. & Ganly, I. (2011) Second primary cancers after an index head and neck cancer: subsite-specific trends in the era of human papillomavirus–associated oropharyngeal cancer. *Journal of clinical oncology*, 29(6), 739.

Moshkovskii, S. (2012) Why do cancer cells produce serum amyloid A acute-phase protein? *Biochemistry (Moscow)*, 77(4), 339-341.

Mu, Y., Li, J., Xie, F., Xu, L. & Xu, G. (2022) Efficacy of autoantibodies combined with tumor markers in the detection of lung cancer. *J Clin Lab Anal*, 36(8), e24504.

Mullighan, C. G. & Downing, J. R. (2009) Global genomic characterization of acute lymphoblastic leukemia. *Seminars in hematology*, 46(1), 3-15.

Mullighan, C. G., Su, X., Zhang, J., Radtke, I., Phillips, L. A., Miller, C. B., Ma, J., Liu, W., Cheng, C. & Schulman, B. A. (2009) Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *New England Journal of Medicine*, 360(5), 470-480.

Murray, A., Chapman, C., Healey, G., Peek, L., Parsons, G., Baldwin, D., Barnes, A., Sewell, H., Fritsche, H. & Robertson, J. (2010) Technical validation of an autoantibody test for lung cancer. *Annals of Oncology*, 21(8), 1687-1693.

Nadeu, F., Royo, R., Massoni-Badosa, R., Playa-Albinyana, H., Garcia-Torre, B., Duran-Ferrer, M., Dawson, K. J., Kulis, M., Diaz-Navarro, A. & Villamor, N. (2022) Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. *Nature Medicine*, 28(8), 1662-1671.

Nagy, B., Bhattoa, H. P., Steiber, Z., Csobán, M., Szilasi, M., Méhes, G., Müller, M., Lázár, J., Kappelmayer, J. & Antal-Szalmás, P. (2014) Serum human epididymis protein 4 (HE4) as a tumor marker in men with lung cancer. *Clinical Chemistry and Laboratory Medicine (CCLM)*, 52(11), 1639-1648.

Nakagawa, K., Noguchi, Y., Uenaka, A., Sato, S., Okumura, H., Tanaka, M., Shimono, M., Ali Eldib, A. M., Ono, T. & Ohara, N. (2005) XAGE-1 expression in non–small cell lung cancer and antibody response in patients. *Clinical cancer research*, 11(15), 5496-5503.

Nalejska, E., Mączyńska, E. & Lewandowska, M. A. (2014) Prognostic and predictive biomarkers: tools in personalized oncology. *Molecular diagnosis & therapy*, 18(3), 273-284.

Nami, B. & Wang, Z. (2018) Genetics and expression profile of the tubulin gene superfamily in breast cancer subtypes and its relation to taxane resistance. *Cancers*, 10(8), 274.

Nardi-Agmon, I. & Peled, N. (2017) Exhaled breath analysis for the early detection of lung cancer: Recent developments and future prospects. *Lung Cancer: Targets and Therapy*, Volume 8, 31-38.

Nasrullah, N., Sang, J., Alam, M. S., Mateen, M., Cai, B. & Hu, H. (2019) Automated Lung Nodule Detection and Classification Using Deep Learning Combined with Multiple Strategies. *Sensors (Basel, Switzerland)*, 19(17), 3722.

Nawa, T., Nakagawa, T., Kusano, S., Kawasaki, Y., Sugawara, Y. & Nakata, H. (2002) Lung cancer screening using low-dose spiral CT: results of baseline and 1-year follow-up studies. *Chest*, 122(1), 15-20.

Nepstad, I., Hatfield, K. J., Grønningsæter, I. S. & Reikvam, H. (2020) The PI3K-Akt-mTOR Signaling Pathway in Human Acute Myeloid Leukemia (AML) Cells. *International Journal of Molecular Sciences*, 21(8). Available online: [Accessed.

Network, N. C. C. (2013) NCCN clinical practice guidelines in oncology (NCCN Guidelines®). Non-small cell lung cancer. Version 2.2013. *http://www. respiratory-thessaly. gr/assets/nscl*, 202, 202013.

Ngamwong, Y., Tangamornsuksan, W., Lohitnavy, O., Chaiyakunapruk, N., Scholfield, C. N., Reisfeld, B. & Lohitnavy, M. (2015) Additive Synergism between Asbestos and Smoking in Lung Cancer Risk: A Systematic Review and Meta-Analysis. *PLoS One*, 10(8), e0135798.

Niemeyer, P., Türeci, Ö., Eberle, T., Graf, N., Pfreundschuh, M. & Sahin, U. (2003) Expression of serologically identified tumor antigens in acute leukemias. *Leukemia research*, 27(7), 655-660.

NLST, T. N. L. S. T. R. T. (2013) Results of initial low-dose computed tomographic screening for lung cancer. *New England Journal of Medicine*, 368(21), 1980-1991.

Nolen, B. M., Langmead, C. J., Choi, S., Lomakin, A., Marrangoni, A., Bigbee, W. L., Weissfeld, J. L., Wilson, D. O., Dacic, S., Siegfried, J. M. & Lokshin, A. E. (2011) Serum biomarker profiles as diagnostic tools in lung cancer. *Cancer Biomark*, 10(1), 3-12.

Nonaka, M., Kataoka, D., Yamamoto, S., Bito, A., Matsuoka, J., Kawada, T. & Takaba, T. (2004) Pre-and post-operative serum carcinoembryonic antigen in primary lung adenocarcinoma. *Ann Thorac Cardiovasc Surg*, 10(5), 281-4.

Nooreldeen, R. & Bach, H. (2021) Current and Future Development in Lung Cancer Diagnosis. *International Journal of Molecular Sciences*, 22(16), 8661.

Nygren, M. K., Døsen-Dahl, G., Stubberud, H., Wälchli, S., Munthe, E. & Rian, E. (2009) β-catenin is involved in N-cadherin–dependent adhesion, but not in canonical Wnt signaling in E2A-PBX1–positive B acute lymphoblastic leukemia cells. *Experimental hematology*, 37(2), 225-233.

Nygren, M. K., Døsen, G., Hystad, M. E., Stubberud, H., Funderud, S. & Rian, E. (2007) Wnt3A activates canonical Wnt signalling in acute lymphoblastic leukaemia (ALL) cells and inhibits the proliferation of B‐ALL cell lines. *British journal of haematology*, 136(3), 400-413.

O'Dowd, E. L., McKeever, T. M., Baldwin, D. R., Anwar, S., Powell, H. A., Gibson, J. E., Iyen-Omofoman, B. & Hubbard, R. B. (2015) What characteristics of primary care and patients are associated with early death in patients with lung cancer in the UK? *Thorax*, 70(2), 161-168.

Okamura, K., Takayama, K., Izumi, M., Harada, T., Furuyama, K. & Nakanishi, Y. (2013) Diagnostic value of CEA and CYFRA 21-1 tumor markers in primary lung cancer. *Lung Cancer*, 80(1), 45-49.

Oken, M. M., Hocking, W. G., Kvale, P. A., Andriole, G. L., Buys, S. S., Church, T. R., Crawford, E. D., Fouad, M. N., Isaacs, C. & Reding, D. J. (2011) Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *Jama*, 306(17), 1865-1873.

Old, L. J. (2007) Cancer is a somatic cell pregnancy. AACR.

Olson, B. M., Frye, T. P., Johnson, L. E., Fong, L., Knutson, K. L., Disis, M. L. & McNeel, D. G. (2010) HLA-A2-restricted T-cell epitopes specific for prostatic acid phosphatase. *Cancer immunology, immunotherapy*, 59, 943-953.

Ost, D. & Fein, A. (2000) Evaluation and management of the solitary pulmonary nodule. *American journal of respiratory and critical care medicine*, 162(3), 782-787.

Ost, D., Fein, A. M. & Feinsilver, S. H. (2003) The solitary pulmonary nodule. *New England Journal of Medicine*, 348(25), 2535-2542.

Ost, D. E. & Gould, M. K. (2012) Decision making in patients with pulmonary nodules. *Am J Respir Crit Care Med*, 185(4), 363-72.

Ouyang, R., Wu, S., Zhang, B., Wang, T., Yin, B., Huang, J., Wei, W., Huang, M., Zhang, M., Wang, Y., Wang, F. & Hou, H. (2021) Clinical value of tumor-associated antigens and autoantibody panel combination detection in the early diagnostic of lung cancer. *Cancer Biomark*, 32(3), 401-409.

Paci, M., Maramotti, S., Bellesia, E., Formisano, D., Albertazzi, L., Ricchetti, T., Ferrari, G., Annessi, V., Lasagni, D., Carbonelli, C., De Franco, S., Brini, M., Sgarbi, G. & Lodi, R. (2009) Circulating plasma DNA as diagnostic biomarker in non-small cell lung cancer. *Lung Cancer*, 64(1), 92-97.

Pakvisal, N., Kongkavitoon, P., Sathitruangsak, C., Pornpattanarak, N., Boonsirikamchai, P., Ouwongprayoon, P., Aporntewan, C., Chantranuwatana, P., Mutirangura, A. & Vinayanuwattikun, C. (2022) Differential expression of immune-regulatory proteins C5AR1, CLEC4A and NLRP3 on peripheral blood mononuclear cells in early-stage non-small cell lung cancer patients. *Sci Rep*, 12(1), 18439.

Pan, J., Yu, L., Wu, Q., Lin, X., Liu, S., Hu, S., Rosa, C., Eichinger, D., Pino, I., Zhu, H., Qian, J. & Huang, Y. (2020) Integration of IgA and IgG Autoantigens Improves Performance of Biomarker Panels for Early Diagnosis of Lung Cancer. *Molecular & Cellular Proteomics*, 19(3), 490-500.

Pan, T., Wang, S., Feng, H., Xu, J., Zhang, M., Yao, Y., Xu, K. & Niu, M. (2022) Preclinical evaluation of the ROCK1 inhibitor, GSK269962A, in acute myeloid leukemia. *Front Pharmacol*, 13, 1064470.

Papaemmanuil, E., Rapado, I., Li, Y., Potter, N. E., Wedge, D. C., Tubio, J., Alexandrov, L. B., Van Loo, P., Cooke, S. L. & Marshall, J. (2014) RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nature genetics*, 46(2), 116-125.

Park, S. H., Lee, A. R., Choi, K., Joung, S., Yoon, J. B. & Kim, S. (2019) TOMM20 as a potential therapeutic target of colorectal cancer. *BMB Rep*, 52(12), 712-717.

Pass, H. I., Beer, D. G., Joseph, S. & Massion, P. (2013) Biomarkers and molecular testing for early detection, diagnosis, and therapeutic prediction of lung cancer. *Thoracic surgery clinics*, 23(2), 211-224.

Pastorino, U., Sverzellati, N., Sestini, S., Silva, M., Sabia, F., Boeri, M., Cantarutti, A., Sozzi, G., Corrao, G. & Marchianò, A. (2019) Ten-year results of the Multicentric Italian Lung Detection trial demonstrate the safety and efficacy of biennial lung cancer screening. *European Journal of Cancer*, 118, 142-148.

Patlak, M. & Nass, S. J. (2013) *Implementing a National Cancer Clinical Trials System for the 21st Century: Second Workshop Summary*.National Academies Press.

Patriquin, L., Merrick, D. T., Hill, D., Holcomb, R. G., Lemieux, M. E., Bennett, G., Karia, B., Rebel, V. I. & Bauer II, T. (2015) Early detection of lung cancer with meso tetra (4-carboxyphenyl) porphyrin-labeled sputum. *Journal of Thoracic Oncology*, 10(9), 1311-1318.

Patz Jr, E. F., Campa, M. J., Gottlin, E. B., Kusmartseva, I., Guan, X. R. & Herndon, J. E. (2007) Panel of serum biomarkers for the diagnosis of lung cancer. *Journal of Clinical Oncology*, 25(35), 5578-5583.

Paul, S., Kantarjian, H. & Jabbour, E. J. (2016) Adult acute lymphoblastic leukemia, *Mayo Clinic Proceedings*. Elsevier.

Peche, L., Scolz, M., Ladelfa, M. F., Monte, M. & Schneider, C. (2012) MageA2 restrains cellular senescence by targeting the function of PMLIV/p53 axis at the PML-NBs. *Cell Death & Differentiation*, 19(6), 926-936.

Peculis, R., Niedra, H. & Rovite, V. (2021) Large Scale Molecular Studies of Pituitary Neuroendocrine Tumors: Novel Markers, Mechanisms and Translational Perspectives. *Cancers*, 13(6), 1395.

Pepe, M. S. (2000) Receiver operating characteristic methodology. *Journal of the American Statistical Association*, 95(449), 308-311.

Pepe, M. S., Etzioni, R., Feng, Z., Potter, J. D., Thompson, M. L., Thornquist, M., Winget, M. & Yasui, Y. (2001) Phases of biomarker development for early detection of cancer. *Journal of the National Cancer Institute*, 93(14), 1054-1061.

Pepe, M. S., Feng, Z., Janes, H., Bossuyt, P. M. & Potter, J. D. (2008) Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *Journal of the National Cancer Institute*, 100(20), 1432-1438.

Petersen, D. L., Krejsgaard, T., Berthelsen, J., Fredholm, S., Willerslev-Olsen, A., Sibbesen, N. A., Bonefeld, C. M., Andersen, M. H., Francavilla, C., Olsen, J. V., Hu, T., Zhang, M., Wasik, M. A., Geisler, C., Woetmann, A. & Odum, N. (2014) B-lymphoid tyrosine kinase (Blk) is an oncogene and a potential target for therapy with dasatinib in cutaneous T-cell lymphoma (CTCL). *Leukemia*, 28(10), 2109-2112.

Petit, C., Gouel, F., Dubus, I., Heuclin, C., Roget, K. & Vannier, J. P. (2016) Hypoxia promotes chemoresistance in acute lymphoblastic leukemia cell lines by modulating death signaling pathways. *BMC Cancer*, 16(1), 746.

Phoenix, T. N., Patmore, D. M., Boop, S., Boulos, N., Jacus, M. O., Patel, Y. T., Roussel, M. F., Finkelstein, D., Goumnerova, L., Perreault, S., Wadhwa, E., Cho, Y.-J., Stewart, C. F. & Gilbertson, R. J. (2016) Medulloblastoma Genotype Dictates Blood Brain Barrier Phenotype. *Cancer Cell*, 29(4), 508-522.

Piccaluga, P. P., Arpinati, M., Candoni, A., Laterza, C., Paolini, S., Gazzola, A., Sabattini, E., Visani, G. & Pileri, S. A. (2011) Surface antigens analysis reveals significant expression of candidate targets for immunotherapy in adult acute lymphoid leukemia. *Leukemia & lymphoma*, 52(2), 325-327.

Piccart-Gebhart, M. J., Procter, M., Leyland-Jones, B., Goldhirsch, A., Untch, M., Smith, I., Gianni, L., Baselga, J., Bell, R., Jackisch, C., Cameron, D., Dowsett, M., Barrios, C. H., Steger, G., Huang, C. S., Andersson, M., Inbar, M., Lichinitser, M., Láng, I., Nitz, U., Iwata, H., Thomssen, C., Lohrisch, C., Suter, T. M., Rüschoff, J., Suto, T., Greatorex, V., Ward, C., Straehle, C., McFadden, E., Dolci, M. S. & Gelber, R. D. (2005) Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med*, 353(16), 1659-72.

Placzek, W. J., Almeida, M. S. & Wüthrich, K. (2007) NMR structure and functional characterization of a human cancer-related nucleoside triphosphatase. *Journal of molecular biology*, 367(3), 788-801.

Ponting, C. P., Oliver, P. L. & Reik, W. (2009) Evolution and functions of long noncoding RNAs. *Cell*, 136(4), 629-641.

Prabhakar, B., Shende, P. & Augustine, S. (2018) Current trends and emerging diagnostic techniques for lung cancer. *Biomedicine & Pharmacotherapy*, 106, 1586-1599.

Pramoonjago, P., Baras, A. S. & Moskaluk, C. A. (2006) Knockdown of Sox4 expression by RNAi induces apoptosis in ACC3 cells. *Oncogene*, 25(41), 5626-5639.

Pui, C.-H. & Evans, W. E. (2006) Treatment of acute lymphoblastic leukemia. *New England Journal of Medicine*, 354(2), 166-178.

Pukkala, E., Martinsen, J. I., Lynge, E., Gunnarsdottir, H. K., Sparén, P., Tryggvadottir, L., Weiderpass, E. & Kjaerheim, K. (2009) Occupation and cancer–follow-up of 15 million people in five Nordic countries. *Acta oncologica*, 48(5), 646-790.

Qi, X., Lv, X., Wang, X., Ruan, Z., Zhang, P., Wang, Q., Xu, Y. & Wu, G. (2021) A New Survival Model Based on Cholesterol Biosynthesis-Related Genes for Prognostic Prediction in Clear Cell Renal Cell Carcinoma. *Biomed Res Int*, 2021, 9972968.

Qian, L., Zhang, W., Lei, B., He, A., Ye, L., Li, X. & Dong, X. (2016) MicroRNA-101 regulates T-cell acute lymphoblastic leukemia progression and chemotherapeutic sensitivity by targeting Notch1. *Oncology Reports*, 36(5), 2511-2516.

Qiao, Y., Li, T., Zheng, S. & Wang, H. (2018) The Hippo pathway as a drug target in gastric cancer. *Cancer Letters*, 420, 14-25.

Qin, C., Wang, Y., Zhao, B., Li, Z., Li, T., Yang, X., Zhao, Y. & Wang, W. (2023) STOML2 restricts mitophagy and increases chemosensitivity in pancreatic cancer through stabilizing PARL-induced PINK1 degradation. *Cell Death & Disease*, 14(3), 191.

Raetz, E. A., Morrison, D., Romanos-Sirakis, E., Gaynon, P., Sposto, R., Bhojwani, D., Bostrom, B. C., Brown, P., Eckroth, E. & Cassar, J. (2014) A phase I study of EZN-3042, a novel survivin messenger ribonucleic acid (mRNA) antagonist, administered in combination with chemotherapy in children with relapsed acute lymphoblastic leukemia (ALL): a report from the therapeutic advances in childhood leukemia and lymphoma (TACL) consortium. *Journal of pediatric hematology/oncology*, 36(6), 458.

Rafei, H., Kantarjian, H. M. & Jabbour, E. J. (2019) Recent advances in the treatment of acute lymphoblastic leukemia. *Leukemia & lymphoma*, 60(11), 2606-2621.

Raghu, G., Amatto, V. C., Behr, J. & Stowasser, S. (2015) Comorbidities in idiopathic pulmonary fibrosis patients: a systematic literature review. *European Respiratory Journal*, 46(4), 1113-1130.

Ramezani-Rad, P., Geng, H., Hurtz, C., Chan, L. N., Chen, Z., Jumaa, H., Melnick, A., Paietta, E., Carroll, W. L., Willman, C. L., Lefebvre, V. & Müschen, M. (2013) SOX4 enables oncogenic survival signals in acute lymphoblastic leukemia. *Blood*, 121(1), 148-155.

Ransohoff, D. F. & Gourlay, M. L. (2010a) Sources of bias in specimens for research about molecular markers for cancer. *Journal of clinical oncology*, 28(4), 698.

Ransohoff, D. F. & Gourlay, M. L. (2010b) Sources of bias in specimens for research about molecular markers for cancer. *J Clin Oncol*, 28(4), 698-704.

Raponi, S., Stefania De Propris, M., Intoppa, S., Laura Milani, M., Vitale, A., Elia, L., Perbellini, O., Pizzolo, G., Foá, R. & Guarini, A. (2011) Flow cytometric study of

potential target antigens (CD19, CD20, CD22, CD33) for antibody-based immunotherapy in acute lymphoblastic leukemia: analysis of 552 cases. *Leukemia & lymphoma*, 52(6), 1098-1107.

Raz, D. J., Zell, J. A., Ou, S. H. I., Gandara, D. R., Anton-Culver, H. & Jablons, D. M. (2007) Natural History of Stage I Non-Small Cell Lung Cancer: Implications for Early Detection. *Chest*, 132(1), 193-199.

Reis, P. P., Drigo, S. A., Carvalho, R. F., Lopez Lapa, R. M., Felix, T. F., Patel, D., Cheng, D., Pintilie, M., Liu, G. & Tsao, M.-S. (2020) Circulating miR-16-5p, miR-92a-3p, and miR-451a in Plasma from Lung Cancer Patients: Potential Application in Early Detection and a Regulatory Role in Tumorigenesis Pathways. *Cancers*, 12(8).

Ren, S., Zhang, S., Jiang, T., He, Y., Ma, Z., Cai, H., Xu, X., Li, Y., Cai, W. & Zhou, J. (2018) Early detection of lung cancer by using an autoantibody panel in Chinese population. *Oncoimmunology*, 7(2), e1384108.

Rho, S. B., Lee, J. H., Park, M. S., Byun, H.-J., Kang, S., Seo, S.-S., Kim, J.-Y. & Park, S.-Y. (2011) Anti-apoptotic protein TCTP controls the stability of the tumor suppressor p53. *FEBS letters*, 585(1), 29-35.

Rinaldi, L., Sepe, M., Delle Donne, R., Conte, K., Arcella, A., Borzacchiello, D., Amente, S., De Vita, F., Porpora, M., Garbi, C., Oliva, M. A., Procaccini, C., Faicchia, D., Matarese, G., Zito Marino, F., Rocco, G., Pignatiello, S., Franco, R., Insabato, L., Majello, B. & Feliciello, A. (2017) Mitochondrial AKAP1 supports mTOR pathway and tumor growth. *Cell Death Dis*, 8(6), e2842.

Robbins, P. F., Lu, Y.-C., El-Gamil, M., Li, Y. F., Gross, C., Gartner, J., Lin, J. C., Teer, J. K., Cliften, P. & Tycksen, E. (2013) Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nature medicine*, 19(6), 747.

Roberts, K. G. (2018) Genetics and prognosis of ALL in children vs adults. *Hematology. American Society of Hematology. Education Program*, 2018(1), 137-145.

Roberts, K. G., Gu, Z., Payne-Turner, D., McCastlain, K., Harvey, R. C., Chen, I.-M., Pei, D., Iacobucci, I., Valentine, M. & Pounds, S. B. (2017) High frequency and poor outcome of Philadelphia chromosome–like acute lymphoblastic leukemia in adults. *Journal of Clinical Oncology*, 35(4), 394.

Román-Gómez, J., Cordeu, L., Agirre, X., Jiménez-Velasco, A., San José-Eneriz, E., Garate, L., Calasanz, M. a. J., Heiniger, A., Torres, A. & Prosper, F. (2006) Epigenetic regulation of Wnt-signaling pathway in acute lymphoblastic leukemia. *Blood*, 109(8), 3462-3469.

Roman-Gomez, J., Jimenez-Velasco, A., Castillejo, J. A., Agirre, X., Barrios, M., Navarro, G., Molina, F. J., Calasanz, M. J., Prosper, F., Heiniger, A. & Torres, A. (2004) Promoter hypermethylation of cancer-related genes: a strong independent prognostic factor in acute lymphoblastic leukemia. *Blood*, 104(8), 2492-2498.

Roolf, C., Richter, A., Konkolefski, C., Knuebel, G., Sekora, A., Krohn, S., Stenzel, J., Krause, B. J., Vollmar, B., Murua Escobar, H. & Junghanss, C. (2018) Decitabine demonstrates antileukemic activity in B cell precursor acute lymphoblastic leukemia with MLL rearrangements. *Journal of Hematology & Oncology*, 11(1), 62.

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. (2015) Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*, 160(1), 48-61.

Rosenberg, S. A., Yang, J. C. & Restifo, N. P. (2004) Cancer immunotherapy: moving beyond current vaccines. *Nature medicine*, 10(9), 909-915.

Ross, K. S., Carter, H. B., Pearson, J. D. & Guess, H. A. (2000) Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. *Jama*, 284(11), 1399-1405.

Rouce, R. H., Shaim, H., Sekine, T., Weber, G., Ballard, B., Ku, S., Barese, C., Murali, V., Wu, M. F., Liu, H., Shpall, E. J., Bollard, C. M., Rabin, K. R. & Rezvani, K. (2016) The TGF-β/SMAD pathway is an important mechanism for NK cell immune evasion in childhood B-acute lymphoblastic leukemia. *Leukemia*, 30(4), 800-11.

Rovatti, P. E., Gambacorta, V., Lorentino, F., Ciceri, F. & Vago, L. (2020) Mechanisms of leukemia immune evasion and their role in relapse after haploidentical hematopoietic cell transplantation. *Frontiers in Immunology*, 11, 147.

Ruan, Y., Kim, H. N., Ogana, H. A., Wan, Z., Hurwitz, S., Nichols, C., Abdel-Azim, N., Coba, A., Seo, S., Loh, Y.-H. E., Gang, E. J., Abdel-Azim, H., Hsieh, C.-L., Lieber, M. R., Parekh, C., Pal, D., Bhojwani, D., Durden, D. L. & Kim, Y.-M. (2021) Preclinical Evaluation of a Novel Dual Targeting PI3Kδ/BRD4 Inhibitor, SF2535, in B-Cell Acute Lymphoblastic Leukemia. *Frontiers in Oncology*, 11.

Rubin, G. D. (2014) Computed tomography: revolutionizing the practice of medicine for 40 years. *Radiology*, 273(2S), S45-S74.

Ruella, M. & Maus, M. V. (2016) Catch me if you can: leukemia escape after CD19-directed T cell immunotherapies. *Computational and structural biotechnology journal*, 14, 357-362.

Saenger, A., Beyrau, R., Braun, S., Cooray, R., Dolci, A., Freidank, H., Giannitsis, E., Gustafson, S., Handy, B. & Katus, H. (2011) Multicenter analytical evaluation of a high-sensitivity troponin T assay. *Clinica chimica acta*, 412(9-10), 748-754.

Sahin, U., Türeci, O., Schmitt, H., Cochlovius, B., Johannes, T., Schmits, R., Stenner, F., Luo, G., Schobert, I. & Pfreundschuh, M. (1995) Human neoplasms elicit multiple specific immune responses in the autologous host. *Proceedings of the National Academy of Sciences*, 92(25), 11810-11813.

Sakuma, K., Sasaki, E., Kimura, K., Komori, K., Shimizu, Y., Yatabe, Y. & Aoki, M. (2018) HNRNPLL stabilizes mRNA for DNA replication proteins and promotes cell cycle progression in colorectal cancer cells. *Cancer Sci*, 109(8), 2458-2468.

Saleh, M., Khalil, M., Abdellateif, M. S., Ebeid, E., Madney, Y. & Kandeel, E. Z. (2021) Role of matrix metalloproteinase MMP-2, MMP-9 and tissue inhibitor of metalloproteinase (TIMP-1) in the clinical progression of pediatric acute lymphoblastic leukemia. *Hematology*, 26(1), 758-768.

Salian, S., Guo, X. Y., Murakami, Y., Kinoshita, T., Kaur, P., Shukla, A., Girisha, K. M., Fujita, M. & Campeau, P. M. (2022) C18orf32 loss-of-function is associated with a neurodevelopmental disorder with hypotonia and contractures. *Hum Genet*, 141(8), 1423-1429.

Sánchez, M., Benegas, M. & Vollmer, I. (2018) Management of incidental lung nodules <8 mm in diameter. *Journal of thoracic disease*, 10(Suppl 22), S2611-S2627.

Sanders, H. R. & Albitar, M. (2010) Somatic mutations of signaling genes in non-small-cell lung cancer. *Cancer Genetics and Cytogenetics*, 203(1), 7-15.

Sato, S., Noguchi, Y., Ohara, N., Uenaka, A., Shimono, M., Nakagawa, K., Koizumi, F., Ishida, T., Yoshino, T. & Shiratori, Y. (2007) Identification of XAGE-1 isoforms: predominant expression of XAGE-1b in testis and tumors. *Cancer Immunity*, 7(1), 5.

Sawabata, N., Ohta, M., Takeda, S.-i., Hirano, H., Okumura, Y., Asada, H. & Maeda, H. (2002) Serum carcinoembryonic antigen level in surgically resected clinical stage I patients with non-small cell lung cancer. *The Annals of thoracic surgery*, 74(1), 174-179.

