Genetic alterations and cancer formation in a European flatfish at sites of different contaminant burdens.

*[†]Adélaïde Lerebours, [‡]Grant D. Stentiford, [‡]Brett P. Lyons, [‡]John P. Bignell, [†]Stéphane A. P. Derocles and [†]Jeanette M. Rotchell**

[†]School of Biological, Biomedical and Environmental Sciences, University of Hull, Cottingham Road, Hull, HU6 7RX, United Kingdom

[‡]Centre for Environment, Fisheries and Aquaculture Science, Weymouth Laboratory, Barrack Road, The Nothe, Weymouth, Dorset DT4 8UB, United Kingdom

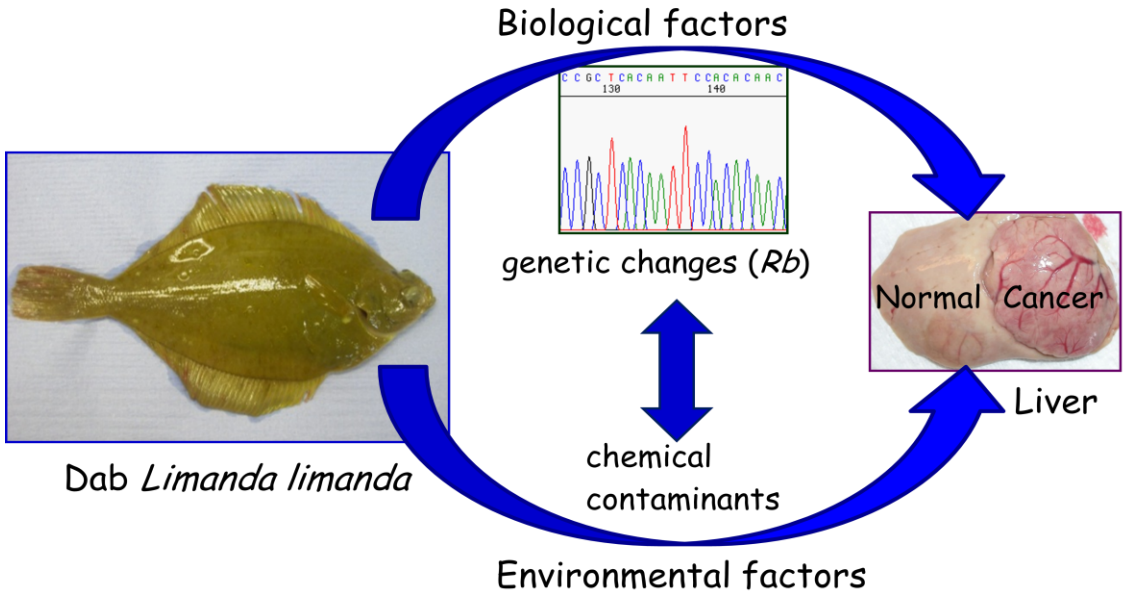
Keywords: fish, liver cancer, *retinoblastoma* gene, genetic changes, contaminants

Corresponding Author

*Jeanette M. Rotchell, phone number: +44 (0) 1482 465333, fax number: +44 (0) 1482 465458, e-mail: J.Rotchell@Hull.ac.uk

Running title: Linking *Rb* genotype, tumour phenotype and contaminant exposure in the flatfish dab.

18 TOC Art


Biological factors
genetic changes (Rb)
Normal Cancer
Liver
chemical
contaminants
Dab Limanda limanda
Environmental factors

19

Abstract

Fish diseases are an indicator for marine ecosystem health since they provide a biological end-point of historical exposure to stressors. Liver cancer has been used to monitor the effects of exposure to anthropogenic pollution in flatfish for many years. The prevalence of liver cancer can exceed 20%. Despite the high prevalence and the opportunity of using flatfish to study environmentally-induced cancer, the genetic and environmental factors driving tumour prevalence across sites are poorly understood. This study aims to define the link between genetic deterioration, liver disease progression, and anthropogenic contaminant exposures in the flatfish dab (*Limanda limanda*). We assessed genetic changes in a conserved cancer gene, *Retinoblastoma* (*Rb*) in association with histological diagnosis of normal, pre-tumour and tumour pathologies in the livers of 165 fish from six sites in the North Sea and English Channel. The highest concentrations of metals (especially cadmium) and organic chemicals correlated with presence of tumour pathology and, with defined genetic profiles of the *Rb* gene, from these sites. Different *Rb* genetic profiles were found in liver tissue near each tumour phenotype, giving insight into the mechanistic molecular-level cause of the liver pathologies. Different *Rb* profiles were also found at sampling sites of differing contaminant burdens. Additionally, profiles indicated that histological 'normal' fish from Dogger sampling locations possessed *Rb* profiles associated with pre-tumour disease. This study highlights an association between *Rb* and specific contaminants (especially cadmium) in the molecular aetiology of dab liver tumourigenesis.

Introduction

Fish diseases represent an indicator of marine ecosystem health since they provide a biological end-point of historical exposure to stressors[1]. Liver pathologies of flatfish including tumours have been used to monitor the effects of exposure to pollution for many years [1-4]. As such they are routinely used in a number of internationally co-ordinated marine monitoring programmes and have been recommended as a key tool for assessing ecosystem health by organisations including the International Council for Exploration of the Sea (ICES) and the Oslo and Paris Convention (OSPAR) Joint Assessments and Monitoring Programme (JAMP) [5].

A high prevalence of dab (*Limanda limanda*) liver tumours, exceeding 20% at some localities in the North Sea, has been reported [6,7]. This prevalence is of interest both in terms of the molecular basis of tumourigenesis, and its ecological implication. Dab is a bottom-dwelling fish particularly sensitive to environmental stressors [4] and can live up to 11 years making it a good indicator of the past history of contamination [8]. It is also widely distributed and highly abundant across the North Sea, Irish Sea and the English Channel [9], facilitating population studies. The genetic structure of dab population is arguably regarded as stable over time, with a life-long residency in sampling regions proposed [9]. This is a fundamental criterion for sentinels of use in biomonitoring programmes. Therefore, the dab offers a unique opportunity to study environmental cancer. While there is debate among the scientific community regarding the impact of such disease on population dynamics [10-13], the underlying genetic and environmental factors driving tumour prevalence across sites are still poorly documented.

Histopathology of tumours and pre-tumours in dab liver are currently diagnosed via a quality assured process involving histological tissue sections generated from wax-embedded samples [14]. Within the UK, such samples are collected and results are reported under the U.K. Clean Seas Environmental Monitoring Programme (CSEMP) [6]. Previous molecular studies using dab have

revealed differences in tumour or surrounding tumour tissues as compared to normal ones, including genetic alterations of cancer genes [15-18], as well as differential gene expression [6,19-21], protein synthesis [22], and metabolic changes [22,23]. Finally, Tysklind et al. (2013), observed significant interactive effects between the genetic structuring of dab populations, environmental contaminants and certain liver pathologies from specific sites in the North and Irish Sea. While some of these studies highlight a role of chemical contaminants in the aetiology of liver pathologies, the precise mechanistic cause and effect relationship, specifically at the sub-cellular / molecular level and how chemicals may interact with genotype to influence tumour development, is still uncharacterised.

Cancer is a multi-factor disease, according to medical studies, resulting from gene-environment interactions. The combination of environmental stressors such as chemicals and the susceptibility of the host can result in alteration of environmentally relevant genes such as mutations in cancer genes. The development of hepatocellular carcinoma (HCC) is a multistep process of transformation of normal cells into malignant cells driven by accumulation of genetic and epigenetic alterations in such genes [24-27].

The *Rb* gene was the first tumour suppressor gene to be characterised [28]. In vertebrates, the *Rb* gene product is a nuclear phosphoprotein that regulates normal cell cycle progression. In humans, *Rb* mutations have been reported in hepatocellular carcinoma (HCC) and RB protein is inactivated in the majority of human cancers [29]. *Rb* alterations have been detected in chemically-induced retinoblastoma in the medaka (*Oryzias latipes*), a laboratory fish model [17]. Dab possess both a similar histopathological liver tumour profile to humans [30] and homologs of human cancer genes [15,16]. It is likely that dab and human share downstream signalling cascades underlying HCC formation; further support for the suitability of this species as a relevant model of environmentally-induced liver cancer.

The present study aims at defining the link between genetic deterioration, visible disease progression and environmental contaminant burdens in a discrete population of flatfish dab [9]. To achieve this, the *Rb* genetic changes and histopathological diagnosis of normal, pre-tumour and tumour in liver of 165 fish collected at four sites at Dogger Bank and two sites in the east English Channel, were assessed. Concentrations of metals (cadmium, Cd; mercury, Hg; lead, Pb; zinc, Zn; copper, Cu) and organic chemicals (polybrominated diphenyl ethers, PBDEs, and polychlorinated biphenyls, PCBs) in the liver of fish from the same sites were analysed in parallel to provide contaminant burden indication.

Material and Methods

Sample Collection

Dab (*Limanda limanda*) were captured at UK CSEMP sites on the Dogger Bank (North Dogger, North East Dogger, Central Dogger and West Dogger), North Sea and the English Channel (Rye Bay and Newhaven) (Table S1) during July 2010, using 30 min tows of a standard Granton trawl aboard the RV *Cefas Endeavour*. These sites are among those used for both ICES and OSPAR statutory monitoring and have been identified as having historically high (Dogger) or low (Rye Bay/Newhaven) prevalence of liver tumours [6, 31]. Upon landing, fish were immediately removed from the catch and placed into flow-through tanks containing aerated seawater. The sex and size (total length) and presence of external signs of disease were noted for each fish using methodology specified by ICES [14]. Otoliths were sampled from each fish and processed for age determination according to Easey & Millner (2008)[32]. Following euthanasia, the body cavity was opened and the liver assessed for the presence of macroscopic liver tumours according to the guidelines set out by Feist et al. (2004)[14]. For each fish (n = 165), a standardised cross section was obtained for histological analysis and placed into 10% neutral

buffered formalin and processed as described in '*Histology/histopathology*'. A part of the liver from the same individual fish (and beside the previous dissected fragment) was also sampled and snap frozen in liquid nitrogen for molecular analysis as described in '*Total RNA, cDNA preparation and* Rb *cDNA isolation'* below.

Chemical concentrations and biomarkers of exposure to polycyclic aromatic hydrocarbons (PAHs) in bile, liver or flesh from fish

Data pertaining to chemical and biomarker analysis was collated from the Marine Environment Monitoring and Assessment National database (www.bodc.ac.uk/projects/uk/merman/), which holds UK data collected to fulfill the UK's mandatory monitoring requirements under the OSPAR Joint Assessments and Monitoring Programme (JAMP). In brief, the measurement of metals, PBDEs and PCBs was performed on 5 pools of livers (flesh for Hg) from 5 fish (representing 25 fish in total) for each site. The fish were from the same trawl as the fish used in the molecular and histology analyses. Chemical analyses were processed using standardised protocols as previously described for metals [33], PBDEs [34], and PCBs [35]. For an indication of exposure to PAHs, bile hydroxypyrene levels and ethoxyresorufin *O*-deethylase (EROD) activities were obtained from a subset of twenty fish (10 males and 10 females) sampled during the same trawls at each site. The livers and gall bladders were collected and analyzed for both EROD and bile measurements following standard protocols published in the ICES Techniques in Marine Environmental Sciences Series (ICES TIMES). EROD activity was determined in liver tissue using a fluorescent assay [36]. Bile samples were analyzed for fluorescent bile metabolites using synchronous fluorescence spectrometry (SFS)[37].

Histology/histopathology

Fish were assessed for grossly visible tumours and histopathological assessment of liver samples from flatfish populations collected under CSEMP. The lesions recorded include those thought to precede the development of benign and malignant lesions such as foci of cellular alteration, non-neoplastic toxicopathic lesions (such as nuclear and cellular polymorphism) and lesions associated with cell death, inflammation and regeneration. Currently, 32 categories of liver lesion are classified under the international Biological Effects Quality Assurance in Monitoring Programmes (BEQUALM) project. The diagnosis of these lesion types in the dab and flounder liver follows the guidelines set out by Feist et al. (2004)[14]. Upon landing, dab of 20 to 30 cm total length from each site in each year were immediately removed from the catch and placed into flow-through tanks containing aerated seawater [14]. The sex, size (total length) and presence of grossly visible signs of disease were recorded for each fish using the methodology specified by the International Council for the Exploration of the Sea (ICES) [38]. Following grossly visible disease assessment, fish were euthanised and, upon opening of the body cavity, the liver was assessed for the presence of visible tumours according to the guidelines set out by Feist et al. (2004)[14]. Liver samples were removed and fixed for 24 h in 10% neutral buffered formalin (NBF) before transfer to 70% industrial methylated spirit (IMS) for subsequent histological assessment. Livers were processed for formalin fixed paraffin embedded histology in a vacuum infiltration processor using standard histological protocols and embedded in paraffin wax. Using a rotary microtome, sections of 3-4 µm were taken and subsequently stained with haematoxylin and eosin (H&E). Slides were examined for microscopic tumours (hepatocellular adenoma and HCC) and pre-tumours (vacuolated foci of cellular alteration (FCA), eosinophilic FCA, basophilic FCA), according to BEQUALM and ICES criteria [14] using a Nikon Eclipse E800 microscope.

