A common low penetrance risk allele for chronic lymphocytic leukaemia in the 3'UTR of *IRF4/MUM1* predicts poor outcome

James M. Allan[1], James R. Bailey[2], Andrew R. Pettitt[3], Robert J. Harris[3], Chris Pepper[4], Chris Fegan[5], Richard S. Houlston[6], Nicola J. Sunter[1], Andrew G. Hall[1], Tryfonia Mainou-Fowler[7], Graham H. Jackson[8], Geoffrey Summerfield[9], Guy Pratt[10*], David J. Allsup[11*].

[1]Northern Institute for Cancer Research, Paul O'Gorman Building, Newcastle University, Newcastle-upon-Tyne. NE2 4HH. UK. [2]Hull York Medical School and University of Hull, Hull. HU16 5JQ. UK. [3]Division of Haematology, University of Liverpool School of Cancer Studies, Liverpool, UK. [4]Department of haematology, School of Medicine, Cardiff University. Cardiff. UK. [5]Cardiff and Vale NHS Trust, Heath Park, Cardiff, CF14 4XW. UK. [6]Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey. SM2 5NG. UK. [7]Haematological Sciences, Leech Building, The Medical School, Newcastle University, Newcastle upon Tyne. UK. [8]Department of Haematology, Royal Victoria Infirmary, Newcastle upon Tyne. NE1 4LP. UK. [9]Department of Haematology, Queen Elizabeth Hospital, Gateshead, Newcastle upon Tyne. NE9 6SX. UK. [10]Department of Haematology, Birmingham Heartlands Hospital, Birmingham. B9 UK. [11]Department of Haematology, Hull Royal Infirmary, Hull. HU3 2JZ. UK.

*GP and DJA contributed equally to this study.

Scientific Section - Neoplasia

Reprints: James M. Allan, Northern Institute for Cancer Research, Medical School, Framlington Place, Newcastle University, Newcastle upon Tyne, NE2 4HH, United Kingdom. James.allan@newcastle.ac.uk

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1038/leu.2009.298

Abstract

High frequency low penetrance risk alleles for chronic lymphocytic leukaemia (CLL) have been identified at 6p25.3 (rs872071, *IRF4/MUM1*), 11q24.1 (rs735665, *GRAMD1B*), 15q23 (rs7176508), 2q37.1 (rs13397985, *SP140*), 2q13 (rs17483466, *ACOXL*) and 19q13.32 (rs11083846, *PRKD2*). Given the often protracted nature of disease course it is plausible that allelic variants predisposing to risk of disease also play a role in progression and overall survival. In order to test this hypothesis, polymorphic status was determined in a case series of 404 patients diagnosed with CLL recruited via 4 clinical centres in the United Kingdom. Mean follow-up time was 2402 days. One hundred and seventy patients had initiated therapy at the time of last follow-up and 123 were deceased. Kaplan-Meier survival analysis and the log-rank test were used to investigate the prognostic significance of constitutional genetic markers. Allelic variants defined by rs735665, rs7176508, rs13397985, rs17483466 and rs11083846 were not significantly association with time from diagnosis to first treatment ($p>0.05$, log-rank test). However, a single nucleotide polymorphism (rs872071) in the 3'UTR of the *interferon regulatory factor 4* (*IRF4*)/*multiple myeloma oncogene 1* (*MUM1*) gene was significantly associated with poor prognosis. Specifically, carriers of the disease-associated allele at the *IRF4* locus had a significantly shorter time from diagnosis to first treatment compared to non-carriers ($p=0.001$). The association with time to first treatment remained significant in multivariate Cox regression analysis that included age, gender, immunoglobulin heavy chain variable region mutational status and stage of disease at diagnosis in the model ($p=0.017$). IRF4 is expressed in germinal centre (GC) and post-GC B-cells, and is important for B-cell maturation and homeostasis. Taken together with the aetiological evidence, these data suggest that a high frequency allelic variant in *IRF4/MUM1* not only affects risk of CLL, but is also an independent predictor of poor prognosis.

Introduction

Chronic lymphocytic leukaemia (CLL) is primarily a disease of the elderly and is characterised by the accumulation of neoplastic B lymphocytes with a characteristic immunophenotype[1,2]. It is the commonest adult leukaemia in the Western world and poses a significant healthcare burden within these countries[3]. The median survival from diagnosis is approximately 10 years[4,5]. However this figure does not reflect the true heterogeneity of the disease, with a proportion of patients remaining asymptomatic and having a survival similar to age matched controls[6]. Others, however, suffer an aggressive illness with a shorter median survival from diagnosis of 2 to 3 years[7].