Sawyers, C. L. (2008) The cancer biomarker problem. *Nature*, 452(7187), 548-552.

Sbirkov, Y., Vergov, B., Mehterov, N. & Sarafian, V. (2022) miRNAs in Lymphocytic Leukaemias—The miRror of Drug Resistance. *International Journal of Molecular Sciences*, 23(9), 4657.

Scanlan, M. J., Welt, S., Gordon, C. M., Chen, Y.-T., Gure, A. O., Stockert, E., Jungbluth, A. A., Ritter, G., Jäger, D. & Jäger, E. (2002) Cancer-related serological recognition of human colon cancer: identification of potential diagnostic and immunotherapeutic targets. *Cancer research*, 62(14), 4041-4047.

Schäfer, M., Oeing, C. U., Rohm, M., Baysal-Temel, E., Lehmann, L. H., Bauer, R., Volz, H. C., Boutros, M., Sohn, D., Sticht, C., Gretz, N., Eichelbaum, K., Werner, T., Hirt, M. N., Eschenhagen, T., Müller-Decker, K., Strobel, O., Hackert, T., Krijgsveld, J., Katus, H. A., Berriel Diaz, M., Backs, J. & Herzig, S. (2016) Ataxin-10 is part of a cachexokine cocktail triggering cardiac metabolic dysfunction in cancer cachexia. *Molecular Metabolism*, 5(2), 67-78.

Schagdarsurengin, U., Paradowska, A. & Steger, K. (2012) Analysing the sperm epigenome: roles in early embryogenesis and assisted reproduction. *Nature Reviews Urology*, 9(11), 609-619.

Schatzkin, A. & Gail, M. (2002) The promise and peril of surrogate end points in cancer research. *Nature Reviews Cancer*, 2(1), 19-27.

Scheuner, D., Song, B., McEwen, E., Liu, C., Laybutt, R., Gillespie, P., Saunders, T., Bonner-Weir, S. & Kaufman, R. J. (2001) Translational control is required for the

unfolded protein response and in vivo glucose homeostasis. *Molecular cell*, 7(6), 1165-1176.

Schmiegel, W., Kreiker, C., Eberl, W., Arndt, R., Classen, M., Greten, H., Jessen, K., Kalthoff, H., Soehendra, N. & Thiele, H. (1985) Monoclonal antibody defines CA 19-9 in pancreatic juices and sera. *Gut*, 26(5), 456-460.

Schneider, J. (2006) Tumor markers in detection of lung cancer. *Advances in clinical chemistry*, 42, 1-41.

Schneider, J., Philipp, M., Velcovsky, H.-G., Morr, H. & Katz, N. (2003) Pro-gastrin-releasing peptide (ProGRP), neuron specific enolase (NSE), carcinoembryonic antigen (CEA) and cytokeratin 19-fragments (CYFRA 21-1) in patients with lung cancer in comparison to other lung diseases. *Anticancer research*, 23(2A), 885.

Schumacher, T. N. & Schreiber, R. D. (2015) Neoantigens in cancer immunotherapy. *Science*, 348(6230), 69-74.

Seemann, M. D., Beinert, T., Fürst, H. & Fink, U. (1999) An evaluation of the tumour markers, carcinoembryonic antigen (CEA), cytokeratin marker (CYFRA 21-1) and neuron-specific enolase (NSE) in the differentiation of malignant from benign solitary pulmonary lesions. *Lung Cancer*, 26(3), 149-155.

Seemann, M. D., Seemann, O., Luboldt, W., Bonél, H., Sittek, H., Dienemann, H. & Staebler, A. (2000) Differentiation of malignant from benign solitary pulmonary lesions using chest radiography, spiral CT and HRCT. *Lung cancer*, 29(2), 105-124.

Seita, J. & Weissman, I. L. (2010) Hematopoietic stem cell: self‐renewal versus differentiation. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 2(6), 640-653.

Seliger, B. & Kellner, R. (2002) Design of proteome‐based studies in combination with serology for the identification of biomarkers and novel targets. *Proteomics*, 2(12), 1641-1651.

Seo, J. & Kim, J. (2018) Regulation of Hippo signaling by actin remodeling. *BMB reports*, 51(3), 151.

Sevin, M., Debeurme, F., Laplane, L., Badel, S., Morabito, M., Newman, H. L., Torres-Martin, M., Yang, Q., Badaoui, B., Wagner-Ballon, O., Saada, V., Sélimoglu-Buet, D., Kraus-Berthier, L., Banquet, S., Derreal, A., Fenaux, P., Itzykson, R., Braun, T., Etienne, G., Berthon, C., Thépot, S., Kepp, O., Kroemer, G., Padron, E., Figueroa, M. E., Droin, N. & Solary, E. (2021) Cytokine-like protein 1–induced survival of monocytes suggests a combined strategy targeting MCL1 and MAPK in CMML. *Blood*, 137(24), 3390-3402.

Shah, N. P., Nicoll, J. M., Bleickardt, E., Nicaise, C., Paquette, R. L. & Sawyers, C. L. (2006) Potent Transient Inhibition of BCR-ABL by Dasatinib Leads to Complete Cytogenetic Remissions in Patients with Chronic Myeloid Leukemia: Implications for Patient Management and Drug Development. American Society of Hematology.

Shah, N. P., Skaggs, B. J., Branford, S., Hughes, T. P., Nicoll, J. M., Paquette, R. L. & Sawyers, C. L. (2007) Sequential ABL kinase inhibitor therapy selects for compound drug-resistant BCR-ABL mutations with altered oncogenic potency. *The Journal of clinical investigation*, 117(9), 2562-2569.

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P. & Stewart, L. A. (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *Bmj*, 350, g7647.

Shang, B., Gao, A., Pan, Y., Zhang, G., Tu, J., Zhou, Y., Yang, P., Cao, Z., Wei, Q. & Ding, Y. (2014) CT45A1 acts as a new proto-oncogene to trigger tumorigenesis and cancer metastasis. *Cell death & disease*, 5(6), e1285-e1285.

Shang, H., Zhang, H., Ren, Z., Zhao, H., Zhang, Z. & Tong, J. (2021) Characterization of the Potential Role of NTPCR in Epithelial Ovarian Cancer by Integrating Transcriptomic and Metabolomic Analysis. *Front Genet*, 12, 695245.

Sharma, D., Newman, T. G. & Aronow, W. S. (2015) Lung cancer screening: history, current perspectives, and future directions. *Archives of medical science: AMS*, 11(5), 1033.

Sharma, S. V., Bell, D. W., Settleman, J. & Haber, D. A. (2007) Epidermal growth factor receptor mutations in lung cancer. *Nature Reviews Cancer*, 7(3), 169-181.

Shen, J., Liu, Z., Todd, N. W., Zhang, H., Liao, J., Yu, L., Guarnera, M. A., Li, R., Cai, L., Zhan, M. & Jiang, F. (2011a) Diagnosis of lung cancer in individuals with solitary pulmonary nodules by plasma microRNA biomarkers. *BMC Cancer*, 11(1), 374.

Shen, J., Su, J., Wu, D., Zhang, F., Fu, H., Zhou, H. & Xu, M. (2015) Growth Inhibition Accompanied by MOB1 Upregulation in Human Acute Lymphoid Leukemia Cells by 3-Deazaneplanocin A. *Biochemical Genetics*, 53(9), 268-279.

Shen, J., Todd, N. W., Zhang, H., Yu, L., Lingxiao, X., Mei, Y., Guarnera, M., Liao, J., Chou, A. & Lu, C. L. (2011b) Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. *Laboratory investigation*, 91(4), 579-587.

Shen, L., Zhang, F., Huang, R., Yan, J. & Shen, B. (2017) Honokiol inhibits bladder cancer cell invasion through repressing SRC-3 expression and epithelial-mesenchymal transition. *Oncology Letters*, 14(4), 4294-4300.

Shu, Y., Zhou, X., Qi, X., Liu, S., Li, K., Tan, J., Liu, Z., Yu, J., Zhang, P. & Zou, L. (2015) β-Arrestin1 promotes the self-renewal of the leukemia-initiating cell-enriched subpopulation in B-lineage acute lymphoblastic leukemia related to DNMT1 activity. *Cancer Letters*, 357(1), 170-178.

Silva, A. C. Q., Vilela, C., Santos, H. A., Silvestre, A. J. D. & Freire, C. S. R. (2020) Recent trends on the development of systems for cancer diagnosis and treatment by microfluidic technology. *Applied Materials Today*, 18, 100450.

Silva, M., Prokop, M., Jacobs, C., Capretti, G., Sverzellati, N., Ciompi, F., van Ginneken, B., Schaefer-Prokop, C. M., Galeone, C. & Marchianò, A. (2018) Long-term active

surveillance of screening detected subsolid nodules is a safe strategy to reduce overtreatment. *Journal of Thoracic Oncology*, 13(10), 1454-1463.

Silva, W. A., Gnjatic, S., Ritter, E., Chua, R., Cohen, T., Hsu, M., Jungbluth, A. A., Altorki, N. K., Chen, Y.-T. & Old, L. J. (2007a) PLAC1, a trophoblast-specific cell surface protein, is expressed in a range of human tumors and elicits spontaneous antibody responses. *Cancer immunity*, 7(1).

Silva, W. A., Gnjatic, S., Ritter, E., Chua, R., Cohen, T., Hsu, M., Jungbluth, A. A., Altorki, N. K., Chen, Y.-T. & Old, L. J. (2007b) PLAC1, a trophoblast-specific cell surface protein, is expressed in a range of human tumors and elicits spontaneous antibody responses. *Cancer Immunity Archive*, 7(1).

Silveira, N. J. F., Varuzza, L., Machado-Lima, A., Lauretto, M. S., Pinheiro, D. G., Rodrigues, R. V., Severino, P., Nobrega, F. G., Silva, W. A., de B Pereira, C. A., Tajara, E. H., Head & Neck Genome Project, G. (2008) Searching for molecular markers in head and neck squamous cell carcinomas (HNSCC) by statistical and bioinformatic analysis of larynx-derived SAGE libraries. *BMC Medical Genomics*, 1(1), 56.

Simpson, A. J., Caballero, O. L., Jungbluth, A., Chen, Y.-T. & Old, L. J. (2005) Cancer/testis antigens, gametogenesis and cancer. *Nature Reviews Cancer*, 5(8), 615-625.

Singh, J., Kumari, S., Arora, M., Verma, D., Palanichamy, J. K., Kumar, R., Sharma, G., Bakhshi, S., Pushpam, D., Ali, M. S., Ranjan, A., Tanwar, P., Chauhan, S. S., Singh, A. & Chopra, A. (2021) Prognostic Relevance of Expression of EMP1, CASP1, and NLRP3 Genes in Pediatric B-Lineage Acute Lymphoblastic Leukemia. *Frontiers in Oncology*, 11.

Sison, E. A. R., Rau, R. E., McIntyre, E., Li, L., Small, D. & Brown, P. (2013) MLL ‐ rearranged acute lymphoblastic leukaemia stem cell interactions with bone marrow stroma promote survival and therapeutic resistance that can be overcome with CXCR 4 antagonism. *British journal of haematology*, 160(6), 785-797.

Sive, J. I., Buck, G., Fielding, A., Lazarus, H. M., Litzow, M. R., Luger, S., Marks, D. I., McMillan, A., Moorman, A. V. & Richards, S. M. (2012) Outcomes in older adults with acute lymphoblastic leukaemia (ALL): results from the international MRC UKALL XII/ECOG 2993 trial. *British journal of haematology*, 157(4), 463-471.

Skaar, J. R., Pagan, J. K. & Pagano, M. (2014) SCF ubiquitin ligase-targeted therapies. *Nature reviews Drug discovery*, 13(12), 889-903.

Smith, S. & Kipling, D. (2004) The role of replicative senescence in cancer and human ageing: utility (or otherwise) of murine models. *Cytogenetic and genome research*, 105(2-4), 455-463.

Snoeckx, A., Reyntiens, P., Desbuquoit, D., Spinhoven, M. J., Van Schil, P. E., van Meerbeeck, J. P. & Parizel, P. M. (2018) Evaluation of the solitary pulmonary nodule: size matters, but do not ignore the power of morphology. *Insights into imaging*, 9(1), 73-86.

Sobin, L. H., Gospodarowicz, M. K. & Wittekind, C. (2011) *TNM classification of malignant tumours*.John Wiley & Sons.

Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S.-i., Watanabe, H., Kurashina, K. & Hatanaka, H. (2007) Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. *Nature*, 448(7153), 561-566.

Soejima, M. & Koda, Y. (2005) Molecular mechanisms of Lewis antigen expression. *Legal Medicine*, 7(4), 266-269.

Soleymani Fard, S., Jeddi Tehrani, M. & Ardekani, A. M. (2012) Prostaglandin E2 induces growth inhibition, apoptosis and differentiation in T and B cell-derived acute lymphoblastic leukemia cell lines (CCRF-CEM and Nalm-6). *Prostaglandins Leukot Essent Fatty Acids*, 87(1), 17-24.

Soma, S., Tsuta, K., Takano, T., Hatanaka, Y., Yoshida, A., Suzuki, K., Asamura, H. & Tsuda, H. (2014) Intratumoral distribution of EGFR-amplified and EGFR-mutated cells in pulmonary adenocarcinoma. *Pathology-Research and Practice*, 210(3), 155-160.

Song, C.-Q., Li, Y., Mou, H., Moore, J., Park, A., Pomyen, Y., Hough, S., Kennedy, Z., Fischer, A. & Yin, H. (2017) Genome-wide CRISPR screen identifies regulators of mitogen-activated protein kinase as suppressors of liver tumors in mice. *Gastroenterology*, 152(5), 1161-1173. e1.

Song, K.-S., Nimse, S. B., Warkad, S. D., Oh, A.-C., Kim, T. & Hong, Y. J. (2019) Quantification of CYFRA 21-1 and a CYFRA 21-1–anti-CYFRA 21-1 autoantibody immune complex for detection of early stage lung cancer. *Chemical Communications*, 55(68), 10060-10063.

Song, K. S., Nimse, S. B., Warkad, S. D., Kim, J. H., Kim, H. J. & Kim, T. (2022) Detection and Quantification of Tp53 and p53-Anti-p53 Autoantibody Immune Complex: Promising Biomarkers in Early Stage Lung Cancer Diagnosis. *Biosensors (Basel)*, 12(2).

Søreide, K. (2009) Receiver-operating characteristic curve analysis in diagnostic, prognostic and predictive biomarker research. *Journal of clinical pathology*, 62(1), 1-5.

Steelman, L. S., Franklin, R. A., Abrams, S. L., Chappell, W., Kempf, C. R., Bäsecke, J., Stivala, F., Donia, M., Fagone, P., Nicoletti, F., Libra, M., Ruvolo, P., Ruvolo, V., Evangelisti, C., Martelli, A. M. & McCubrey, J. A. (2011) Roles of the Ras/Raf/MEK/ERK pathway in leukemia therapy. *Leukemia*, 25(7), 1080-1094.

Steinhardt, A. A., Gayyed, M. F., Klein, A. P., Dong, J., Maitra, A., Pan, D., Montgomery, E. A. & Anders, R. A. (2008) Expression of Yes-associated protein in common solid tumors. *Human pathology*, 39(11), 1582-1589.

Strimbu, K. & Tavel, J. A. (2010) What are biomarkers? *Current Opinion in HIV and AIDS*, 5(6), 463.

Su, M., Guo, J. & Huang, J. (2017) Meta-analysis of the correlation between the rs17401966 polymorphism in kinesin family member 1B and susceptibility to hepatitis B virus related hepatocellular carcinoma. *Clin Mol Hepatol*, 23(2), 138-146.

Sun, N., Sun, S., Gao, Y., Li, Y., Lu, Z., Yuan, Z., Che, Y., Huang, J., Mao, S. & Lei, Y. (2020a) Utility of isocitrate dehydrogenase 1 as a serum protein biomarker for the early detection of non‐small‐cell lung cancer: A multicenter in vitro diagnostic clinical trial. *Cancer science*, 111(5), 1739-1749.

Sun, Y.-Q., Li, S.-Q., Zhao, X.-S. & Chang, Y.-J. (2020b) Measurable residual disease of acute lymphoblastic leukemia in allograft settings: how to evaluate and intervene. *Expert Review of Anticancer Therapy*, 20(6), 453-464.

Surapaneni, R., Singh, P., Rajagopalan, K. & Hageboutros, A. (2012) Stage I lung cancer survivorship: risk of second malignancies and need for individualized care plan. *Journal of Thoracic Oncology*, 7(8), 1252-1256.

Swensen, S. J., Viggiano, R. W., Midthun, D. E., Müller, N. L., Sherrick, A., Yamashita, K., Naidich, D. P., Patz, E. F., Hartman, T. E. & Muhm, J. R. (2000) Lung nodule enhancement at CT: multicenter study. *Radiology*, 214(1), 73-80.

Swerdlow, S. H., Campo, E., Pileri, S. A., Harris, N. L., Stein, H., Siebert, R., Advani, R., Ghielmini, M., Salles, G. A. & Zelenetz, A. D. (2016) The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood, The Journal of the American Society of Hematology*, 127(20), 2375-2390.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H. & Bork, P. (2019) STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research*, 47(D1), D607-D613.

Takada, K., Zhu, D., Bird, G. H., Sukhdeo, K., Zhao, J. J., Mani, M., Lemieux, M., Carrasco, D. E., Ryan, J., Horst, D., Fulciniti, M., Munshi, N. C., Xu, W., Kung, A. L., Shivdasani, R. A., Walensky, L. D. & Carrasco, D. R. (2012) Targeted disruption of the BCL9/β-catenin complex inhibits oncogenic Wnt signaling. *Sci Transl Med*, 4(148), 148ra117.

Takeuchi, S., Kasamatsu, A., Yamatoji, M., Nakashima, D., Endo-Sakamoto, Y., Koide, N., Takahara, T., Shimizu, T., Iyoda, M. & Ogawara, K. (2017) TEAD4-YAP interaction regulates tumoral growth by controlling cell-cycle arrest at the G1 phase. *Biochemical and biophysical research communications*, 486(2), 385-390.

Tammemagi, M. C., Mayo, J. R. & Lam, S. (2013) Cancer in pulmonary nodules detected on first screening CT. *The New England journal of medicine*, 369(21), 2060.

Tan, Y., Wang, X., Song, H., Zhang, Y., Zhang, R., Li, S., Jin, W., Chen, S., Fang, H., Chen, Z. & Wang, K. (2021) A PML/RARα direct target atlas redefines transcriptional deregulation in acute promyelocytic leukemia. *Blood*, 137(11), 1503-1516.

Tan, Y., You, H., Wu, C., Altomare, D. A. & Testa, J. R. (2010) Appl1 is dispensable for mouse development, and loss of Appl1 has growth factor-selective effects on Akt signaling in murine embryonic fibroblasts. *Journal of Biological Chemistry*, 285(9), 6377-6389.

Tan, Y. S., Kim, M., Kingsbury, T. J., Civin, C. I. & Cheng, W.-C. (2014) Regulation of RAB5C is important for the growth inhibitory effects of MiR-509 in human precursor-B acute lymphoblastic leukemia. *PloS one*, 9(11), e111777.

Tanner, N. T. & Silvestri, G. A. (2015) Screening for lung cancer using low-dose computed tomography. Are we headed for DANTE's paradise or inferno? : American Thoracic Society.

Tanton, H., Sewastianik, T., Seo, H. S., Remillard, D., Pierre, R. S., Bala, P., Aitymbayev, D., Dennis, P., Adler, K., Geffken, E., Yeoh, Z., Vangos, N., Garbicz, F., Scott, D., Sethi, N., Bradner, J., Dhe-Paganon, S. & Carrasco, R. D. (2022) A novel β-catenin/BCL9 complex inhibitor blocks oncogenic Wnt signaling and disrupts cholesterol homeostasis in colorectal cancer. *Sci Adv*, 8(17), eabm3108.

Tasian, S. K., Hurtz, C., Wertheim, G. B., Bailey, N. G., Lim, M. S., Harvey, R. C., Chen, I.-M., Willman, C. L., Astles, R. & Zebrowski, A. (2017) High incidence of Philadelphia chromosome-like acute lymphoblastic leukemia in older adults with B-ALL. *Leukemia*, 31(4), 981-984.

Taylor, G. S., Jia, H., Harrington, K., Lee, L. W., Turner, J., Ladell, K., Price, D. A., Tanday, M., Matthews, J. & Roberts, C. (2014) A recombinant modified vaccinia ankara vaccine encoding Epstein–Barr virus (EBV) target antigens: a phase I trial in UK patients with EBV-positive cancer. *Clinical Cancer Research*, 20(19), 5009-5022.

Taylor, M. R., Slavov, D., Gajewski, A., Vlcek, S., Ku, L., Fain, P. R., Carniel, E., Di Lenarda, A., Sinagra, G. & Boucek, M. M. (2005) Thymopoietin (lamina‐associated polypeptide 2) gene mutation associated with dilated cardiomyopathy. *Human mutation*, 26(6), 566-574.

Taylor, R., Najafi, F. & Dobson, A. (2007) Meta-analysis of studies of passive smoking and lung cancer: effects of study type and continent. *International journal of epidemiology*, 36(5), 1048-1059.

Terasawa, K., Sagae, S., Toyota, M., Tsukada, K., Ogi, K., Satoh, A., Mita, H., Imai, K., Tokino, T. & Kudo, R. (2004) Epigenetic Inactivation of TMS1/ASC in Ovarian Cancer. *Clinical Cancer Research*, 10(6), 2000-2006.

Tesfai, Y., Ford, J., Carter, K. W., Firth, M. J., O'Leary, R. A., Gottardo, N. G., Cole, C. & Kees, U. R. (2012) Interactions between acute lymphoblastic leukemia and bone marrow stromal cells influence response to therapy. *Leukemia Research*, 36(3), 299-306.

Thomas, D. A. & Massagué, J. (2005) TGF-β directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. *Cancer cell*, 8(5), 369-380.

Thomas, K. & Gould, M. (2008) Diagnosis and staging of non-small cell lung cancer. UptoDate.

Thomas, R. M., Nechamen, C. A., Mazurkiewicz, J. E., Ulloa-Aguirre, A. & Dias, J. A. (2011) The adapter protein APPL1 links FSH receptor to inositol 1, 4, 5-trisphosphate production and is implicated in intracellular Ca2+ mobilization. *Endocrinology*, 152(4), 1691-1701.

Thomas, S. N., Tong, Z., Stebe, K. J. & Konstantopoulos, K. (2009) Identification, characterization and utilization of tumor cell selectin ligands in the design of colon cancer diagnostics. *Biorheology*, 46(3), 207-225.

Thomas, S. N., Zhu, F., Schnaar, R. L., Alves, C. S. & Konstantopoulos, K. (2008) Carcinoembryonic antigen and CD44 variant isoforms cooperate to mediate colon carcinoma cell adhesion to E-and L-selectin in shear flow. *Journal of biological chemistry*, 283(23), 15647-15655.

Tokuishi, K., Yamashita, S.-i., Ohbo, K. & Kawahara, K. (2012) Splice variant HE4-V3 expression is associated with favorable prognosis in pulmonary adenocarcinoma. *Tumor Biology*, 33(1), 103-109.

Tomashefski Jr, J., Connors Jr, A., Rosenthal, E. & Hsiue, I.-L. (1990) Peripheral vs central squamous cell carcinoma of the lung. A comparison of clinical features, histopathology, and survival. *Archives of pathology & laboratory medicine*, 114(5), 468-474.

Toumbis, M., Rasidakis, A., Passalidou, E., Dimitroulis, J., Gaga, M., Alchanatis, M., Orphanidou, D. & Jordanoglou, J. (1995) Diagnostic usefulness of 5 tumor-markers in patients with primary lung-cancer. *Oncology reports*, 2(6), 1135-1140.

Tsim, S., O'Dowd, C. A., Milroy, R. & Davidson, S. (2010) Staging of non-small cell lung cancer (NSCLC): A review. *Respiratory Medicine*, 104(12), 1767-1774.

Tsumatori, G., Ozeki, Y., Takagi, K., Ogata, T. & Tanaka, S. (1999) Relation between the Serum E‐Selectin Level and the Survival Rate of Patients with Resected Non‐small Cell Lung Cancers. *Japanese journal of cancer research*, 90(3), 301-307.

Tufman, A., Tian, F. & Huber, R. M. (2013) Can microRNAs improve the management of lung cancer patients? A clinician's perspective. *Theranostics*, 3(12), 953-963.

Tulinsky, L., Dzian, A., Matakova, T. & Ihnat, P. (2022) Overexpression of the miR-143/145 and reduced expression of the let-7 and miR-126 for early lung cancer diagnosis. *J Appl Biomed*, 20(1), 1-6.

Tyagi, P. & Mirakhur, B. (2009) MAGRIT: the largest-ever phase III lung cancer trial aims to establish a novel tumor-specific approach to therapy. *Clinical lung cancer*, 10(5), 371-374.

Umano, Y., Tsunoda, T., Tanaka, H., Matsuda, K., Yamaue, H. & Tanimura, H. (2001) Generation of cytotoxic T cell responses to an HLA-A24 restricted epitope peptide derived from wild-type p53. *British journal of cancer*, 84(8), 1052-1057.

Urieli-Shoval, S., Linke, R. P. & Matzner, Y. (2000) Expression and function of serum amyloid A, a major acute-phase protein, in normal and disease states. *Current opinion in hematology*, 7(1), 64-69.

Usami, Y., Wu, Y. & Göttlinger, H. G. (2015) SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature*, 526(7572), 218-23.

van den Berk, L. C., van der Veer, A., Willemse, M. E., Theeuwes, M. J., Luijendijk, M. W., Tong, W. H., van der Sluis, I. M., Pieters, R. & den Boer, M. L. (2014) Disturbed CXCR 4/CXCL 12 axis in paediatric precursor B‐cell acute lymphoblastic leukaemia. *British journal of haematology*, 166(2), 240-249.

Van Everdink, W., Baranova, A., Lummen, C., Tyazhelova, T., Looman, M., Ivanov, D., Verlind, E., Pestova, A., Faber, H. & van der Veen, A. (2003) RFP2, c13ORF1, and FAM10A4 are the most likely tumor suppressor gene candidates for B-cell chronic lymphocytic leukemia. *Cancer genetics and cytogenetics*, 146(1), 48-57.

Varelas, X., Miller, B. W., Sopko, R., Song, S., Gregorieff, A., Fellouse, F. A., Sakuma, R., Pawson, T., Hunziker, W. & McNeill, H. (2010) The Hippo pathway regulates Wnt/β-catenin signaling. *Developmental cell*, 18(4), 579-591.

Veeck, J. & Dahl, E. (2012) Targeting the Wnt pathway in cancer: The emerging role of Dickkopf-3. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1825(1), 18-28.

Veenstra, C., Karlsson, E., Mirwani, S. M., Nordenskjöld, B., Fornander, T., Pérez-Tenorio, G. & Stål, O. (2019) The effects of PTPN2 loss on cell signalling and clinical outcome in relation to breast cancer subtype. *Journal of Cancer Research and Clinical Oncology*, 145, 1845-1856.

Verma, V., Simone, C. B. & Werner-Wasik, M. (2017) Acute and late toxicities of concurrent chemoradiotherapy for locally-advanced non-small cell lung cancer. *Cancers*, 9(9), 120.

Vicente-Manzanares, M., Zareno, J., Whitmore, L., Choi, C. K. & Horwitz, A. F. (2007) Regulation of protrusion, adhesion dynamics, and polarity by myosins IIA and IIB in migrating cells. *The Journal of cell biology*, 176(5), 573-580.

Vicioso, Y., Gram, H., Beck, R., Asthana, A., Zhang, K., Wong, D. P., Letterio, J. & Parameswaran, R. (2019) Combination Therapy for Treating Advanced Drug-Resistant Acute Lymphoblastic Leukemia. *Cancer Immunology Research*, 7(7), 1106-1119.

Vigneron, N. (2015) Human tumor antigens and cancer immunotherapy. *BioMed research international*, 2015.

Voon, P. J. & Kong, H. L. (2011) Tumour genetics and genomics to personalise cancer treatment. *Annals of the Academy of Medicine-Singapore*, 40(8), 362.