Total RNA isolation, cDNA synthesis and Rb *cDNA isolation from individual fish*

153 For each fish an additional sample of liver (approximately 20 mg) was removed from near the sample
154 used in histology analysis, for parallel molecular analyses, specifically isolation of the *Rb* cDNA. Total
155 RNAs were extracted using the High Pure RNA Tissue kit (Roche Diagnostics Ltd, West Sussex, U.K.)
156 according to the supplier's instructions. RNA quality (integrity of 18S and 28S ribosomal bands) was
157 evaluated by electrophoresis on a 1% agarose-formaldehyde gel. First strand cDNAs were synthesized
158 from 1 µg of total RNA using the SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen Ltd, Paisley,
159 U. K.) and according to the supplier's instructions.

160 Three overlapping parts of the coding sequence of the *Rb* cDNA: RbA1, RbA2 and RbB,
161 containing the region of functional importance were amplified. Primer pairs used to amplify the region
162 between 620 and 1942 bp of the *Rb* cDNA (Accession number: **AY973250**) are described in Table S3
163 (contained in Supplemental Information). One µL of the reverse transcribed product was used as a
164 template for subsequent polymerase chain reaction (PCR) in a 25 µL final volume using 2.5 units of the
165 Expand High Fidelity^{PLUS} enzyme (Roche Diagnostics Ltd, West Sussex, U.K.), primers at a final
166 concentration of 1 µM and following the supplier's protocol. PCR reactions were performed using the
167 following programme: one cycle at 94°C for 2 min and 40 amplification cycles at 94°C for 30 s, 60°C
168 (RbA1) or 65°C (RbA2 and RbB) for 30 s, and 72°C for 1 min. 10 µL of each PCR product were then
169 forward and reverse sequenced commercially (Macrogen, Amsterdam, Netherlands). Both strands for
170 each overlapping fragment were assembled using the sequence-editing software CodonCode Aligner
171 version 4.0. Sequences were aligned using ClustalW 1.81.

172

173 *Statistical analysis*

174 Statistical analyses were performed using R 3.0 (R Development Core Team 2013). The distribution
175 of different tumour stages and genetic profiles among sites, and the relation between the genetic

profiles and tumour stages were first analysed by correspondence analyses, using the "dudi.coa" function (ade4 package). The distribution of chemicals among sites was assessed by a principal component analysis (PCA), using the "dudi.pca" function (ade4 package). The effect of the site, genotype, sex and age of fish on the presence (pre-tumour and tumour) - absence (normal) of tumour was also tested using generalized linear models (GLIM). All of these factors were included in the model. Statistical analyses were performed using GLIM (Poisson family, link log), with the anova.glm function in R. The best-fit model was selected using Akaike information criterion (AIC). Full explanation of the models used to derive Figures 1-4 are given in Supplemental Information as Supplemental Methods, SM1.

Results

Fish biometric distribution relative to locality

The size and weight ranges for the fish used in this study are provided in Table S2 in the Supplementary Information section. In terms of the biometric data for the 165 fish sampled in this study there were significant differences in the composition of the individuals at specific sampling locations as follows. Fish sampled at North Dogger/Central Dogger were significantly larger/smaller than other sampling sites (Table S2). Fish sampled at Dogger sites were also significantly older than fish sampled at Newhaven (Table S2). However, no significant differences between fish sampled at all the sampling sites were evident for Fulton Condition Index, liver weight or hepatosomatic index (HSI) (Table S2). PCA statistical analysis of all the factors subsequently indicated a significant effect of site, genotype and age of fish on the presence-absence of liver tumours (GLIM, site: $p = 0.006$; genotype: $p = 0.028$; age: $p = 0.0007$; sex: $p = 0.057$). We shall thus present the results in the order of site/locality, phenotype, genotype, age and sex.

Distribution of metals, PCBs and PBDEs relative to locality

The concentrations of contaminants in dab liver differed significantly by site (Table S4a-c) and this dataset has been used to produce a PCA plot to characterise the distribution of individual chemicals in relation to site (Figure 1). For instance, the liver of fish sampled from Newhaven was characterised by relatively low levels of PCB contamination (Figure 1; Table S4c), whereas that of fish sampled from North Dogger was characterised by high concentrations of Cd (406 $\pm$ 122 μg/kg liver tissue)(Figure 1; Table S4a). Associations between different chemical contaminants are presented in Table S5. Principal component analysis showed the following highlights: the liver of fish from Rye Bay was characterised by contamination with the greatest number, and highest concentrations, of PCB congeners (particularly CB101, 105, 110, 138, 153 and 187)(Table S4c); fish from Newhaven less so (though PCBs still formed the dominant profile)(Table S4c); those from Central, West and North East Dogger being weakly associated to metals, PBDEs and PCB contamination; and those from North Dogger being most associated to metals (with the highest association for Cd) (Figure 1; Table S4a-c).

Sampling site-specific distribution of tumour phenotypes

The occurrence of normal, pre-tumour (including all FCA types), and tumour liver phenotype form a gradient progressing from the Newhaven to North Dogger sites (Figure 2; Figure S1). Correspondence analysis revealed a gradient as follows: normal livers were mostly found in fish sampled at the Newhaven site (81%) and then at Rye Bay (67%) and North East Dogger (66%) to a lesser extent (Figure 2; Figure S1). This latter site also contained fish displaying pre-tumours (24%), whilst this pathology also dominated in fish from the West Dogger (31%) and Central Dogger sites (36%)(Figure 2; Figure S1). In terms of prevalence, tumours were most prevalent in the livers of fish from the North

Dogger site (20%) (Figure 2; Figure S1). North Dogger was thus characterised by high Cd levels (406 ± 122 μg/kg liver tissue) and high liver tumour prevalence (20%).

Different Rb *genetic profiles are found between sites and tumour phenotypes*

Rb genetic profiles were characterised in fish samples from six sites within a North Sea and English Channel dab population. Four nucleotides were found to be changed in the *Rb* coding sequence at 996 bp (G to A), 1088 bp (T to C), 1514 bp (G to T) and 1592 bp (G to T) leading to 17 different genetic profiles annotated from A to Q (Table 1). All of these changes occurred within the *Rb* sequence encoding the functionally important and conserved A and B domains.

Differing *Rb* profiles were associated with fish captured at different North Sea and English Channel locations (Figure S2). Correspondence analysis (Figure 3) revealed three groupings: one associates fish from Newhaven, Rye Bay, Central Dogger, North East Dogger with profiles A, B, C, E, H, P and Q; a second associates fish from West Dogger with profiles D, F, G and I; and the third associates fish from North Dogger with profiles L, M, N and O (Figure 3, Table 1).

Additionally, several *Rb* profiles were identified in livers of fish displaying normal, pre-tumour and tumour phenotypes. Correspondence analysis (Figure 4) showed that five *Rb* profiles; A, D, I, Q and P were associated with normal liver phenotype, ten profiles; B, C, E, F, H, J, K, M, N and O are associated with liver pre-tumour stages, and profiles G and L are associated with a liver tumour phenotype (Figure 4). The differences in these *Rb* profiles hinge around only four nucleotide positions of the *Rb* sequence (Table 1). On close examination of the *Rb* gene status at samples from West Dogger, genotypes seen in pre-tumour fish (profiles C and D, Table 1) are also seen in normal fish from that site, giving an indication that normal fish from that site on a pathogenesis trajectory to liver tumour (Figures 2- 4, Table 1).

Age and sex

The age of fish has a significant effect on the liver phenotype (normal and tumour) (GLIM1, $p = 0.0007$, see Supplementary Information, SM1, for full statistics). Fish from Dogger Bank are significantly older than fish from Newhaven ($p < 0.05$, Supplementary Information, Table S2). However, the age of fish from a given site displaying normal and tumour phenotypes is similar (GLIM2, $p = 0.0756$, see Supplementary Information, SM1, for full statistics). The sex of fish has no effect on the phenotype observed (normal and tumour) using the number of fish sampled in this study (GLIM1, $p = 0.06$ see Supplementary Information, SM1, for full statistics).

In summary, we link the presence of liver tumours in dab to specific contaminant classes and *Rb* gene status in liver tissue next to that used in histology, providing a potential mechanism for future characterisation and prediction of disease prevalence in such populations.

Discussion

For the first time, this study provides a link between genetic deterioration, visible disease progression and specific environmental contaminant profiles in discrete populations of marine fish. Specifically, we are the first to link genetic profiles (using the *Rb* gene) to histopathological diagnosis of normal, pre-tumour and tumour, in liver tissue of the same individual fish from different sampling sites. These sampling sites have also been characterised in terms of predominant contaminant classes present in the fish liver tissue, thus providing an indication of the potential causality in generation of differing *Rb* genetic profiles. Such profiles also indicate that normal fish from the Dogger Bank also possess *Rb* profiles associated with pre-tumour disease (Figure 2, Table 1) suggesting that such fish are possibly heading towards liver tumours.

Characteristic Rb *profiles are associated with disease phenotype*

In terms of *Rb* genetic profiles, four nucleotide positions were altered, corresponding to a region of functional importance of the *Rb* gene, leading to 17 genetic profiles (Table 1). *Rb* profiles were not randomly distributed, with specific profiles associated with both sampling site (Figure 3) and liver phenotype (Figure 4). Of the *Rb* gene alterations characterised (Table 1), several were similar to those found in tumours sampled from a different dab population in the Irish Sea from a previous study [15]. The exception is one change occurring at 996 bp, corresponding to a G/G to G/A change, which has not been identified previously.

Regarding the precise molecular-level biological mechanisms of cause (pollutant-induced mutational activation/inactivation of key genes) and effect (pre-tumour and tumour liver phenotypes), understanding the implications of these *Rb* allele zygosity patterns (contained in Table 1) are key. For instance, focussing on *Rb* profile L (Table 1), which associates with both tumour phenotype (Figure 4) and North Dogger sampling site (Figure 3), this entails heterozygosity at two of the four nucleotide positions and a homozygous alteration at another (1592 bp). For the transitional, pre-tumour phenotype, the *Rb* profiles E, F, J, K, and O all display homozygous T allele at position 1592 bp. Such alterations in an established tumour suppressor gene may reflect driving steps in the multi-stage progression towards the tumour endpoint (as evidenced in rodent studies by Wang et al. (2012)[39]) and as such require further biochemical characterisation.

Of important note is the lack of any homozygous A/A detected at position 996 bp of the *Rb* sequence (Table 1) in any of the 165 fish analysed. The latter nucleotide alteration would theoretically lead to a change of amino acid involving a lysine (K) instead of glutamic acid (E). The glutamic acid (with polar acid properties) to lysine (with polar basic properties) alteration also occurs within the

functionally conserved Domain A of the protein that is responsible for a key LxCxE motif and transcription factor binding [40]. This theoretical change is identified as lethal phenotype $Rb^{-/-}$ in mice embryos [41]. The existence of such phenotype in dab may have already had, or could have future, repercussions at the population level and is of interest from the perspective of population sustainability of the dab.

Related to the lethality and phenotype discussion is age, an important cofactor involved in the epidemiology of tumour development. The analyses show that the age of fish is a potentially confounding factor. In general, fish are older at Dogger Bank than at Newhaven (Table S2). In this study, no significant differences between the age of fish displaying a normal or a tumour phenotype at each site were observed. However, the limited number of fish and associated age classes make it difficult to demonstrate clear links with tumour formation in our study. Since tumourigenesis is typically a multi-stage event involving several gene activation/inactivation events, one would expect older fish to display a higher prevalence of pre-tumour and tumour phenotypes. Taking into account previously published work, dab with HCA (a pre-tumour phenotype) were found in older age classes sampled from North Dogger Bank, yet no cases of HCC (actual tumour phenotype) were observed in fish of age >5 yr at this site [7]. Thus adding weight to the notion of an $Rb^{-/-}$ lethal phenotype.