The ability to accurately predict disease course at diagnosis would lead to improvements in patient management. To this end, patient and disease phenotypic characteristics have proven useful in predicting progression and prognosis in CLL, and have facilitated the development of disease staging systems[8,9]. Disease-specific genetic alterations such as *ATM* and *P53* mutation, mutational status of the immunoglobulin heavy-chain variable region (IgV_H), and expression of CD38, ZAP-70 and MCL-1 have also proven useful in predicting patient prognosis[10-13]. Despite these advances accurately predicting patient outcome in CLL remains challenging, and evidence suggests the existence of other yet to be identified contributing factors.

In addition to somatically acquired alterations it is likely that constitutional genetic variation also plays a role in defining disease progression and patient outcome, and there is evidence that polymorphic variation in genes important in regulating cellular response to cytotoxic chemotherapy, such as *TNF-α*, *P53*, *BAX, POLB* and *BCL-2*[14-17], can impact on outcome post-therapy. However, given the often protracted nature of disease course it is clear that factors regulating transformation to an aggressive phenotype, to the point where treatment is indicated, are also important in defining patient outcome. By extrapolating this model we can hypothesise that genetic variation important in determining risk of developing CLL might also

play a role in regulating progression to an aggressive phenotype and patient prognosis. A strong familial basis to CLL indicates a role for inherited genetic variation in defining risk of disease[18]. However, failure to identify a single disease-associated locus suggests that risk is defined predominantly by the co-inheritance of high frequency low-penetrance alleles. Consistent with a polygenic disease model, putative low penetrance risk alleles for CLL have been identified at 2q13 (rs17483466, *ACOXL*), 2q37.1 (rs13397985, *SP140*), 6p25.3 (rs872071, *IRF4/MUM1*), 11q24.1 (rs735665, *GRAMD1B*), 15q23 (rs7176508) and 19q13.32 (rs11083846, *PRKD2*)[19]. Of these, the strongest statistical evidence for association with disease risk was found at the *interferon regulatory factor 4* (*IRF4*)/*multiple myeloma oncogene* (*MUM1*) locus at 6p25.3. Given a role in disease aetiology it is plausible that these genetic variants also play a role in driving disease progression and affecting patient outcome. In order to test this hypothesis we determined polymorphic status at these six susceptibility loci in a case series of 404 CLL patients and investigated association with treatment-free survival and overall survival.

Patients, Materials and Methods.

Subjects

Peripheral blood samples were collected from 404 individuals of European origin resident in the UK diagnosed with CLL between 1970 and 2006 who attended 4 clinical centres in the United Kingdom (Birmingham Heartlands Hospital, Hull Royal Infirmary, Liverpool Royal University Hospital and The University of Wales College of Medicine Hospital). All patients had typical CLL mature lymphocytes expressing CD5, CD19, CD23 and clonally restricted surface immunoglobulin. The proportion of incident and prevalent cases varied between clinical center. However, peripheral blood for DNA extraction was taken between 1993 and 2006 for all cases (Hull 1993–1996; Birmingham 1998–2006; Cardiff 1998–2006; Liverpool 2000–2006). For prevalent cases originally diagnosed before 1998, CLL was confirmed via disease re-staging using diagnostic criteria applicable at the time of sample collection. Clinical staging was based on the Binet system. Hence, for all incident and prevalent cases a diagnosis of CLL has been confirmed in accordance with the current WHO classification guidelines [ref]. One hundred and seventy patients had received treatment for CLL at the time of last follow-up. Treatment was initiated in patients with symptomatic disease and those with progression (lymphocyte doubling time of less than 12 months). Choice of chemotherapy was at the discretion of local treating clinicians and was in accordance with national guidelines [ref]. Chemotherapy regimens employed were chlorambucil, oral cyclophosphamide, oral and intravenous fludarabine, (FC) fludarabine and cyclophosphamide, FCR (fludarabine, cyclophosphamide and rituximab), CHOP (cyclophophamide, doxorubicin, vincristine and prednisolone), alemtuzumab and both autologous and allogeneic haematopoietic stem cell transplantation. Patients were followed from diagnosis until last known follow-up or death. Mean follow-up time was 2402 days and 123 patients were deceased at the time of last follow-up.

Ethics.

Collection of blood samples and clinicopathological information from cases and controls was undertaken with informed consent and relevant ethical review board approval in accordance with the

tenets of the Declaration of Helsinki, and also following institutional review by the respective Hospital Trust Research and Development Committee.