Wahidi, M. M., Govert, J. A., Goudar, R. K., Gould, M. K. & McCrory, D. C. (2007) Evidence for the treatment of patients with pulmonary nodules: when is it lung cancer?: ACCP evidence-based clinical practice guidelines. *Chest*, 132(3), 94S-107S.

Walker, B. A., Leone, P. E., Chiecchio, L., Dickens, N. J., Jenner, M. W., Boyd, K. D., Johnson, D. C., Gonzalez, D., Dagrada, G. P. & Protheroe, R. K. (2010) A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. *Blood, The Journal of the American Society of Hematology*, 116(15), e56-e65.

Walsh, G. L. (2016) Pancreatic cancer patients and thoracic surgeons: Could/should their clinic templates ever overlap? *The Journal of thoracic and cardiovascular surgery*, 1(152), 83-84.

Walter, J. E., Heuvelmans, M. A., de Jong, P. A., Vliegenthart, R., van Ooijen, P. M., Peters, R. B., ten Haaf, K., Yousaf-Khan, U., van der Aalst, C. M. & de Bock, G. H. (2016) Occurrence and lung cancer probability of new solid nodules at incidence screening with low-dose CT: analysis of data from the randomised, controlled NELSON trial. *The Lancet Oncology*, 17(7), 907-916.

Walter, S., Weinschenk, T., Stenzl, A., Zdrojowy, R., Pluzanska, A., Szczylik, C., Staehler, M., Brugger, W., Dietrich, P.-Y. & Mendrzyk, R. (2012) Multipeptide immune response to cancer vaccine IMA901 after single-dose cyclophosphamide associates with longer patient survival. *Nature medicine*, 18(8), 1254-1261.

Walz, A., Ooms, A., Gadd, S., Gerhard, D., Smith, M., GuidryáAuvil, J. M., Meerzaman, D., Chen, Q.-R., Hsu, C. & Yan, C. (2015) Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. *Cancer cell*, 27(2), 286-297.

Wan, L., Liu, Q., Liang, D., Guo, Y., Liu, G., Ren, J., He, Y. & Shan, B. (2021) Circulating tumor cell and metabolites as novel biomarkers for early-stage lung cancer diagnosis. *Frontiers in Oncology*, 11, 630672.

Wan, Z. H., Ma, Y. H., Jiang, T. Y., Lin, Y. K., Shi, Y. Y., Tan, Y. X., Dong, L. W. & Wang, H. Y. (2019) Six2 is negatively correlated with prognosis and facilitates epithelial-mesenchymal transition via TGF-β/Smad signal pathway in hepatocellular carcinoma. *Hepatobiliary Pancreat Dis Int*, 18(6), 525-531.

Wang, C., Zhu, X., Feng, W., Yu, Y., Jeong, K., Guo, W., Lu, Y. & Mills, G. B. (2016a) Verteporfin inhibits YAP function through up-regulating 14-3-3σ sequestering YAP in the cytoplasm. *American journal of cancer research*, 6(1), 27.

Wang, F., Demir, S., Gehringer, F., Osswald, C. D., Seyfried, F., Enzenmüller, S., Eckhoff, S. M., Maier, T., Holzmann, K., Debatin, K.-M., Wirth, T., Meyer, L. H. & Ushmorov, A. (2018a) Tight regulation of FOXO1 is essential for maintenance of B-cell precursor acute lymphoblastic leukemia. *Blood*, 131(26), 2929-2942.

Wang, G., Yu, X., Xia, J., Sun, J., Huang, H. & Liu, Y. (2021a) MicroRNA-9 restrains the sharp increase and boost apoptosis of human acute myeloid leukemia cells by adjusting the Hippo/YAP signaling pathway. *Bioengineered*, 12(1), 2906-2914.

Wang, H., Song, C., Ding, Y., Pan, X., Ge, Z., Tan, B.-H., Gowda, C., Sachdev, M., Muthusami, S. & Ouyang, H. (2016b) Transcriptional regulation of JARID1B/KDM5B histone demethylase by ikaros, histone deacetylase 1 (HDAC1), and casein kinase 2 (CK2) in B-cell acute lymphoblastic leukemia. *Journal of Biological Chemistry*, 291(8), 4004-4018.

Wang, J., Jiang, D., Li, Z., Yang, S., Zhou, J., Zhang, G., Zhang, Z., Sun, Y., Zhang, Z., Li, X., Tao, L., Shi, J., Lu, Y., Zheng, L., Song, C. & Yang, K. (2020a) BCAP31, a cancer/testis antigen-like protein, can act as a probe for non-small-cell lung cancer metastasis. *Scientific Reports*, 10(1), 4025.

Wang, J., Jiang, W., Zhang, T., Liu, L., Bi, N., Wang, X., Hui, Z., Liang, J., Lv, J., Zhou, Z., Xiao, Z., Feng, Q., Chen, D., Yin, W. & Wang, L. (2018b) Increased CYFRA 21-1, CEA and NSE are Prognostic of Poor Outcome for Locally Advanced Squamous Cell Carcinoma in Lung: A Nomogram and Recursive Partitioning Risk Stratification Analysis. *Translational Oncology*, 11(4), 999-1006.

Wang, J., Xu, P., Hao, Y., Yu, T., Liu, L., Song, Y. & Li, Y. (2021b) Interaction between DNMT3B and MYH11 via hypermethylation regulates gastric cancer progression. *BMC Cancer*, 21(1), 914.

Wang, K., Song, X., Wang, S., Li, X., Zhang, Z., Xie, L. & Song, X. (2022a) Plasma SNORD42B and SNORD111 as potential biomarkers for early diagnosis of non-small cell lung cancer. *J Clin Lab Anal*, 36(11), e24740.

Wang, N., Bai, X., Wang, X., Wang, D., Ma, G., Zhang, F., Ye, J., Lu, F. & Ji, C. (2023) A Novel Fatty Acid Metabolism-Associated Risk Model for Prognosis Prediction in Acute Myeloid Leukaemia. *Current Oncology*, 30(2), 2524-2542.

Wang, R., Wang, G., Zhang, N., Li, X. & Liu, Y. (2013) Clinical evaluation and cost-effectiveness analysis of serum tumor markers in lung cancer. *BioMed research international*, 2013.

Wang, R. J., Wu, P., Cai, G. X., Wang, Z. M., Xu, Y., Peng, J. J., Sheng, W. Q., Lu, H. F. & Cai, S. J. (2014) Down-regulated MYH11 expression correlates with poor prognosis in stage II and III colorectal cancer. *Asian Pac J Cancer Prev*, 15(17), 7223-8.

Wang, S., Peng, L., Xu, W., Zhou, Y., Zhu, Z., Kong, Y., Leung, S., Wang, J., Yan, X. & Mi, J.-Q. (2022b) Preclinical characterization and comparison between CD3/CD19 bispecific and novel CD3/CD19/CD20 trispecific antibodies against B-cell acute lymphoblastic leukemia: targeted immunotherapy for acute lymphoblastic leukemia. *Frontiers of Medicine*, 16(1), 139-149.

Wang, T. & Hong, W. (2002) Interorganellar regulation of lysosome positioning by the Golgi apparatus through Rab34 interaction with Rab-interacting lysosomal protein. *Molecular biology of the cell*, 13(12), 4317-4332.

Wang, W., Chen, D., Chen, W., Xin, Z., Huang, Z., Zhang, X., Xi, K., Wang, G., Zhang, R., Zhao, D., Liu, L. & Zhang, L. (2020b) Early Detection of Non-Small Cell Lung

Cancer by Using a 12-microRNA Panel and a Nomogram for Assistant Diagnosis. *Frontiers in Oncology*, 10.

Wang, X., Zhi, X., Yang, Z., Tian, H., Li, Y., Li, M., Zhao, W., Zhang, C., Wang, T., Liu, J., Shen, D., Zheng, C., Zhao, D., Yang, S., Qi, J., Xin, H., Stojadinovic, A., Avital, I., Lee, L. J., Rao, J. & Zhang, W. (2017) A novel serum based biomarker panel has complementary ability to preclude presence of early lung cancer for low dose CT (LDCT). *Oncotarget*, 8(28).

Wang, Y., Tong, X., Li, G., Li, J., Deng, M. & Ye, X. (2012) Ankrd17 positively regulates RIG‐I‐like receptor (RLR)‐mediated immune signaling. *European journal of immunology*, 42(5), 1304-1315.

Wang, Y., Wo, Y., Lu, T., Sun, X., Liu, A., Dong, Y., Du, W., Su, W., Huang, Z. & Jiao, W. (2021c) Circ-AASDH functions as the progression of early stage lung adenocarcinoma by targeting miR-140-3p to activate E2F7 expression. *Transl Lung Cancer Res*, 10(1), 57-70.

Wang, Z., Ying, M., Wu, Q., Wang, R. & Li, Y. (2016c) Overexpression of myosin VI regulates gastric cancer cell progression. *Gene*, 593(1), 100-109.

Waragai, M., Nagamitsu, S., Xu, W., Li, Y. J., Lin, X. & Ashizawa, T. (2006) Ataxin 10 induces neuritogenesis via interaction with G‐protein β2 subunit. *Journal of neuroscience research*, 83(7), 1170-1178.

Ward, J. P., Gubin, M. M. & Schreiber, R. D. (2016) The role of neoantigens in naturally occurring and therapeutically induced immune responses to cancer. *Advances in immunology*, 130, 25-74.

Warner, J. R. & McIntosh, K. B. (2009) How common are extraribosomal functions of ribosomal proteins? *Molecular cell*, 34(1), 3-11.

Watkins, P. A., Maiguel, D., Jia, Z. & Pevsner, J. (2007) Evidence for 26 distinct acyl-coenzyme A synthetase genes in the human genomes. *Journal of lipid research*, 48(12), 2736-2750.

Weeraratne, S. D., Amani, V., Neiss, A., Teider, N., Scott, D. K., Pomeroy, S. L. & Cho, Y.-J. (2011) miR-34a confers chemosensitivity through modulation of MAGE-A and p53 in medulloblastoma. *Neuro-oncology*, 13(2), 165-175.

Wei, A. H., Roberts, A. W., Spencer, A., Rosenberg, A. S., Siegel, D., Walter, R. B., Caenepeel, S., Hughes, P., McIver, Z., Mezzi, K., Morrow, P. K. & Stein, A. (2020) Targeting MCL-1 in hematologic malignancies: Rationale and progress. *Blood Rev*, 44, 100672.

Wei, C., Wu, S., Li, X., Wang, Y., Ren, R., Lai, Y. & Ye, J. (2013) High expression of FER tyrosine kinase predicts poor prognosis in clear cell renal cell carcinoma. *Oncology letters*, 5(2), 473-478.

Welch, H. G., Prorok, P. C., O'Malley, A. J. & Kramer, B. S. (2016) Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness. *New England Journal of Medicine*, 375(15), 1438-1447.

Wen, J. L., Wen, X. F., Li, R. B., Jin, Y. C., Wang, X. L., Zhou, L. & Chen, H. X. (2015) UBE3C promotes growth and metastasis of renal cell carcinoma via activating Wnt/β-catenin pathway. *PLoS One*, 10(2), e0115622.

Wen, Y., Zhou, X., Lu, M., He, M., Tian, Y., Liu, L., Wang, M., Tan, W., Deng, Y. & Yang, X. (2019) Bclaf1 promotes angiogenesis by regulating HIF-1α transcription in hepatocellular carcinoma. *Oncogene*, 38(11), 1845-1859.

Weon, J. L. & Potts, P. R. (2015) The MAGE protein family and cancer. *Current opinion in cell biology*, 37, 1-8.

White, L. S., Soodgupta, D., Johnston, R. L., Magee, J. A. & Bednarski, J. J. (2018) Bclaf1 Promotes Maintenance and Self-Renewal of Fetal Hematopoietic Stem Cells. *Blood*, 132, 1269.

Whitehurst, A. W., Xie, Y., Purinton, S. C., Cappell, K. M., Swanik, J. T., Larson, B., Girard, L., Schorge, J. O. & White, M. A. (2010) Tumor antigen acrosin binding protein normalizes mitotic spindle function to promote cancer cell proliferation. *Cancer research*, 70(19), 7652-7661.

Wieskopf, B., Demangeat, C., Purohit, A., Stenger, R., Gries, P., Kreisman, H. & Quoix, E. (1995) Cyfra 21-1 as a biologic marker of non-small cell lung cancer: evaluation of sensitivity, specificity, and prognostic role. *Chest*, 108(1), 163-169.

Wilkinson, S. (2019) ER‐phagy: shaping up and destressing the endoplasmic reticulum. *The FEBS journal*, 286(14), 2645-2663.

Wille, M. M., Dirksen, A., Ashraf, H., Saghir, Z., Bach, K. S., Brodersen, J., Clementsen, P. F., Hansen, H., Larsen, K. R. & Mortensen, J. (2016) Results of the randomized Danish lung cancer screening trial with focus on high-risk profiling. *American journal of respiratory and critical care medicine*, 193(5), 542-551.

Winer-Muram, H. T., Jennings, S. G., Tarver, R. D., Aisen, A. M., Tann, M., Conces, D. J. & Meyer, C. A. (2002) Volumetric growth rate of stage I lung cancer prior to treatment: serial CT scanning. *Radiology*, 223(3), 798-805.

Wong, P., Yeoh, C., Ahmad, A., Chelala, C., Gillett, C., Speirs, V., Jones, J. & Hurst, H. (2014) Identification of MAGEA antigens as causal players in the development of tamoxifen-resistant breast cancer. *Oncogene*, 33(37), 4579-4588.

Woo, H.-H., Yi, X., Lamb, T., Menzl, I., Baker, T., Shapiro, D. J. & Chambers, S. K. (2011) Posttranscriptional suppression of proto-oncogene c-fms expression by vigilin in breast cancer. *Molecular and cellular biology*, 31(1), 215-225.

Wood, D. E., Kazerooni, E. A., Baum, S. L., Eapen, G. A., Ettinger, D. S., Hou, L., Jackman, D. M., Klippenstein, D., Kumar, R. & Lackner, R. P. (2018) Lung cancer

screening, version 3.2018, NCCN clinical practice guidelines in oncology. *Journal of the National Comprehensive Cancer Network*, 16(4), 412-441.

Woods, K., Pasam, A., Jayachandran, A., Andrews, M. C. & Cebon, J. (2014) Effects of epithelial to mesenchymal transition on T cell targeting of melanoma cells. *Frontiers in oncology*, 4, 367.

Wu, H.-Y., Pan, Y.-Y., Kopylov, A. T., Zgoda, V., Ma, M.-C., Wang, C.-H., Su, W.-C., Lai, W.-W., Cheng, P.-N. & Liao, P.-C. (2020a) Assessment of Serological Early Biomarker Candidates for Lung Adenocarcinoma by using Multiple Reaction Monitoring-Mass Spectrometry. *Proteomics. Clinical applications*, 14(4), e1900095.

Wu, J., Lu, Y., Qin, A., Qiao, Z. & Jiang, X. (2017) Overexpression of RAB34 correlates with poor prognosis and tumor progression in hepatocellular carcinoma. *Oncol Rep*, 38(5), 2967-2974.

Wu, Q., Yu, L., Lin, X., Zheng, Q., Zhang, S., Chen, D., Pan, X. & Huang, Y. (2020b) Combination of serum miRNAs with serum exosomal miRNAs in early diagnosis for non-small-cell lung cancer. *Cancer management and research*, 12, 485.

Wu, R., Yang, H., Wan, J., Deng, X., Chen, L., Hao, S. & Ma, L. (2018) Knockdown of the Hippo transducer YAP reduces proliferation and promotes apoptosis in the Jurkat leukemia cell. *Molecular Medicine Reports*, 18(6), 5379-5388.

Wu, Y., Jing, H. & Zhang, J. (2022) MicroRNA-340 and MicroRNA-450b-5p: Plasma Biomarkers for Detection of Non-Small-Cell Lung Cancer. *J Environ Public Health*, 2022, 8024700.

Xia, L.-l., Tang, Y.-b., Song, F.-f., Xu, L., Ji, P., Wang, S.-j., Zhu, J.-m., Zhang, Y., Zhao, G.-p. & Wang, Y. (2016) DCTPP1 attenuates the sensitivity of human gastric cancer cells to 5-fluorouracil by up-regulating MDR1 expression epigenetically. *Oncotarget*, 7(42), 68623.

Xiao, F., Liu, D., Guo, Y., Shi, B., Song, Z., Tian, Y. & Liang, C. (2013) Novel and convenient method to evaluate the character of solitary pulmonary nodule-comparison of three mathematical prediction models and further stratification of risk factors. *PloS one*, 8(10), e78271.

Xie, M., Li, T., Li, N., Li, J., Yao, Q., Han, W. & Ruan, G. (2015) VSTM-v1, a potential myeloid differentiation antigen that is downregulated in bone marrow cells from myeloid leukemia patients. *J Hematol Oncol*, 8, 25.

Xing, L., Su, J., Guarnera, M. A., Zhang, H., Cai, L., Zhou, R., Stass, S. A. & Jiang, F. (2015) Sputum microRNA biomarkers for identifying lung cancer in indeterminate solitary pulmonary nodules. *Clinical Cancer Research*, 21(2), 484-489.

Xing, S., Zeng, T., Xue, N., He, Y., Lai, Y.-z., Li, H.-l., Huang, Q., Chen, S.-l. & Liu, W.-l. (2019) Development and validation of tumor-educated blood platelets integrin alpha 2b (ITGA2B) RNA for diagnosis and prognosis of non-small-cell lung cancer through RNA-seq. *International journal of biological sciences*, 15(9), 1977.

Xu, N., Li, Y. L., Li, X., Zhou, X., Cao, R., Li, H., Li, L., Lu, Z. Y., Huang, J. X., Fan, Z. P., Huang, F., Zhou, H. S., Zhang, S., Liu, Z., Zhu, H. Q., Liu, Q. F. & Liu, X. L. (2016) Correlation between deletion of the CDKN2 gene and tyrosine kinase inhibitor resistance in adult Philadelphia chromosome-positive acute lymphoblastic leukemia. *J Hematol Oncol*, 9, 40.

Xu, X., Wang, Y., Bryce, N. S., Tang, K., Meagher, N. S., Kang, E. Y., Kelemen, L. E., Köbel, M., Ramus, S. J. & Friedlander, M. (2021) Targeting the actin/tropomyosin cytoskeleton in epithelial ovarian cancer reveals multiple mechanisms of synergy with anti-microtubule agents. *British Journal of Cancer*, 125(2), 265-276.

Xu, Y.-Z., Zhang, H., Yang, Q.-Z., Sun, R.-L., Wang, K., Shi, Y.-Y. & Wang, Z. (2020) Integrated regulatory-metabolic network model reveals critical mechanism and potential targets for Hepatocellular Carcinoma.

Xu, Y., Xu, X., Ni, X., Pan, J., Chen, M., Lin, Y., Zhao, Z., Zhang, L., Ge, N. & Song, G. (2023) Gene-based cancer-testis antigens as prognostic indicators in hepatocellular carcinoma. *Heliyon*, 9(3).

Xue, Y., Cong, W., Xie, S., Shu, J., Feng, G. & Gao, H. (2018) Folate-receptor-positive circulating tumor cells as an efficacious biomarker for the diagnosis of small pulmonary nodules. *Journal of cancer research and therapeutics*, 14(7), 1620.

Yanez, L., Bermudez, A., Richard, C., Bureo, E. & Iriondo, A. (2009) Successful induction therapy with decitabine in refractory childhood acute lymphoblastic leukemia. *Leukemia*, 23(7), 1342-1343.

Yang, B., O'Herrin, S. M., Wu, J., Reagan-Shaw, S., Ma, Y., Bhat, K. M., Gravekamp, C., Setaluri, V., Peters, N. & Hoffmann, F. M. (2007) MAGE-A, mMage-b, and MAGE-C proteins form complexes with KAP1 and suppress p53-dependent apoptosis in MAGE-positive cell lines. *Cancer research*, 67(20), 9954-9962.

Yang, B., Ren, N., Guo, B., Xin, H. & Yin, Y. (2020a) Measuring serum human epididymis secretory protein autoantibody as an early biomarker of lung cancer. *Transl Cancer Res*, 9, 735-41.

Yang, C., Wu, S., Wu, X., Zhou, X., Jin, S. & Jiang, H. (2019a) Silencing circular RNA UVRAG inhibits bladder cancer growth and metastasis by targeting the microRNA-223/fibroblast growth factor receptor 2 axis. *Cancer Science*, 110(1), 99-106.

Yang, F., Jing, F., Li, Y., Kong, S., Zhang, S., Huo, Y., Huang, X. & Yu, S. (2022) Plasma lncRNA LOC338963 and mRNA AP3B2 are upregulated in paraneoplastic Lambert-Eaton myasthenic syndrome. *Muscle Nerve*, 66(2), 216-222.

Yang, F., Zhou, X., Miao, X., Zhang, T., Hang, X., Tie, R., Liu, N., Tian, F., Wang, F. & Yuan, J. (2014a) MAGEC2, an epithelial-mesenchymal transition inducer, is associated with breast cancer metastasis. *Breast cancer research and treatment*, 145, 23-32.

Yang, J., Lin, L., Jiang, W., Wu, J. & Lin, X. (2020b) Early diagnosis of lung cancer in the elderly using four tumor markers and serum ferritin for better surgical management. *Asian journal of surgery*, 43(11), 1088-1089.

Yang, W. L., Wei, L., Huang, W. Q., Li, R., Shen, W. Y., Liu, J. Y., Xu, J. M., Li, B. & Qin, Y. (2014b) Vigilin is overexpressed in hepatocellular carcinoma and is required for HCC cell proliferation and tumor growth. *Oncol Rep*, 31(5), 2328-2334.

Yang, Y., Mallampati, S., Sun, B., Zhang, J., Kim, S.-B., Lee, J.-S., Gong, Y., Cai, Z. & Sun, X. (2013) Wnt pathway contributes to the protection by bone marrow stromal cells of acute lymphoblastic leukemia cells and is a potential therapeutic target. *Cancer letters*, 333(1), 9-17.

Yang, Z., Qi, W., Sun, L., Zhou, H., Zhou, B. & Hu, Y. (2019b) DNA methylation analysis of selected genes for the detection of early-stage lung cancer using circulating cell-free DNA. *Advances in Clinical and Experimental Medicine*, 28(3), 355-360.

Yarchoan, M., Johnson, B. A., Lutz, E. R., Laheru, D. A. & Jaffee, E. M. (2017) Targeting neoantigens to augment antitumour immunity. *Nature Reviews Cancer*, 17(4), 209-222.

Yin, J.-Y., Zhang, J.-T., Zhang, W., Zhou, H.-H. & Liu, Z.-Q. (2018) eIF3a: a new anticancer drug target in the eIF family. *Cancer letters*, 412, 81-87.

Yip, R., Yankelevitz, D. F., Hu, M., Li, K., Xu, D. M., Jirapatnakul, A. & Henschke, C. I. (2016) Lung cancer deaths in the National Lung Screening Trial attributed to nonsolid nodules. *Radiology*, 281(2), 589-596.

Yokoyama, A., Kohno, N., Kondo, K., Ueda, S., Hirasawa, Y., Watanabe, K., Takada, Y. & Hiwada, K. (1998) Comparative evaluation of sialylated carbohydrate antigens, KL-6, CA19-9 and SLX as serum markers for interstitial pneumonia. *Respirology (Carlton, Vic.)*, 3(3), 199-202.

Yonemori, K., Tateishi, U., Uno, H., Yonemori, Y., Tsuta, K., Takeuchi, M., Matsuno, Y., Fujiwara, Y., Asamura, H. & Kusumoto, M. (2007) Development and validation of diagnostic prediction model for solitary pulmonary nodules. *Respirology*, 12(6), 856-862.

Yoshida, M., Horiguchi, H., Kikuchi, S., Iyama, S., Ikeda, H., Goto, A., Kawano, Y., Murase, K., Takada, K. & Miyanishi, K. (2019) miR-7977 inhibits the Hippo-YAP signaling pathway in bone marrow mesenchymal stromal cells. *PloS one*, 14(3), e0213220.

Yoshimura, K., Suzuki, Y., Inoue, Y., Tsuchiya, K., Karayama, M., Iwashita, Y., Kahyo, T., Kawase, A., Tanahashi, M. & Ogawa, H. (2020) CD200 and CD200R1 are differentially expressed and have differential prognostic roles in non-small cell lung cancer. *Oncoimmunology*, 9(1), 1746554.

Young, R. P., Duan, F., Chiles, C., Hopkins, R. J., Gamble, G. D., Greco, E. M., Gatsonis, C. & Aberle, D. (2015) Airflow limitation and histology shift in the National Lung Screening Trial. The NLST-ACRIN cohort substudy. *American journal of respiratory and critical care medicine*, 192(9), 1060-1067.

Yu, D., Du, K., Liu, T. & Chen, G. (2013a) Prognostic Value of Tumor Markers, NSE, CA125 and SCC, in Operable NSCLC Patients. *International Journal of Molecular Sciences*, 14(6), 11145-11156.

Yu, F.-X., Zhang, Y., Park, H. W., Jewell, J. L., Chen, Q., Deng, Y., Pan, D., Taylor, S. S., Lai, Z.-C. & Guan, K.-L. (2013b) Protein kinase A activates the Hippo pathway to modulate cell proliferation and differentiation. *Genes & development*, 27(11), 1223-1232.

Yu, K., Wang, J., Lu, T., Ma, D., Wei, D., Guo, Y., Cheng, B., Wang, W. & Fang, Q. (2019) Overexpression of heme oxygenase‐1 in microenvironment mediates vincristine resistance of B‐cell acute lymphoblastic leukemia by promoting vascular endothelial growth factor secretion. *Journal of Cellular Biochemistry*, 120(10), 17791-17810.

Yu, L., Todd, N. W., Xing, L., Xie, Y., Zhang, H., Liu, Z., Fang, H., Zhang, J., Katz, R. L. & Jiang, F. (2010) Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers. *International journal of cancer*, 127(12), 2870-2878.

Yu, W., Wang, L., Liu, S., Liu, Y., Wang, S. & Sun, X. (2023) Combination of serum ACSL4 levels and low-dose 256-slice spiral CT exhibits the potential in the early screening of lung cancer. *Medicine (Baltimore)*, 102(5), e32733.

Yu, Z., Zhang, G., Yang, M., Zhang, S., Zhao, B., Shen, G. & Chai, Y. (2017) Systematic review of CYFRA 21-1 as a prognostic indicator and its predictive correlation with clinicopathological features in Non-small Cell Lung Cancer: A meta-analysis. *Oncotarget*, 8(3), 4043.

Yuan, Z., Wang, L., Hong, S., Shi, C. & Yuan, B. (2022) Diagnostic value of HSP90α and related markers in lung cancer. *J Clin Lab Anal*, 36(6), e24462.

Yuki, K., Cheng, N., Nakano, M. & Kuo, C. J. (2020) Organoid models of tumor immunology. *Trends in immunology*, 41(8), 652-664.

Zammarchi, F., Corbett, S., Adams, L., Tyrer, P. C., Kiakos, K., Janghra, N., Marafioti, T., Britten, C. E., Havenith, C. E. & Chivers, S. (2018) ADCT-402, a PBD dimer–containing antibody drug conjugate targeting CD19-expressing malignancies. *Blood, The Journal of the American Society of Hematology*, 131(10), 1094-1105.

Zang, R., Li, Y., Jin, R., Wang, X., Lei, Y., Che, Y., Lu, Z., Mao, S., Huang, J., Liu, C., Zheng, S., Zhou, F., Wu, Q., Gao, S., Sun, N. & He, J. (2019) Enhancement of diagnostic performance in lung cancers by combining CEA and CA125 with autoantibodies detection. *Oncoimmunology*, 8(10), e1625689.

Zeigerer, A., Gilleron, J., Bogorad, R. L., Marsico, G., Nonaka, H., Seifert, S., Epstein-Barash, H., Kuchimanchi, S., Peng, C. G. & Ruda, V. M. (2012) Rab5 is necessary for the biogenesis of the endolysosomal system in vivo. *Nature*, 485(7399), 465-470.