Sex is also considered a confounding factor in the epidemiology of flatfish tumour development [40]. In our study, using a relatively small sample size of 165 fish (n = 11-37 at each sampling site), using the statistical approach described, no influence of sex was detected for any of the variables investigated but this is undermined by low numbers of males at certain sites (Table S2). N, W and C Dogger, in particular, has bigger and older fish, and the majority are females, which may in turn be due to relatively low numbers of animals sampled during current study. In previous work, focusing on age primarily as a confounding factor, yet importantly using very large dataset, evidence suggested that (despite some

314 significant differences between the mean age of fish sampled from specific sites) the mean age of all
315 male (5.3 yr) and all female (4.8 yr) fish sampled during the programme was similar, and relevantly,
316 data demonstrated a very similar prevalence of specific diseases in male and female dab[7].
317
318 *Characteristic* Rb *profiles are associated with sampling site*
319 Focussing on sampling sites, of particular interest are the results from North Dogger where fish livers
320 exhibit the highest prevalence (20%) of advanced stage tumour (Figure 2; Figure S1), possess specific
321 *Rb* genetic profiles (Figure 3), and display a high concentration of Cd (406 ± 122 μg/kg liver
322 tissue)(Figure 1; Table S4a). While site-specific disease profiles have been reported between sampling
323 years [6], these results highlight North Dogger Bank as a site of concern for prevalence of carcinogenesis
324 and involvement of Cd. Cd is a heavy metal with no essential role in organisms, classified as a human
325 carcinogen by the International Agency for Research on Cancer, and induces cancer in several
326 organs/tissue of animals by multiple direct and indirect mechanisms [43-45]. The liver is a target organ of
327 Cd toxicity in animals including fish [42]. Cd is a weak genotoxic chemical that inhibits DNA damage
328 repair pathways [46] and apoptosis induced by toxicants [47]. Cd co-exposure thus enhances the
329 carcinogenic potential, or may act as a promoter, of other genotoxic chemicals, such as PAHs
330 previously identified in the molecular aetiology of liver carcinogenesis in Atlantic killifish (*Fundulus*
331 *heteroclitus*) [48], to cause cancer. This is particularly relevant for dab populations that are chronically
332 exposed to a mixture of environmental contaminants such as the case at Dogger Bank. While the PAH
333 levels are not characterised in this study, the levels of hydroxpyrene and EROD activity (124 ± 52 ng/g
334 and 83 ± 58 pmol/min/mg protein respectively at North Dogger, Table S6) indicate that PAHs are
335 present but at levels significantly lower than the reference sites (for instance 124 ± 52 ng/g, 124 ± 52

pmol/min/mg protein for Newhaven)(Table S6). Further work involving controlled laboratory exposure is required to confirm the exposure-effect relationship.

Wider implications of Rb *involvement in fish tumour pathologies*

In terms of wider implications and utility of this work, there are two to consider: development of an early warning system and 'mutator phenotype'. Genetic modifications can occur earlier than microscopic histopathological changes in the tumourigenesis process. Here we have linked for the first time, *Rb* profiles in samples dissected from tissue located beside liver tissue, in the same individual fish, displaying a particular liver phenotype (Figure 4). Profile data also indicates that normal fish from Dogger sampling locations also possess *Rb* profiles associated with pre-tumour disease, providing an indication that such fish are heading towards development of a liver tumour. Relating *Rb* profiles to specific early neoplastic pre-tumour phenotype (different FCAs) may be used to predict future tumour prevalence likelihoods and is subject of a current study. A limitation of the study to highlight, however, is that the molecular analysis was conducted using liver tissue next to, yet not the exact same, liver tissue sample used for histopathology assessment. Inherent in such an approach is the scope for false negatives/positives, and that tissues of the same liver may show heterogeneity of cell type. More recently, a laser capture microdissection technique to address this limitation has been optimised in dab [49]. Nonetheless, this work associates *Rb* profile status with liver pathology. In addition, a second mechanism of possible RB interaction, via regulators of chromatin structure including methyltransferases, may be involved [20-21, 50]. Taken together our results and those from the literature highlight possible involvement of *Rb* in both genetic and epigenetic mechanisms in the aetiology of dab liver tumourigenesis.

Mutations in critical cell cycle control genes such as *Rb* represent a cellular defect that may catalyse the accumulation of further mutations, characteristic of a 'mutator phenotype' accelerating the disease process [51-52]. The genetic instability found in our study reflects the accumulation of DNA damage which is a key event driving the tumourigenesis process. In the absence of normal *Rb* gene, genomic instability and chromosomal aberrations are allowed to accumulate leading to tumour initiation, progression and metastasis [53]. The prevalence of cancer in most fish populations is extremely low with background levels similar to those seen in terrestrial wild animal populations and humans [7]. The high prevalence of HCA and other liver tumour types in dab and other marine flatfish populations from coastal environments [3,7,42,54] may be accounted for by the mutator phenotype theory. Herein we also show that the flatfish model provides an opportunity to study the mechanistic molecular etiology, including the relative contributing factors from the environment and the genotype, in the multi-step initiation and progression of vertebrate liver cancer.

This work represents a novel approach attempting to link genetic causes (by contaminant-induced damage in a conserved gene) to population-level biological endpoints (high prevalence of liver tumours). We assessed genetic changes in a key cancer gene, *Rb*, and made a histopathological diagnosis of normal, pre-tumour and tumour in the livers of 165 fish collected at four sites at Dogger Bank and two sites in the east English Channel. Four genetic changes were found within the *Rb* sequence at functionally important sites. Characteristic *Rb* genetic profiles were found in samples beside the tissue exhibiting different tumour phenotypes, giving insight into the mechanistic molecular-level cause of the observed liver pathologies, as well as a possible early warning tool for regulatory authorities. Characteristic *Rb* profiles were also found for sampling sites with differing contaminant burdens. This study highlights the involvement of *Rb* and specific contaminants (particularly cadmium) in the molecular aetiology of dab liver tumourigenesis.

Acknowledgements

The authors wish to acknowledge the Marine Environment Monitoring and Assessment National Database (MERMAN) for the provision of the supporting chemical contaminant data used in this study and recognise the support of The Department of Environment, Fisheries, and Rural Affairs (Defra) for funding the Clean Seas Environmental Programme, and specifically, contract #SLA22G (to JB) for completion of field sampling, histopathology and otolith assessment. Technical support in relation to reading otoliths was provided by Glenn Saunders, Mark Etherton and Suzy End, while support for chemical analysis was provided by Thi Bolam and Jon Barber (all Cefas).

Supporting Information Available

Tumour phenotype prevalence data and distribution of Rb genetic alleles at each sampling location are supplied as additional Figures. The sampling site coordinates, biometric data, analytical chemistry data plus correlation associations among chemical contaminants, and biomarkers of PAH exposure (hydroxyprene levels and EROD activities) are also supplied as additional Tables. The primers used for the isolation for the *Rb* cDNA are also available as an additional Table. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

1. Stentiford, G. D.; Longshaw, M.; Lyons, B. P.; Jones, G.; Green, M.; Feist, S. W. Histopathological biomarkers in estuarine fish species for the assessment of biological effects of contaminants. *Mar. Environ. Res*. **2003**, *55(2)*, 137-159.
2. Lyons ,B. P.; Stentiford, G. D.; Green, M.; Bignell, J.; Bateman, K.; Feist, S. W.; et al. DNA adduct analysis and histopathological biomarkers in European flounder (*Platichthys flesus*) sampled from UK estuaries. *Mutat. Res. Fund. Mol. Mech. Mutagen*. **2004**, *552(1-2)*, 177-186.
3. Malins, D. C.; Krahn, M. M.; Brown, D. W.; Rhodes, L. D.; Myers, M. S.; McCain, B. B.; et al. Toxic-chemicals in marine sediment and biota from Mukilteo, Washington - relationships with hepatic neoplasms and other hepatic lesions in English Sole (*Parophrys vetulus*). *JNCI* **1985**, *74(2)*, 487-494.
4. Vethaak, A. D.; ap Rheinallt T. Fish disease as a monitor for marine pollution: the case of the North Sea. *Rev. Fish Biol. Fisheries* **1992**, *2(1)*, 1-32.
5. Lyons, B. P.; Thain, J. E.; Stentiford, G. D.; Hylland, K.; Davies, I. M.; Vethaak, A. D. Using biological effects tools to define Good Environmental Status under the European Union Marine Strategy Framework Directive. *Mar. Poll. Bull.* **2010**, *60(10)*, 1647-1651.
6. Stentiford, G. D.; Bignell, J. P.; Lyons, B. P.; Feist, S. W. Site-specific disease profiles in fish and their use in environmental monitoring. *Mar. Ecol. Prog. Ser*. **2009**, *381*, 1-15.
7. Stentiford, G. D.; Bignell, J. P.; Lyons, B. P.; Thain, J. E.; Feist, S. W. Effect of age on liver pathology and other diseases in flatfish: Implications for assessment of marine ecological health status. *Mar. Ecol. Prog. Ser.* **2010**, *411*, 215-230.
8. Henderson, P. A. On the variation in dab *Limanda limanda* recruitment: A zoogeographic study. *J. Sea Res.* **1998**, *40(1-2)*, 131-142.

9. Tysklind, N.; Taylor, M. I.; Lyons, B. P.; Goodsir, F.; McCarthy, I. D.; Carvalho, G. R. Population genetics provides new insights into biomarker prevalence in dab (*Limanda limanda* L.): a key marine biomonitoring species. *Evol. Appl.* **2013**, *6(6)*, 891-909.
10. De Maagd, P. G .-J.: Vethaak, A. D. (1998). Biotransformation of PAHs and their carcinogenic effects in fish. In *Handbook of Environmental Chemistry*; Neilson, A. H. (Ed.); Springer: Stockholm, Sweden. 1998; pp. 265-310.
11. Grizzle, J. M.; Goodwin, A. E. Neoplasms and related lesions. In: *Fish diseases and disorders Volume 2 Non-infectious disorders*; Leatherland J. F., Woo, P. T. K., Eds.; CABI Publishing: Wallingford, U.K. 1998. pp. 37–104.
12. McCallum, H. 2008. Tasmanian devil facial tumour disease: lessons for conservation biology. *Trends Ecol. Evol*. **2008**, *23(11)*, 631-637.
13. McAloose, D.; Newton, A. L. Wildlife cancer: a conservation perspective. *Nat. Rev. Cancer* **2009**, *9(7)*, 517-526.
14. Feist, S. W.; Lang, T.; Stentiford, G. D.; Kohler, A. 2004. Biological effects of contaminants: use of liver pathology of the European flatfish dab (*Limanda limanda* L). *ICES TIMES* **2004**, *38*, 42.
15. du Corbier, F. A.; Stentiford, G. D.; Lyons, B. P.; Rotchell, J. M. Isolation of the retinoblastoma cDNA from the marine flatfish dab (*Limanda limanda*) and evidence of mutational alterations in liver tumors. *Environ. Sci. Technol*. **2005**, *39(24)*, 9785-9790.
16. Rotchell, J. M.; Stagg, R. M.; Craft, J. A. Chemically-induced genetic damage in fish: isolation and characterization of the dab (*Limanda limanda*) ras gene. *Mar. Poll. Bull.* **1995**, *31(4-12)*, 457-459.

17. Rotchell, J. M.; Blair, J. B.; Shim, J. K.; Hawkins, W. E.; Ostrander, G. K. Cloning of the Retinoblastoma cDNA from the Japanese medaka (*Oryzias latipes*) and preliminary evidence of mutational alterations in chemically-induced retinoblastomas. *Gene* **2001**, *263(1-2)*, 231-237.
18. Rotchell, J. M.; du Corbier, F. A; Stentiford, G. D.; Lyons, B. P.; Liddle, A. R.; Ostrander, G. K. A novel population health approach: Using fish *retinoblastoma* gene profiles as a surrogate for humans. *Comp. Biochem. Physiol. - C Toxicol. Pharmacol.* **2009**, *149(2)*, 134-140.
19. Small, H. J.; Williams, T. D.; Sturve, J.; Chipman, J. K.; Southam, A. D.; Bean, T. P.; et al. Gene expression analyses of hepatocellular adenoma and hepatocellular carcinoma from the marine flatfish *Limanda limanda*. *Dis. Aquat. Org*. **2010**, *88(2)*, 127-141.
20. Mirbahai, L.; Southam, A. D.; Sommer, U.; Williams, T. D.; Bignell, J. P.; Lyons, B. P.; et al. Disruption of DNA methylation via s-adenosylhomocysteine is a key process in high incidence liver carcinogenesis in fish. *J. Proteome Res.* **2013**, *12(6)*, 2895-2904.
21. Mirbahai, L.; Yin, G.; Bignell, J. P.; Li, N.; Williams, T. D.; Chipman, J. K. DNA methylation in liver tumorigenesis in fish from the environment. *Epigenet.* **2011**, *6(11)*, 1319-1333.
22. Stentiford, G. D.; Johnson, P. J.; Martin, A.; Wenbin, W.; Ward, D. G.;Viant, M.; Lyons, B. P.; Feist, S.W. Liver tumours in wild flatfish: a histopathological, proteomic and metabolomic study. *OMICS: J. Integrat. Biol.* **2005**, *9*, 281-299.
23. Southam, A. D.; Easton, J. M.; Stentiford, G. D.; Ludwig, C.; Arvanitis, T. N.; Viant, M R. Metabolic changes in flatfish hepatic tumours revealed by NMR-based metabolomics and metabolic correlation networks. *J. Proteome Res.* **2008**, *7(12)*, 5277-5285.
24. Baylin, S. B.; Ohm, J. E. 2006. Epigenetic gene silencing in cancer - A mechanism for early oncogenic pathway addiction? *Nat. Rev. Cancer* **2006**, *6(2)*, 107-116.