IgV_H gene mutation

RNA was extracted using the Qiagen RNeasy mini kit (Qiagen, Crawley, UK) and complementary DNA was synthesised using the Reverse-iT kit according to the manufacturer's instructions (ABgene, Epsom, UK). The IgV_H gene was amplified using PCR as previously described[20], and sequenced using BIG dye terminator sequencing method according to the manufacturer's instructions (Applied Biosystems, Warrington, UK). Sequences were compared to germline using IgBLAST (National Centre for Biotechnology Information, Bethesda, Maryland; http://ncbi.nih.gov/igblast/) and those with homology of 98% or higher were regarded as unmutated, and those with less then 98% homology were regarded as mutated[21,22].

ZAP70 and CD38 expression

ZAP70 expression was assessed using flow cytometry following incubation of peripheral blood mononuclear cells with anti-CD3, anti-CD5, anti-CD19 and anti-ZAP70 monoclonal antibodies. Cells were analysed by flow cytometry, as previously described [Francis et al, 2006; Pepper et al, 2008], and at least 10,000 events were scored for each sample. The cut-off for ZAP70 positivity was 20%. CD38 expression was determined by triple colour immunophenotyping using ant-CD5, anti-CD38 and anti-CD19 monoclonal antibodies, as previously described [Francis et al, 2006; Pepper et al, 2008]. The cut-off for CD38 positivity was 20%.

SNP genotyping.

DNA was extracted from samples using either standard phenol/chloroform techniques or the Qiamp DNA extraction kit (Qiagen, Crawley, UK) and genotyped by competitive allele-specific PCR

KASPar chemistry (KBiosciences, Hoddesdon, UK) for 6 single nucleotide polymorphisms (rs872071, rs11083846, rs7176508, rs13397985, rs735665 and rs17483466) previously associated with risk of developing CLL[23]. Primer and probe sequences for each variant were as follows: XXX. Genotyping quality control was determined via the inclusion of blinded replicate samples, together with direct sequencing of 5% of samples to test for assay accuracy. Assay reproducibility and accuracy was greater than 99% for all 6 SNP assays. Buccal DNA samples were available for 44 cases, and genotype concordance between peripheral blood and buccal samples was >95% for all 6 polymorphisms tested. As such, we are confident that genotyping accuracy is not significantly affected by allelic loss in contaminating leukemic lymphocytes. The distribution of genotypes for all six polymorphisms investigated in this study was not significantly from that expected under Hardy-Weinberg Equilibrium, suggesting that the data was not affected by systematic bias (data not shown).

Statistical analysis.

The primary survival end-point in this study was treatment-free survival, defined as time from diagnosis to initiation of CLL-specific treatment or last follow-up. The secondary end-points in this study were overall survival, defined as time from diagnosis to death or last follow-up. Comparisons of treatment-free survival and overall survival between polymorphic groups were performed using the Kaplan-Meier method and significance was determined using the log-rank test. Multivariate analysis using Cox regression was used to test for independency of polymorphic variants and established prognostic markers in CLL, with age included in the model as a continuous variable. The Chi-squared test was used to examine the distribution of categorical variables as a function of genotype status. Differences in the expression of ZAP70 and CD38 by polymorphism genotype were examined using a Wilcoxon-type test for trend (Cuzick, 1985). All statistical analyses were conducted using either SPSS 15.0 or STATA 10. A p-value of <0.05 was considered statistically significant.

Results

Relationship between clinicopathologic characteristics and outcome.

The mean age at diagnosis of the 404 CLL patients recruited to this study was 66 years (range 24-100 years), which included 240 males and 164 females, and the mean follow-up time from diagnosis was 2402 days (6.6 years). Two hundred and sixty-eight, 40 and 42 patients presented with Binet stage A, B and C disease, respectively.

In order to determine whether the cases recruited to this multicentre study were representative of typical CLL we ascertained outcome as a function of established prognostic markers in CLL. IgV_H mutation status significantly predicted for overall survival, where cases with mutated IgV_H survived longer than those with germline IgV_H (n=351, p<0.001)(Figure 1A). Stage of disease at diagnosis was also significantly associated with overall survival, where patients with stage A disease survived significantly longer than those with stage B or C disease (n=350, p<0.001)(Figure 1B). Patients 70 years or older at diagnosis had a significantly shorter overall survival compared to patients aged 69 years or younger (n=357, p<0.001)(data not shown). Males had a moderately shorter overall survival compared to females, but this was not statistically significant (n=404, p=0.236)(data not shown). Unmutated IgV_H, stage B or C disease and older age at diagnosis were also significantly associated with shorter treatment-free survival (data not shown). These data confirm that established prognostic markers for CLL predict outcome in the cohort of patients recruited to this study.

Relationship between polymorphism status and prognosis.