Zeng, Y., He, Y., Yang, F., Mooney, S. M., Getzenberg, R. H., Orban, J. & Kulkarni, P. (2011) The cancer/testis antigen prostate-associated gene 4 (PAGE4) is a highly intrinsically disordered protein. *Journal of Biological Chemistry*, 286(16), 13985-13994.

Zhai, J. S., Song, J. G., Zhu, C. H., Wu, K., Yao, Y. & Li, N. (2016) Expression of APPL1 is correlated with clinicopathologic characteristics and poor prognosis in patients with gastric cancer. *Curr Oncol*, 23(2), e95-e101.

Zhang, H., Peng, C., Hu, Y., Li, H., Sheng, Z., Chen, Y., Sullivan, C., Cerny, J., Hutchinson, L., Higgins, A., Miron, P., Zhang, X., Brehm, M. A., Li, D., Green, M. R. & Li, S. (2012) The Blk pathway functions as a tumor suppressor in chronic myeloid leukemia stem cells. *Nature genetics*, 44(8), 861-871.

Zhang, J. & Oak, J. (2023) Challenges of Detecting Measurable/Minimal Disease in Acute Leukemia, *Seminars in Diagnostic Pathology*. Elsevier.

Zhang, L., Wang, G., Chen, S., Ding, J., Ju, S., Cao, H. & Tian, H. (2016) Depletion of thymopoietin inhibits proliferation and induces cell cycle arrest/apoptosis in glioblastoma cells. *World J Surg Oncol,* 14(1), 267.

Zhang, Q., Zhang, J., Fu, Z., Dong, L., Tang, Y., Xu, C., Wang, H., Zhang, T., Wu, Y. & Dong, C. (2019) Hypoxia-induced microRNA-10b-3p promotes esophageal squamous cell carcinoma growth and metastasis by targeting TSGA10. *Aging (Albany NY)*, 11(22), 10374.

Zhang, S. (2017) *Regulation of FYN phosphorylation by the PTPN23 tumor suppressor phosphatase in breast tumorigenesis*State University of New York at Stony Brook.

Zhang, W., Tao, H., Chen, X., Sugimura, H., Wang, J. & Zhou, P. (2017a) High expression of Wls is associated with lymph node metastasis and advanced TNM stage in gastric carcinomas. *Pathol Int*, 67(3), 141-146.

Zhang, X., Li, J., Wang, Y., Liu, M., Liu, F., Zhang, X., Pei, L., Wang, T., Jiang, D., Wang, X., Zhang, J. & Dai, L. (2022a) A Diagnostic Model With IgM Autoantibodies and Carcinoembryonic Antigen for Early Detection of Lung Adenocarcinoma. *Frontiers in immunology*, 12, 728853.

Zhang, X., Tan, J., Chen, Y., Ma, S., Bai, W., Peng, Y. & Shi, G. (2022b) Identification of serum MiRNAs as candidate biomarkers for non-small cell lung cancer diagnosis. *BMC Pulm Med*, 22(1), 479.

Zhang, Y., Li, H., Cao, R., Sun, L., Wang, Y., Fan, S., Zhao, Y., Kong, D., Cui, L., Lin, L., Wang, K., Li, Y. & Zhou, J. (2017b) Suppression of miR-708 inhibits the Wnt/β-catenin signaling pathway by activating DKK3 in adult B-all. *Oncotarget*, 8(38), 64114-64128.

Zhang, Y., Shen, H., Withers, H. G., Yang, N., Denson, K. E., Mussell, A. L., Truskinovsky, A., Fan, Q., Gelman, I. H. & Frangou, C. (2017c) VGLL4 selectively represses YAP-dependent gene induction and tumorigenic phenotypes in breast cancer. *Scientific reports*, 7(1), 6190.

Zhang, Y., Yu, Y., Su, X. & Lu, Y. (2021) HOXD8 inhibits the proliferation and migration of triple-negative breast cancer cells and induces apoptosis in them through regulation of AKT/mTOR pathway. *Reprod Biol*, 21(4), 100544.

Zhang, Z., Zhou, L., Xie, N., Nice, E. C., Zhang, T., Cui, Y. & Huang, C. (2020) Overcoming cancer therapeutic bottleneck by drug repurposing. *Signal transduction and targeted therapy*, 5(1), 113.

Zhao, Z., Liu, M., Xu, Z., Cai, Y., Peng, B., Liang, Q., Yan, Y., Liu, W., Kang, F., He, Q., Hong, Q., Zhang, W., Li, J., Peng, J. & Zeng, S. (2022) Identification of ACSF gene family as therapeutic targets and immune-associated biomarkers in hepatocellular carcinoma. *Aging*, 14(19), 7926-7940.

Zhong, L., Coe, S. P., Stromberg, A. J., Khattar, N. H., Jett, J. R. & Hirschowitz, E. A. (2006) Profiling Tumor-Associated Antibodies for Early Detection of Non-small Cell Lung Cancer. *Journal of Thoracic Oncology*, 1(6), 513-519.

Zhong, M., Zhang, Y., Pan, Z., Wang, W., Zhang, Y., Weng, Y., Huang, H., He, Y. & Liu, O. (2021) Clinical Utility of Circulating Tumor Cells in the Early Detection of Lung Cancer in Patients with a Solitary Pulmonary Nodule. *Technology in Cancer Research & Treatment*, 20.

Zhou, H., Yao, Y., Li, Y., Guo, N., Zhang, H., Wang, Z., Chen, Y. & Dai, G. (2022) Identification of Small Nucleolar RNA SNORD60 as a Potential Biomarker and Its Clinical Significance in Lung Adenocarcinoma. *Biomed Res Int*, 2022, 5501171.

Zhou, M., Kong, Y., Wang, X., Li, W., Chen, S., Wang, L., Wang, C. & Zhang, Q. (2021a) LC-MS/MS-based quantitative proteomics analysis of different stages of non-small-cell lung cancer. *BioMed Research International*, 2021.

Zhou, P. L., Wu, Z., Zhang, W., Xu, M., Ren, J., Zhang, Q., Sun, Z. & Han, X. (2021b) Circular RNA hsa_circ_0000277 sequesters miR-4766-5p to upregulate LAMA1 and promote esophageal carcinoma progression. *Cell death & disease*, 12(7), 676.

Zhou, Z., Hu, T., Xu, Z., Lin, Z., Zhang, Z., Feng, T., Zhu, L., Rong, Y., Shen, H. & Luk, J. M. (2015) Targeting Hippo pathway by specific interruption of YAP‐TEAD interaction using cyclic YAP‐like peptides. *The FASEB Journal*, 29(2), 724-732.

Zhu, C., Li, L. & Zhao, B. (2015) The regulation and function of YAP transcription co-activator. *Acta biochimica et biophysica Sinica*, 47(1), 16-28.

Zhu, F., Yan, P., Zhang, J., Cui, Y., Zheng, M., Cheng, Y., Guo, Y., Yang, X., Guo, X. & Zhu, H. (2019) Deficiency of TPPP2, a factor linked to oligoasthenozoospermia, causes subfertility in male mice. *J Cell Mol Med*, 23(4), 2583-2594.

Zhu, J., Weng, Y., Wang, F. & Zhao, J. (2022) LINC00665/miRNAs axis-mediated collagen type XI alpha 1 correlates with immune infiltration and malignant phenotypes in lung adenocarcinoma. *Open Medicine*, 17(1), 1259-1274.

Zhu, X. N., Wei, Y. S., Yang, Q., Liu, H. R., Zhi, Z., Zhu, D., Xia, L., Hong, D. L., Yu, Y. & Chen, G. Q. (2023) FBXO22 promotes leukemogenesis by targeting BACH1 in MLL-rearranged acute myeloid leukemia. *J Hematol Oncol*, 16(1), 9.

Chapter 9: Appendices

9.1 Appendix I. Protocol for the Systematic Literature Review

Objectives

To report the diagnostic performance of biomarkers for NSCLCs including the sensitivity and specificity of tumour antigens, autoantibodies, miRNA, mRNA and ctDNA as diagnostic biomarkers for lung cancer

Description of the condition

Lung cancer is the leading cause of cancer mortality world-wide with high incidence.

Why it is important to do this review?

A previous systematic literature review by Yang et al. (2019) focussed on autoantibodies as diagnostic biomarkers for lung cancer. The review was limited to autoantibodies only and using ELISA as the method of detection. This review focused on biomarkers that can facilitate early detection of lung cancer including antigens, autoantibodies, mRNA, miRNA, and tumour cell markers using different methods as listed in the research article.

Inclusion criteria

- Human studies
- Mentioned the method of analysis
- Human subjects more than 10 to calculate the statistics
- Mentioned the cut-off value of the antigen and healthy volunteers' inclusion
- Measurement of biomarkers in body fluids including blood (serum or plasma), urine, sputum, pleural effusion, and bronchial lavage
- Studies that focused on NSCLC and high-risk individuals are considered for further analysis

Exclusion criteria

- Animal studies
- Cell-line studies
- Workshop and conferences
- Reviews
- Unrelated studies such as those that use antigen as a vaccine or immunotherapy
- Omics data such as proteomic and transcriptomics that haven't mentioned the sensitivity /specificity/AUC of biomarkers identified
- SCLC was excluded as it constituted about 20% of LC

Search methods for identification of studies

All searches performed using the following keywords:-

(cancer* or tumor* or tumour* or neoplasm* or carcinoma* or malignancy*) AND (lung* or pulmonary) AND (antigen* OR protein* OR RNA* OR ctDNA* OR miRNA* OR cell surface marker* OR inflammatory cell*) AND (early detection OR early diagnosis OR early biomarker OR early marker)

Electronic searches

Searched the following online databases:-

PubMed

MEDLINE

CINAHL Complete

Scopus

Web of science

Cochrane Library

Clinical trial.gov

Data collection and analysis

Selection of studies

Duplicates of studies were removed from all databases using Endnote. Based on the title and abstract, all retrieved studies were screened and assessed for relevance by two independent contributors. Then, the remaining studies were screened by compatibility of the full texts with the inclusion and exclusion criteria. The points of disagreement between the contributors were addressed by a third contributor.

Risk of bias

QUADAS was used to assess risk of bias. It was based on comparing the accuracy of both the index test and the reference in the population for a diagnosis of the lung cancer. The index test refers to the test of interest and the reference test refers to the existing diagnostic test that has been historically the standard test for the diagnosis. The following questions were answered for each study:

- ✓ Were a consecutive or random sample of patients enrolled?
- ✓ Was a case control design avoided?
- ✓ Did the study avoid inappropriate exclusions?
- ✓ Were the index test results interpreted without knowledge of the results of the reference standard? If a threshold was used, was it pre-specified?
- ✓ Was the reference standard likely to correctly classify the target condition?
- ✓ Were the reference standard results interpreted without knowledge of the results of the index test?
- ✓ Was there an appropriate interval between index test and reference standard?
- ✓ Did all patients receive the same reference standard?
- ✓ Were all patients included in the analysis?

The items will be provided as yes, no, unclear or not applicable.

Data extraction and management

The data was extracted from the remaining studies independently the following data from published reports using a standardized form.

- ✓ Histological type, and tumour stage
- ✓ Study methodology: techniques used.
- ✓ Specimen type including blood or sputum
- ✓ Participants: sample size
- ✓ Outcomes
- o For sensitivity extracted.
  Sensitivity = number of true-positive results x100%/ number of true-positive results + false-negative results number
- o For specificity extracted,
  Specificity = number of true-negative results x 100% /number of true-negative results + false-positive results number

Table 9.1 Risk of bias assessment of the studies meeting inclusion and exclusion criteria

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Abou-Zeid et al. (2023) (Abou-Zeid et al., 2023)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Yes (histopathological exam)	Y	Low	Low	Unclear	Y (likely)	Y	Unclear	High	Low
Ajona et al., 2021(Ajona et al., 2021)	N	N	Y	High	Low	Y	NA	Low	Low	Unclear	Y (likely)	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Bigbee et al. (2012) (Bigbee et al., 2012)	N	N	Y	High	Low	Y	NA	Low	Low	Yes (pathology review)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Carrozi et al. (2017) (Carozz	Y	N (cases from interve	Y	High	Low	Unclear	NA	Unclear	Low	Unclear	Y	Unclear	Low	Unclear	Y (likely)	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
i et al., 2017)		ntion arm)																	
Cazzoli et al. (2013) (Cazzoli et al., 2013)	N	N	Y	High	Low	Unclear	NA	unclear	Low	Unclear	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Chen et al. (2018) (Chen et al., 2018b)	Unclear	N	Y	Unclear	Low	Unclear	NA	unclear	Low	Y (pathologically confirmed)	Y (likely)	Low	Low	Unclear	Unclear	Y	Unclear	Unclear	Low
Chen et al. (2020) (Chen et al., 2020a)	N	N	Y	High	Low	Unclear	NA	unclear	Low	Unclear	Y	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Chen et al. (2021) (Chen et al., 2021a)	N	N	Y	High	Low	Unclear	No, used ROC to find best cutoff	High	Low	Y (mass biopsy, surgical resection)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
D'Ambrosi et al. (2023) (D'Ambrosi et al., 2023)	N	N	Y	High	Low	unclear	NA	Unclear	Low	Y (histological analysis)	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Dong et al. (2021a) (Dong et al., 2021a)	N	N	Y	High	Low	unclear	NA	Unclear	Low	Unclear	Y	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Dong et al. (2021b) (Dong et al., 2021b)	N	N	Y	High	Low	unclear	NA	Unclear	Low	Unclear	Y	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Doseeva et al. (2015) (Doseeva et al., 2015)	N	N	Y	High	Low	Y	NA	low	Low	Y (surgical pathology)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Dou et al. (2018) (Dou et al., 2018)				High	Low			Unclear	Low			Low	Low				Unclear	High	Low
Du et al. (2018) (Du et al., 2018)	Y (looks like consecutive)	Y	Y	Low	Low	Y	Yes	Low	Low	Y	Y	Low	Low	Unclear	Y	Y	Unclear	Unclear	Low
Duan et al. (2021) (Duan et al., 2021)	N	N	Y	High	Low	unclear	NA	Unclear	Low	Unclear	Y	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Ezzatifar et al. (2022) (Ezzatifar et al., 2022)	Y	Y	Y	Low	Low	unclear	No	Unclear	Low	Y (histopathological examination)	Y	Low	Low	Unclear	Unclear	Y	Unclear	Unclear	Low
Fahrmann et al. (2022)	Y	Y	Y	Low	Low	unclear	NA	Unclear	Low	Y (histopatholog	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	Unclear	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
(Fahrmann et al., 2022)										ically confirmed)									
Fan et al. (2018) (Fan et al., 2018)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (pathologically confirmed)	Unclear	Unclear	Low	Unclear	Y	Y	Unclear	High	Low
Farlow et al. (2010a) (Farlow et al., 2010a)	N	N	Y	High	Low	Y	NA	Low	Low	Y (surgical pathology reports)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Farlow et al. (2010b) (Farlow et al., 2010b)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (surgical resection, pathological confirmation)	Y	Low	Low	unclear	Y	Y	Unclear	High	Low
Gao et al.	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (pathol	Y	Unclear	Low	unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
(2015) (Gao et al., 2015b)										ogical report)									
Gasparri et al. (2023) (Gasparri et al., 2023)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Unclear	Y	Unclear	Low	unclear	Y	Y	Unclear	High	Low
Goebel et al. (2019) (Goebel et al., 2019)	N	N	Y	high	Low	Yes	NA	Low	Low	Unclear	Y	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Hennessey et al. (2012) (Hennessey et al., 2012)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Unclear	Y	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Higgins et al. (2012) (Higgins et al., 2012)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (confirmed histopathologically at time of surgery)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Hua et al. (2022) (Hua et al., 2022)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (pathologically confirmed)	Y	Unclear	Low	Unclear	Y	Y	Unclear	High	Low
Huo et al. (2020) (Huo et al., 2020)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (symptoms, Xray, ling CT etc)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Jeong et al. (2021) (Jeong et al., 2021)	N	N	Y	High	Low	unclear	NA	Unclear	Low	Unclear	Y (likely)	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Jiang et al. (2021) (Jiang et al., 2021)	N	N	Y	High	Low	Unclear	No, used Youden's index to determine cutoff	High	Low	Unclear	Y (likely)	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Jiang et al. (2022) (Jiang et al., 2022b)				High	Low	Unclear		High	Low			Unclear	Low				Unclear	High	
Joseph et al. (2012) (Joseph et al., 2012)	N	N	Y	High	Low	Y	NA	Low	Low	Y (verified by pathology following resection)	Y	Low	Low	Unclear	Yes	Yes	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Jung et al. (2016) (Jung et al., 2017)	N	N	Y	High	Low	Y	NA	Low	Low	Y (CT imaging, pathologic/clinical staging)	Y	Low	Low	Y (collected within 4 weeks and before treatment)	Y	Y	Low	High	Low
Kupert et al. (2011) (Kupert et al., 2011b)	N	N	Y	High	Low	Y (deidentified)	N	High	Low	Y (underwent resection, pathologically confirmed	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Lai et al. (2022) (Lai et al., 2022)	Unclear	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (pathologically confirmed)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Lastwika et al. (2019) (Lastwika et al., 2019)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (radiologist blinded to clinical and histologic findings)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Leung et al. (2020) (Leung et al., 2020)	Unclear	N	Y	High	Low	Y	NA	Low	Low	Y (pathology reports of surgical resection)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Li et al. (2019) (Li et al., 2019a)	N	N	Y	High	Low	unclear	NA	Unclear	Low	Unclear	Y (likely)	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Li et al. (2022) (Li et	Y	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (pathological	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
al., 2022)										analysis)									
Li et al. (2023) (Li et al., 2023a)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Unclear	Y	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Lin et al. (2017) (Lin et al., 2017)	N	N	Y	High	Low	Y	N	High	Low	Y (CT scan, pathologic analysis)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Liu et al. (2020) (Liu et al., 2020)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (radiographic exam and histological confirmation	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Lowe et al. (2014) (Lowe et al., 2014)	N	N	Y	High	Low	Y	NA	Low	Low	Y (pathologically confirmed)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Ma et al. (2017) (Ma et al., 2017)	Unclear	Y (retrospective cohort)	Y	Unclear	Low	Unclear	No, they used Youden index	High	Low	Y (pathologically confirmed)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Ma et al. (2021) (Ma et al., 2021)	Y (retrospective and prospective cohort)	Y	Y	Low	Low	Unclear	N	High	Low	Y (pathologically or histologically confirmed)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Meng et al. (2023) (Meng et al., 2023)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (histologically confirmed)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Mu et al. (2022) (Mu et al., 2022)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (pathology examination)	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Nolen et al. (2011) (Nolen et al., 2011)	Y	N	Y	High	Low	Y	NA	Low	Low	Y (biopsy-proven)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Ouyang et al. (2021) (Ouyang et al., 2021)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (pathology results after surgery)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Paci et al. (2009) (Paci et al., 2009)	N	N	Y	High	Low	Unclear	N	High	Low	Y	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Pakvisal et al. (2022) (Pakvisal et al., 2022)	N	N	Y	High	Low	Unclear	N	High	Low	Y	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Pan et al. (2020)	Y (cohort study)	Y	Y	Low	Low	Y	No, they used	High	Low	Y (histopatholog	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
(Pan et al., 2020)							optimal cut-off values			ical confirmation)									
Reis et al. (2020) (Reis et al., 2020)	N	N	Y	High	Low	Y	NA	Low	Low	unclear (just stated confirmed diagnosis)	Y (likely)	Unclear	Low	Unclear	Y	Y	Unclear	High	Low
Ren et al. (2017) (Ren et al., 2018)	N	N	Y	High	Low	Unclear	Y	Unclear	Low	Y (histopathologically confirmed)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Showe et al. (2009)	N	N	Y	High	Low	Unclear	Y	Unclear	Low	Y (histopathologically confirmed)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Song et al. (2019)	Unclear	N	Y	High	Low	unclear	NA	Unclear	Low	Unclear	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
(Song et al., 2019)																			
Song et al. (2022) (Song et al., 2022)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Unclear	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Sun et al. (2020) (Sun et al., 2020a)	Y	N	Y	High	Low	Y	NA	Low	Low	Y (histopathologically confirmed)	Y (likely)	Low	Low	Unclear	Y	Y	Unclear	High	Low
Tulinsky et al. (2022) (Tulinsky et al., 2022)	N	N	Y	High	Low	Y	NA	Low	Low	Y	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Wan et al. (2021) (Wan et al., 2021)	N	N	Y	High	Low	Y	NA	Low	Low	Y	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Wang et al. (2013) (Wang et al., 2013)	N	N	N	High	Low	unclear	No	Unclear	Low	Y	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Wang et al. (2017) (Wang et al., 2017)	N	N (looks like case control)	Y	High	Low	Y	NA	Low	Low	Y (histopathological analysis)	Yes	Low	Low	unclear	Y (kind of)	Y	Unclear	High	Low
Wang et al. (2020) (Wang et al., 2020b)	Y (likely)	Y (likely)	Y	Low	Low	unclear	Yes	Unclear	Low	Y (histopathological exam)	Y (likely)	Low	Low	Unclear	Y	N	High	High	Low
Wang et al. (2022) (Wang et al., 2022a)	Y	Y	Y	Low	Low	Unclear	No	Unclear	Low	Y	Y	Low	Low	Unclear	Y	Y	Unclear	Unclear	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Wiesko pf et al. (1995) (Wiesk opf et al., 1995)	Unclear (likely consec utive for cancer patient s, unclear for control)	N	Y	High	Low	Unclear	NA	Uncle ar	Low	Y (clinical , chest radiogr aphy etc)	Yes	Low	Low	Uncl ear	Y	Y	Unclear	High	Low
Wu et al. (2020a) (Wu et al., 2020a)	N	N	Y	High	Low	Unclear	NA	Uncle ar	Low	Y	Y	Low	Low	Uncl ear	Y	Y	Unclear	High	Low
Wu et al. (2020b) [118]	N	N (likely case control)	Y	High	Low	Unclear	NA	Uncle ar	Low	Unclear (just said met diagnos tic standar ds)	Yes	Uncl ear	Low	Uncl ear	Y	Y	Unclear	High	Low
Wu et al.	N	N	Y	High	Low	Unclear	NA	Uncle ar	Low	Y	Y	Uncl ear	Low	Uncl ear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
(2022) (Wu et al., 2022)																			
Xing et al. (2015) (Xing et al., 2015)	Y (all patients selected)	N	Y	High	Low	Y	NA	Low	Low	Y (histopathologic exam)	Y (likely)	Low	Low	Unclear	Y (likely)	Y	Unclear	High	Low
Xing et al. (2019) (Xing et al., 2019)	N	N	Y	High	Low	Y	No, they used Youden index as cut-off value	High	Low	Y (pathologically diagnosed)	Y (likely)	Low	Low	Unclear	Y (likely)	Y	Unclear	High	Low
Xue et al. [135]	N	N	Y	High	Low	Y	N	High	Low	Y	Y	Low		Unclear	Y	Y	Unclear	High	Low
Yang et al. (2019) (Yang	N	N	Y	High	Low	Y	NA	Unclear	Low	Y	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
et al., 2019b)																			
Yang et al. (2020a) (Yang et al., 2020a)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (surgical resection)	Y	Low	Low	Unclear	Unclear	Y	Unclear	High	Low
Yang et al. (2020b) (Yang et al., 2020b)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y (pleural effusion etc)	Y (likely)	Low	Low	Unclear	Y (likely)	Y	Unclear	High	Low
Yu et al. (2023) (Yu et al., 2023)	Y	Y	Y	Low	Low	Unclear	N	High	Low	Y	Y	Low	Low	Unclear	Y (likely)	Y	Unclear	High	Low
Yuan et al. (2022) (Yuan et al., 2022)	Y	Y	Y	Low	Low	Unclear	NA	Unclear	Low	Y	Y	Low	Low	Unclear	Y	Y	Unclear	High	Low
Zang et al. (2019)	Y	N	Y	High	Low	Unclear	No, they used	High	Low	Y (possibl	Y (likely)	Low	Low	Unclear	Y (likely)	Y	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
(Zang et al., 2019)							YI to establish cut-off			y Xray and CT)									
Zhang et al. (2022a) (Zhang et al., 2022b)	N	N	Y	High	Low	Unclear	No, used Youden's index to determine cutoff	High	Low	Unclear	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Zhang et al. (2022b) (Zhang et al., 2022a)	N	N	Y	High	Low	Unclear	N	High	Low	Y	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Zhong et al. (2006) (Zhong et al., 2006)	N (only controls were random)	N	Y	High	Low	unclear	NA	Unclear	Low	unclear	Yes	Unclear	Low	Unclear	unclear	Yes	Unclear	High	Low

Author (year)	Was a consecutive or random sample of patients enrolled?	Was a case control design avoided?	Did the study avoid inappropriate exclusions?	Domain 1: Patient selection	Were there concerns that the included patients and setting do not match the review question?	Were the index test results interpreted without knowledge of the results of the reference	If a threshold was used, was it pre-specified?	Domain 2: Index test	Were there concerns that the index test, its conduct, or its interpretation differ from the review question?	Was the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Domain 3: Reference standard	Are there concerns that the target condition as defined by the reference standard does not match the question?	Was there an appropriate interval between index test and reference	Did all patients receive the same reference standard?	Were all patients included in the analysis?	Domain 4: Flow and timing	Risk of bias	Applicability
Zhong et al. (2021) (Zhong et al., 2021)	N	Y (no control)	Y	High	Low	Y	Y	Low	Low	Unclear	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low
Zhou et al. (2022) (Zhou et al., 2022)	N	N	Y	High	Low	Unclear	NA	Unclear	Low	Y	Unclear	Unclear	Low	Unclear	Unclear	Y	Unclear	High	Low

Table 9.2 CTAs expression in lung cancer using Kaplan Meier plot for the survival

Gene	Probe set	survival for Low expression cohort (months)	survival for high expression cohort (months)	P-value
DNAJB11	223054_at	119.87	52	8.40E-12
MAGEA1	207325_x_at	86.27	48.6	1.40E-11
SSX2IP	203015_s_at	91	52	2.80E-11
DDX12	213378_s_at	89	52	1.70E-10
(DNAJB14)	222850_s_at	52	111	1.20E-09
MAGEA3	209942_x_at	86.27	49.97	2.70E-09
DDX11/KRG2	208149_x_at	88.7	54	1.10E-08
(GAGE3)/CT4.3	207663_x_at	89	54.2	1.10E-07
MAGEA12	210467_x_at	84	52	2.70E-07
GAGE1/4/7/11	207086_x_at	88	56	6.00E-07
GAGE12 E	207086_x_at	88	56	6.00E-07
TPTE/CT44	220205_at	80.03	59	1.30E-05
SAGE	220793_at	79.5	56.5	2.00E-05
MAGEA10	210295_at	86.27	57.33	2.40E-05
DDX10/HRH-J8	204977_at	79.54	57	8.70E-05
NA88A/VENTXP1	216726_at	81.2	61.2	0.0001
TEX15/CT42	221448_s_at	79.87	59	0.0001
DNAJB2 (HSPF3)	202500_at	62	74	0.0002
MORC1/CT33	220850_at	79.27	57	0.0003
LDHC/CT32/	207022_s_at	78	62.2	0.0004
DDX13 (SKIV2L)	203727_at	81	59.11	0.0004
(LDHC)/CT32	207022_s_at	78	62.2	0.0004
MAGE-C1	206609_at	79.5	61.2	0.0006
CAGE1	1563787_a_at	91	62	0.0008
TSP50	220126_at	81	56.7	0.0009
CTAGE-1	220957_at	79.27	61.3	0.0009
PAGE-4	205564_at	76	60.73	0.001
SSX3	211670_x_at	78.5	62.2	0.0012
SYCP1	206740_x_at	79.87	60	0.0018
NXF2/CT39	220257_x_at	79.87	62.2	0.0021
SSX1	206626_x_at	78	64.1	0.0023
DNAJB4	203811_s_at	75.43	62.47	0.0035
SGY-1/CT34/	220284_at	76	59	0.0053
MEGEA2	214603_at	74	59.53	0.0058
FATE/CT43	231573_at	86.27	63	0.0085
FATE1	231573_at	86.27	63	0.0085
(GPATCH2)	239768_x_at	69	89	0.0094
SSX2	216471_x_at	76	63.3	0.01
SPO11/ CT35	222259_s_at	76	62.3	0.0185
(DNAJB13)	230936_at	70	90	0.0188
PLU-1/ KDM5B	211202_s_at	63	77.6	0.019