25. Feinberg, A. P.; Ohlsson, R.; Henikoff, S. The epigenetic progenitor origin of human cancer. *Nat. Rev. Genet.* **2006**, *7(1)*, 21-33.

26. Jones, P. A.; Baylin, S. B. The epigenomics of cancer. *Cell* **2007**, *128(4)*, 683-692.

27. Tischoff, I.; Tannapfel, A. DNA methylation in hepatocellular carcinoma. *World J. Gastroenterol.* **2008**, *14(11)*, 1741-1748.

28. Friend, S. H.; Bernards, R.; Rogelj, S.; Weinberg, R. A.; Rapaport, J. M.; Albert, D. M.; et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* **1986**, *323(6089)*, 643-646.

29. Burkhart, D. L.; Sage, J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat. Rev. Cancer* **2008**, *8(9)*, 671-682.

30. Stern, H. M.; Zon, L. I. Cancer genetics and drug discovery in the zebrafish. *Nat. Rev. Cancer* **2003**, *3(7)*, 533-539.

31. OSPAR, 2010. Quality Status Report 2010. OSPAR Commission. London 176pp.

32. Easey, M. W.; Millner, R. S. Improved methods for the preparation and staining of thin sections of fish otoliths for age determination. *Sci. Ser. Tech. Rep., Cefas Lowestoft*, **2008**, *143*, 1-12.

33. Jones, B. R.; Laslett, R. E. Methods for analysis for trace metals in marine and other samples. *Aquat. Environ. Prot: Analyt. Meth., MAFF Direct. Fish. Res. Lowestoft*, **1994**, *11*, 1-29.

34. De Boer, J.; Allchin, C.; Law, R.; Zegers, B.; Boon, J. P. Method for the analysis of polybrominated diphenylethers in sediments and biota. *TrAC* **2001**, *20(10)*, 591-599.

35. Allchin, C. R.; Kelly, C. A.; Portmann, J. E. Methods of analysis for chlorinated hydrocarbons in marine and other samples. *Aquat. Environ. Prot.: Analyt. Meth. MAFF Direct. Fish. Res*. Lowestoft **1989**, *6*, 1-25.

36. Stagg, R.; McIntosh, A. Biological effects of contaminants: determination of CYP1A-dependent monooxygenase activity in dab by fluorimetric measurement of EROD activity. ICES

Techniques in Marine Environmental Sciences 23. **1998**, International Council for the Exploration of the Sea, Copenhagen, Denmark. 16 pp.

37. Ariese, F.; Beyer, J.; Jonsson, G.; Visa, C. P.; Krahn, M. M. Review of analytical methods for determining metabolites of polycyclic aromatic compounds (PACs) in fish bile. ICES Techniques in Marine Environmental Sciences 39. International Council for the Exploration of the Sea, **2005**, Copenhagen, Denmark, 41 pp.
38. Bucke, D.; Vethaak, A. D.; Lang, T.; Mellergaard, S. Common diseases and parasites of fish in the North Atlantic: training guide for identification. ICES Techniques in Marine Environmental Sciences 19, **1996**, ICES, Copenhagen, Denmark.
39. Wang, B.; Hikosaka, K.; Sultana, N.; Sharkar, M. T. K.; Noritake, H.; Kimura, W.; Wu, Y. X.; Kobayashi, Y.; Uezato, T.; Miura, N. Liver tumour formation by a mutant retinoblastoma protein in transgenic mice is caused by up regulation of c-Myc target genes. *Biochem. Biophys. Res. Comm.* **2012**, *417*, 601-606.
40. Lee, J. O.; Russo, A. A.; Pavletich, N. P. Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7. *Nature* **1998**, *391*, 859-865.
41. Riley, D. J.; Liu, C. Y.; Lee, W. H. Mutations of N-terminal regions render the retinoblastoma protein insufficient for functions in development and tumor suppression. *Mol. Cell. Biol.* **1997**, *17(12)*, 7342-7352.
42. Vethaak, A. D.; Jol, J. G.; Pieters, J. P. F. Long-term trends in the prevalence of cancer and other major diseases among flatfish in the southeastern North Sea as indicators of changing ecosystem health. E*nviron. Sci. Technol.* **2009**, *43(6)*, 2151-2158.

43. Achanzar, W. E.; Webber, M. M.; Waalkes, M. P. Altered apoptotic gene expression and acquired apoptotic resistance in cadmium-transformed human prostate epithelial cells. *Prostate* **2002**, *52(3)*, 236-244.

44. Waisberg, M.; Joseph, P.; Hale, B.; Beyersmann, D. Molecular and cellular mechanisms of cadmium carcinogenesis. *Toxicol* **2003**, *192(2-3)*, 95-117.

45. Joseph, P. Mechanisms of cadmium carcinogenesis. *Toxicol. Appl. Pharmacol*. **2009**, *238(3)*, 272-279.

46. Giaginis, C.; Gatzidou, E.; Theocharis, S. DNA repair systems as targets of cadmium toxicity. *Toxicol. Appl. Pharmacol.* **2006**, *213(3)*, 282-290.

47. Mukherjee, J. J.; Gupta, S. K.; Kumar, S. Inhibition of benzopyrene diol epoxide-induced apoptosis by cadmium(II) is AP-1-independent: role of extracelluler signal related kinase. *Chem. Biol. Interact.* **2008**, *172(1)*, 72-80.

48. Wills, L. P.; Jung, D.; Koehrn, K.; Zhu, S.; Willett, K. L.; Hinton, D. E.; et al. Comparative chronic liver toxicity of benzo[a]pyrene in two populations of the Atlantic killifish (*Fundulus heteroclitus*) with different exposure histories. *Environ. Health Perspect.* **2010**, *118(10)*, 1376-1381.

49. Lerebours, A.; Bignell, J. P.; Stentiford, G. D.; Feist, S. W.; Lyons, B. P.; Rotchell, J. M. Advanced diagnostics applied to fish liver tumours: relating pathology to underlying molecular aetiology. *Mar. Poll. Bull*. **2013**, *72*, 94-98.

50. Robertson, K. D.; Ait-Si-Ali, S.; Yokochi, T.; Wade, P. A.; Jones, P. L.; Wolffe, A. P. DNMT1 forms a complex with RB, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat. Genet.* **2000**, *25(3)*, 338-342.

51. Hussain, S. P.; Harris, C. C. Molecular epidemiology of human cancer. *Toxicol. Lett.* **1998**, *102-103*, 219-225.

52. Loeb, K. R.; Loeb, L. A. Significance of multiple mutations in cancer. *Carcinogenesis* **2000**, *21(3)*, 379-385.

53. Di Fiore, R.; D'Anneo, A.; Tesoriere, G.; Vento, R. RB1 in cancer: Different mechanisms of RB1 inactivation and alterations of pRb pathway in tumorigenesis. *J. Cell. Physiol*. **2013**, *228(8)*, 1676-1687.

54. Kranz, H.; Dethlefsen, V. Liver anomalies in dab (*Limanda limanda*) from the southern North Sea with special consideration given to neoplastic lesions. *Dis. Aquat. Org*. **1990**, *9*, 171-185.

Figure and Table Legends

Figure 1. Principal component analysis showing the association between concentrations of chemicals in liver of fish and sampling site (*n* = 30 pools of 5 fish). Axis1 represents 60% of variance. Axis2 represents 17% of variance.

Figure 2. Correspondence analysis showing the distribution of phenotypes (normal, pre-tumour, tumour) across North Sea/English Channel sampling sites (*n* = 165). Axis1 represents 95% of variance. Axis2 represents 5% of variance.

Figure 3. Correspondence analysis showing the distribution of *Rb* genotypes across North Sea/English Channel sampling sites (*n* = 165). Axis1 represents 38% of variance. Axis2 represents 29% of variance.

Figure 4. Correspondence analysis showing the association between *Rb* genotypes and liver histopathological phenotypes (*n* = 165 fish). Axis1 represents 60% of variance. Axis2 represents 40% of variance.

Table 1. Spectrum of *Rb* genetic profiles identified in a North Sea/English Channel dab population from differing localities (*n* = 165 individual fish).

Figure 1

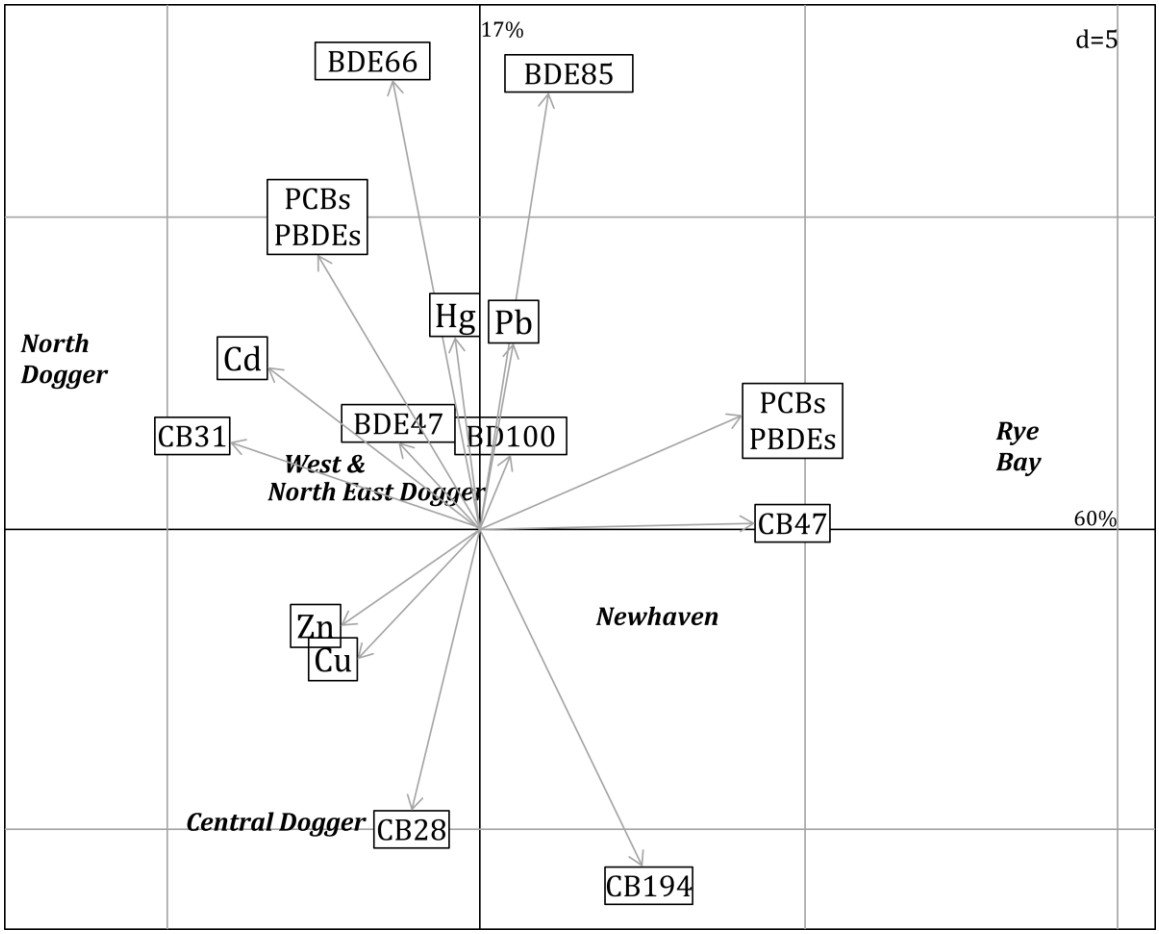

17%
d=5
BDE66
BDE85
PCBs
PBDEs
Hg
Pb
North
Dogger
Cd
CB31
BDE47
BD100
West &
North East Dogger
PCBs
PBDEs
Rye
Bay
CB47
60%
Zn
Cu
Newhaven
Central Dogger
CB28
CB194

Figure 2


5.2
%
d=0.2
tumour
North East Dogger
West Dogger
normal
Newhaven
North Dogger
94.8 %
pretumour
Rye Bay
Central Dogger