Our primary hypothesis predicts that risk alleles for CLL would also predispose to more rapid onset of symptomatic disease requiring treatment. In order to test this we investigated the association between polymorphism genotype and time from diagnosis to the initiation of treatment (treatment-free survival). Specifically, for each polymorphism we tested carriers of the disease risk-associated allele and compared treatment-free survival to non-carriers. Carrier status for the CLL risk allele was not significantly associated with treatment-free survival for 5 of the 6 polymorphisms tested (rs17483466,

p=0.146; rs13397985, p=0.253; rs735665, p=0.211; rs7176508, p=0.672; rs11083846, p=0.448)(Figure 1A-E). However, carriers of the CLL risk allele defined by rs872071 in the *interferon regulatory factor 4* (*IRF4*) gene had a significantly shorter treatment-free survival compared to non-carriers (p=0.001)(Figure 1F). When examined by individual genotype, heterozygotes and homozygotes for the CLL-risk allele at *IRF4* both had a significantly shorter treatment-free survival compared to non-carriers (Figure 2A), but there was no significant difference between these two carrier groups, suggesting that the risk allele defined by rs872071 acts dominantly with respect to disease progression. Carrier status for the CLL-risk allele was also associated with a shorter time from the initiation of treatment to death or last follow-up, although this was not statistically significant (p=0.23)(date not shown).

Despite a highly significant difference in treatment-free survival and a modest difference in survival post-treatment, there was no significant difference in overall survival between carriers and non-carriers for the risk allele defined by rs872071 (p=0.29)(Figure 3B). As a potential explanation of this discrepancy we identified 51 patients who died before the indication of treatment for their leukaemia, which included several patients known to have died of causes apparently independent of their leukaemia (including second primary cancers and cardiac failure) or any directly related co-morbidities. When these 51 early-death patients were excluded from the analysis, carrier status for the *IRF4* disease-associated variant significantly predicted overall survival in the remaining patients (n=353, p=0.028)(data not shown).

We also tested the disease associated variant for all 6 polymorphisms in recessive and additive genetic models but none were significantly associated with either treatment-free survival or overall survival (data not shown). Taken together, these data suggest that CLL risk-associated alleles at 2q13 (rs17483466, *ACOXL*), 2q37.1 (rs13397985, *SP140*), 11q24.1 (rs735665, *GRAMD1B*), 15q23 (rs7176508) and 19q13.32 (rs11083846, *PRKD2*)[24] do not affect CLL prognosis. By contrast, the disease associated polymorphism identified at 6p25.3 (*IRF4*) also predicts CLL prognosis.

IRF4 genotype predicts prognosis independent of IgV_H gene mutation

We next examined the demographic and clinical characteristics of carriers (n=348 (86.4%)) for the *IRF4* risk allele and compared these to non-carriers (n=55, (13.6%)). There was no significant difference in gender and age at diagnosis between carriers and non-carriers (Table 1). There was also no significant difference between carriers and non-carriers with respect to Binet stage of disease at diagnosis or IgV_H mutational status (Table 1). Multivariate analysis using Cox regression including age, stage of disease at diagnosis, IgV_H mutational status and gender in the model confirmed that carrier status for the IRF4 disease-associated risk allele was an independent indicator of treatment-free survival (p=0.017).

Evidence suggests that CLL aetiology is influenced by antigenic stimulation, possibly involving a plethora of different antigens as indicated by the heterogeneous nature of IgV_H gene usage. It is conceivable that *IRF4* status predisposes to disease development and subsequent disease progression ~~via chronic exposure to specific antigens~~. In order to test this we compared IgV_H gene usage between carriers and non-carriers of the disease-associated variant. Data on IgV_H gene usage was available from 235 carriers and 39 non-carriers of the disease-associated allele at the *IRF4* locus. The most commonly used IgV_H gene segments were 4-34, 1-69, 3-7, 3-23, 3-30 and 3-21, accounting for 42% (114/274) of cases, but the distribution of IgV_H gene usage was not significantly different between carriers and non-carriers (Figure 3A), suggesting that ~~differences in exposure to specific~~ antigenic stimulation ~~are~~ unlikely to be responsible for the poor prognosis seen in carriers.

Relationship between *IRF4* genotype and status of CD38 and ZAP70.