LAGE1	215733_x_at	73.3	64.1	0.025
TAF7L	220325_at	76	63.4	0.0254
TDRD1/CT41.1	221018_s_at	74	65.1	0.0284
PAGE-1	206897_at	73.2	65	0.0299
MAGE-C2	215932_at	74	64.1	0.0326
(DNAJB1)/Sis1/	200666_s_at	72.33	65.57	0.0366
LUZP4/CT28	220665_at	73.3	65	0.0461
(DNAJB12)	202867_s_at	63.4	73.2	0.0556
(DNAJB8)	237284_at	88	72	0.0578
HCA661/CT30/	207385_at	76.3	62.47	0.0656
MAGEA11	210503_at	73.2	63	0.069
SSX4	208586_s_at	75.43	65.2	0.0718
SPA17	205406_s_at	68.6	71.27	0.0897
DDX19A/FLJ11126	202578_s_at	72.33	67	0.1016
(DDX18)	208897_s_at	74	65.57	0.1223
MAGE-C3	216592_at	72.33	65.2	0.1291
(DDX17)	208151_x_at	73.3	64.1	0.1478
HAGE	220004_at	68.6	70	0.4235
SYCP2	206546_at	68.67	72.33	0.4689
FTHL17/CT38	224379_at	75.43	79.27	0.5335
LIPI	242178_at	76.3	73.3	0.6202
SYCP2L	236337_at	76	79.87	0.7598
SYCP3	1553599_a_at	76	79.27	0.8441
BRDT	206787_at	68	70.3	0.8794
ADAM2	207664_at	69.93	68.6	0.9338
TPX1/CT36	210262_at	68.7	70.6	0.9629
LAGE-1b	207337_at	70.5	68.6	0.99
CASC5/ D40	220247_at	68.1	72.33	0.9977
TPX2	210052_s_at	96.2	42	<1e-16

Table 9.3 Analysis of genes expression identified from SEREX and their association with the survival using BloodSpot

Antigen	Probe set	High expression	Survival
SIX2	206510-at	Mature B-ALL with t(8;14) p<0.01	0.324
	206511-s-at	Mature B-ALL with t(8;14) p<0.01	0.36
TUBA3C	210527-x-at	Pro-B-ALL t(11q23)/MLL-NS	0.97
TCF7L2	212759-s-at	ALL t(12;21) p<0.001	0.937
	212761-at	ALL t(12;21) p<0.001	0.729
	216511-s-at	ALL t(12;21)/ Pre-B-ALL t(9;22) p<0.001	0.7
	216037-x-at	ALL t(12;21) p<0.001	0.92
	236094-at	ALL t(12;21)/ Pre-B-ALL t(9;22) p<0.001	0.916
	216035-x-at	ALL t(12;21) p<0.001	0.198
	212762-s-at	ALL t(12;21) p<0.001	0.279
WLS	**228949-at**	**Healthy bone marrow**	**0.0045**
	228950-s-at	Healthy bone marrow	0.434
	221958-s-at	Healthy bone marrow	0.32
SEZ6L2	238406-x-at	B-ALL ALL with t(8;14)-NS	0.124
	238404-x-at	B-ALL ALL with t(8;14)-NS	0.652

	233337-s-at	Healthy bone marrow	0.841
	218720-x-at	Healthy bone marrow	0.063
	223458-at	B-ALL ALL with t(8;14)-NS	0.971
WNT2b	206459-s-at	B-ALL ALL with t(8;14)-NS	0.548
	206458-s-at	B-ALL ALL with t(8;14)-NS	0.695
DNMT1	201697-s-at	ALL with t(1;19)- p<0.001	0.275
HNRPLL	1554453-at	Healthy bone marrow- p<0.001	0.537
	236104-at	Healthy bone marrow- p<0.001	0.67
	225386-s-at	Healthy bone marrow- p<0.001	0.45
	225385-s-at	Healthy bone marrow- p<0.001	0.409
MYO6	210480-s-at	All types of B-ALL p<0.01 except ALL t(8;14)/ Pro-B-ALL t(11q23)/MLL -NS	0.996
	203215-s-at	All types of B-ALL p<0.01 except ALL t(8;14)/ Pro-B-ALL t(11q23)/MLL -NS	0.0819
	203216-s-at	All types of B-ALL p<0.01 except ALL t(8;14)/ Pro-B-ALL t(11q23)/MLL -NS	0.617
FLNB	208614-s-at	All types of B-ALL p<0.001 except ALL t(8;14)- p<0.05	0.686
	208613-s-at	All types of B-ALL p<0.001 except ALL t(8;14)- p<0.05	0.849
TIMP1	201666-at	All types of B-ALL p<0.001 except ALL t(8;14)- NS	0.414
ATL3	223452-s-at	All types of B-ALL p<0.001	0.939
	224893-at	All types of B-ALL p<0.001	0.649
	223453-s-at	All types of B-ALL p<0.001	0.228
KIF5B	201991-s-at	ALL t(12;21)/ Pre-B-ALL not(9;22)/hyperdiploid p<0.001	0.928
	224662-at	ALL t(12;21)/ Pre-B-ALL not(9;22)/hyperdiploid p<0.001	0.271
	201992-s-at	**ALL t(12;21)/ Pre-B-ALL no t(9;22)/hyperdiploid p<0.001**	**0.00661**
TMA7	221791-s-at	B-ALL with t(1;19)/ t(12;21)- p<0.05	0.721
PTPN1	202716-at	B-ALL ALL with t(8;14)-NS	0.578
	217686-at	B-ALL ALL with t(8;14)-NS	0.115
	217689-at	B-ALL ALL with t(8;14)-NS	0.887
	239526-x-at	B-ALL ALL with t(8;14)-NS	0.516
TMCO3	220240-s-at	Healthy bone marrow- p<0.05	0.578
	230317-x-at	Healthy bone marrow- p<0.05	0.861
	226050-at	Healthy bone marrow- p<0.05	0.941
	220241-at	ALL with t(1;19)- p<0.001	0.562
TMP3		Not found	
UBE2D2	201343-at	ALL subtypes of B-ALL-p<0.001/ B-ALL with t(8;14) p<0.01	0.0996
	201345-s-at	ALL subtypes of B-ALL-p<0.001/ B-ALL with t(8;14) p<0.01	0.4
	201344-at	ALL subtypes of B-ALL-p<0.001/ B-ALL with t(8;14) p<0.01	0.156
MYL12B	221474-at	Pro-B-ALL t(11q23)/MLL- p<0.01	0.238
C10ORF82	231609-at	Healthy bone marrow-NS	0.238
DNMT1	201697-s-at	ALL with t(1;19)- p<0.001	0.275
HNRPLL (SRRF)	1554453-at	Healthy bone marrow- p<0.001	0.537
	236104-at	Healthy bone marrow- p<0.001	0.67
	225386-s-at	Healthy bone marrow- p<0.001	0.45
	225385-s-at	Healthy bone marrow- p<0.001	0.409
MYO6	210480-s-at	All types of B-ALL p<0.01 except ALL t(8;14)/ Pro-B-ALL t(11q23)/MLL -NS	0.996
	203215-s-at	All types of B-ALL p<0.01 except ALL t(8;14)/ Pro-B-ALL t(11q23)/MLL -NS	0.0819
	203216-s-at	All types of B-ALL p<0.01 except ALL t(8;14)/ Pro-B-ALL t(11q23)/MLL -NS	0.617
FLNB	208614-s-at	All types of B-ALL p<0.001 except ALL t(8;14)- p<0.05	0.686
	208613-s-at	All types of B-ALL p<0.001 except ALL t(8;14)- p<0.05	0.849
ATL3	223452-s-at	All types of B-ALL p<0.001	0.939
	224893-at	All types of B-ALL p<0.001	0.649
	223453-s-at	All types of B-ALL p<0.001	0.228
AASDH	235435-at	ALL hyperdiploid p<0.001	0.748
	228041-at	ALL hyperdiploid p<0.001	0.801
FER	227579-at	All types of B-ALL p<0.001 except ALL t(8;14)/ t(1;19)- NS	0.66
	206412-at	All types of B-ALL p<0.001 except ALL t(8;14)/ t(1;19)- NS	0.892
	232064-at	All types of B-ALL p<0.001 except ALL t(8;14)/ t(1;19)- NS	0.91
MCL1	214057-at	ALL t(8;14)- NS	0.593
	200798-x-at	ALL t(8;14)- NS	0.272
	214056-at	ALL t(8;14)- NS	0.157
	227175-at	ALL t(8;14)- NS	0.301
	200797-s-at	ALL t(8;14)- NS	0.808
	200796-s-at	ALL t(8;14)- NS	0.648
ANKRD17	225852-at	All types of B-ALL p<0.001	0.811
	212211-at	All types of B-ALL p<0.001	0.545
DKFZp686C152 13		Not found	
SERINC3	**221472-at**	**Pro-B-ALL t(11q23)/MLL/ Pre-B-ALL t(9;22) p<0.05**	**0.0246**
	221473-x-at	Pro-B-ALL t(11q23)/MLL/ Pre-B-ALL t(9;22) p<0.05	0.313

	221471-at	Pro-B-ALL t(11q23)/MLL/ Pre-B-ALL t(9;22) p<0.05	0.791
	211769-x-at	Pro-B-ALL t(11q23)/MLL/ Pre-B-ALL t(9;22) p<0.05	0.205
KRT31	206677-at	ALL t(8;14)- NS	0.484
TRAP1	201391-at	Healthy bone marrow-NS	0.368
	239622-at	Healthy bone marrow-NS	0.813
NTPCR	223272-s-at	All types of B-ALL p<0.001	0.534
	226813-at	All types of B-ALL p<0.001	0.631
TOMM20L		Not found	
A-COL04217		Not found	
C5ORF48 (TEX43)	237428-at	All types of B-ALL p<0.001	0.899
ACAD11197		Not found	
MIP	226863-at	Healthy bone marrow-NS	0.825
TPT1	212869-x-at	All types of B-ALL p<0.05 except except ALL t(8;14)-NS	0.686
	212284-x-at	All types of B-ALL p<0.05 except except ALL t(8;14)-NS	0.515
	211943-x-at	All types of B-ALL p<0.05 except except ALL t(8;14)-NS	0.703
	216520-s-at	All types of B-ALL p<0.05 except except ALL t(8;14)-NS	0.282
	214327-x-at	All types of B-ALL p<0.05 except except ALL t(8;14)-NS	0.779
PG0247		Not found	
HOXD8	231906-at	Pro-B-ALL t(11q23)/MLL/ Pre-B-ALL t(9;22) / Pre-B-ALL no t(9;22) -NS	0.576
FLJ21112		Not found	
CCDC89	1553228-at	Mature B-ALL with t(8;14)-NS	0.102
HMG20 (UBC)	**211296-x-at**	**Pre-B-ALL t(9;22) / Pre-B-ALL no t(9;22) -NS**	**0.000204**
	208980-s-at	Pre-B-ALL t(9;22) / Pre-B-ALL no t(9;22) -NS	0.102
TMPO	224944-at	ALL t(12;21) p<0.001	0.982
	209753-s-at	ALL t(12;21) p<0.001	0.245
	237863-at	ALL t(12;21) p<0.001	0.475
	209754-s-at	ALL t(12;21) p<0.001	0.882
	203432-at	ALL t(12;21) p<0.001	0.646
LINC00661	1555988-a-at	Mature B-ALL with t(8;14)-NS	0.782
	237242-at	Mature B-ALL with t(8;14)-NS	0.645
PCAT6		Not found	
HDLBP	200643-at	Mature B-ALL with t(8;14)-NS	0.403
	221767-x-at	Mature B-ALL with t(8;14)-NS	0.475
	219674-s-at	Mature B-ALL with t(8;14)-NS	0.623
	222916-s-at	Mature B-ALL with t(8;14)-NS	0.956
	225012-at	Mature B-ALL with t(8;14)-NS	0.204
	235624-at	Mature B-ALL with t(8;14)-NS	0.0597
SUN1	212074-at	All types of B-ALL p<0.001 except ALL t(8;14)- NS	0.609
	214169-at	All types of B-ALL p<0.001 except ALL t(8;14)- NS	0.723
	230210-at	All types of B-ALL p<0.001 except ALL t(8;14)- NS	0.502
	206487-at	All types of B-ALL p<0.001 except ALL t(8;14)- NS	0.274
MYH11	239307-at	Healthy bone marrow-NS	0.689
	201497-x-at	Healthy bone marrow-NS	0.36
	201496-x-at	Healthy bone marrow-NS	0.14
	207961-x-at	Healthy bone marrow-NS	0.0925
	201495-x-at	Healthy bone marrow-NS	0.888
	1568760-at	Healthy bone marrow-NS	0.544
ROCK1	**213044-at**	**All types of B-ALL p<0.05 except ALL t(8;14)/ ALL t(1;19)/ Pre-B-ALL t(9;22)- NS**	**0.00147**
	214578-s-at	**All types of B-ALL p<0.05 except ALL t(8;14)/ ALL t(1;19)/ Pre-B-ALL t(9;22) - NS**	**0.00578**
	230239-at	All types of B-ALL p<0.001 except ALL t(8;14)/ ALL t(1;19)- NS	0.68
	235854-x-at	All types of B-ALL p<0.001 except ALL t(8;14)/ ALL t(1;19)- NS	0.738
B9D1	205662-at	Healthy bone marrow-NS	0.499
	210535-at	Healthy bone marrow-NS	0.925
	210534-s-at	Healthy bone marrow-NS	0.345
RAB5C	201136-s-at	Healthy bone marrow-NS	0.591
	201140-s-at	Healthy bone marrow-NS	0.0548
RAB34	1555630-a-at	All types of B-ALL p<0.05 except ALL t(12;21)- NS	0.999
	224710-at	All types of B-ALL p<0.05 except ALL t(12;21)- NS	0.882
APPL1	222538-s-at	Pro-B-ALL t(11q23)/MLL p<0.001	0.403
	218158-s-at	Pro-B-ALL t(11q23)/MLL p<0.001	0.489
CLIC4	221881-s-at	ALL hyperdiploid-NS	0.433
	201560-at	ALL hyperdiploid-NS	0.143
	201559-s-at	ALL hyperdiploid-NS	0.127
TRAP1	201391-at	Healthy bone marrow-NS	0.368

	239622-at	Healthy bone marrow-NS	0.813
IDI1	**208881-x-at**	**All types of B-ALL p<0.01 except ALL t(8;14)- NS**	**0.000785**
	204615-x-at	**All types of B-ALL p<0.01 except ALL t(8;14)- NS**	**p<0.0001**
	242065-x-at	All types of B-ALL p<0.01 except ALL t(8;14)- NS	0.877
UBE3C	243519-at	Healthy bone marrow-NS	0.39
	201817-at	Healthy bone marrow-NS	0.302
	1554794-a-at	Healthy bone marrow-NS	0.161
	1555405-at	Healthy bone marrow-NS	0.878
	1554793-at	Healthy bone marrow-NS	0.951
ZNF676		Not found	
SURF2	205224-at	ALL t(8;14)/Pro-B-ALL t(11q23)/MLL -p<0.001	0.422
TEX38		Not found	
TPPP2	231140-at	ALL t(8;14)-NS	0.287
	1559785-at	ALL t(8;14)-NS	0.402
AKAP1	201675-at	All types of B-ALL p<0.001 except ALL t(12;21)- NS	0.708
	210626-at	All types of B-ALL p<0.001 except ALL t(12;21)- NS	0.591
	210625-s-at	All types of B-ALL p<0.001 except ALL t(12;21)- NS	0.417
	201674-s-at	All types of B-ALL p<0.001 except ALL t(12;21)- NS	0.847
UVRAG-DT		Not found	
FBXO22	225736-at	ALL with t(1;19)-NS	0.419
	225737-s-at	ALL with t(1;19)-NS	0.231
	225734-at	ALL with t(1;19)-NS	0.721
	219638-at	ALL with t(1;19)-NS	0.642
LOC338963	1563878-a-at	Healthy bone marrow-NS	0.249
KIF1B	209234-at	Healthy bone marrow- p<0.001	0.161
	226968-at	Healthy bone marrow- p<0.001	0.218
	225878-at	Healthy bone marrow- p<0.001	0.0562
	228657-at	Healthy bone marrow- p<0.001	0.319
BCLAF1	214499-s-at	All types of B-ALL p<0.05 except ALL t(12;21)/ (1;19)- NS	0.312
	201101-s-at	All types of B-ALL p<0.05 except ALL t(12;21)/ (1;19)- NS	0.364
	229454-at	All types of B-ALL p<0.05 except ALL t(12;21)/ (1;19)- NS	0.868
	239897-at	All types of B-ALL p<0.05 except ALL t(12;21)/ (1;19)- NS	0.696
	201084-s-at	All types of B-ALL p<0.05 except ALL t(12;21)/ (1;19)- NS	0.707
	201083-s-at	All types of B-ALL p<0.05 except ALL t(12;21)/ (1;19)- NS	0.333
CUL1	**238509-at**	**ALL with t(1;19)/ ALL hyperdiploid - p<0.01**	**0.0422**
	207614-s-at	**ALL with t(1;19)/ ALL hyperdiploid - p<0.01**	**0.0161**
LINC00251		Not found	
C18orf32	224957	Healthy bone marrow- p<0.001	0.697
PARL	**218271-s-at**	**Healthy bone marrow- p<0.01**	**0.00178**
	228881-at	Healthy bone marrow- p<0.01	0.549
DKK3	214247-s-at	Healthy bone marrow-NS	0.933
	221127-s-at	All types of B-ALL p<0.001	0.423
	221126-at	Healthy bone marrow-NS	0.451
	230508-at	ALL t(8;14)/ Pre-B-ALL no t(9;22) -NS	0.1
	202196-s-at	ALL t(8;14)/ Pre-B-ALL no t(9;22) -NS	**0.0233**
PRM1	206358-at	ALL t(8;14)-NS	0.199
MRFAP1	226091-s-at	All types of B-ALL p<0.05 except ALL t(8;14)- NS	0.739
EIF3A	200596-s-at	Healthy bone marrow- p<0.001	0.742
	200596-s-at	Healthy bone marrow- p<0.001	0.444
	200597-at	Healthy bone marrow- p<0.001	0.163
RPL28	2000003-s-at	All types of B-ALL p<0.05	0.486
	213223-at	All types of B-ALL p<0.05	0.0652
MYH10	212372-at	ALL t(12;21)- p<0.001	0.978
	213067-at	ALL t(12;21)- p<0.001	0.372
PTPN23	223150-s-at	ALL t(8;14)- p<0.001	0.406
	223149-s-at	All types of B-ALL p<0.001	**0.0378**

Table 9.4 Expression of identified SEREX genes and their function in both normal and cancer tissues. Only antigens identified during this project were analysed for their expression and function in health and disease

	Gene	Function and expression in normal tissue	Expression in cancer
UOH-ALL-17	SIX2/ UOH-ALL-17	It belongs to homeobox genes and encodes proteins that firstly identified to be involved in limbs and eye development. High expression is normally found in salivary tissue, prostate and stomach. It also plays transcriptional regulator role (Walz et al., 2015)	It is upregulated in HCC and associated with poor prognosis. It enhances metastasises and EMT via activation of TGF-β/Smad signalling (Wan et al., 2019)
UOH-ALL-20	TUBA3C/ UOH-ALL-20	It encodes for microtubules assembly, and it is strictly expressed in the testis (Jordan, 2002).	TUBA3C is upregulated in breast cancer and it is a good biomarker for taxane sensitivity as its level decreased in chemo-resistance cases (Nami & Wang, 2018). TUBA3C is highly increased in NSCLC with poor survival with p-value=0.007 (Djureinovic et al., 2016)
UOH-ALL-23	TCF7L2	It encodes for high motility (HMG) box-containing transcription factor and it is enriched in many tissues such as fat, ovarian, stomach. It plays an important in glucose homeostasis and acts as transcriptional factor in Wnt pathways (Hatzis et al., 2008)	It is overexpressed in CML and associated with increased transcription of Wnt related genes such as cyclin D. It also cooperates with MYC and promotes tumour progression (Desterke et al., 2020)
UOH-ALL-19	WLS	It enables Wnt protein activation and involves in protein transport and positive regulation of cell communication. It is expressed in many tissues such as kidney, Golgi apparatus, endothelial tissues, gall bladder (Bänziger et al., 2006)	It is highly expressed in gastric cancer and associated with advanced metastasis (Zhang et al., 2017a) It overexpressed in cB-ALL and associated with poor prognosis and relapse. It involves in proliferation and anti-apoptotic activity via regulating GSK3β activation (Chiou et al., 2014)
UOH-ALL-32	KIF5B	It belongs the motor protein involved in mitochondria transport and movement of synaptic vesicle. It is found in all most tissues (Hirokawa & Tanaka, 2015).	It downregulated with increased expression of miR 125b), miR-99a and miR-100 in cB-ALL resistant to vincristine (Akbari Moqadam et al., 2013).

UOH-ALL-6	TMA7	It is involved in the epithelial-mesenchymal interactions have an important role in folliculomorphogenesis and mature hair follicle cycling (Yang et al., 2005).	HSPC016 expression is associated with aggressive dermal papilla cells (DPC) (Song et al., 2012).
UOH-ALL-7	PTPN1	It encodes non-receptor type PTPN that regulates pre-mRNA splicing and assembly of small ribonucleo-protein. It is found in most all tissues (Hendriks & Pulido, 2013)	It is a negative regulator for JAK/STAT signalling pathway that have been found reduced or truncated or inactive in classical Hodgkin lymphoma and associated TKI (Karaca Atabay et al., 2022)
UOH-ALL-21	SEZ6L2	It is found on the cell surface and encodes a seizure related protein. It is normally expressed in lung, adrenal, brain, testis (Konyukh et al., 2011).	Overexpression of SEZ6L2 has been found in breast cancer and associated with TNM stage. SEZ6L2 expression is regulated by upstream transcription factor 1 (USF1). SEZ6L2 knockdown result in significant decreased of BC growth and metastasis (Chen et al., 2022a)
UOH-ALL-10	WNT2B	It has low tissue specificity and finds in ECM, fibroblast, peritubular and Leydig cells. It involves in cell regulation and developmental processes (Bänziger et al., 2006).	Wnt2b is decreased in B-ALL (p=0.033) (Memarian et al., 2012)
UOH-ALL-15	DNMT1	This gene is encoded for the enzyme that transfer methyl group to cytosine in DNA. DNA methylation is important for epigenetic regulation. It is widely expressed in highly proliferate cells such as spermatocytes, trophoblast cells, and acts a transcription factor (Moore et al., 2013)	It (Liu et al., 2022b)is highly expressed in liver cancer and associated with poor prognosis (p=0.001). High expression was found in cB-ALL and increased the disease burden (Haque & Vaiselbuh, 2022)
UOH-ALL-24	HNRNPLL (SRRF)	It encodes an RNA-binding protein that a key regulator of alternative splicing including CD45/PTPRC, and STAT5A. it is widely expressed in many tissues including bone marrow, retina, adrenal gland, stomach in ubiquitous form (Han et al., 2010).	HNRNPLL overexpression is found in colorectal cancer, and it acts a protein stabiliser for mRNA that are encoded for the regulators DNA replication. Whereas it acts a metastasis inhibitor via modulating the alternative splicing of CD44 during EMT (Sakuma et al., 2018).
UOH-ALL-25	MYO6	It is encoded a reverse-direction motor protein and belongs to actin motor family with ATPase activity. It involves in many intracellular processes including structural integrity, vesicular membrane trafficking, regulates actin cytoskeleton via control	MYO6 is overexpressed in gastric cancer and associated with the poor survival. MYO6 knockdown decrease colony formation and proliferation of cancer cells via inducing G0/G1 cell cycle arrest. Further analysis, it reveals increased in cell-cycle inhibitor p21,

		the association of septins with actin. It is widely found in transport membrane, skeletal muscle, kidney, retina, and intestine (Krendel & Mooseker, 2005).	and apoptosis related proteins Bax and caspase3 as well as decreased cyclin A, Bcl-2 expression, and cyclin D1 that are significantly induced apoptosis and cell migration inhibition (Wang et al., 2016c).
UOH-ALL-27	FLNB	It belongs to the filamin family, and the encoded protein connects cell membrane components to the actin cytoskeleton. It involves in cell migration, morphology, and integrin activity. It is found in many tissues such as thyroid, pancreatic cell, enterocytes, and adrenal cells (Baldassarre et al., 2009).	It is overexpressed in liver cancer and associated with poor prognosis. FLNB promotes metastasis and angiogenesis via inducing endothelial cell motility and interacting with Rac-1 and Vav2. This complex interaction leads VEGF activation and thus increases the angiogenesis processes (del Valle-Pérez et al., 2010).
UOH-ALL-29	TIMP1	It encodes a natural inhibitor of the matrix metalloproteinase and belongs to TIMP family. It functions in degradation and regulation of the extracellular matrix, and it also promotes cell proliferation via increasing cytokines, hormone, and antiapoptotic protein. It involves in signalling pathways via CD63 and ITGB1 activation. It is expressed in GI tract, salivary, testis, cervix, adipose tissue (Egeblad & Werb, 2002).	It is highly expressed in cB-ALL and associated poor over survival. TIMP1 and CD34+CD38− CSC could be useful biomarkers for B-ALL diagnosis and correlated to MRD (Saleh et al., 2021).
UOH-ALL-33	AASDH	It is encoded a non-ribosome peptide synthesase enzyme. It involves in lipid metabolism and modified gene and protein following transcription and translational. It is found in all tissues such as breast, retina, thymus, tonsil, and pancreas (Watkins et al., 2007).	It is upregulated in HCC and associated with shorter survival and poor outcome (Zhao et al., 2022). It has been reported that circ-AASDH is upregulated in lung adenocarcinoma and promotes tumour progression via sponging to miR 140-3p and inhibits its action on E2F7 transcription (Wang et al., 2021c).
UOH-ALL-44	FER	It is encoded a non-transmembrane receptor tyrosine kinase protein that belongs to FPS/FES family. It functions in regulating cell adhesion and growth factor receptors which mediates signalling between the cell surface and cytoskeleton. It also acts a downstream of EGFR and enhances NF-KB activation. They are multiple variants which are produced by alternative splicing and the related pseudogene is in	It is highly increased in renal cancer and associated with poor prognosis. It knockdowns resulting in cell cycle arrest in G0/G1 phase (Wei et al., 2013).