Figure 3

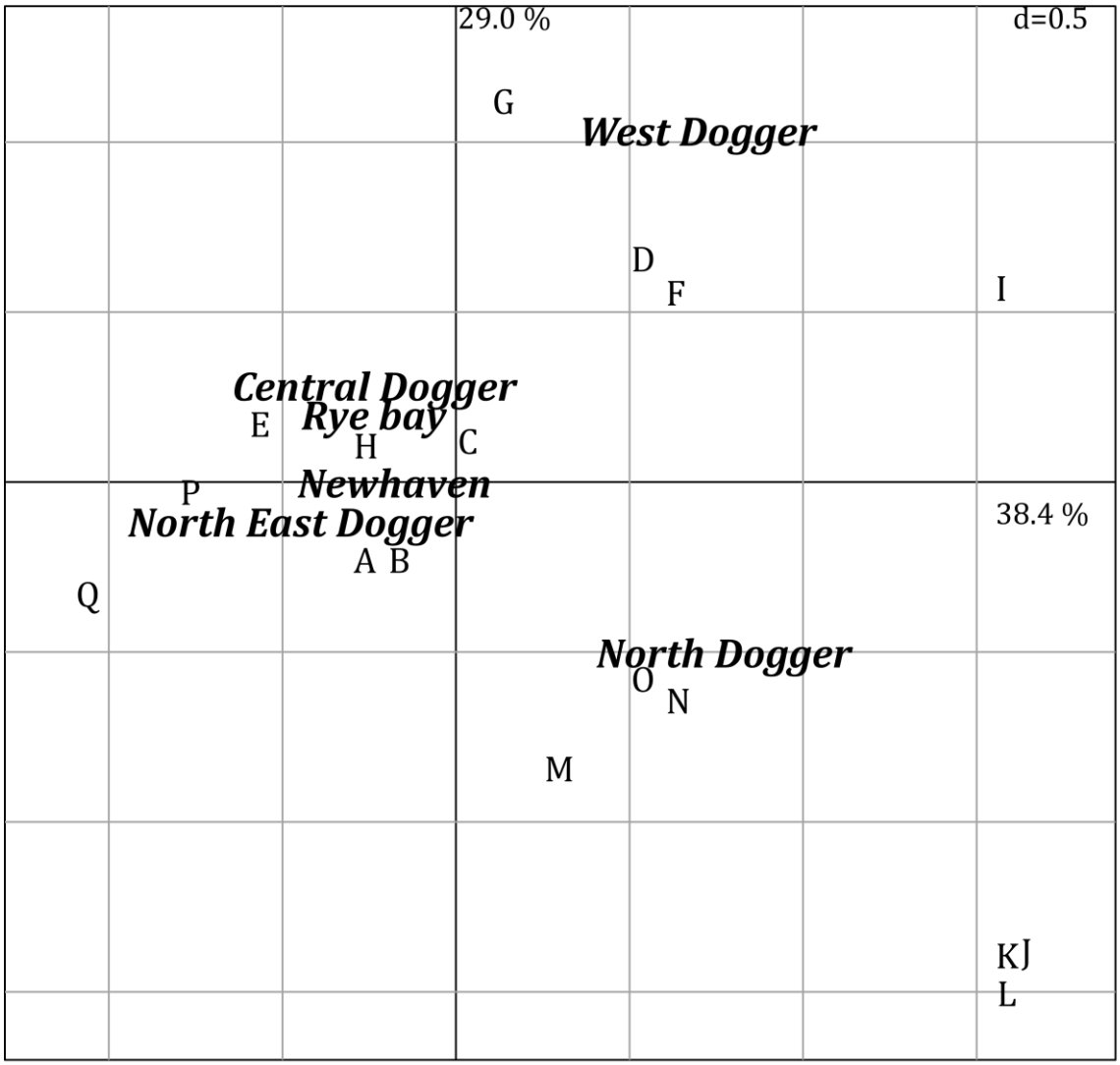

29.0 %
d=0.5
G
West Dogger
D
F
I
Central Dogger
Rye bay
E
H
C
Newhaven
P
North East Dogger
38.4 %
A B
Q
North Dogger
O
N
M
KJ
L

Figure 4

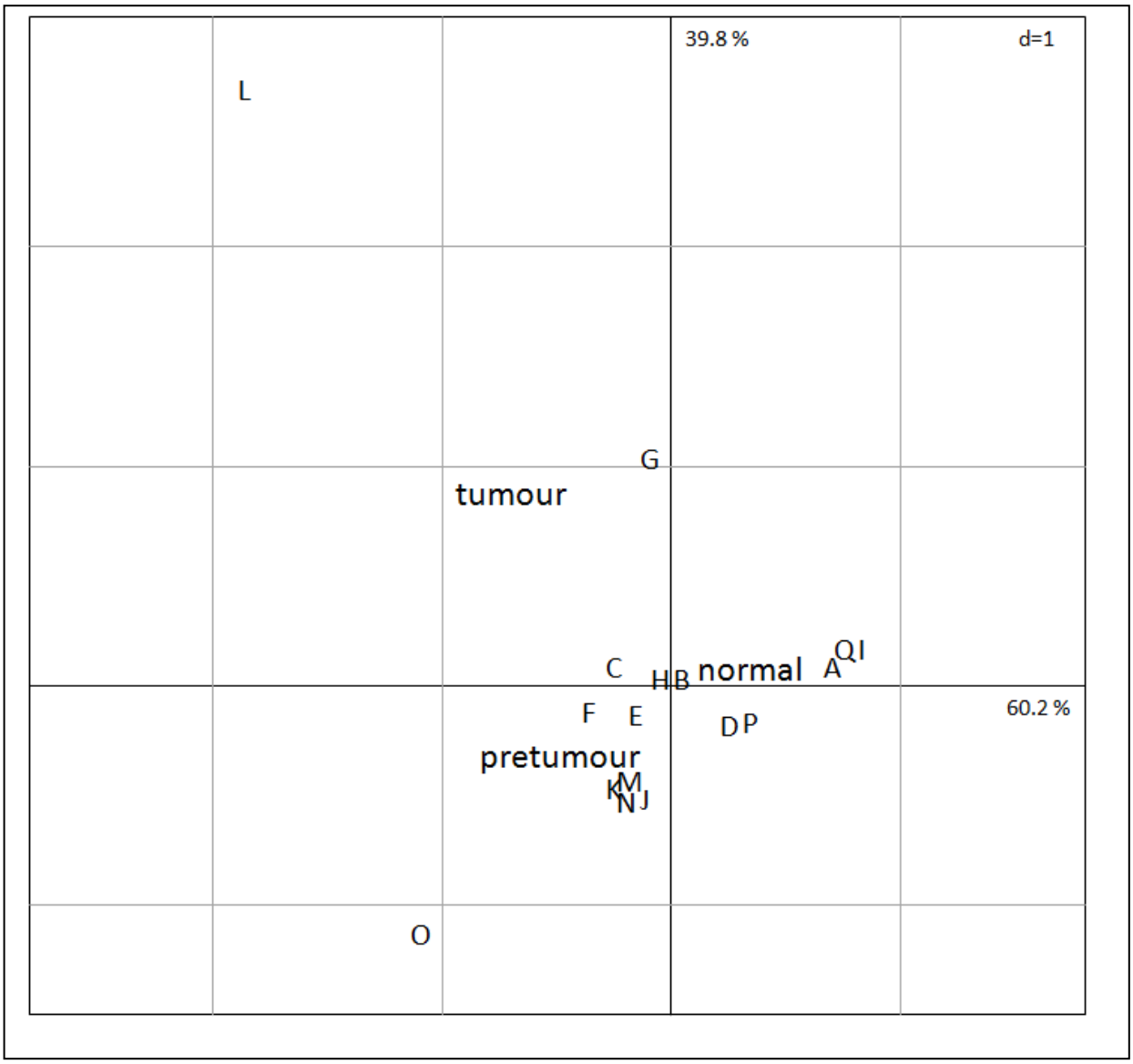


Table 1

Profile name	*Rb* cDNA genetic changes			
	996 bp	1088 bp	1514 bp	1592 bp
A	G	T	G	G
B	G	T/C	G/T	G/T
C	G/A	T/C	G/T	G/T
D	G/A	T	G	G
E	G	C	T	T
F	G/A	C	T	T
G	G/A	T/C	G	G
H	G	T/C	G	G
I	G/A	T	G/T	G/T
J	G/A	C	G/T	T
K	G	C	G/T	T
L	G/A	C	G/T	G/T
M	G	C	G	G
N	G	T/C	G	G/T
O	G	T/C	T	T
P	G	C	G/T	G/T
Q	G	T	G	G/T

Supplemental Information Cover Sheet

Authors: Adélaïde Lerebours, Grant D. Stentiford, Brett P. Lyons, John P. Bignell, Stéphane A. P. Derocles and Jeanette M. Rotchell.

Manuscript Title: Genetic alterations and cancer formation in a European flatfish at sites of different contaminant burdens.

Summary: 21 pages: 1 Methods detail document, 2 Figures, and 6 Tables

Supplemental Information List of Figures and Tables

Figure S1. Prevalence (%) of normal, pretumour and tumour liver phenotypes of fish, collected during this study, at each sampling site.

Figure S2. The main *Rb* allele genetic profiles (%) found in each fish liver tissues collected at West Dogger (n = 16) North Dogger (n = 11), North East Dogger (n = 41), Central Dogger (n = 11), Newhaven (n = 37) and Rye Bay (n = 38).

Table S1. Geographic coordinates of the CSEMP sampling sites.

Table S2. Biometric data for body length (L), body weight (W_B), Fulton Index Condition (FIC) (FIC = W_B/L^3 x 100), liver weight (W_L), Hepatosomatic Index (HSI) (HSI = W_L/W_B x 100), age and sex ratio (m/f). Data are expressed as mean ± SD, n = 165 fish, (with the exception of n = 119 fish for liver biometry). The significance of body length, body weight and age at the different sites was assessed using the contrast method (Hastie & Pregibon 1992) with the Esticon function in R (DoBy package, Hojsgaard 2004). Values with different letters are significantly different (GLIM, contrats method, P-

641 value < 0.05). No significant differences between sites were found for HSI and FIC. No model was
642 validated for the liver weight. NA, not available.

643 **Table S3.** Oligonucleotide sequences of primer pairs used to amplify and sequence the *Rb* cDNA
644 region between 620 and 1942 bp. [a]Forward primer. [b]Reverse primer.

645 **Table S4a.** Concentrations of Zn (mg/kg), Cu (µg/kg), Pb (µg/kg) and Cd (µg/kg) measured in liver of
646 fish and Hg (µg/kg) measured in the flesh of fish, *n* = 5 pools of 5 fish, and expressed as mean,
647 standard deviation (SD), and standard error of the mean (SEM).

648 **Table S4b**. Concentrations of PBDEs (µg/kg) measured in liver of fish (*n* = 5 pools of 5 fish, mean,
649 SD, SEM).

650 **Table S4c.** Concentrations of PCBs (µg/kg) measured in liver of fish (*n* = 5 pools of 5 fish, mean, SD,
651 SEM).

652 **Table S5**. Matrix of correlation showing the Pearson's correlation coefficients for metals, PBDE and
653 PCB congeners using mean concentrations. The matrix was conducted prior to the PCA in order to
654 identify any correlations between chemicals. For instance, the coefficients show that CB153 is highly
655 correlated to CB101, CB105, CB110, CB118, CB128, CB138, CB141, CB149 and CB151 (shown as
656 shaded in the Table).

657 **Table S6.** Levels of hydroxypyrene (ng/g, mean ± SD) measured in the bile of fish sampled from Rye
658 Bay (*n* = 15), Newhaven (*n* = 13), North East Dogger (*n* = 19), North Dogger (*n* = 17), West Dogger
659 (*n* = 15) and Central Dogger (*n* = 15) and EROD (pmol/min/mg protein, mean ± SD) activities in the
660 liver of fish from Rye Bay (*n* = 20), Newhaven (*n* = 20), North East Dogger (*n* = 20), North Dogger (*n*
661 =19), West Dogger (*n* = 20) and Central Dogger (*n* = 20). The significance of EROD and PYR1OH at
662 the different sites was assessed using the contrast method (Hastie & Pregibon 1992) with the Esticon

663 function in R (DoBy package, Hojsgaard 2004). Values with different letters are significantly different
664 (GLIM, contrasts method, *P*-value < 0.05).
665
666 **Table Statistics References**
667 Hastie, T. J.; Pregibon, D. (1992) *Generalized linear models.* In: *Statistical Models in S*. Chambers, J.
668 M., Hastie, T. J. Eds.; Wadsworth & Brooks/Cole Publishing Inc. Belmont, USA;1992.
669 Hojsgaard, S. Statistical inference in context specific interaction models for contingency tables.
670 *Scandinavian Journal of Statistics*, **2004**, *31*, 143– 158.
671

Supplemental Information

SM1. Methods detail for GLIM statistical model approach

The effect of the site, genotype, sex and age of fish on the presence (pre-tumour and tumour) or absence (normal) of tumour was tested using generalized linear models (GLIM). All of these factors were included in the model. Statistical analyses were performed by presence–absence of tumour for each fish captured using GLIM (binomial family, link logit), with the anova.glm function in R. Statistical analyses were performed using GLIM (Poisson family, link log), with the anova.glm function in R. The best-fit model was selected using Akaike information criterion (AIC) and the normality of residuals has been tested for all models. Statistical analyses were performed using R 3.0 (R Development Core Team 2013). The model estimates data is included below.