Given the power of ZAP70 and CD38 expression as prognostic markers in CLL [refs] we sought to determine their relationship with *IRF4* genotype. Carriers of the CLL-risk allele in *IRF4* were significantly more likely than non-carriers to have CLL positive for ZAP70 (p=0.039) and CD38 (p=0.025)(Table 1) (positivity for ZAP70 and CD38 is defined by 20% or more cells positive for expression (see materials and methods)). Data on the percentage of leukaemic lymphocytes positive for ZAP70 and CD38 were available for X and X patients, respectively. When examined using this

expression data, *IRF4* genotype was significantly correlated with percentage of cells positive for ZAP70 in an allele dose-dependent fashion (p=0.007), but not with percentage of cells positive for CD38 (p=0.47). Specifically, the mean percentage of cells positive for ZAP70 was X, X and X for homozygotes of the CLL risk allele (G/G), heterozygotes (G/A) and homozygotes for the non-risk allele (A/A), respectively. been proposed as a regulator of CD38 expression, although the precise relationship remains to be defined. Carriers of the CLL disease-associated allele in IRF4 were more likely to be positive for CD38 expression (p=0.014)(table 2).

Discussion.

Acknowledgements.

This work was supported by Leukaemia Research, London, (grant number 06002). The authors declare no competing financial interests. The authors thank Simon Kometa (Newcastle University) for statistical advice.

Abbreviations

CLL, chronic lymphocytic leukaemia; TFS, treatment-free survival; OS, overall survival; SNP, single nucleotide polymorphism

Author Contributions statement.

J.M.A., D.J.A., G.P. and R.S.H. developed the study hypothesis. J.M.A. established the study, performed research, analysed data and wrote the manuscript. N.J.S. performed research. A.G.H. developed the Newcastle Haematology Biobank, incorporating the Newcastle-based CLL study. T.M.-F., G.H.J., G.S., R.J.H., A.R.P., D.J.A., J.R.B., G.P., C.P. and C.F. developed protocols for recruitment of individuals with CLL and sample acquisition, performed sample collection of cases and associated demographic, diagnostic and prognostic data. GP and DJA contributed equally to this study. All authors contributed to the final paper.

Table and figure legends.

Table 1 - Comparison of patient demographic and clinicopathologic characteristics by carrier status of the CLL risk allele in *IRF4* defined by rs872071.

Table 2 - Cox regression model of treatment-free survival. Age was included in the model as a continuous variable. HR indicates hazard ratio; and 95% CI, 95% confidence intervals.

Figure 1 – Kaplan-Meier curves for overall survival stratified by IgV_H mutation status (A), Binet stage of disease at diagnosis (B), and age at diagnosis (C).

Figure 2 – Kaplan-Meier curves for treatment-free survival stratified by rs17483466 (A), rs13397985 (B), rs735665 (C), rs7176508 (D), rs11083846 (E) and rs872071 (F). Carriers include patients with at least one copy of the risk allele for development of CLL. Non-carriers include patients homozygous for the non-risk allele for the development of CLL. p values were calculated using the log-rank test.

Figure 3 – Kaplan-Meier curves for treatment-free survival by genotype and overall survival by carrier status for the CLL risk allele in *IRF4*. Carriers and non-carriers are as defined in the legend to figure 2. p values were calculated using the log-rank test.

Figure 4 – Association between *IRF4* polymorphic status and IgV_H gene usage (A), percentage of cells positive for ZAP70 (B) and percentage of cells positive for CD38 (C) in CLL. In (A) IgV_H gene segment usage is stratified by carrier status for the CLL risk allele in *IRF4*. The 6 most frequently used gene segments are individually presented. Figure is based on data from In (B) and (C) the percentage of cells positive for ZAP70 and CD38, respectively, is stratified by IRF4 genotype. The length of the boxes represent the interquartile range, and the whiskers represent the 1.5 interquartile range. The boxes are

divided by the median value for each group. The p value was derived using a Wilcoxon-type test for trend.

Reference List

(1) Oscier D, Fegan C, Hillmen P et al. Guidelines on the diagnosis and management of chronic lymphocytic leukaemia. Br J Haematol. 2004;125:294-317.

(2) Cheson BD, Bennett JM, Grever M et al. National Cancer Institute-sponsored Working Group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. Blood. 1996;87:4990-4997.

(3) Landis SH, Murray T, Bolden S, Wingo PA. Cancer statistics, 1998. CA Cancer J Clin. 1998;48:6-29.

(4) Binet JL, Auquier A, Dighiero G et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. Cancer. 1981;48:198-206.

(5) Rai KR, Sawitsky A, Cronkite EP et al. Clinical staging of chronic lymphocytic leukemia. Blood. 1975;46:219-234.

(6) Dighiero G. CLL Biology and Prognosis. Hematology Am Soc Hematol Educ Program. 2005;278-284.

(7) Dohner H, Fischer K, Bentz M et al. p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. Blood. 1995;85:1580-1589.

(8) Binet JL, Auquier A, Dighiero G et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. Cancer. 1981;48:198-206.

(9) Rai KR, Sawitsky A, Cronkite EP et al. Clinical staging of chronic lymphocytic leukemia. Blood. 1975;46:219-234.