		chromosome X. it is found in many tissues such as parathyroid, testis, smooth muscle, ovary, and skin (Greer, 2002).	
UOH-ALL-46	MCL1	It encodes an antiapoptotic protein that belongs to the Bcl2 family. It has three isoforms: isoform 1 inhibits apoptosis and enhances cell survival whereas variant 2 and 3 induce apoptosis and cell death. It is mainly enriched in bone marrow and expressed in many tissues such as liver, urinary bladder, and appendix (Czabotar et al., 2014).	It is upregulated in AML and associated with poor prognosis and treatment resistance. Direct inhibitors of MCL1 are challenging as it inhibits it function in normal tissues may interfere with safety profiles (Wei et al., 2020).
UOH-ALL-47	ANKRD17 (FLJ22206, GTAR, KIAA0697, MASK2)	It belongs to ankyrin repeat-containing protein. It involves in DNA replication control and innate immune pathways against virus and bacteria. It also regulates blood vessel formation and circular system. It is expressed in all tissues such as neurons and bone marrow (Wang et al., 2012).	It is overexpressed in urinary bladder cancer and its knockdown results in inhibit cancer proliferation. Ankyrin 17 is a mask protein that is required for YAP transport to the nucleus as overexpression of YAP upregulates gene transcription and leads to tumour progression (Dong et al., 2016).
UOH-ALL-42	SERINC3	It encodes a transmembrane protein that is required for inhibiting lentiviruses infectivity as it impairs viral penetration to the cytoplasm. It is also involved in cell cycle regulation. It is found in many tissues such as liver, placenta, thymus, and parathyroid and it's enriched in testis (Usami et al., 2015).	It acts a tumour suppressor, and it is downregulated in ovarian cancer. Methylation of SERINC3 enhances tumour escape from apoptosis and increases the survival (Terasawa et al., 2004).
UOH-ALL-36	KRT 31	It belongs to keratin gene family. It expresses skin, hair, squamous epithelial cell, endometrial ciliated cells. It involves in cell homeostasis (Hansson et al., 2003).	KRT 31 is downregulated in head and neck squamous carcinoma (Silveira et al., 2008).
UOH-ALL-34	TOMM20	It encodes a translocase of outer mitochondrial membrane 20L receptor which belongs TOM complex. It plays an important in mitochondrial protein transport. It is enriched in testis and small amount is found GI, endocrine tissues (Chacinska et al., 2009).	TOMM20 is overexpressed in colon cancer and associated with high metastatic rate. Its knockdown results in a significant reduction of tumour proliferation and metastasis (Park et al., 2019).
UOH-ALL-35	NTPCR	It encodes a non-specific nucleoside triphosphatase. It involves in basic cellular processes that hydrolyses ATP and	NTPCR is significantly decreased in ovarian cancer. Increasing expression of NTPCR inhibits G2/M phase and consequently cell

		energy production and expresses in all tissues such as skeletal muscle and pancreas (Placzek et al., 2007).	invasion and migration. It has a target STAT3, and GATAS2 and binds to them acting as tumour suppressor (Shang et al., 2021).
UOH-ALL-30	ATL3	It encodes a GTPase integral membrane protein and involves in biogenesis of endoplasmic reticulum network. They are various variants produced by the alternative splicing. It is highly enriched in testis and is expressed in different levels in many tissues such as skin, thymus, adipose tissues, and colon (Wilkinson, 2019).	Loss of ATL3 leads to reduction of axonal lysosomes and autophagy. This results in axon deformity and neuropathy (Behrendt et al., 2021).
UOH-ALL-16	ATXN10	It encodes an ataxin 10 protein which involves in neuron differentiation and survival via activating Ras-MAP kinase pathway. It also has a role in maintaining homeostasis and intracellular glycosylation and is found in all tissues (Waragai et al., 2006).	It is overexpressed in cancer cachexia which a multifactorial process that leads to loss of adipose tissues and skeletal muscles and associated with 30% of cancer morbidity. Ataxin 10 along with a panel of cytokines including Adamts like 4, chemokine ligand 2, syntaxin 7, multiple inositol polyphosphate phosphatase1, bridging integrator1, and glucosidase alpha acid are used as a signature of cachexia in colorectal cancer. Measuring serum level of ataxin 10 may be used a predictive biomarker for cardiomyopathy in cachexic patients (Schäfer et al., 2016).
UOH-ALL-58	LINC00661	It belongs to long non-coding RNA (lncRNAs) that lacks protein translation with function of regulating gene transcription. It is restricted expressed in testis (Ponting et al., 2009).	It is overexpressed in Cholangiocarcinoma (CCA) and associated with lower overall survival along with other lncRNAs including CHRM3.AS2 and MIR205HG. LINC00661 overexpression is involved in autophagy dysfunction (Liu et al., 2021b).
UOH-ALL-50	UBC	UBC is encoded for a polyubiquitin protein that is conjugated to monomers or polymers. Ubiquitination has an important role in regulation of signalling pathways, DNA repair, protein degradation, and kinase modification. It is found in many cells such as skeletal muscle, tongue, gall bladder, and bone marrow (Marinovic et al., 2000). UBC reduction plays a pivotal	As UBC involves in cell cycle control, downregulation of UBC is associated with decrease in cell cycle components such as CDK1. Increased expression of UBC enhances the proliferation of MSC and accompanied with proliferative genes in cell cycle (Kim et al., 2018). It is overexpressed in renal cancer and associated with poor prognosis (Alchahin et al., 2022).

		role in replicative senescence of bone marrow mesenchymal stromal cells.	
UOH-ALL-65	ROCK1	It encodes a serine/threonine kinase protein that binds to GTP-bound form of Rho upon activation. It functions in regulating cell polarity and cytoskeletal organization, adhesion, and motility. It expresses in bone marrow, lung, urinary bladder, and adipose tissues (Julian & Olson, 2014).	ROCK1 is overexpressed in AML and associated with poor survival. ROCK1 knockdown enhances tumour cell apoptosis and significantly inhibits the blast proliferation. It has been suggested that miR 340-5p downregulation is the main cause of ROCK1 upregulation. This confirms that ROCK1 may be a therapeutic target in AML (Liu et al., 2019).
UOH-ALL-73	IDI1 isopentenyl-diphosphate delta isomerase 1	It is encoded a peroxisomally-localised enzyme that involves in cholesterol synthesis. It catalyses farnesyl diphosphate synthesis by converting isopentenyl diphosphate (IPP) to dimethylallyl diphosphate (DMAPP). It expresses in many cells such as breast, adrenal gland, skeletal muscle, liver, and bone marrow (Laupitz et al., 2004).	IDI1 is downregulated in clear cell renal carcinoma and associated with the adverse outcome (Qi et al., 2021).
UOH-ALL-59	PCAT6	PCAT6 is belongs to lncRNAs and expresses in salivary, skin, ovary and highly expressed in testis (Du et al., 2013).	PCAT6 increases the metastasis of lung cancer. In addition, overexpression of PCAT6 involves EMT transition via regulating miR-326/KLF pathway. PCAT6 Knockdown significantly inhibits tumour proliferation and M2 polarization (Chen et al., 2022c).
UOH-ALL-60	HDLBP	It involves in regulating excess cholesterol levels and sterol metabolism in cells and encodes for the protein binding to HDL. It also induces formation of heterochromatin via binding RNA. It expresses in most tissues in various levels and is enriched in pancreas and placenta (Chiu et al., 1997).	It is overexpressed in hepatocellular carcinoma and is involved in tumour growth and proliferation. Knockdown of vigilin leads to inhibit carcinogenesis and sensitized to cisplatin (Yang et al., 2014b). However, it acts a tumour suppressor in breast cancer as it binds to the 3' untranslated mRNA region of the c-fms proto-oncogene. The c-fms proto-oncogene is encoded for the receptor of macrophage colony stimulating factor (CSF1) which is function as a surface tyrosine kinase. This interaction inhibits translation of CSF1 and destabilises it (Woo et al., 2011).

UOH-ALL-62	SUN1	SUN1 encodes a nuclear envelope protein with SUN domain and belongs to the unc-84 homolog family. It plays a role in migration and nuclear anchorage. There are different variants that produced by the alternative splicing. It expresses ubiquitously in many cells such as skin, endometrium, fat, and testis (Li & Noegel, 2015).	It acts a tumour suppressor and decrease its expression in colon cancer. High expression of SUN1 is associated with good prognosis. It reduces the expression of brain derived neurotrophic factor (BDNF) and therefore inhibiting BDNF/tropomyosin-related kinase B (TrkB) signalling. It also increases the acetylation of methyl-CpG binding protein 2 (MeCP2) and associates with SIRT1 for interfering with the BDNF promoter (Liu et al., 2021a).
UOH-ALL-67	RAB5C	It encodes a Rab protein and belongs to a small GTPase of the Ras superfamily. It involves in docking, endocytosis, and vesicular transport. It is found many cells such as bone marrow, kidney, skin, urinary bladder, colon, rectum, and pancreas (Zeigerer et al., 2012).	It is overexpressed in B-ALL and its knockdown inhibit the growth. Its expression is regulated by miR509 that is a tumour suppressor and decreased in B-ALL (Tan et al., 2014).
UOH-ALL-68	RAB34	it encodes Rab protein that belongs to Golgi-bound member of the secretory pathway. It functions in micropinocytosis activation and lysosomes repositioning. It has many different variants which are produced by the alternative splicing of the transcript. It has low tissue specificity and expresses in many cells such as theca cell, peritubular cells, bone marrow, thyroid, and renal cells (Wang & Hong, 2002).	It is overexpressed in HCC and correlated with poor prognosis. RAB34 knockdown results in a significant reduced proliferation and metastasis rate via arrest of G1 phase and mesenchymal-epithelial transition (Wu et al., 2017).
UOH-ALL-69	APPL1	It is encoded an adapter protein that has a regulatory role in cell proliferation, immune response and cell trafficking via binding to membrane receptors, signalling proteins. APPL1 interacts with RAB5A and regulates signalling pathway of cell proliferation. It also inhibits Fc-gamma receptor-mediated phagocytosis via PI3K/Akt signalling interaction (Tan et al., 2010). It involves in TGFBR1 signalling. It has wide distribution in ribosomes, pancreas, thyroid, and bone marrow (Thomas et al., 2011).	APPL1 expression is correlated to TNM stage of gastric cancer and poor prognosis (Zhai et al., 2016). It activates Akt pathway via acting as a scofflod protein and phosphorylates Akt that leads to activation of Cdc42-associated kinase 1 that promote tumour growth. APPL1 induces angiogenesis via activating of nuclear factor κB (Hupalowska et al., 2012).

UOH-ALL-71	CLIC4	It encodes CLIC4 protein that belongs to the p64 family. It regulates cellular processes such as cell volume regulation, intracellular pH maintenance, and cell membrane stabilisation and transport. It expresses in many cells including pancreas, skeletal muscle, heart, B-cell, endothelial cells, and kidney (Malik et al., 2012).	It is highly expressed in AML and associated with adverse prognosis. CLIC4 is regulated by TGF-β and c-Myc. It also induces the NF-κB-dependent activation of hypoxia-inducible factor (HIF) that promotes tumour proliferation and microenvironment (Huang et al., 2020).
UOH-ALL-72	TRAP1	It encodes a mitochondrial chaperone protein which belongs to the HSP90 family. It involves in regulating cellular stress response as it has ATPase activity interacting with TNF1. It has various transcript variants that are produced by the alternative splicing. It is found in many cells such as heart, skeletal muscle, liver, tongue, bone marrow, and adrenal glands (Kang et al., 2007).	It is overexpressed in ulcerative colitis and proceeded to colon cancer. It was found that oxidative stress induces the production of TRAP1 which promotes neoplastic change and progression (Chen et al., 2014a). Upregulation of TRAP1 in T-ALL leads to apoptosis resistance and chemotherapy resistance (Ariës et al., 2018).
UOH-ALL-52	TMPO	This gene encodes multiple distinct LEM domain-containing protein isoforms that are produced by the alternative splicing. It is involved in several cellular functions such as chromatin organization, regulation of gene expression, and cell cycle control. There are three isoforms: alpha, beta, and gamma isoform. It is widely expressed ubiquitously in many tissues such as bone marrow, lymph nodes, appendix, and ovaries (Taylor et al., 2005).	It is overexpressed in glioblastoma. TMPO knockdown results in enhanced apoptosis and inhibited cell proliferation via increasing the cleavage of caspase3 and PARP protein (Zhang et al., 2016)
UOH-ALL-53	HOXD8	It belongs to the homeobox family which is a highly conservated transcription factor and plays a role in morphogenesis and cellular processes. It is expressed in the kidney, ovary, testis, and adrenal (Banerjee-Basu & Baxevanis, 2001).	It acts a tumour suppressor inducing apoptosis. It is downregulated in TNBC and increased its expression results in decreased cancer proliferation via inhibiting AKT/mTOR pathway (Zhang et al., 2021)
UOH-ALL-56	TPT1	It encodes a regulatory protein that involves in the control of cell growth and proliferation including protein synthesis, cell division, and apoptosis. It also plays an important role in	It is upregulated in diffuse large B-cell lymphoma (DLBCL) and the knockdown of TCPT leads to a reduction of cell-mediated

		mitotic and meiotic cell divisions through phosphorylation and binding to microtubules. It is found in many tissues such as the urinary bladder, thyroid, bone marrow, lung, ovary, and lymph node (Rho et al., 2011).	adhesion via regulating the p-CREB/BCL-2 signaling pathway (He et al., 2015).
UOH-ALL-66	B9D1	It encodes a B9 domain-containing protein that a component of the tectonic-like complex. It involves in ciliogenesis and hedgehog signalling. It is found in ciliated cells, endometrium, spermatids, respiratory tract, and renal cells (Garcia-Gonzalo et al., 2011).	B9D1 deletion is found in Meckel syndrome (Hopp et al., 2011). It acts as tumour suppressor gene and is inactivated in liver cancer. Its inhibition associated RAS/MAPK signalling activation promoting liver tumour in mice (Song et al., 2017).
UOH-ALL-63	MYH11	It encodes a smooth muscle myosin protein. It utilises the ATP hydrolysis for conversion chemical energy into mechanical energy that mainly acts as a contractile protein. It is highly expressed in intestine, smooth muscle, Sertoli cells, and peritubular cells (Milewicz & Kwartler, 2012).	It is downregulated in colon cancer and associated with poor survival (Wang et al., 2014). Furthermore, DNA methyltransferase 3b (DNMT3B) inhibits the expression of MYH11 via DNA methylation. Loss of MYH11 reduces its inhibitory effect on TNFRSF14 transcription and thus promoting tumour progression in gastric cancer (Wang et al., 2021b).
UOH-ALL-41	UBE3D	It encodes an E3 ubiquitin-protein ligase that transfers ubiquitin to the target gene. Gene ubiquitination mainly promotes their degradation by proteasome. It also enables cyclin binding activity. It is widely expressed in tissues such as neuron, and oligodendrocytes (Huang et al., 2015).	It is highly expressed in breast cancer. UBE3D stabilizes CPSF73 and protects it from degradation. CPSF73 plays an important role in maintaining cancer proliferation and recurrence especially self-renewal cells. Knockdown of UBE3D depletes CPSF73 that resulting in the inhibition of tumour growth (Liu et al., 2022a).
UOH-ALL-43	TEX43	TEX43 involves in sperm motility. It is highly enriched in testis and small amount in colon.	Not found
UOH-ALL-49	CCDC89	CCDC89 involves in cilium organisation and sperm motility.	Not found
UOH-ALL-97	CUL1	It belongs to E3 ubiquitin ligase that involved in ubiquitination of cell cycle components and related proteins involved in signal transduction. It is broadly found in most tissues and highly enriched in the testis (Skaar et al., 2014)	It is upregulated in osteosarcoma and associated with poor prognosis via inducing EMT and wnt pathway activation as well as miR-377-3p inhibition (Liang et al., 2021).

UOH-ALL-79	PTPN23	It is similar to PTPN1 (Hendriks & Pulido, 2013)	It is a tumour suppressor gene and its loss in BC associated with poor survival. Induction of PTPN23 inhibits FYN kinase and B-catenin (Zhang, 2017)
UOH-ALL-100	DKK3	It belongs to dickkopf family that inhibits wnt activation. It is mainly expressed in many tissues with high expression in heart and brain (Lee et al., 2009)	It is downregulated in head and neck cancer and correlated to shorter survival (Katase et al., 2020)
UOH-ALL-1	TPM3	It belongs to actin-binding protein regulating other binding proteins via copolymerisation. It is enriched in bone marrow, appendix, and found ubiquitous in other tissues (De Paula et al., 2009)	It (Xu et al., 2021) is overexpressed in ovarian cancer. Targeting of TPM3 uses ATM-3507 synergistically with anti-microtubules such as vinorelbine inducing cell death.
UOH-ALL-2	TMCO3	It is encoded of the monovalent cation: proton antiporter 2 (CPA2) regulating proton (Na, K) transport across cellular membranes (Van Everdink et al., 2003)	It acts as an oncogene, its upregulation found in HCC and associated with poor prognosis. Knockdown of TMCO3 results in the increase of p53/p21 inducing cell apoptosis as well as reduction of EMT (Dai et al., 2022).
UOH-ALL-3	C10orf82	It is uncharacterised protein and recently suggested as CTA. It is highly enriched in the testis and small expression found in the ovary (Almutairi et al., 2022).	Its expression correlated to the overall survival to ovarian cancer (Almutairi et al., 2022)
UOH-ALL-4	UBE2D2	It belongs E3 ubiquitin ligases and their substrates p53, and peroxisomal biogenesis factor 5 (PEX5). It is expressed in many tissues such as bone marrow, thyroid (Metzger et al., 2012).	Circ_UBE2D2 upregulation found in breast cancer and poor prognosis. It targets miR-200a-3p and its downregulation increases tamoxifen sensitivity (Hu et al., 2020b).
UOH-ALL-77	UBE3C	It is similar to UBE2D2 (Metzger et al., 2012).	It is upregulated in renal carcinoma via activating wnt pathway and associated with metastasis (Wen et al., 2015)
UOH-ALL-5	MYL12B	It encoded of a non-muscle myosin II regulating myosin assembly. It is found in all tissues lung, colon (Vicente-Manzanares et al., 2007).	MYL12B upregulation is associated with melanoma resistant. Inhibition of MYL12B results inducing ROS and pro-survival pathways leading to cell arrest and apoptosis.
UOH-ALL-90	MYH10	It belongs a member of the myosin superfamily. It involved in regulating in cell motility and cytokinesis. It found in many tissues lung (Vicente-Manzanares et al., 2007).	MYH10 loss is associated with HCC metastasis and poor prognosis via the activation of EGFR pathway (Jin et al., 2021).

UOH-ALL-92	RPL28	It encodes the 60S large subunit components of ribosomal protein L28E family. It catalysed protein synthesis and found in many tissues in spleen, ovary (Warner & McIntosh, 2009).	RPL28 overexpression found colorectal cancer and associated with shorter survival via high expression of extracellular matrix (Labriet et al., 2019).
UOH-ALL-89	EIF3A	It enables RNA binding activity regulating translation initiation factor activity mainly DNA synthesis (Yin et al., 2018).	Upregulation of eIF3a mediates chemo-resistance of SCLC via inducing of cellular senescence. EIF3A Knockdown inhibits tumour progression and chemotherapy resistance via reversing TGF-β1(Chen et al., 2021b)
UOH-ALL-93	MRFAP1	It is a nuclear protein that involved in cellular growth and maintained normal histone modification (Larance et al., 2012).	MRFAP1 is downregulated in gastric cancer and its overexpression decreases cancer proliferation. As it is highly expressed in response to inhibition of neddylation activation enzyme (NAE) which plays a crucial role in protein turnover (Hu et al., 2018).
UOH-ALL-101	PRM1	It is involved in DNA packaging and spermatid development. It is restricted found in the testis (Schagdarsurengin et al., 2012)	PRM1 is highly upregulated in colon cancer and correlated to the stage and prognosis. Inhibition of PRM1 results in the decrease cancer proliferation (Chen et al., 2018c).
UOH-ALL-96	PARL	It is a key regulator of mitochondrial homeostasis regulating apoptosis and remodelling. It is found in the most tissues (Meissner et al., 2011).	It is downregulated in pancreatic cancer and associated with chemo-resistance (Qin et al., 2023).
UOH-ALL-99	C18orf32	It is involved in the activation of NF-kappa-B pathway. It is found ubiquitinated in all tissues (Matsuda et al., 2003).	C18orf32 loss is associated with the autosomal recessive neurodevelopmental disorder with hypotonia and contractures (Salian et al., 2022)
UOH-ALL-98	LINC00251	It belongs to long non-coding RNA and strictly expressed in the testis. It has synonym C8orf25; NCRNA00251 (Karol et al., 2016)	LINC00251 polymorphism with other genes (PMP7, DOK5) increase the risk of osteonecrosis in children with B-ALL (Maamari et al., 2020)
UOH-ALL-84	BCLAF1	It is encoded a transcriptional repressor interacting with BCL2 proteins and inducing apoptosis. It is expressed ubiquitously in most tissues (Lee et al., 2012)	BCLAF1 is upregulated in AML and associated shorter survival as well as low miR-194-5p. Its knockdown induces blast differentiation (White et al., 2018)
UOH-ALL-81	KIF1B	It is similar KIF5B (Hirokawa & Tanaka, 2015)	KIF1B polymorphism increases the risk of HCC (Su et al., 2017)

UOH-ALL-85	LOC338963	It belongs to long non-coding RNA and strictly expressed in the testis (Karol et al., 2016)	It is highly expressed in Lambert-Eaton myasthenic syndrome (LEMS) which is an autoimmune neuromuscular junction disorder (Yang et al., 2022)
UOH-ALL-86	FBXO22 (FBX22; FISTC1)	It is encoded a member of the F-box protein that functions in ubiquitination of transcriptional target catalysing their degradation. It is found in the most tissues (Skaar et al., 2014).	It promotes AML progression especially with MLL subtype sustaining leukaemia stem cell (LSC). Targeting FBXO22 increases BACH1 and eradicate LSC (Zhu et al., 2023).
UOH-ALL-83	TPPP2	It is involved in tubulin binding activity and regulating the sperm motility. It is mainly found in testis and salivary gland (Zhu et al., 2019). (P18; CT152; C14orf8; p25beta)	TPPP2 deficiency is related to male infertility (Zhu et al., 2019). It is associated with poor prognosis in hepatocellular cancer (Xu et al., 2023).
UOH-ALL-88	AKAP1	It belongs to anchor proteins that involves in the signalling pathway of cAMP dependent and in confining RNA to a specific cellular compartment. Thus, it regulates oxidative metabolism, biogenesis, and cell survival. It is found in many tissues (Affaitati et al., 2003)	AKAP1 triggers glioblastoma growth and is a transcriptional target of Myc. Downregulation of AKAP1 is coupled with the assembled of sestrin2 which a leucine sensor and mTOR inhibitor on the mitochondria inhibiting tumour growth (Rinaldi et al., 2017)
UOH-ALL-80	SURF2	This gene has a bidirectional promoter with surfeit 1 (SURF1). It is expressed in many tissues and acts as a cargo receptor (Lennard et al., 1994)	It is down-stream regulator of the oncogene, superoxide dismutase 2 (SOD2), which is highly expressed in head and neck cancer and associated radiotherapy resistance (Jung et al., 2019)
UOH-ALL-87	UVRAG-DT	It is a ncRNA and highly expressed in testis and fat. It is involved in regulating smooth muscle veins (Kimura et al., 2006)	circUVRAG (Yang et al., 2019a) is upregulated in bladder cancer and associated with poor survival. Its silencing inhibits cancer progression and metastasis via targeting miR-223/ fibroblast growth factor receptor 2 (FGFR2)
UOH-ALL-78	ZNF676	It is involved in transcription regulation enabling DNA-binding transcription factor, RNA polymerase II specific DNA binding activity as well as telomerase haemostasis. It has a broad expression in many tissues (Mangino et al., 2012).	ZNF676 overexpression is found in pituitary neuroendocrine tumour and correlated to invasive tumour (Peculis et al., 2021)
UOH-ALL-57	MIP	It has a crucial role in maintaining lens homeostasis and transparency (Lo et al., 2014) as well as intracellular	MIP(Khan et al., 2021) is involved in promoting cancer growth and proliferation via activating the PI3K/Akt signalling in breast cancer

		communication. It is found in eye and other tissues such as testis, liver, brain and adrenal.	

Table 9.5 Geng et al., 2012 (GSE38403) analysis of their expression and association with the survival using BloodSpot

Gene	Probe set	High expression	Survival
FLT4	234379-at	ALL subtypes except B-ALL with t (8;14), t (1;19) - p<0.001	0.0713
	210316-at	ALL subtypes except B-ALL with t (8;14), t (1;19) - p<0.001	0.0353
	229902-at	ALL subtypes except B-ALL with t (8;14), t (1;19) - p<0.001	0.0107
MRC1	204438-at	ALL subtypes - p<0.001/ B-ALL with t (8;14)- p<0.05	0.00117
BMP2	205290-s-at	ALL hyperdiploid p<0.001	0.347
	205289-at	ALL hyperdiploid/ ALL t (12;21) p<0.001	0.485
ROR1	205805-s-at	B-ALL with t (1;19) p<0.001	0.289
	211057-at	B-ALL with t (1;19) p<0.001	0.326
	232060-at	B-ALL with t (1;19) p<0.001	0.358
P2RY5		Not found	
CTGF	209101-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.863
ADAM23	206046-at	B-ALL with t (8;14) p<0.001	0.785
	213808-at	Healthy bone marrow	0.963
	244463-at	Healthy bone marrow	0.679
	1559268-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.549
APBB2	212972-x-at	B-ALL with t (1;19) p<0.001	0.178
	212985-at	B-ALL with t (1;19) p<0.001	0.157
	213419-at	B-ALL with t (1;19) p<0.001	0.181
	216747-at	Pre-B-ALL t (9;22) p<0.001	0.683
	216750-at	B-ALL with t (8;14) p<0.001	0.261
	40148-at	B-ALL with t (1;19) p<0.001	0.632
	212970-at	B-ALL with t (1;19) p<0.001	0.374
EMP1	201324-at	ALL subtypes except c-/Pre-B-ALL no t (9;22)- p<0.001/B-ALL with t (12;21)- p<0.01	0.0343
	201325-s-at	ALL subtypes except c-/Pre-B-ALL no t (9;22)- p<0.001/B-ALL with t (12;21)- p<0.01	0.0115
	1564796-at	Pre-B-ALL t (9;22) p<0.001	0.731
	213895-at	Pre-B-ALL t (9;22)/ ALL hyperdiploid p<0.001	0.754
PTPN14	226282-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.176
	244533-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.239
	242321-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.29
	205503-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.62
S100A16	227998-at	ALL hyperdiploid/ c-/Pre-B-ALL t (9;22)- p<0.001	0.232
PRO2646		Not found	
FLJ00125		Not found	
FCRL5	224406-s-at	B-ALL with t (8;14) - p<0.001	0.402

	231647-s-at	B-ALL with t (8;14) - p<0.001	0.766
	224404-s-at	B-ALL with t (8;14) - p<0.001	0.65
	224405-at	B-ALL with t (8;14) - p<0.001	0.95
	1555799-at	B-ALL with t (8;14) - p<0.001	0.57
KIAA0226L (C13orf18)	219471-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.513
	44790-s-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.454
DPPA4	241199-x-at	B-ALL with t (8;14) - p<0.001	0.913
	232985-s-at	Healthy bone marrow-NS	0.0823
	219651-at	Healthy bone marrow-NS	0.0194
KDM5D(HY-67)	206700-s-at	Pro-B-ALL t(11q23) /MLL- p<0.05	0.645
SPATA1(SPAP1)	221057-at	Pro-B-ALL t(11q23) /MLL- p<0.05	0.793
CYTL1(C17)	219837-s-at	ALL subtypes of B-ALL-p<0.001	0.00945
PLVAP(FELS)	221529-s-at	ALL subtypes except B-ALL with t (8;14), t (12;21), t (1;19) - p<0.001	0.0445

Table 9.6 CTA associated with survival from CTA database 11 of 89 were associated with survival

Gene	Probe set	High expression	Survival
MAGEA4	214254-at	Mature B-ALL ALL with t (8;14)- p<0.05	0.00404
TEX101	223906-s-at	B-ALL with t (8;14)-NS	0.00539
GPATCH2	243704-at	ALL t (1;19) p<0.001	0.009
CTAGE5	204055-s-at	ALL hyperdiploid p<0.001	0.0129
RQCD1	213179-at	ALL t (1;19) p<0.001	0.0137
DNAJB2	202500-at	c-/Pre-B-ALL t (9;22) p<0.01	0.0183
DNAJB4	203810-at	ALL t (12;21)-NS	0.0185
MORC3	213000-at	ALL hyperdiploid-NS/ ALL t (12;21) p<0.01	0.0232
SYCP2L	236337-at	ALL with t (1;19) p<0.001	0.0246
CTCFL	1552368-at	B-ALL with t (8;14)-NS	0.028
NR6A1	211402-x-at	Healthy bone marrow-p<0.01	0.0345
MAGEA6	214612-x-at	c-/Pre-B-ALL t (9;22)-NS	0.0468
MORC4	219038-at	ALL hyperdiploid-NS	0.0473

Table 9.7 Protein identified from protoarray by Jordenset al. 2020 that associated with the survival

Antigen	Probe set	High expression	survival
BMX	206464-at	Healthy bone marrow-p<0.001	0.0461
SEPT9	1559025-at	ALL subtypes of B-ALL-p<0.001	0.0205