GLIM1: Effect of the site, genotype, sex and age of fish on the presence (pre-tumour and tumour) - absence (normal) of tumour.

```
> glm1=glm(pheno~site+geno*age+sex, data=myd, family=binomial)

Analysis of Deviance Table
Model: binomial, link: logit
Response: pheno
Terms added sequentially (first to last)
```

Table showing model estimates for GLIM1.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			133	178.99		
site	4	14.3605	129	164.63	0.0062292	**
geno	16	28.4180	113	136.21	0.0281688	*
age	1	11.5611	112	124.65	0.0006734	***
sex	1	3.6345	111	121.02	0.0565942	.
geno:age	10	8.3497	101	112.67	0.5947242	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

GLIM2: Effect of site, phenotype and sex on the age of the fish.

```
> glm2=glm(age~site*sex+pheno, data=myd, family=poisson)
Analysis of Deviance Table
Model: poisson, link: log
Response: age
Terms added sequentially (first to last)
```

Table showing model estimates for GLIM2.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			**133**	**71.693**		
site	**4**	**33.847**	**129**	**37.846**	**8.011e-07**	*******
sex	**1**	**1.176**	**128**	**36.671**	**0.2783**	
pheno	**1**	**3.157**	**127**	**33.514**	**0.0756**	**.**
site:sex	**3**	**1.333**	**124**	**32.181**	**0.7213**	

Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**

Figure S1. Prevalence (%) of normal, pretumour and tumour liver phenotypes of fish, collected during this study, at each sampling site.

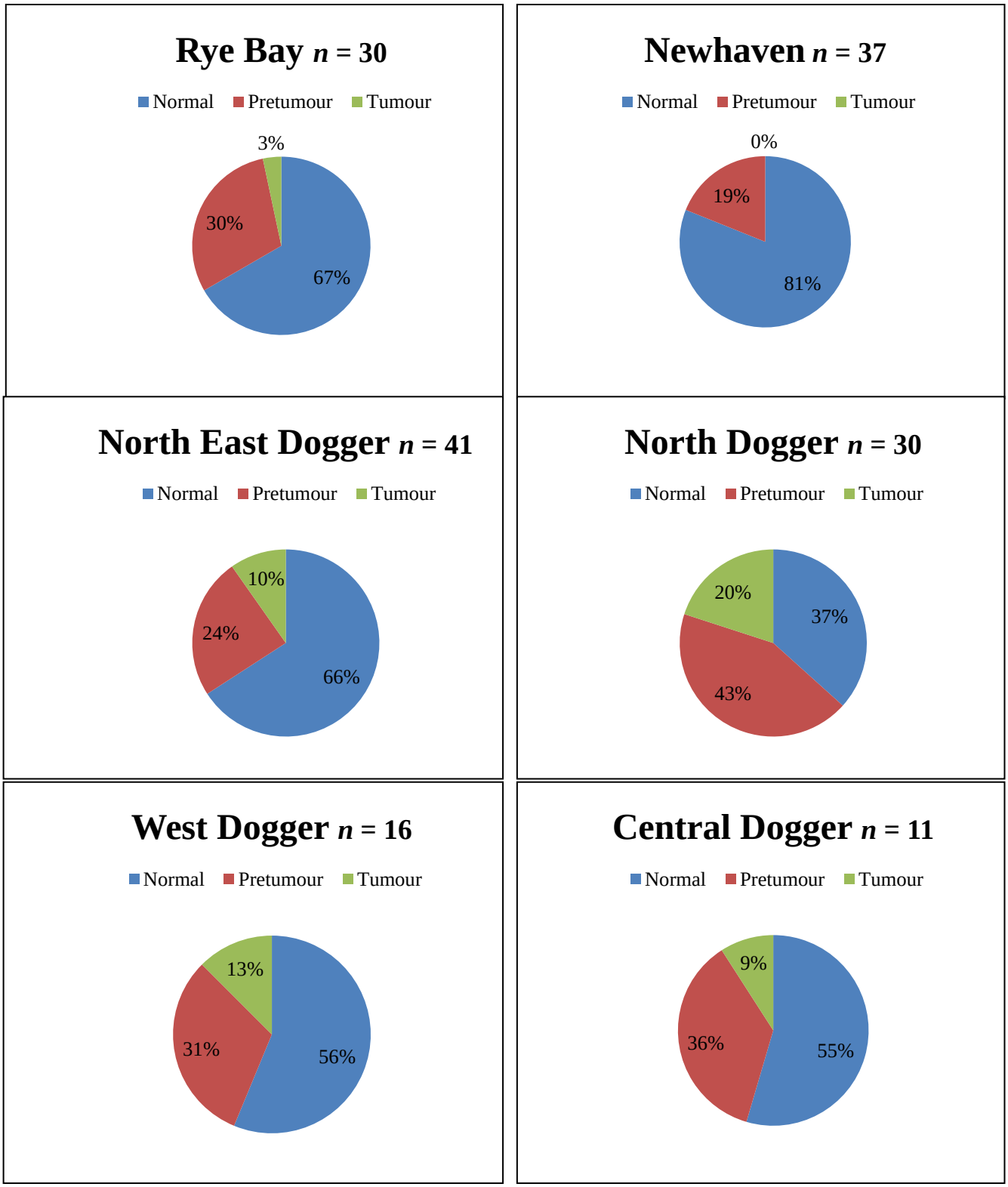


Figure S2. The main *Rb* allele genetic profiles (%) found in each fish liver tissues collected at West Dogger (n = 16) North Dogger (n = 11), North East Dogger (n = 41), Central Dogger (n = 11), Newhaven (n = 37) and Rye Bay (n = 38).

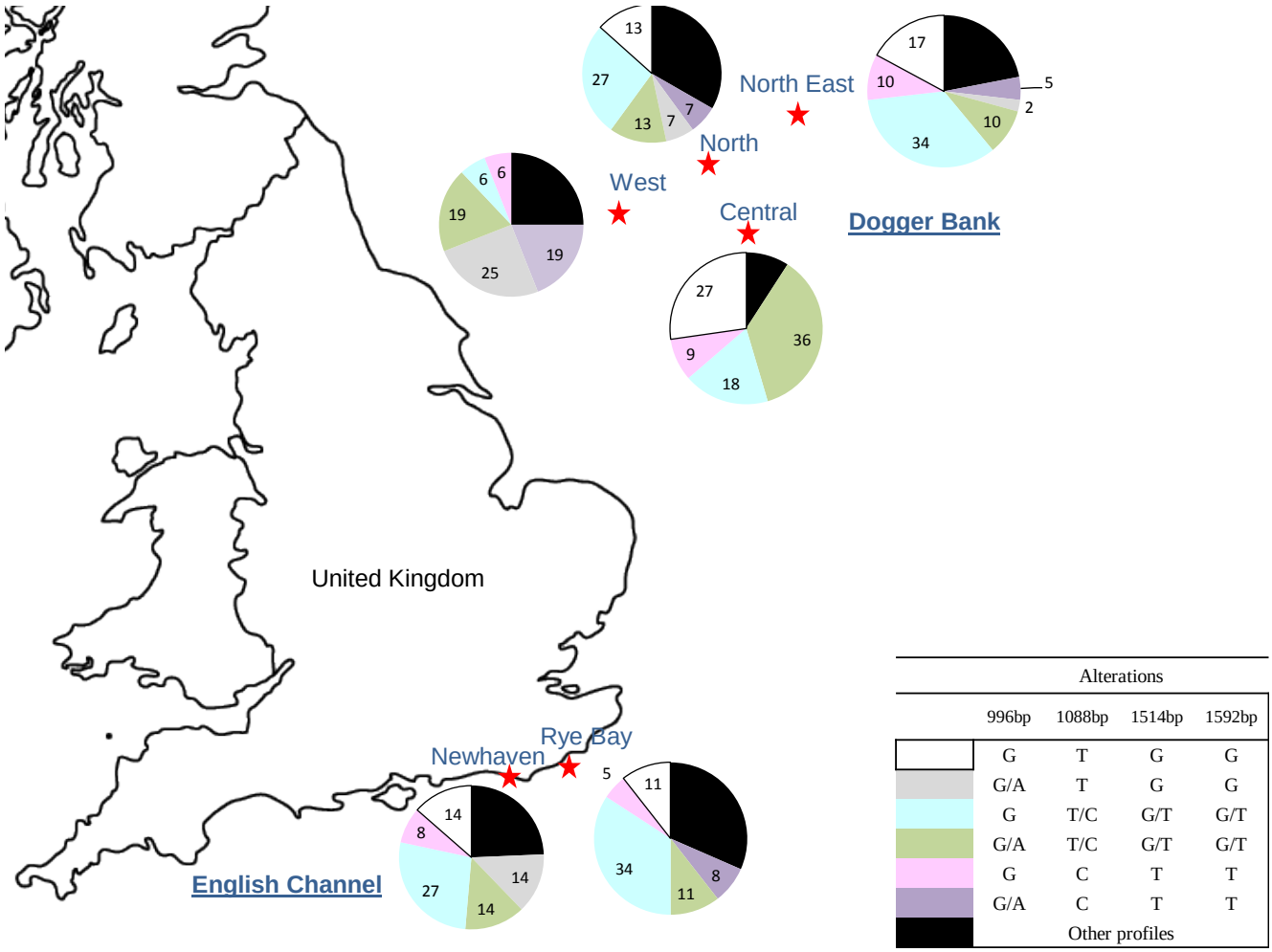


Table S1. Geographic coordinates of the CSEMP sampling sites.

Site	Latitude, Longitude
Newhaven	50° 45.59’ N, 00° 00.00’E
Rye Bay	50° 46.74’ N, 00° 46.83’E
Central Dogger Bank	54° 30.00’ N, 02° 42.53’E
North Dogger Bank	55° 04.08’ N, 02° 05.40’E
North East Dogger Bank	55° 18.05’ N, 02° 53.82’E
West Dogger Bank	54° 46.76’ N, 01° 17.69’E

Table S2. Biometric data for body length (L), body weight (W_B), Fulton Index Condition (FIC) (FIC = W_B/L^3 x 100), liver weight (W_L), Hepatosomatic Index (HSI) (HSI = W_L/W_B x 100), age and sex ratio (m/f). Data are expressed as mean ± SD, *n* = 165 fish, (with the exception of *n* = 119 fish for liver biometry). The significance of body length, body weight and age at the different sites was assessed using the contrast method (Hastie & Pregibon 1992) with the Esticon function in R (DoBy package, Hojsgaard 2004). Values with different letters are significantly different (GLIM, contrats method, P-value < 0.05). No significant differences between sites were found for HSI and FIC. No model was validated for the liver weight. NA, not available.

	Rye Bay	Newhaven	North East Dogger	North Dogger	West Dogger	Central Dogger
Body length (cm)	23.2 ± 2.3[a]	22.7 ± 1.6 [a]	23.2 ± 3.2[a]	27.1 ± 2.9[b]	23.6 ± 3.1[a]	20.8 ± 3.6[c]
Body weight (g)	135 ± 44 [a]	127 ± 30 [a]	124 ± 56 [a]	193 ± 58[b]	121 ± 45 [a]	91 ± 39[c]
Fulton Index Condition	1 ± 0.1	1 ± 0.1	0.9 ± 0.1	0.9 ± 0.1	0.9 ± 0.1	0.9 ± 0.1
Liver weight (g)	1.6 ± 0.5	1.9 ± 0.6	1.2 ± 0.6	1.1 ± 0.5	1 ± 0.4	1.1 ± 0.5
HSI	1.5 ± 0.3	1.7 ± 0.6	1.2 ± 0.5	1.1 ± 0.4	1.1 ± 0.3	1.1 ± 0.3
Age (year)	NA	3.2 ± 1[a]	5.3 ± 1.1[b]	6 ± 1.5 [b]	5.5 ± 1[b]	4.5 ± 1.6[b]
Sex ratio m/f	0.58	0.68	0.46	0.07	0.00	0.22

Table S2. Supporting statistics information

GLIM 3: Effects of site, sex and age on the length (l) of fish.

```
> glm3=glm(l~site+sex*age, data=myd)

Analysis of Deviance Table
Model: gaussian, link: identity
Response: l
```

Table showing the model estimates for GLIM 3.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			**133**	**151924**		
site	**4**	**50002**	**129**	**101922**	**< 2.2e-16**	*******
sex	**1**	**13605**	**128**	**88317**	**6.407e-09**	*******
age	**1**	**31772**	**127**	**56545**	**< 2.2e-16**	*******

```
sex:age  1     5686       126      50858 0.0001745 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

GLIM 4: Effects of site, sex and age on the weight (w) of fish.

```
> glm4=glm.nb(w~site+sex*age, data=myd)

Analysis of Deviance Table
Model: Negative Binomial (15.8388), link: log
Response: w
```

Table showing the model estimates for GLIM 4.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			133	330.99	
site	4	92.081	129	238.91	< 2.2e-16 ***
sex	1	27.452	128	211.46	1.610e-07 ***
age	1	62.628	127	148.83	2.496e-15 ***
sex:age	1	12.495	126	136.33	0.000408 ***

GLIM 5: Effects of site and sex on the hepatosomatic index (hsi).