(10) Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999;94:1848-1854.

(11) Dohner H, Fischer K, Bentz M et al. p53 gene deletion predicts for poor survival and non-response to therapy with purine analogs in chronic B-cell leukemias. Blood. 1995;85:1580-1589.

(12) Dohner H, Stilgenbauer S, Benner A et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000;343:1910-1916.

(13) Austen B, Powell JE, Alvi A et al. Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. Blood. 2005;106:3175-3182.

(14) Nuckel H, Frey UH, Bau M et al. Association of a novel regulatory polymorphism (-938C>A) in the BCL2 gene promoter with disease progression and survival in chronic lymphocytic leukemia. Blood. 2007;109:290-297.

(15) Saxena A, Moshynska O, Sankaran K, Viswanathan S, Sheridan DP. Association of a novel single nucleotide polymorphism, G(-248)A, in the 5'-UTR of BAX gene in chronic lymphocytic leukemia with disease progression and treatment resistance. Cancer Lett. 2002;187:199-205.

(16) Starczynski J, Pepper C, Pratt G et al. Common polymorphism G(-248)A in the promoter region of the bax gene results in significantly shorter survival in patients with chronic lymphocytic Leukemia once treatment is initiated. J Clin Oncol. 2005;23:1514-1521.

(17) Sellick GS, Wade R, Richards S et al. Scan of 977 nonsynonymous SNPs in CLL4 trial patients for the identification of genetic variants influencing prognosis. Blood. 2008;111:1625-1633.

(18) Sellick GS, Goldin LR, Wild RW et al. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. Blood. 2007;110:3326-3333.

(19) Di Bernardo MC, Crowther-Swanepoel D, Broderick P et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2008.

(20) Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999;94:1848-1854.

(21) Damle RN, Wasil T, Fais F et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood. 1999;94:1840-1847.

(22) Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999;94:1848-1854.

(23) Di Bernardo MC, Crowther-Swanepoel D, Broderick P et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2008.

(24) Di Bernardo MC, Crowther-Swanepoel D, Broderick P et al. A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. Nat Genet. 2008.

Table 1

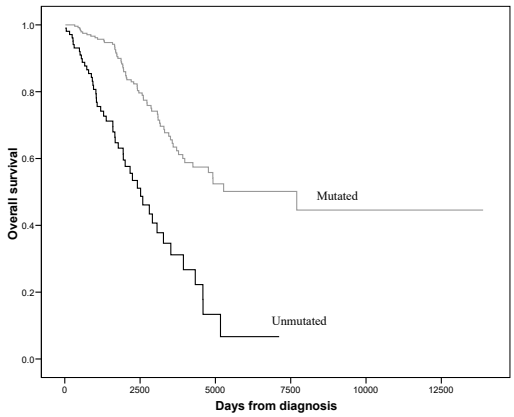
Variable	Carriers of the disease risk allele	Non-carriers of the disease risk allele	*P*
Total, no. (%)	449 (100)	55 (100)	
Sex, no. (%)			
Female	139 (39.8)	25 (45.5)	0.430
Male	210 (60.2)	30 (54.5)	
Age at diagnosis, no. (%)			
Less than 60 y	77 (24.8)	12 (25.5)	0.397
60 to 69 y	102 (32.8)	11 (23.4)	
over 70 y	132 (42.4)	24 (51.1)	
Binet stage at diagnosis, no. (%)			
A	229 (75.3)	40 (85.1)	0.221
B	38 (12.5)	2 (4.3)	
C	37 (12.2)	5 (10.6)	
IgV_H Mutational status, no. (%)			
Unmutated	94 (31)	10 (20.4)	0.131
Mutated	209 (69)	39 (79.6)	
CD38 expression, no. (%)			
20% or more	88 (37.1)	6 (16.7)	0.016
Less than 20%	149 (62.9)	30 (83.3)	
ZAP70 expression, no. (%)			
20% or more	97 (39.0)	9 (25)	0.105
Less than 20%	152 (61.0)	27 (75)	
Lymphocyte doubling time, no. (%)			
Less than12 months	62 (35.6)	3 (12.5)	0.024
Greater than12 months	112 (64.4)	21 (87.5)	
Treatment status, no. (%)			
Untreated	188 (53.9)	45 (81.8)	<0.001
Treated	161 (46.1)	10 (18.2)	
Deceased, no. (%)			
No	237 (67.9)	43 (78.2)	0.125
Yes	112 (32.1)	12 (21.8)	