Table 9.8 LAA from Iacobucci & Mullighan, 2017

Antigen	Probe set	High expression	survival
HOMER1	226651-at	ALL subtypes of B-ALL-p<0.05	0.000581
HOMER3	215489-x-at	Healthy bone marrow- p<0.001	0.0311
CDKN2B	207530-s-at	ALL t (8;14)-NS	0.0193
IKZF1	1565817-at	ALL t (12;21)/ t (8;14)-p<0.001/t (1;19)- p<0.05	0.0304

Table 9.9 Expression analysis of the main components of hippo pathway and their association with the survival using BloodSpot

Gene	Probe set	High expression	p-value
LATS1	227772-at	ALL hyperdiploid p<0.001	0.671
	219813-at	ALL hyperdiploid p<0.001	0.495
	1570425-s-at	B-ALL with t (8;14)-NS	0.477
	1570231-at	c-/Pre-B-ALL no t (9;22)-NS	0.337
LATS2	227013-at	Pro-B-ALL t(11q23)/MLL/ ALL hyperdiploid p<0.001	0.0562
	230348-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.456
	223380-s-at	Pro-B-ALL t(11q23)/MLL/ ALL hyperdiploid p<0.001	0.321
	223379-s-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.948
YAP1	224895-at	ALL t (1;19)-NS	0.858
	224894-at	Healthy bone marrow p<0.01	0.0825
	213342-at	Healthy bone marrow -NS	0.968
TEAD1	1553322-s-at	B-ALL with t (8;14) p<0.01	0.218
	224955-at	B-ALL with t (8;14) p<0.01	0.356
	214600-at	ALL t (12;21)-NS	0.906
TEAD2	243766-s-at	B-ALL with t (8;14)-NS	0.392
	226408-at	B-ALL with t (8;14)-NS	0.544
	238322-s-at	B-ALL with t (8;14)-NS	0.37
	238323-at	B-ALL with t (8;14)-NS	0.292
	238321-at	B-ALL with t (8;14)-NS	0.244
TEAD3	209454-s-at	B-ALL with t (8;14)- p<0.001	0.275
TEAD4	204281-at	B-ALL with t (8;14), t (1;19), c-/Pre-B-ALL t (9;22) - p<0.001	0.119
	41037-at	B-ALL with t (8;14), t (1;19), c-/Pre-B-ALL t (9;22) - p<0.001	0.0223
TAZ	37278-at	ALL hyperdiploid-NS	0.464
	203977-at	ALL hyperdiploid-NS	0.927
MST1	216320-x-at	Healthy bone marrow p<0.001	0.262
	205614-x-at	Healthy bone marrow p<0.001	0.593
MST2		Not found	
SAV1	218276-s-at	ALL t (1;19)- p<0.001	0.116
	222573-s-at	ALL t (1;19)/ c-/Pre-B-ALL t (9;22)- p<0.001	0.245
	234491-s-at	ALL t (1;19)- p<0.001	0.0885
MOB1a	201299-s-at	Pro-B-ALL t(11q23) /MLL p<0.001	0.625
	214812-s-at	ALL t (1;19)- p<0.001	0.986
	201298-s-at	ALL t (1;19)- p<0.001	0.552
	201297-s-at	ALL t (1;19)- p<0.001	0.66
MOB1b	225997-at	Healthy bone marrow p<0.001	0.0145
MOBKL2A		Not found	

AURKA	208079-s-at	Healthy bone marrow p<0.001	0.0641
	204092-s-at	Healthy bone marrow p<0.001	0.0978
	208080-at	ALL hyperdiploid p<0.001	0.274
AURKB	239219-at	B-ALL with t (8;14)-NS	0.415
	209464-at	ALL t (1;19)-NS	0.0781
WWTR1	202132-at	B-ALL with t (8;14)-NS	0.548
	202134-s-at	B-ALL with t (8;14)-NS	0.786
	202133-at	c-/Pre-B-ALL t (9;22)-NS	0.639
STK4	236259-at	Healthy bone marrow p<0.001	0.377
	225364-at	B-ALL with t (8;14) p<0.01	0.512
	243981-at	ALL t (12;21) p<0.001	0.33
	223746-at	Healthy bone marrow p<0.001	0.256
	211085-s-at	Healthy bone marrow p<0.001	0.891
	205411-at	Healthy bone marrow p<0.001	0.653
	1569791-at	Healthy bone marrow p<0.001	0.616
STK3	204068-at	ALL hyperdiploid p<0.001	0.672
	211078-s-at	ALL hyperdiploid p<0.001	0.0684
AMOTL2	203002-at	c-/Pre-B-ALL t (9;22)-NS	0.635
TAOK1	227454-at	ALL subtypes except B-ALL with t (8;14), - p<0.001/ALL t (1;19)- p<0.05	0.65
	231193-s-at	ALL subtypes except B-ALL with t (8;14), - p<0.001/ALL t (1;19)- p<0.05	0.971
	224778-s-at	ALL subtypes except B-ALL with t (8;14), - p<0.001/ALL t (1;19)- p<0.05	0.468
	224769-at	ALL subtypes except B-ALL with t (8;14), - p<0.001/ALL t (1;19)- p<0.05	0.239
	238420-at	ALL subtypes except B-ALL with t (8;14), - p<0.001/ALL t (1;19)- p<0.05	0.08
	216310-at	ALL subtypes except B-ALL with t (8;14), - p<0.001/ALL t (1;19)- p<0.05	0.0328
TAOK2	204877-s-at	B-ALL t (12;21) - p<0.05	0.269
	204878-s-at	B-ALL t (12;21) - p<0.05	0.0446
	204986-s-at	B-ALL t (12;21) - p<0.05	0.111
TAOK3	220761-s-at	B-ALL t (12;21) - p<0.001	0.549
	221508-at	B-ALL t (12;21) - p<0.001	0.811
FRMD6	225464-at	Healthy bone marrow p<0.001	0.00172
	225481-at	ALL subtypes of B-ALL-p<0.001	0.0939
NF2	204991-s-at	B-ALL with t (8;14)-NS	0.192
WWC1	216074-x-at	B-ALL with t (8;14)-NS	0.469
Wnt components			
Gene	Probe set	High expression	p-value
RYK	202853-s-at	ALL t (12;21)-NS	0.968
	216976-s-at	Healthy bone marrow-NS	0.498
	214172-x-at	ALL t (12;21)-NS	0.404
Wnt1	208570-at	Pro-B-ALL t(11q23) /MLL-NS	0.692
Wnt2B	206459-s-at	B-ALL with t (8;14)-NS	0.548
	206458-s-at	B-ALL with t (8;14)-NS	0.695
Wnt2	205648-at	Pro-B-ALL t(11q23) /MLL-NS	0.0974
LEF1	221558-s-at	ALL t (1;19)- p<0.001	0.883
	221557-s-at	ALL t (1;19)- p<0.001	0.853

	210948-s-at	ALL t (1;19)- p<0.001	0.326
Wnt3A		Not found	
WNT5A	205990-s-at	B-ALL with t (8;14) p<0.001	0.78
	231227-at	B-ALL with t (8;14) p<0.001	0.532
	213425-at	B-ALL with t (8;14) p<0.001	0.992
WNT10B	206213-at	B-ALL with t (8;14)-NS	0.655
WNT16B		Not found	
Wnt16	221113-s-at	ALL t (1;19)- p<0.001	0.923
	224022-x-at	ALL t (1;19)- p<0.001	0.971
Wnt11	206737-at	B-ALL with t (8;14)-NS	0.532
WNT10A	229154-at	B-ALL with t (8;14) p<0.05	0.59
	223709-s-at	B-ALL with t (8;14) p<0.05	0.18
Wnt10B	206213-at	B-ALL with t (8;14)-NS	0.655
Wnt14		Not found	
FZD8	216587-s-at	c-/Pre-B-ALL t (9;22)- p<0.05	0.637
	227405-s-at	ALL t (12;21) p<0.05	0.937
	224325-at	ALL t (12;21) p<0.05	0.0795
FZD9		Not found	
SFRP1	228413-s-at	B-ALL with t (8;14)-NS	0.483
	202036-s-at	Healthy bone marrow-NS	0.587
	202035-s-at	Healthy bone marrow-NS	0.473
	202037-s-at	Healthy bone marrow p<0.001	0.335
SFRP2	223121-s-at	Healthy bone marrow p<0.001	0.46
	223122-s-at	Healthy bone marrow p<0.001	0.725
SFRP4	204051-s-at	B-ALL with t (8;14) p<0.01	0.333
	204052-s-at	B-ALL with t (8;14) p<0.01	0.696
SFRP5	207468-s-at	B-ALL with t (8;14)-NS	0.486
DKK1	204602-at	Pro-B-ALL t(11q23) /MLL- p<0.001	0.128
DKK2	224199-at	Healthy bone marrow p<0.01	0.338
	219908-at	Pro-B-ALL t(11q23) /MLL- p<0.01	0.671
DKK3	214247-s-at	Healthy bone marrow-NS	0.933
	221127-s-at	Healthy bone marrow p<0.001	0.423
	21126-at	B-ALL with t (8;14)-NS	0.451
	230508-at	B-ALL with t (8;14)-NS	0.1
	202196-s-at	B-ALL with t (8;14)-NS	0.0233
DKK4	206619-at	B-ALL with t (8;14)-NS	0.637
DKKL1	220284-at	B-ALL with t (8;14) p<0.05	0.109
WIF-1		Not found	
TCF4	212382-at	ALL subtypes- p<0.001	0.00214
	203753-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.0215
	212385-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.00603
	222146-s-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.0577
	213891-s-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.0292
	212386-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.0172
	212387-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.012
	228837-at	ALL t (1;19)- p<0.001	0.215
TCF3	213730-x-at	ALL t (12;21)/ c-/Pre-B-ALL t (9;22)- p<0.001	0.448
	213731-s-at	B-ALL with t (8;14) c-/Pre-B-ALL no t (9;22)- p<0.001	0.142
	210776-x-at	ALL t (12;21)/ c-/Pre-B-ALL t (9;22)- p<0.001	0.955

	213811-x-at	ALL t (12;21)/ c-/Pre-B-ALL t (9;22)- p<0.001	0.893
	213732-at	B-ALL with t (8;14) c-/Pre-B-ALL no t (9;22)- p<0.001	0.635
	215260-s-at	ALL subtypes- p<0.001	0.0297
	213809-x-at	B-ALL with t (8;14) p<0.001	0.436
	216645-at	c-/Pre-B-ALL no t (9;22) c-/Pre-B-ALL t (9;22)- p<0.001	0.642
	209153-s-at	ALL subtypes- p<0.001	0.369
	216647-at	B-ALL with t (8;14) p<0.001	0.852
	228052-x-at	B-ALL with t (8;14) p<0.001	0.581
	209152-s-at	ALL subtypes- p<0.001	0.494
	209151-x-at	ALL t (1;19)- p<0.001	0.643
TCF1		Not found	
Wnt6	221608-at	B-ALL with t (8;14)-NS	0.552
	221609-s-at	B-ALL with t (8;14)-NS	0.431
	222086-s-at	Pro-B-ALL t(11q23) /MLL/ c-/Pre-B-ALL t (9;22) - p<0.001	0.6
	71933-at	ALL hyperdiploid- p<0.001	0.584
Wnt5B	223557-s-at	B-ALL with t (8;14)-NS	0.434
	221029-s-at	B-ALL with t (8;14)-NS	0.224
	230299-s-at	B-ALL with t (8;14)-NS/ Pro-B-ALL t(11q23)/MLL- p<0.05	0.183
FZD3	239082-at	ALL t (1;19)- p<0.001	0.538
	227524-at	B-ALL with t (8;14)/ ALL t (1;19)- p<0.001	0.751
	219683-at	ALL t (1;19)- p<0.001	0.695
	227499-at	ALL t (1;19)- p<0.001	0.991
LRP5	209468-at	ALL subtypes except B-ALL with t (8;14) - p<0.001	0.662
	229591-at	Pro-B-ALL t(11q23) /MLL- p<0.001	0.388
LRP6	205606-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.0766
	225745-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.949
	34697-at	ALL hyperdiploid/ ALL t (1;19)- p<0.001	0.372
GSK3β		Not found	
CK1α		Not found	
WIF1	204712-at	Healthy bone marrow p<0.01	0.659
SFRP5	207468-s-at	B-ALL with t (8;14)-NS	0.486
DVL3	201907-x-at	ALL subtypes except B-ALL with t (8;14), hyperdiploid - p<0.001	0.118
	201908-at	ALL subtypes except B-ALL with t (8;14), t (9;22) - p<0.001	0.0279
HDAC1	201209-at	ALL t (1;19)/ Pro-B-ALL t(11q23)/MLL - p<0.001	0.233

Table 9.10 Expression analysis of the main component of TGFβ and their association with the survival using BloodSpot

Antigen	Probe set	High expression	Survival
TGFBR2	207334-s-at	ALL with t (1;19)- p<0.01	0.291
	208944-at	ALL with t (1;19)- p<0.01	0.626
TGFBR1	236561-at	ALL with t (12;21)- p<0.01	0.463

	224793-s-at	ALL with t (12;21)- p<0.01	0.132
	206943-at	ALL with t (12;21)- p<0.01	0.94
TGFB1	203085-s-at	Healthy bone marrow p<0.001	0.000832
	203084-at	c-/Pre-B-ALL t (9;22)/ c-/Pre-B-ALL no t (9;22)- p<0.001	0.456
SMURF1	212666-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)/ ALL hyperdiploid -NS	0.747
	232665-x-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)/ ALL hyperdiploid -NS	0.0262
	215458-s-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)/ ALL hyperdiploid -NS	0.333
	212668-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)/ ALL hyperdiploid -NS	0.734
	1559426-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)/ ALL hyperdiploid -NS	0.15
SMURF2	205596-s-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS	0.162
	227489-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS	0.0208
	232020-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS	0.0305
SMAD3	218284-at	ALL subtypes of B-ALL- p<0.05	0.0317
	205398-s-at	ALL subtypes of B-ALL- p<0.05	0.005
	205397-x-at	ALL subtypes of B-ALL- p<0.05	0.137
	205396-at	ALL subtypes of B-ALL- p<0.05	<0.0001

Table 9.11 Expression analysis of the main component of FOXO pathways and their association with the survival using BloodSpot

Gene	Probe set	High expression	p-value
FOXO1	202724-s-at	ALL subtypes of B-ALL except Pro-B-ALL t(11q23)/MLL -p<0.001	0.599
	202723-s-at	ALL subtypes of B-ALL except Pro-B-ALL t(11q23)/MLL -p<0.001	0.835
	228484-s-at	ALL subtypes of B-ALL except Pro-B-ALL t(11q23)/MLL -p<0.001	0.0309
FKHR-L1		Not found	
FOXO3	231548-at	Healthy bone marrow p<0.01	0.319
FOXO3a		Not found	
FOXO3b		Not found	
FOXO4	205451-at	ALL hyperdiploid-NS	0.105
FOXM1	202580-x-at	ALL t (1;19)-NS	0.777
FOXO6	239657-x-at	ALL subtypes of B-ALL except Pro-B-ALL/MLL, hyperdiploid -p<0.001/ ALL t (12;21) -p<0.05	0.0364
FOXG1	207658-s-at	B-ALL with t (8;14), c-/Pre-B-ALL no t (9;22)-NS	0.231
	206018-at	B-ALL with t (8;14), c-/Pre-B-ALL no t (9;22)-NS	0.378

Table 9.12 Expression analysis of the main components of B cell signalling and their association with the survival using BloodSpot

Gene	Probe set	High expression	p-value
IGLL1	206660-at	ALL hyperdiploid/ ALL t (12;21) -p<0.001	0.304
Igβ *(CD79B)*	205297-s-at	ALL subtypes of B-ALL-p<0.001	0.694
	1555748-x-at	ALL subtypes of B-ALL-p<0.001	0.00539
	1555746-at	ALL subtypes of B-ALL-p<0.001	0.698
IGHM	212827-at	Healthy bone marrow p<0.01	0.0966
	209374-s-at	Healthy bone marrow p<0.01	0.807
	216491-x-at	Healthy bone marrow p<0.01	0.0761
SLP65 (Blank)	207655-s-at	ALL subtypes of B-ALL-p<0.001	0.083
Igll1	206660-at	ALL t (12;21)/ t (1;19) -p<0.001	0.304
Igα (CD79A)	205549-s-at	ALL subtypes of B-ALL-p<0.001	0.558
	1555779-a-at	ALL subtypes of B-ALL-p<0.001	0.788
PLCG2 (PLCγ2)	204613-at	Healthy bone marrow p<0.001	0.704
	1563263-at	Healthy bone marrow p<0.001	0.969
BTK	205504-at	ALL subtypes of B-ALL-p<0.001	0.848
VAV1	206219-s-at	ALL hyperdiploid/ ALL t (12;21) -p<0.001/ Pro-B-ALL/MLL- p<0.01	0.0465
VAV2	226063-at	Pro-B-ALL/MLL- p<0.001	0.904
VAV3	224221-s-at	ALL subtypes of B-ALL except t (8;14), hyperdiploid -p<0.001/ Pro-B-ALL/MLL /ALL t (12;21) -p<0.01	0.0896
	218807-at	ALL subtypes of B-ALL except t (8;14), hyperdiploid -p<0.001/ Pro-B-ALL/MLL /ALL t (12;21) -p<0.01	0.13
	218806-s-at	ALL subtypes of B-ALL except t (8;14), hyperdiploid -p<0.001/ Pro-B-ALL/MLL /ALL t (12;21) -p<0.01	0.00597
VAV3-AS1		Not found	
BLK	206255-at	ALL subtypes of B-ALL-p<0.001	0.39
	210934-at	ALL subtypes of B-ALL-p<0.001	0.0272
LYN	202625-at	B-ALL with t (8;14) -p<0.001	0.0279
	202627-s-at	B-ALL with t (8;14) -p<0.001	0.108
	210754-s-at	B-ALL with t (8;14) -p<0.001	0.0551
ZAP70	214032-at	c-/Pre-B-ALL t (9;22)- p<0.001	0.973
	1555613-a-at	c-/Pre-B-ALL t (9;22)- p<0.001	0.45
VPREB1	221349-at	ALL subtypes of B-ALL-p<0.001	0.764
VPREB2		Not found	
VPREB3	220068-at	ALL subtypes of B-ALL-p<0.001	0.83
SYK	226068-at	Healthy bone marrow-NS	0.263
	207540-s-at	Healthy bone marrow-NS	0.228
	244023-at	Healthy bone marrow-NS	0.582
	209269-s-at	Healthy bone marrow-NS	0.00757
IGLL3P	215946-x-at	ALL t (8;14)/ t (1;19) -NS	0.0902
IGLL3		Not found	

Table 9.13 Expression of the upstream regulators and association with the survival using BloodSpot

Antigen	Probe set	High expression	Survival
VEGFA	212171-x-at	Healthy bone marrow p<0.05	0.628
	210513-s-at	Healthy bone marrow p<0.05	0.736
	211527-x-at	Healthy bone marrow p<0.05	0.938
	210512-s-at	Healthy bone marrow p<0.05	0.323
IL-6	205207-at	B-ALL with t (8;14)-NS	0.545
IL-13	207844-at	B-ALL with t (8;14)-NS	0.721
P38MAPK		Not found	
ERBB2	216836-s-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS	0.652
	210930-s-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS	0.143
	234354-x-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS	0.393
IFNA1	208375-at	B-ALL with t (8;14)- p<0.05	0.48
PGR	228554-at	B-ALL with t (8;14)-NS	0.173
KLF6	211610-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS, ALL with t (1;19)-NS	0.531
	1555832-s-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS, ALL with t (1;19)-NS	0.343
	208961-s-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS, ALL with t (1;19)-NS	0.896
	208960-s-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS, ALL with t (1;19)-NS	0.362
	224606-at	ALL subtypes of B-ALL- p<0.001 except B-ALL with t (8;14)-NS, ALL with t (1;19)-NS	0.823
CD3		Not found	
Wnt3A		Not found	
NOTCH1	218902-at	c-/Pre-B-ALL no t (9;22)- p<0.01	0.245
	223508-at	c-/Pre-B-ALL no t (9;22)- p<0.01	0.939
PTGE2		Not found	

Table 9.14 Previous studied DEG microarray (Genget al. 2012) in leukaemia

Gene	Disease	Function
FLT4	AML	Overexpression of FLT4 was found in AML promoting tumour progression, survival, and NK dysfunction. FLT4 inhibitor (Lee *et al*., 2022), MAZ51, reduced immunosuppression restoring T-cell and NK function and increased IFN-γ.
CYTL1(C17)	CMML	It overexpressed in chronic myelomonocytic leukaemia (CMML) and involved in the activation of MEK that triggers resistance to apoptosis. Inhibition of both MCL1 and MAPK (Sevin *et al*., 2021) results in a decreased leukemic burden and apoptosis.
EMP1	cB-ALL	It overexpressed in B-ALL and associated with poor prognosis. It is involved in the activation of Src kinase family and hippo pathway. EMP1 inhibition induced cell cycle arrest and apoptosis (Ariës *et*

		al., 2014). Overexpression of EMP1 was found in glucocorticoids resistance (Singh *et al*., 2021)
PLVAP	MB	Plasmalemma vesicle associated protein (PLVAP) is a member of diaphragmed endothelial fenestrations and increased in tumour endothelium (Mellberg *et al*., 2009). It overexpressed in medulloblastoma (MB) that involved in angiogenesis and Wnt pathways mediates resistance to anti- PD-L1 (Phoenix *et al*., 2016)
MRC1	aB-ALL	CD206 (Dander *et al*., 2021) was prevalently expressed in B-ALL that is involved in sustaining leukaemic niche survival and proliferation
BMP-2	c B-ALL	It increased and involved in leukemic microenvironment with high expression of *CTGF* associated with poor prognosis (Tesfai *et al*., 2012)
DPPA4	NSCLC	It was upregulated in NSCLC and associated with poor prognosis. DPPA4 knockdown (Li et al., 2019c) inhibited tumour proliferation via targeting glycolysis and lower lactate dehydrogenase B.

Table 9.15 Previous studied DEG GSE13204 (Kohlmann et al., 2008) in cancer

Gene	Disease	Function
LIN7A	AML	LIN7A (He *et al*., 2023) downregulated in AML due to hypermethylation and associated with poor clinical outcomes. LIN7A is suggested as a predictive biomarker for decitabine treatment in AML. Its expression synergises the antiapoptic efficacy of decitabine and cytarabine especially in patients with AML t (8;21).
LILRA3	B-NHL	LILRA3 (Low *et al*., 2013) downregulated in B-non-Hodgkin's lymphoma (B-NHL).
VSTM1	AML	VSTM1 (Xie *et al*., 2015) downregulated in AML due to promoter methylation. SIRL-1 restoration inhibits leukaemia growth suggesting a potential target for leukaemia diagnosis and treatment. Signal inhibitory receptor on leukocytes-1 (SIRL-1) inhibits myeloid effector functions such as ROS.
P2RY2	AML	It upregulated in AML via the activation of PI3k/Akt. Inhibition of P2RY2/Akt signalling (Lin *et al*., 2022) potentiates the antileukemia effect of Selinexor. Selinexor is an FDA drug that targets the nuclear exportin XPO1 for diffuse B-cell lymphoma
PTGS2	cB-ALL	PTGS2 upregulated in cB-ALL and associated with poor survival. PTGS2 and HK3 (Gao et al., 2015a) were suggested as biomarkers for B-ALL prognosis.
CYP4F2	NSCLC	CYP4F2 (Chen et al., 2022b) mediates immune-invasion of NSCLC via inducing PD-L1, IL-6, TGFβ in cancer-associated fibroblasts (CAF). High CYP4F2 tumours were associated with more resistant to anti-PD-1 treatment. CYP4F2 inhibition potentiates the anti-PD-1 therapy in mouse models suggesting CYP4F2-dependent Arachidonic Acid (AA) mediates cancer invasion

Table 9.16 Previous studied Wnt pathway in leukaemia

Gene	Disease	Function
LEF1	aB-ALL	It was highly expressed in more than 25% of B-ALL and overexpression of LEF1 was associated with poor prognosis (Kühnl *et al*., 2011)
NKD2	T-ALL	It inhibits Wnt signalling (Hinze *et al*., 2018)
LGR6	T-ALL	It acts as a negative regulator of Wnt signalling (Hinze *et al*., 2018)
GSK3	T-ALL	It acts as tumour suppressor upon its activation results in the decrease of PI3K/AKT pathway (Falà *et al*., 2008)
Wnts	CLL	High expression of Wnt3, Wnt5b, Wnt6, Wnt10a, Wnt14, and Wnt16, as well as the Wnt receptor Fzd3(Lu *et al*., 2004) and associated with decrease of apoptosis.
Wnt16b	B-ALL	It is upregulated in B-ALL with t (1;19) and high expression of TCF4, Dvl2 and β-catenin found. Induced apoptosis results from inhibition of Wnt16b (Mazieres *et al*., 2005)
Wnt3a	B-ALL	Expression of Wnt3a caused increase of β-catenin and it was not affect cell survival and proliferation (Nygren *et al*., 2007)
Wnt5a	B-ALL	It has anti-proliferative activity via enhancing GSK3β-independent β-catenin degradation. It also inhibits TCF-mediated transcription acting as a negative regulator of the Wnt/β-catenin pathway (Liang *et al*., 2003)
Wnt16, Fzd3	B-ALL	Hypermethylation of Wnt antagonists SFRP, WIF1, and Dkk3 induces over-activation of Wnt signalling (Martin *et al*., 2008).
DVL3	AML	Protein arginine methyltransferase 5 (PRMT5) upregulated in AML and sustained leukaemia stem cells (LSC) that are resistant to TKIs. PRMT5 inhibition results in depletion of DVL3 and diminish Wnt pathway that involves in maintaining LSC (Jin *et al*., 2016)
TCF4	T-ALL	It overexpressed in T-ALL and regulated expression of BIRC5 indicating that TCF4 is involved in tumour progression. TCF4 is positively regulated by ANRIL and negatively regulated by miR-7-5p (Li et al., 2020a).
FZD6	T-ALL	Overexpression of FZD6 is associated with increased of WNT10B rearrangement and it involved in the induction of Wnt signalling. Knock-down of FZD6/ WNT10B could be used to target leukaemic stem cells (Cassaro *et al*., 2021).