```
> glm5=glm(hsi~site+sex, data=myd, family=poisson)

Analysis of Deviance Table
Model: poisson, link: log
Response: hsi
```

Table showing the model estimates for GLIM 5

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			118	19.640	
site	5	5.2988	113	14.341	0.3805
sex	1	0.1181	112	14.223	0.7311

GLIM 6: Effects of site, sex and age on Fulton Index Condition (fulton)

```
> glm6=glm(fulton~site+sex+age, data=myd, family=poisson)

Analysis of Deviance Table
Model: poisson, link: log
```

Response: fulton

Table showing the model estimates for GLIM 6.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			**133**	**0.020116**	
site	**4**	**0.0064169**	**129**	**0.013699**	**1.0000**
sex	**1**	**0.0000384**	**128**	**0.013661**	**0.9951**
age	**1**	**0.0000051**	**127**	**0.013655**	**0.9982**

851 **Table S3.** Oligonucleotide sequences of primer pairs used to amplify and sequence the *Rb* cDNA
852 region between 620 and 1942 bp. [a]Forward primer. [b]Reverse primer.

853

854

Region of *Rb* cDNA amplified	Primer sequence (5'-3')
RbA1: 620 bp – 1070 bp	AATCAGAGCTGCCATGACCT[a]
	CCAGGGGAAACAAACATCTG[b]
RbA2: 983 bp – 1437 bp	GGCAGCATATGGAGAGAGCGG[a]
	GAGCAGGCGGCTGGGTTGG[b]
RbB: 1350 bp – 1942 bp	CGTCCGGGCCATCGTGTCTT[a]
	ACGTTGTTGCTGCCAGGCACA[b]

855

856

Table S4a. Concentrations of Zn (mg/kg), Cu (µg/kg), Pb (µg/kg) and Cd (µg/kg) measured in liver of fish and Hg (µg/kg) measured in the flesh of fish, *n* = 5 pools of 5 fish, and expressed as mean, standard deviation (SD), and standard error of the mean (SEM).

		Rye Bay	Newhaven	North East Dogger	North Dogger	West Dogger	Central Dogger
Zn	mean	21.2	18	20.2	23.8	22.8	25.4
	sd	2.8	2.3	1.9	3.3	1.8	1.1
	sem	1.2	1.0	0.9	1.5	0.8	0.5
Cu	mean	3480	3600	3120	3980	3940	4040
	sd	1062	735	390	2217	1081	631
	sem	475	329	174	992	483	282
Hg	mean	80	38	66	72	104	68
	sd	21.2	8.4	11.4	20.5	25.1	8.4
	sem	9.5	3.7	5.1	9.2	11.2	3.7
Pb	mean	112	18	102	40	252	34
	sd	173	8.4	57.6	0	168	8.9
	sem	77.4	3.7	25.8	0	75.3	4
Cd	mean	122	48	388	406	382	276
	sd	41.5	17.9	84.7	112	89.8	68.8
	sem	18.5	8	37.9	49.9	40.2	30.8

Table S4b. Concentrations of PBDEs (µg/kg) measured in liver of fish (*n* = 5 pools of 5 fish, mean, SD, SEM).

		Rye Bay	Newhaven	North East Dogger	North Dogger	West Dogger	Central Dogger
BD100	mean	0.59	0.82	0.99	0.52	0.87	0.52
	sd	0.12	0.07	0.52	0.17	0.23	0.17
	sem	0.05	0.03	0.23	0.08	0.10	0.08
BD138	mean	0.33	0.31	0.38	0.50	0.40	0.31
	sd	0.06	0.08	0.07	0.16	0.11	0.04
	sem	0.03	0.04	0.03	0.07	0.05	0.02
BD153	mean	0.33	0.32	0.38	0.48	0.41	0.32
	sd	0.05	0.10	0.07	0.16	0.10	0.06
	sem	0.02	0.06	0.03	0.07	0.05	0.02
BD154	mean	0.39	0.41	0.46	0.59	0.49	0.38
	sd	0.07	0.08	0.10	0.19	0.14	0.06
	sem	0.03	0.04	0.04	0.08	0.06	0.03
BD183	mean	0.26	0.26	0.32	0.41	0.33	0.27
	sd	0.05	0.07	0.06	0.14	0.09	0.04
	sem	0.03	0.03	0.03	0.06	0.04	0.02
BDE28	mean	0.16	0.16	0.19	0.25	0.20	0.15
	sd	0.04	0.05	0.04	0.08	0.05	0.03
	sem	0.02	0.03	0.02	0.04	0.02	0.01
BDE47	mean	2.52	3.18	4.04	2.76	4.40	2.70
	sd	0.44	0.25	2.47	1.20	1.18	0.58
	sem	0.20	0.13	1.10	0.54	0.53	0.26
BDE66	mean	0.25	0.24	0.29	0.37	0.31	0.24
	sd	0.05	0.06	0.04	0.12	0.08	0.04
	sem	0.02	0.03	0.02	0.06	0.04	0.02
BDE85	mean	0.15	0.20	0.17	0.23	0.18	0.15
	sd	0.03	0.10	0.03	0.08	0.06	0.02
	sem	0.01	0.05	0.01	0.03	0.02	0.01
BDE99	mean	0.22	0.23	0.26	0.33	0.28	0.22
	sd	0.03	0.05	0.04	0.11	0.06	0.03
	sem	0.01	0.02	0.02	0.05	0.03	0.01

869 **Table S4c.** Concentrations of PCBs (µg/kg) measured in liver of fish (*n* = 5 pools of 5 fish, mean, SD,

870 SEM).

		Rye Bay	Newhaven	North East Dogger	North Dogger	West Dogger	Central Dogger
	mean	6.66	4.00	3.08	2.16	3.16	2.42
CB101	sd	1.13	1.62	1.25	0.86	0.62	0.64
	sem	0.50	0.81	0.56	0.39	0.28	0.29
	mean	2.66	1.48	1.25	0.80	1.10	0.91
CB105	sd	0.38	0.26	0.68	0.32	0.37	0.33
	sem	0.17	0.13	0.30	0.15	0.17	0.15
	mean	3.76	2.18	1.22	0.82	1.24	1.01
CB110	sd	0.67	0.91	0.44	0.26	0.28	0.24
	sem	0.30	0.46	0.20	0.12	0.13	0.11
	mean	9.92	5.40	4.70	2.96	4.12	3.34
CB118	sd	1.32	1.55	2.78	1.26	1.26	1.29
	sem	0.59	0.78	1.24	0.56	0.57	0.58
	mean	2.82	1.60	1.43	0.74	1.31	0.78
CB128	sd	0.51	0.46	0.86	0.34	0.75	0.48
	sem	0.23	0.23	0.38	0.15	0.33	0.22
	mean	17.20	9.55	8.98	5.14	9.32	5.80
CB138	sd	3.11	2.86	5.97	2.06	3.34	2.33
	sem	1.39	1.43	2.67	0.92	1.50	1.04
	mean	0.56	0.35	0.31	0.16	0.24	0.19
CB141	sd	0.13	0.09	0.17	0.08	0.05	0.03
	sem	0.06	0.04	0.08	0.03	0.02	0.02
	mean	3.92	2.58	1.37	0.91	1.54	1.06
CB149	sd	0.79	0.75	0.55	0.32	0.40	0.25
	sem	0.35	0.38	0.25	0.14	0.18	0.11
	mean	1.84	1.09	0.82	0.62	0.97	0.63
CB151	sd	0.51	0.50	0.43	0.21	0.32	0.18
	sem	0.23	0.25	0.19	0.10	0.14	0.08
	mean	22.60	11.30	11.14	6.16	11.78	7.28
CB153	sd	5.86	3.61	7.78	2.49	4.93	3.47
	sem	2.62	1.80	3.48	1.11	2.20	1.55
	mean	0.75	0.49	0.37	0.27	0.33	0.29
CB156	sd	0.13	0.17	0.15	0.09	0.10	0.02
	sem	0.06	0.08	0.07	0.04	0.04	0.01
	mean	2.08	1.14	0.73	0.39	0.77	0.64
CB170	sd	0.43	0.30	0.51	0.14	0.32	0.08
	sem	0.19	0.15	0.23	0.06	0.15	0.03
	mean	4.38	2.53	1.94	0.84	1.56	1.03
CB180	sd	1.27	0.62	1.60	0.38	0.95	0.65
	sem	0.57	0.31	0.71	0.17	0.43	0.29
	mean	7.14	4.00	2.92	1.88	3.68	1.90
CB187	sd	1.06	1.20	1.63	0.59	0.94	0.70
	sem	0.47	0.60	0.73	0.27	0.42	0.31
	mean	0.89	0.75	0.28	0.16	0.55	1.01
CB194	sd	0.13	0.17	0.26	0.06	0.26	0.09
	sem	0.06	0.09	0.11	0.02	0.12	0.04

871

		Rye Bay	Newhaven	North East Dogger	North Dogger	West Dogger	Central Dogger
CB28	mean	0.40	0.42	0.45	0.36	0.43	0.46
	sd	0.05	0.05	0.15	0.08	0.13	0.08
	sem	0.02	0.02	0.07	0.03	0.06	0.04
CB31	mean	0.23	0.27	0.34	0.36	0.30	0.31
	sd	0.01	0.04	0.08	0.10	0.05	0.02
	sem	0.01	0.02	0.03	0.04	0.02	0.01
CB44	mean	0.36	0.39	0.40	0.47	0.38	0.43
	sd	0.06	0.04	0.08	0.13	0.06	0.03
	sem	0.03	0.02	0.04	0.06	0.02	0.01
CB47	mean	0.60	0.44	0.31	0.25	0.31	0.28
	sd	0.12	0.21	0.10	0.07	0.14	0.07
	sem	0.05	0.11	0.04	0.03	0.06	0.03
CB49	mean	0.92	0.65	0.47	0.38	0.42	0.42
	sd	0.20	0.32	0.17	0.10	0.13	0.07
	sem	0.09	0.16	0.08	0.04	0.06	0.03
CB52	mean	1.34	1.01	0.95	0.75	0.81	0.83
	sd	0.23	0.42	0.34	0.19	0.28	0.17
	sem	0.10	0.21	0.15	0.08	0.13	0.08
CB66	mean	1.50	0.98	0.78	0.54	0.71	0.62
	sd	0.33	0.38	0.27	0.17	0.24	0.15
	sem	0.15	0.19	0.12	0.08	0.11	0.07

872

873

Table S5. Matrix of correlation showing the Pearson's correlation coefficients for metals, PBDE and PCB congeners using mean concentrations. The matrix was conducted prior to the PCA in order to identify any correlations between chemicals. For instance, the coefficients show that CB153 is highly correlated to CB101, CB105, CB110, CB118, CB128, CB138, CB141, CB149 and CB151 (shown as shaded in the Table).

	Zn	Cu	Hg	Pb	Cd	BD100	BD138	BD153	BD154	BD183	BDE28	BDE47
Zn	1.00	0.73	0.53	0.09	0.57	-0.61	0.28	0.28	0.28	0.35	0.35	-0.23
Cu	0.73	1.00	0.30	0.01	0.19	-0.62	0.17	0.17	0.17	0.39	0.39	-0.22
Hg	0.53	0.30	1.00	0.86	0.56	0.06	0.33	0.33	0.33	0.02	0.02	0.37
Pb	0.09	0.01	0.86	1.00	0.38	0.51	0.18	0.18	0.18	-0.30	-0.30	0.71
Cd	0.57	0.19	0.56	0.38	1.00	0.13	0.80	0.80	0.80	0.44	0.44	0.47
BD100	-0.61	-0.62	0.06	0.51	0.13	1.00	0.04	0.04	0.04	-0.50	-0.50	0.88