Table 2

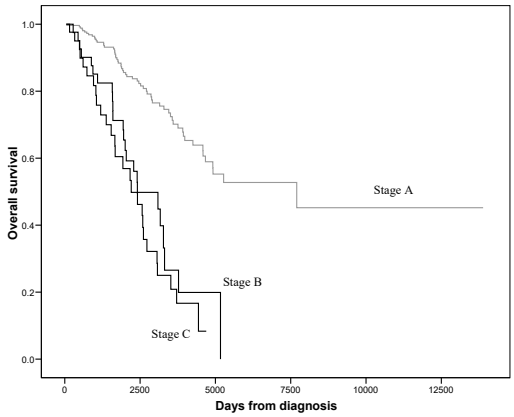
Variable	HR (95% CI)	Standard error	*P*
Sex			
Female			
Male	1.17 (0.84-1.6)	0.17	0.360
Age	1.01 (0.99-1.03)	0.01	0.081
Binet stage at diagnosis			
A			
B	6.35 (4.16-9.69)	0.22	<0.001
C	9.10 (5.75-14.41)	0.23	<0.001
IgV_H Mutational status, no. (%)			
Mutated			
Unmutated	2.17 (1.50-3.14)	0.19	<0.001
IRF4 status at rs872071			
Non-carriers			
Carriers	2.38 (1.20-4.71)	0.35	0.013

Figure 1

(A)


Overall survival
1.0
0.8
0.6
0.4
0.2
0.0
Mutated
Unmutated
0
2500
5000
7500
10000
12500
Days from diagnosis

(B)


Overall survival
1.0
0.8
0.6
0.4
0.2
0.0
Stage A
Stage B
Stage C
0
2500
5000
7500
10000
12500
Days from diagnosis

(C)

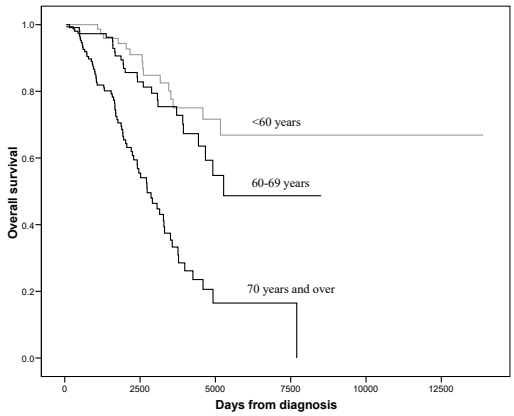
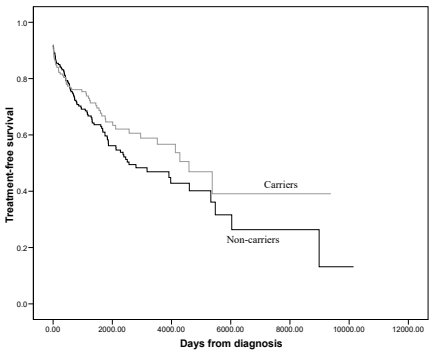
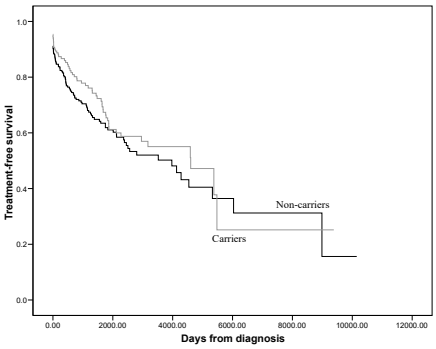

Overall survival
1.0
0.8
0.6
0.4
0.2
0.0
<60 years
60-69 years
70 years and over
0
2500
5000
7500
10000
12500
Days from diagnosis