Table 9.17 Previous studied of hippo pathway in cancer

Gene/miRNA	Disease	Role
YAP1	CML	Overexpression of YAP1 increases tumour aggressiveness. Its knock down leads to apoptosis of leukaemia cells. The target genes c-Myc and survivin as result of low YAP1 expression
YAP/TAZ	CML	It interacts physically with β-catenin or Dvl that resulting in regulation of β-catenin levels. Increasing of β-catenin also results from Dvl phosphorylation from knock down of TAZ (Varelas *et al*., 2010). In the absence of Wnt signalling, TAZ interacts with the TrCP E3 ligase complex is worked as a scaffold for β-catenin phosphorylation by GSK3. Wnt3a causes TAZ dephosphorylation and stabilization and this leads to translocation of TAZ in the

		nucleus more easily and regulate gene expression (Azzolin *et al*., 2012).
Mob1	T-ALL	MOB1B upregulated in T-ALL cell line treated with 3-deazaneplanocin A (DZNep) and inhibited cell growth (Shen *et al*., 2015).
LATS2	AML	It is over-expressed in patients with AML (Gholami *et al*., 2014). The MST2-ETV6 fusion gene acts as an oncogene in AML patients with t (8;12) translocation (Allegra *et al*., 2021).
LATS2	B-ALL	Down-regulation of LATS2 is observed in ALL patients and is correlated to promoter region methylation in leukaemic cells (Jimenez-Velasco *et al*., 2005).
MST1	T-ALL	Loss of MST1 enhances T-cell concomitantly with mutagenic stimuli. Chromosomal instability and faster growing lymphomas is seen in mice with MST1 deletion (Wu *et al*., 2018).
KIBBRA	T-ALL	KIBBRA is heavily methylated and involves in the disease development (Wu *et al*., 2018).
YAP	CLL	High expression of YAP is found in CLL with methylated WWC1 in one third of patients and associated with poor outcome (Höffken *et al*., 2021).
MOBKL2A	MCL	Downregulated MOBKL2A involved in pathogenesis of mantle cell lymphoma (Hartmann *et al*., 2010)
FRMD6	PC	FRMD6 is an upstream regulator of hippo-pathway that maintaining tissue regeneration. Low expression of FRMD6 with aberrant hypermethylation found in prostate cancer and act as a tumour suppressor gene. Knockout of FRMD6 increases cell viability (Haldrup *et al*., 2021).
miR-550-1	AML	Low levels of miR-550-1 occurs due to hypermethylation of TAZ and reduces its stability. Thus it act as a tumour suppressor (Hu et al., 2020a)
miR-9	AML	Downregulated miR-9 activates YAP signalling and reduces the apoptosis (Wang et al., 2021a)
miR-7977	AML	It inhibits hippo pathway and decreases hippo kinase expression (YAP, TEAD, MST1) (Chorzalska *et al*., 2017).
miR-181a	CML	Low expression of miR-181a decreases YAP activation (Yoshida *et al*., 2019)
TAZ	CML	Overexpression of TAZ is associated imatinib mesylate resistance (Li *et al*., 2016)
TAZ/ AURKB	CML	High expression of both found CML patients at advanced phases (Li *et al*., 2016)
TAOK1	B-ALL	It forms a fusion protein PAX5–TAOK1 acts as a competitive inhibitor of wild type PAX5 (tumour suppressor) (Coyaud *et al*., 2010)
TAOK3	BC	It overexpressed associated with poor prognosis and drug resistance via action of NF-κB pathways (Lai *et al*., 2020)

Table 9.18 Previous LAA (Iacobucci & Mullighan, 2017) and DEG from protoarray studied in cancer

Gene	Disease	Function

HOMER3	AML/ MDS	HOMER3 is a scaffold protein that act as a transcriptional factor controlling the growth and differentiation in some organs. It decreased in AML with cytogenetic associated with unfavourable outcome while it's over expressed in AML samples with normal cytogenetics. Upregulation of HOMER3 may induce apoptosis and decrease proliferation via arrest of e G2/M phase (Li et al., 2013b).
CDKN2B (p15)	B-ALL	Homozygous deletion of p15 affects overall survival in aB-ALL but not children. The methylated gene was not associated with survival (Kim et al., 2009a). P15 deletion was associated with long term poor prognosis and high relapse to TKIs in Philadelphia chromosome patients (Xu *et al*., 2016)
SEPT9	AML	k*lf*

Table 9.19 Previous studied FOXO components

Gene	Disease	Function
FOXO1	BCP-ALL	FOXO1 is a key regulator for differentiation of pro– and pre–B-cell. Depletion of FOXO1 causes decreased CCND3 expression leading to growth arrest and apoptosis. FOXO1 knock down results in reduction of mTORC1 activity (Dharaneeswaran *et al*., 2014). However, the report suggests that FOXO1 has an oncogenic activity and involves in PI3K-AKT and RAS-ERK pathways (Wang et al., 2018a).
FOXO1	cB-ALL	Increased FOXO1 and Akt signalling is associated with upstream Ikaros isoform 6 (Ik6). This results in chemotherapy resistance and poor prognosis (Han *et al*., 2017)
FOXO3	aB-ALL	Downregulation of FOXO3 found in Ph-positive ALL only and no other subtypes. It acts as a tumour suppressor and restoration of it by using proteasome inhibitors (bortezomib) inhibit leukemic growth (Du & Chen, 2013).
FOXM1	B-ALL	It upregulated in B-ALL and associated with poor prognosis. FOXM1 expression is negatively controlled by FOXO3a. It is involved in cell survival, proliferation, and colony formation in leukemogenesis. FOXM1 inhibition could be target for the disease management with TKIS (Buchner *et al*., 2015)

Table 9.20 Previous studied B-cell signalling components

Gene	Disease	Function
SYK	B-ALL	Inhibiting of SYK results induce apoptosis and reduce tumour burden (Köhrer *et al*., 2016)
BLK	CML	*B lymphoid kinase* (*Blk*) acts as a tumour suppressor for leukemic stem cells (LSCs) and downregulated in CML via C myc activation. BLK didn't affect the normal hematopoietic stem cells (HSCs) and regulates LSC via PAX5 and p27 pathway. Increased BLK activity reduces the disease burden (Zhang *et al*., 2012)
BLK	CTCL	BLK is an oncogene in Cutaneous T-cell lymphoma (CTCL). Knock-down of BLK reduce the tumour growth (Petersen *et al*., 2014).

Table 9.21 Studied upstream regulators in leukaemia

Gene	Disease	Function
VEGF	B-ALL model	VEGF overexpressed in B-ALL model due to both cyclin-dependent kinase CDK6 and independent CDK6 mechanisms promoting lymphoid cancers. VEGF (Yu *et al*., 2019) also promotes vincristine resistance via upregulation of PI3K/Akt pathway. FOXO3 is an upstream regulator of VEGF causing its inhibition while FOXO1 enhance VEGF overexpression and enhancing tumour progression.
IL-6	aB-ALL	IL-6 upregulated in aB-ALL and associated with CNS involvement. Admission of CART (CD19) combined with short RNA silencing IL-6 was entered phase 1 clinical trial (Chen et al., 2020b).
P38MAPK	B-ALL	High expression was found in B-ALL modulating bone marrow stromal proliferation and survival. P38MAPK inhibitor (Gaundar *et al*., 2009) has significantly reduced leukaemia burden reducing proliferative cytokines CXCL12, IL-6 and VEGF upregulation
ErbB2	cB-ALL	ErbB2 upregulated in cB-ALL and promotes tumour proliferation and apoptosis inhibition (Derakhshan *et al*., 2022).
IFN-α	cB-ALL	Treatment with IFN-α (Sun et al., 2020b) induced B-ALL remission combined with lymphocytes infusion following allogenic transplant reducing MDR
KLF6	AML	KLF6 (DeKelver *et al*., 2013) upregulated in AML due to pro-oncogenic fusion protein RUNX1-ETO enhancing leukaemia progression
CD3	B-ALL model	CD3 (Wang et al., 2022b) overexpressed in B-ALL and has been suggested to be targeted combined with CD19 and CD20 for relapsed B-ALL
PGE2	B-ALL cell line	PGE2 (Soleymani Fard *et al*., 2012) induced growth inhibition and apoptosis of B-ALL
Notch1	B-ALL model	Notch1 activation (Hauer *et al*., 2011) enhances self-renewal of LC both dependent of Rag1 mutation is involved in leukaemia initiating and independent via downregulating CDKN2A -cell differentiation mutation is found in 30% of T-ALL associated with IKZF1 deletion
Wnt3a	B-ALL	Expression of Wnt3a caused increases of β-catenin and it did not affect cell survival and proliferation (Nygren *et al*., 2007)

Table 9.22 CTA expression for the survival using BloodSpot

Antigen	Probe set	High expression	survival
ADAM2	207664-at	Pro-B-ALL t(11q23)/MLL-NS	0.181
BAGE	1555369-at	c-/Pre-B-ALL t(9;22)-NS	0.929
BAGE	207712-at	ALL t(12;21)/ Mature B-ALL ALL with t(8;14)-NS	0.92
BAGE	1555603-at	Healthy bone marrow-NS	0.344
BAGE	1555605-x-at	Healthy bone marrow-NS	0.063
BRDT	206787-at	B-ALL with t(8;14)-NS/ ALL t(12;21) p<0.05	0.919
CAGE1	1563787-a-at	Pro-B-ALL t(11q23)/MLL-NS	0.572
CAGE1	1563787-a-at	Pro-B-ALL t(11q23)/MLL-NS	0.572
CTAGE1	220957-at	ALL t(1;19)-NS	0.373

CTAGE5	215930-s-at	ALL hyperdiploid p<0.001	0.228
CTAGE5	**204055-s-at**	**ALL hyperdiploid p<0.001**	**0.0129**
CTCFL	1552368-at	B-ALL with t(8;14)-NS	0.028
DDX1	201241-at	Mature B-ALL ALL with t(8;14)/ ALL t(1;19)/ Pro-B-ALL t(11q23)/MLL-NS	0.0667
DDX10	204977-at	ALL hyperdiploid/ Pro-B-ALL t(11q23)/MLL p<0.001	0.396
DDX10	1563522-at	ALL t(12;21)p<0.001	0.319
DDX11	208159-x-at	c-/Pre-B-ALL t(9;22)/ ALL with t(8;14)-NS	0.423
DDX11	208149-x-at	c-/Pre-B-ALL t(9;22)/ ALL with t(8;14)-NS	0.737
DDX11	232816-s-at	c-/Pre-B-ALL t(9;22)/ ALL with t(8;14)-NS	0.814
DDX11	210206-s-at	c-/Pre-B-ALL t(9;22)/ ALL with t(8;14)-NS	0.563
DDX11L2	223777-at	Healthy bone marrow p<0.001	0.52
DDX17	213998-s-at	ALL t(12;21)p<0.001	0.92
DDX17	230180-at	ALL t(12;21)p<0.001	0.762
DDX17	208719-s-at	Pro-B-ALL t(11q23)/MLL p<0.001	0.968
DDX17	208151-x-at	ALL t(12;21)/Pro-B-ALL t(11q23)/MLL p<0.001	0.942
DDX18	208895-s-at	Mature B-ALL with t(8;14) p<0.001	0.424
DDX18	208897-s-at	ALL t(1;19) p<0.001	0.31
DDX18	208896-at	Mature B-ALL with t(8;14) p<0.001	0.209
DDX18	205763-s-at	ALL t(1;19)/ Healthy bone marrow p<0.001	0.198
DDX19A	202578-s-at	Pro-B-ALL t(11q23)/MLL-NS	0.459
DDX19A	202577-s-at	ALL t(1;19)-NS	0.178
DDX19A	1570128-at	c-/Pre-B-ALL t(9;22)/ c-/Pre-B-ALL no t(9;22)-NS	0.194
DNAJB1	200664-s-at	Healthy bone marrow p<0.001	0.202
DNAJB11	223045-at	Healthy bone marrow p<0.001	0.25
DNAJB12	214338-at	Mature B-ALL ALL with t(8;14)-NS	0.123
DNAJB13	230936-at	Mature B-ALL ALL with t(8;14) p<0.05	0.993
DNAJB13	1552976-at	Pro-B-ALL t(11q23)/MLL p<0.05	0.861
DNAJB14	226399-at	ALL hyperdiploid p<0.001	0.861
DNAJB14	219237-s-at	ALL t(12;21)/ c-/Pre-B-ALL t(9;22) p<0.001	0.46
DNAJB14	222850-s-at	c-/Pre-B-ALL t(9;22) p<0.001	0.863
DNAJB2	202500-at	c-/Pre-B-ALL t(9;22) p<0.01	0.0183
DNAJB4	203811-s-at	ALL t(12;21)-NS	0.139
DNAJB4	203810-at	ALL t(12;21)-NS	0.0185
DNAJB8	237284-at	Mature B-ALL ALL with t(8;14) p<0.001	0.617
FATE1	231573-at	ALL t(12;21) p<0.001	0.301
FTHL17	224379-at	Healthy bone marrow/ ALL t(1;19) p<0.01	0.59
GAGE1	208283-at	Healthy bone marrow p<0.01	0.531
GAGE3	207663-x-at	Mature B-ALL with t(8;14) p<0.001	0.363
GPATCH2	243704-at	ALL t(1;19) p<0.001	0.009
LDHC	207022-s-at	Healthy bone marrow/ ALL t(12;21) p<0.001	0.504
LIPI	242178-at	Pro-B-ALL t(11q23)/MLL-NS/ ALL t(12;21) p<0.01	0.624
MAGEA1	207325-x-at	Mature B-ALL ALL with t(8;14)-NS	0.797
MAGEA10	210295-at	Mature B-ALL ALL with t(8;14)-p<0.05	0.282
MAGEA11	210503-at	Pro-B-ALL t(11q23)/MLL/ Mature B-ALL with t(8;14)-p<0.05	0.0643
MAGEA12	210467-x-at	c-/Pre-B-ALL t(9;22)/ c-/Pre-B-ALL no t(9;22)- NS	0.172
MAGEA3	209942-x-at	Healthy bone marrow -NS	0.0548
MAGEA4	214254-at	Mature B-ALL ALL with t(8;14)-NS	0.00404
MAGEA5	1553585-a-at	Pro-B-ALL t(11q23)/MLL-NS/ ALL t(1;19)- p<0.05	0.847
MAGEA5	214642-x-at	c-/Pre-B-ALL t(9;22)-NS/ ALL t(1;19)- p<0.05	0.0634

MAGEA6	214612-x-at	c-/Pre-B-ALL t(9;22)-NS	0.0468
MAGEA7	210274-at	B-ALL ALL with t(8;14)-NS/ ALL t(12;21) p<0.01	0.679
MAGEB1	207534-at	B-ALL with t(8;14)-NS	0.442
MAGEB18	1552913-at	Healthy bone marrow-NS	0.689
MAGEB2	206218-at	c-/Pre-B-ALL t(9;22)-NS	0.472
MAGEB3	207579-at	B-ALL with t(8;14)-NS/ ALL t(12;21) p<0.05	0.556
MAGEC1	206609-at	Pro-B-ALL t(11q23)/MLL-NS	0.642
MAGEC2	220062-s-at	ALL t(1;19) p<0.01	0.445
MAGEC2	215932-at	B-ALL with t(8;14)-NS/ ALL t(1;19) p<0.01	0.646
MAGEC3	216592-at	B-ALL with t(8;14)-NS	0.54
MAGEE1	229286-at	ALL hyperdiploid-NS	0.407
MAGEE1	1556047-s-at	ALL hyperdiploid-NS	0.316
MAGEE2	1553254-at	Healthy bone marrow-NS	0.611
MORC1	220850-at	ALL t(12;21) p<0.001	0.368
MORC3	213000-at	ALL hyperdiploid-NS/ ALL t(12;21) p<0.01	0.0232
MORC4	219038-at	ALL hyperdiploid-NS	0.0473
NR6A1	211402-x-at	Healthy bone marrow/ ALL t(12;21) p<0.001	0.0345
NXF2	220981-x-at	Healthy bone marrow/ ALL t(12;21) p<0.001	0.133
PAGE1	206897-at	Pro-B-ALL t(11q23)/MLL-NS	0.295
PAGE2B	231307-at	Healthy bone marrow/ ALL t(12;21) p<0.001	0.33
PAGE4	205564-at	B-ALL ALL with t(8;14)-NS/ ALL t(12;21) p<0.001	0.641
PAGE5	236152-at	Healthy bone marrow/ ALL t(12;21) p<0.001	0.121
PASD1	240687-at	c-/Pre-B-ALL t(9;22)- p<0.01	0.957
PRAME	204086-at	B-ALL with t(8;14) p<0.001	0.806
RQCD1	213179-at	ALL t(1;19) p<0.001	0.0137
SAGE1	220793-at	Pro-B-ALL t(11q23)/MLL-NS	0.718
SAGE1	220793-at	Pro-B-ALL t(11q23)/MLL-NS	0.718
SPA17	205406-s-at	ALL t(1;19)-NS/ ALL t(12;21) p<0.001	0.098
SPANXC	220217-x-at	ALL t(12;21) p<0.001	0.272
SPO11	222259-s-at	c-/Pre-B-ALL t(9;22)-NS	0.838
SSX1	206626-x-at	ALL with t(12;21) p<0.001	0.108
SSX1	206627-x-at	Pro B-ALL t(11q23)/MLL p<0.001	0.108
SSX3	211732-x-at	c-/Pre-B-ALL not(9;22)c-/ Pro-B-ALL t(11q23)/MLL-NS	0.523
SSX3	211670-x-at	Mature B-ALL with t(8;14)NS	0.771
SSX3	207666-x-at	ALL t(12;21)-NS	0.701
SSX5	208528-x-at	Mature B-ALL ALL with t(8;14)p<0.001	0.705
SYCP1	206746-x-at	Mature B-ALL ALL with t(8;14)/ ALL with t(1;19)-NS	0.899
SYCP1	216917-s-at	Mature B-ALL ALL with t(8;14)/ Pre-B-ALL t(9;22)-NS	0.0834
SYCP2	206546-at	Mature B-ALL ALL with t(8;14)-NS	0.201
SYCP2	237920-at	ALL with t(1;19)-NS	0.582
SYCP2L	236337-at	ALL with t(1;19) p<0.001	0.0246
SYCP3	1553599-a-at	Pro-B-ALL with t(11q23)/MLL p<0.001	0.852
SYCP3	241861-at	ALL with t(1;19) p<0.001	0.263
TAF7L	220325-at	B-ALL ALL with t(8;14)-NS/ ALL t(1;19)- p<0.01	0.169
TAF7L	224380-s-at	Healthy bone marrow/ ALL t(1;19) p<0.01	0.0906
TDRD1	221018-s-at	Mature B-ALL ALL with t(8;14)-NS	0.166
TEX101	223906-s-at	B-ALL with t(8;14)-NS/ ALL t(1;19) p<0.001	0.00539
TEX15	221448-s-at	c-/Pre-B-ALL t(9;22)-NS	0.466
TEX15	232760-at	Pro-B-ALL t(11q23)/MLL-NS	0.316
TPTE	220205-at	ALL t(1;19) -NS	0.846

TPX2	210052-s-at	Healthy bone marrow/ ALL t(12;21) p<0.001	0.373
XAGE3	236040-at	Healthy bone marrow p<0.001	0.764
XAGE3	236040-at	Healthy bone marrow/ ALL t(12;21) p<0.001	0.764
XAGE4	1565454-at	Healthy bone marrow p<0.01	0.719
XAGE4	1565454-at	Healthy bone marrow	0.719

Table 9.23 LAA expression for the survival using BloodSpot

Antigen	Probe set	High expression	Survival
WT1	206067-s-at	Pro-B-ALL t(11q23)/MLL p<0.001	0.349
	216953-s-at	Pro-B-ALL t(11q23)/MLL p<0.001	0.145
BMX	206464-at	Healthy bone marrow p<0.001	0.0461
ERG	213541-s-at	ALL hyperdiploid/ALL t(12;21) p<0.001	0.147
	241926-s-at	ALL hyperdiploid/ALL t(12;21) p<0.001	0.541
	222079-at	ALL hyperdiploid/ALL t(12;21) p<0.001	0.601
	211626-x-at	ALL hyperdiploid/ALL t(12;21) p<0.001	0.679
PAX5	221969-at	ALL t(1;19) p<0.001	0.823
	206802-at	Pro-B-ALL t(11q23)/MLL/ B-ALL with t(8;14) p<0.001	0.231
IKZF1	25039-s-at	ALL t(12;21) p<0.001	0.9
	227346-at	ALL t(12;21) p<0.001	0.177
	216901-s-at	Healthy bone marrow p<0.001	0.982
	227344-at	Pro-B-ALL t(11q23)/MLL /ALL t(12;21) p<0.001	0.268
	220704-at	Pro-B-ALL t(11q23)/MLL /ALL t(12;21) p<0.001	0.841
	205038-at	ALL t(12;21) p<0.001	0.702
	1565816-at	Pro-B-ALL t(11q23)/MLL p<0.001	0.296
	1565818-s-at	Pro-B-ALL t(11q23)/MLL /ALL t(12;21) p<0.001	0.942
	1565817-at	ALL t(12;21) p<0.001	0.0304
CDKN2A	209644-x-at	ALL t(12;21)/ B-ALL with t(8;14) p<0.001	0.283
	211156-at	B-ALL with t(8;14) p<0.001	0.599
	207039-at	ALL t(12;21) p<0.001	0.53
CDKN2B	207530-s-at	B-ALL with t(8;14) p<0.05	0.0193
	236313-at	ALL t(12;21) p<0.05	0.761
CDKN2C	211792-s-at	B-ALL with t(8;14) p<0.001	0.305
	204159-at	ALL hyperdiploid/B-ALL with t(8;14) p<0.001	0.97
CDKN2D	210240-s-at	ALL t(1;19)/ Healthy bone marrow p<0.001	0.09
	213586-at	Pro-B-ALL t(11q23)/MLL p<0.001	0.481
CDKN2AIP	218929-at	ALL t(12;21)/ ALL t(1;19) p<0.001	0.884
CDKN2AIPNL	1554348-s-at	c-/Pre-B-ALL t(9;22) p<0.001	0.915
	233006-at	B-ALL with t(8;14) p<0.001	0.61
NRAS	202647-s-at	Pro-B-ALL t(11q23)/MLL/ c-/Pre-B-ALL no t(9;22) p<0.01	0.31
	224985-at	B-ALL with t(8;14) p<0.01	0.232
KRAS	214352-s-at	ALL t(1;19) p<0.001	0.681
	204009-s-at	ALL t(1;19) p<0.001	0.0149
	204010-s-at	Pro-B-ALL t(11q23)/MLL p<0.001	0.332
	1559204-x-at	c-/Pre-B-ALL t(9;22) p<0.001	0.109
	1559203-s-at	c-/Pre-B-ALL t(9;22)/ c-/Pre-B-ALL no t(9;22) p<0.001	0.731
GAK	40225-at	ALL t(12;21) p<0.001	0.465
	202281-at	ALL hyperdiploid/ALL t(12;21) p<0.001	0.623
	202280-at	B-ALL with t(8;14) p<0.001	0.205
HCAR1	224131-at	ALL hyperdiploid-NS	0.436

HOMER1	226651-at	B-ALL ALL with t(8;14)-NS/ ALL t(12;21) p<0.001	0.000581
HOMER3	215489-x-at	Healthy bone marrow/ ALL t(12;21) p<0.001	0.0311
TSPAN15	218693-at	c-/Pre-B-ALL t(9;22)-NS	0.062
MYCBP2	1557370-s-at	ALL hyperdiploid/ ALL t(12;21) p<0.001	0.521
	201959-s-at	ALL t(12;21) p<0.001	0.615
	201960-s-at	ALL hyperdiploid/ ALL t(12;21)/ Pro-B-ALL t(11q23)/MLL p<0.001	0.165
CDC7	204510-at	ALL t(1;19) -NS	0.881
SEP15	200902-at	ALL t(12;21)/ Pro-B-ALL t(11q23)/MLL p<0.001	0.0475
WWOX	223868-s-at	Mature B-ALL with t(8;14)NS	0.356
	210695-s-at	ALL with t(1;19) NS	0.18
	223747-x-at	Mature B-ALL with t(8;14)NS	0.687
	219077-s-at	ALL with t(1;19) NS	0.783
	221147-x-at	Mature B-ALL with t(8;14)NS	0.164

Bold text is used to indicate the gene, probe, subset and survival association (P<0.05)

Table 9.24 CTAs examined weren't found BloodSpot

NY-ESO-1 (CTAG1B)	PAGE2,3	C21orf99	SEMG1	LUZP4
N-RAGE	MAGEB10,	CDCA1	SPAG1,4,8.9,17	KIAA0100
LAGE-1a, 1b	MAGEB10-PS,	CSAG1	SPANXA1	JARID1B
HAGE	MAGEB16,	CSAG2	SPANXB1	HORMAD1,2
SPANX,	MAGEB16-PS1,	ZNF1	SPANXD	IGSF11
CT47A1-11	MAGEB17,	TDRD4	SPANXE	IGF2BP3
CT47B1	MAGEB17-PS	TEX14,15	SPANXN1,2,3,4,5	FLJ36144
CT66	NYSAR35	TFDP3	SPINLW1	TSPY1
CT69	TPX1	CT56	SSX4,6,7,9	FAM133A
CT70	SGY-1	TPTE	RHOXF2	BAGE2,3,4,5
CTAG1B	MORC	TSP50	POTEA,B,C,D,E,F,G,H	ACRBP
CTAGE-2	PLU-1	TSPY3	PTPN20A	ACTL8
CT146	HCA661	NA88	OTOA	ATAD2
CT45,	D40	XAGE1a	NXF2B	Piwil2, PIWIL1
CT45A1,2,3,4,5,6	HOMTES85	MAGEA2, 2B	MMA1b	XAGE1,2a, 2b,5,
CT7	MMA1a	MAGEA7-PS	MPHOSPH1	XAGE-3b
CT49	OYTES1	MAGEA9,9PS	LAGE-1b	LDHC
	TSP50			LOC196993

Table 9.25 Expression protoarray genes by Jordenset al. 2020 using BloodSpot

Gene name	Probe set	High expression	survival
WARS	200629-at	Mature B-ALL ALL with t(8;14)p<0.05/ ALL t(12;21) p<0.001	0.313
	200628-s-at	Healthy bone marrow / ALL t(12;21) p<0.001	0.534
WARS2	222734-at	Pro-B-ALL t(11q23)/MLL-NS	0.163
	218766-s-at	Pro-B-ALL t(11q23)/MLL-NS/ ALL t(1;19) p<0.001	0.859
	1561297-at	Pro-B-ALL t(11q23)/MLL-NS/ ALL t(1;19) p<0.001	0.553
VGLL4	212399-s-at	ALL t(12;21) p<0.001	0.722
	214004-s-at	ALL t(12;21) p<0.001	0.117
TOX2	228737-at	Pro-B-ALL t(11q23)/MLL - p<0.01	0.5
SEPT9	41220-at	ALL hyperdiploid - p<0.001/ ALL t(12;21) p<0.001	0.446

	208657-s-at	ALL hyperdiploid-NS/ ALL t(12;21) p<0.001	0.764
	207425-s-at	ALL hyperdiploid - p<0.001/ ALL t(12;21) p<0.001	0.969
	1559025-at	ALL hyperdiploid-NS/ ALL t(12;21) p<0.001	0.0205
RAB3IL1	219579-at	B-ALL with t(8;14)-NS/ ALL t(12;21) p<0.001	0.807
NAT1	214440-at	ALL t(12;21)-NS	0.278
MUC20	1558220-at	ALL with t(1;19)-NS/ ALL t(12;21) p<0.001	0.0125
	226622-at	ALL with t(1;19)-NS/ ALL t(12;21) p<0.001	0.498
	230043-at	ALL with t(1;19)- p<0.001/ ALL t(12;21)-p<0.001	0.695
	243774-at	ALL with t(1;19)- p<0.001/ ALL t(12;21)-p<0.001	0.325
	231941-s-at	ALL with t(1;19)- p<0.001/ ALL t(12;21)-NS	0.496
LMX1A	1553541-at	c-/Pre-B-ALL t(9;22)-NS	0.637
IGLL1	206660-at	ALL with t(1;19)- p<0.001	0.304
GCC1	218912-at	ALL with t(1;19)- p<0.01/ ALL hyperdiploid-p<0.01	0.951
	243306-s-at	B-ALL with t(8;14)-NS/ ALL hyperdiploid-p<0.01	0.917
	243437-at	B-ALL with t(8;14)-NS/ c-/Pre-B-ALL no t(9;22)- p<0.001	0.118
GAK	40225-at	ALL t(12;21) p<0.001	0.465
	202281-at	ALL hyperdiploid/ALL t(12;21) p<0.001	0.623
	202280-at	B-ALL with t(8;14) p<0.001	0.205
DCTPP1	218069-at	B-ALL with t(8;14) p<0.001	0.238
CDCA3	223307-at	Healthy bone marrow-NS / ALL t(12;21) p<0.001	0.706
	221436-s-at	Healthy bone marrow-NS / ALL t(12;21) p<0.001	0.0577
CDC42EP1	204693-at	Healthy bone marrow-NS / ALL t(12;21) p<0.001/ ALL with t(1;19)- p<0.01	0.802
BMX	206464-at	Healthy bone marrow/ ALL t(12;21) -p<0.001	0.0461
APOBEC3A	210873-x-at	B-ALL with t(8;14)- p<0.05	0.67
ACOX1	209600-s-at	ALL hyperdiploid - p<0.001	0.106
	227962-at	ALL hyperdiploid - p<0.001	0.573
	213501-at	ALL hyperdiploid - p<0.001	0.793
	209601-at	B-ALL with t(8;14)- p<0.05/ c-/Pre-B-ALL no t(9;22)- p<0.01	0.196
	207656-s-at	B-ALL with t(8;14)- p<0.05/ c-/Pre-B-ALL no t(9;22)- p<0.05	0.345

Bold text is used to indicate the gene, probe, subset and survival association.