Figure 2

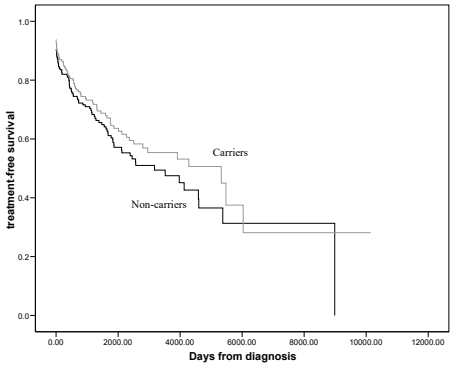
(A) rs17483466, p=0.146



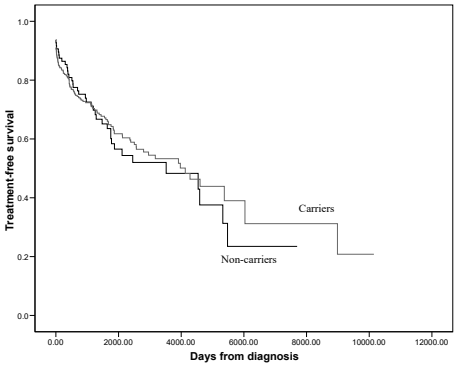
(B) rs13397985, p=0.253



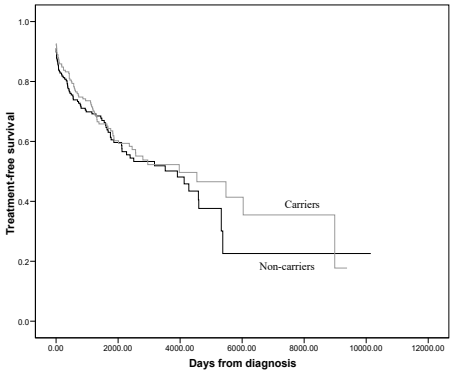
(C) rs735665, p=0.211



(D) rs7176508, p=0.672



(E) rs11083846, p=0.448



(F) rs872071, p=0.001

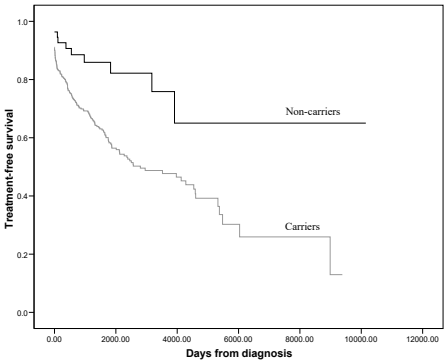
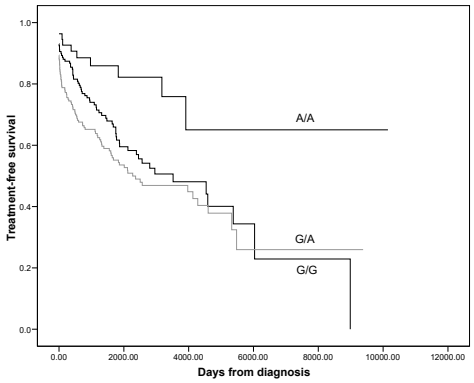


Figure 3

(A)


Treatment-free survival
1.0
0.8
0.6
0.4
0.2
0.0
A/A
G/A
G/G
0.00
2000.00
4000.00
6000.00
8000.00
10000.00
12000.00
Days from diagnosis

(B)

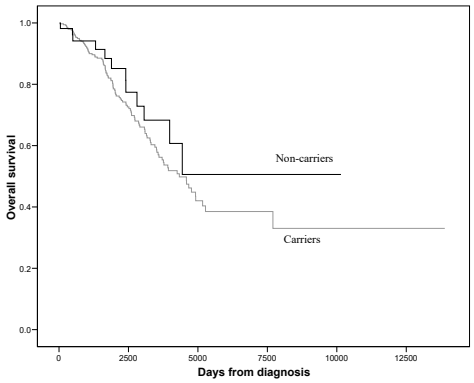
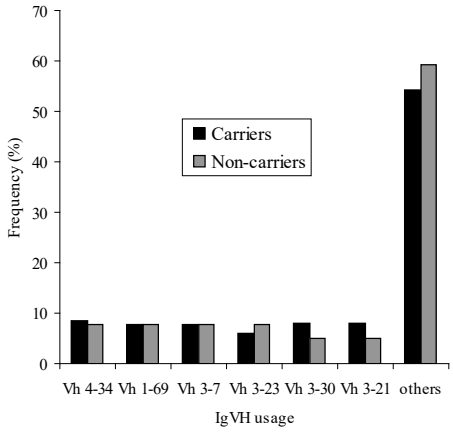
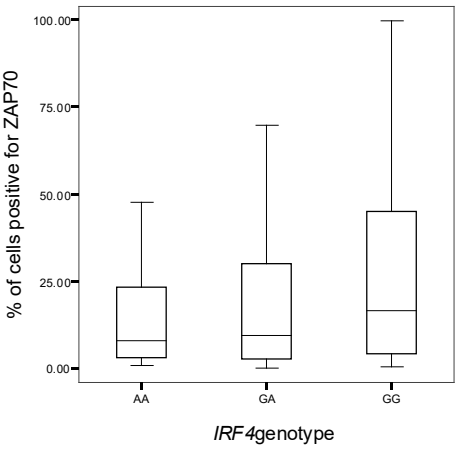

Overall survival
1.0
0.8
0.6
0.4
0.2
0.0
Non-carriers
Carriers
0
2500
5000
7500
10000
12500
Days from diagnosis

Figure 4

(A)



(B) p=0.011



(C) p=0.467