BD138	0.28	0.17	0.33	0.18	0.80	0.04	1.00	1.00	1.00	0.80	0.80	0.30
BD153	0.28	0.17	0.33	0.18	0.80	0.04	1.00	1.00	1.00	0.80	0.80	0.30
BD154	0.28	0.17	0.33	0.18	0.80	0.04	1.00	1.00	1.00	0.80	0.80	0.30
BD183	0.35	0.39	0.02	-0.30	0.44	-0.50	0.80	0.80	0.80	1.00	1.00	-0.30
BDE28	0.35	0.39	0.02	-0.30	0.44	-0.50	0.80	0.80	0.80	1.00	1.00	-0.30
BDE47	-0.23	-0.22	0.37	0.71	0.47	0.88	0.30	0.30	0.30	-0.30	-0.30	1.00
BDE66	0.23	0.02	0.46	0.25	0.61	-0.10	0.87	0.87	0.87	0.76	0.76	0.06
BDE85	-0.65	-0.47	0.08	0.33	-0.02	0.50	0.40	0.40	0.40	0.20	0.20	0.36
BDE99	0.15	-0.04	0.48	0.48	0.87	0.43	0.89	0.89	0.89	0.45	0.45	0.67
CB101	-0.46	-0.45	0.04	0.15	-0.68	0.01	-0.55	-0.55	-0.55	-0.42	-0.42	-0.30
CB105	-0.47	-0.52	0.00	0.10	-0.65	0.02	-0.54	-0.54	-0.54	-0.41	-0.41	-0.32
CB110	-0.48	-0.41	-0.09	0.02	-0.77	-0.07	-0.60	-0.60	-0.60	-0.39	-0.39	-0.40
CB118	-0.45	-0.51	0.02	0.11	-0.64	0.00	-0.53	-0.53	-0.53	-0.40	-0.40	-0.32
CB128	-0.53	-0.55	0.05	0.21	-0.63	0.14	-0.53	-0.53	-0.53	-0.48	-0.48	-0.18
CB138	-0.46	-0.51	0.15	0.30	-0.57	0.15	-0.50	-0.50	-0.50	-0.48	-0.48	-0.14
CB141	-0.58	-0.56	-0.22	-0.09	-0.75	-0.01	-0.56	-0.56	-0.56	-0.36	-0.36	-0.40
CB149	-0.55	-0.42	-0.12	0.04	-0.80	0.00	-0.62	-0.62	-0.62	-0.43	-0.43	-0.34
CB151	-0.46	-0.39	0.12	0.26	-0.65	0.07	-0.51	-0.51	-0.51	-0.42	-0.42	-0.20
CB153	-0.40	-0.48	0.20	0.32	-0.53	0.11	-0.49	-0.49	-0.49	-0.46	-0.46	-0.16
CB156	-0.50	-0.52	-0.14	-0.05	-0.71	-0.06	-0.54	-0.54	-0.54	-0.33	-0.33	-0.43
CB170	-0.42	-0.39	0.03	0.13	-0.70	-0.04	-0.60	-0.60	-0.60	-0.44	-0.44	-0.34
CB180	-0.55	-0.55	-0.04	0.12	-0.69	0.10	-0.57	-0.57	-0.57	-0.46	-0.46	-0.25
CB187	-0.46	-0.41	0.14	0.28	-0.62	0.09	-0.50	-0.50	-0.50	-0.43	-0.43	-0.18
CB194	0.13	0.21	-0.02	-0.05	-0.63	-0.33	-0.90	-0.90	-0.90	-0.64	-0.64	-0.41
CB28	0.26	-0.24	-0.16	-0.22	0.31	0.12	-0.16	-0.16	-0.16	-0.32	-0.32	0.08
CB31	0.31	0.44	-0.12	-0.26	0.59	-0.15	0.77	0.77	0.77	0.77	0.77	0.12
CB44	0.35	0.39	0.02	-0.30	0.44	-0.50	0.80	0.80	0.80	1.00	1.00	-0.30
CB47	-0.43	-0.41	-0.05	0.06	-0.76	-0.03	-0.71	-0.71	-0.71	-0.53	-0.53	-0.36
CB49	-0.61	-0.52	-0.28	-0.13	-0.82	-0.02	-0.59	-0.59	-0.59	-0.35	-0.35	-0.41
CB52	-0.56	-0.68	-0.16	-0.04	-0.63	0.07	-0.50	-0.50	-0.50	-0.37	-0.37	-0.33
CB66	-0.55	-0.53	-0.09	0.06	-0.73	0.06	-0.61	-0.61	-0.61	-0.47	-0.47	-0.29

881

	BDE66	BDE85	BDE99	CB101	CB105	CB110	CB118	CB128	CB138	CB141	CB149	CB151
Zn	0.23	-0.65	0.15	-0.46	-0.47	-0.48	-0.45	-0.53	-0.46	-0.58	-0.55	-0.46
Cu	0.02	-0.47	-0.04	-0.45	-0.52	-0.41	-0.51	-0.55	-0.51	-0.56	-0.42	-0.39
Hg	0.46	0.08	0.48	0.04	0.00	-0.09	0.02	0.05	0.15	-0.22	-0.12	0.12
Pb	0.25	0.33	0.48	0.15	0.10	0.02	0.11	0.21	0.30	-0.09	0.04	0.26
Cd	0.61	-0.02	0.87	-0.68	-0.65	-0.77	-0.64	-0.63	-0.57	-0.75	-0.80	-0.65
BD100	-0.10	0.50	0.43	0.01	0.02	-0.07	0.00	0.14	0.15	-0.01	0.00	0.07
BD138	0.87	0.40	0.89	-0.55	-0.54	-0.60	-0.53	-0.53	-0.50	-0.56	-0.62	-0.51
BD153	0.87	0.40	0.89	-0.55	-0.54	-0.60	-0.53	-0.53	-0.50	-0.56	-0.62	-0.51
BD154	0.87	0.40	0.89	-0.55	-0.54	-0.60	-0.53	-0.53	-0.50	-0.56	-0.62	-0.51
BD183	0.76	0.20	0.45	-0.42	-0.41	-0.39	-0.40	-0.48	-0.48	-0.36	-0.43	-0.42
BDE28	0.76	0.20	0.45	-0.42	-0.41	-0.39	-0.40	-0.48	-0.48	-0.36	-0.43	-0.42
BDE47	0.06	0.36	0.67	-0.30	-0.32	-0.40	-0.32	-0.18	-0.14	-0.40	-0.34	-0.20
BDE66	1.00	0.54	0.73	-0.10	-0.08	-0.16	-0.07	-0.09	-0.06	-0.14	-0.21	-0.07
BDE85	0.54	1.00	0.45	0.36	0.34	0.30	0.34	0.41	0.40	0.36	0.34	0.42
BDE99	0.73	0.45	1.00	-0.51	-0.50	-0.61	-0.49	-0.44	-0.39	-0.57	-0.61	-0.45
CB101	-0.10	0.36	-0.51	1.00	0.99	0.99	0.99	0.99	0.98	0.96	0.98	0.99
CB105	-0.08	0.34	-0.50	0.99	1.00	0.98	1.00	0.99	0.98	0.97	0.97	0.97
CB110	-0.16	0.30	-0.61	0.99	0.98	1.00	0.98	0.96	0.95	0.98	0.99	0.97
CB118	-0.07	0.34	-0.49	0.99	1.00	0.98	1.00	0.99	0.98	0.97	0.97	0.97
CB128	-0.09	0.41	-0.44	0.99	0.99	0.96	0.99	1.00	0.99	0.95	0.96	0.98
CB138	-0.06	0.40	-0.39	0.98	0.98	0.95	0.98	0.99	1.00	0.92	0.94	0.98
CB141	-0.14	0.36	-0.57	0.96	0.97	0.98	0.97	0.95	0.92	1.00	0.97	0.92
CB149	-0.21	0.34	-0.61	0.98	0.97	0.99	0.97	0.96	0.94	0.97	1.00	0.97
CB151	-0.07	0.42	-0.45	0.99	0.97	0.97	0.97	0.98	0.98	0.92	0.97	1.00
CB153	-0.03	0.37	-0.38	0.98	0.97	0.94	0.98	0.99	1.00	0.90	0.93	0.98
CB156	-0.09	0.33	-0.56	0.98	0.99	0.98	0.98	0.96	0.93	0.99	0.97	0.94
CB170	-0.15	0.28	-0.57	1.00	0.99	0.99	0.99	0.98	0.97	0.95	0.98	0.98
CB180	-0.13	0.39	-0.50	0.99	0.99	0.98	0.99	1.00	0.98	0.98	0.98	0.97
CB187	-0.06	0.43	-0.42	0.99	0.97	0.96	0.97	0.98	0.99	0.92	0.96	1.00
CB194	-0.74	-0.59	-0.86	0.46	0.43	0.50	0.43	0.40	0.40	0.38	0.49	0.44
CB28	-0.34	-0.63	0.00	-0.40	-0.32	-0.41	-0.33	-0.34	-0.35	-0.32	-0.44	-0.49
CB31	0.42	0.00	0.58	-0.86	-0.87	-0.83	-0.86	-0.88	-0.89	-0.79	-0.83	-0.85
CB44	0.76	0.20	0.45	-0.42	-0.41	-0.39	-0.40	-0.48	-0.48	-0.36	-0.43	-0.42
CB47	-0.29	0.18	-0.66	0.98	0.97	0.98	0.97	0.96	0.95	0.95	0.98	0.95
CB49	-0.19	0.35	-0.62	0.95	0.95	0.98	0.95	0.93	0.89	0.99	0.98	0.91
CB52	-0.07	0.37	-0.46	0.94	0.97	0.94	0.97	0.95	0.92	0.98	0.92	0.89
CB66	-0.18	0.34	-0.55	0.99	0.99	0.99	0.99	0.99	0.97	0.98	0.99	0.97

890

	CB153	CB156	CB170	CB180	CB187	CB194	CB28	CB31	CB44	CB47	CB49	CB52	CB66
Zn	-0.40	-0.50	-0.42	-0.55	-0.46	0.13	0.26	0.31	0.35	-0.43	-0.61	-0.56	-0.55
Cu	-0.48	-0.52	-0.39	-0.55	-0.41	0.21	-0.24	0.44	0.39	-0.41	-0.52	-0.68	-0.53
Hg	0.20	-0.14	0.03	-0.04	0.14	-0.02	-0.16	-0.12	0.02	-0.05	-0.28	-0.16	-0.09
Pb	0.32	-0.05	0.13	0.12	0.28	-0.05	-0.22	-0.26	-0.30	0.06	-0.13	-0.04	0.06
Cd	-0.53	-0.71	-0.70	-0.69	-0.62	-0.63	0.31	0.59	0.44	-0.76	-0.82	-0.63	-0.73
BD100	0.11	-0.06	-0.04	0.10	0.09	-0.33	0.12	-0.15	-0.50	-0.03	-0.02	0.07	0.06
BD138	-0.49	-0.54	-0.60	-0.57	-0.50	-0.90	-0.16	0.77	0.80	-0.71	-0.59	-0.50	-0.61
BD153	-0.49	-0.54	-0.60	-0.57	-0.50	-0.90	-0.16	0.77	0.80	-0.71	-0.59	-0.50	-0.61
BD154	-0.49	-0.54	-0.60	-0.57	-0.50	-0.90	-0.16	0.77	0.80	-0.71	-0.59	-0.50	-0.61
BD183	-0.46	-0.33	-0.44	-0.46	-0.43	-0.64	-0.32	0.77	1.00	-0.53	-0.35	-0.37	-0.47
BDE28	-0.46	-0.33	-0.44	-0.46	-0.43	-0.64	-0.32	0.77	1.00	-0.53	-0.35	-0.37	-0.47
BDE47	-0.16	-0.43	-0.34	-0.25	-0.18	-0.41	0.08	0.12	-0.30	-0.36	-0.41	-0.33	-0.29
BDE66	-0.03	-0.09	-0.15	-0.13	-0.06	-0.74	-0.34	0.42	0.76	-0.29	-0.19	-0.07	-0.18
BDE85	0.37	0.33	0.28	0.39	0.43	-0.59	-0.63	0.00	0.20	0.18	0.35	0.37	0.34
BDE99	-0.38	-0.56	-0.57	-0.50	-0.42	-0.86	0.00	0.58	0.45	-0.66	-0.62	-0.46	-0.55
CB101	0.98	0.98	1.00	0.99	0.99	0.46	-0.40	-0.86	-0.42	0.98	0.95	0.94	0.99
CB105	0.97	0.99	0.99	0.99	0.97	0.43	-0.32	-0.87	-0.41	0.97	0.95	0.97	0.99
CB110	0.94	0.98	0.99	0.98	0.96	0.50	-0.41	-0.83	-0.39	0.98	0.98	0.94	0.99
CB118	0.98	0.98	0.99	0.99	0.97	0.43	-0.33	-0.86	-0.40	0.97	0.95	0.97	0.99
CB128	0.99	0.96	0.98	1.00	0.98	0.40	-0.34	-0.88	-0.48	0.96	0.93	0.95	0.99
CB138	1.00	0.93	0.97	0.98	0.99	0.40	-0.35	-0.89	-0.48	0.95	0.89	0.92	0.97
CB141	0.90	0.99	0.95	0.98	0.92	0.38	-0.32	-0.79	-0.36	0.95	0.99	0.98	0.98
CB149	0.93	0.97	0.98	0.98	0.96	0.49	-0.44	-0.83	-0.43	0.98	0.98	0.92	0.99
CB151	0.98	0.94	0.98	0.97	1.00	0.44	-0.49	-0.85	-0.42	0.95	0.91	0.89	0.97
CB153	1.00	0.93	0.97	0.97	0.98	0.42	-0.34	-0.89	-0.46	0.94	0.87	0.91	0.96
CB156	0.93	1.00	0.97	0.98	0.93	0.40	-0.33	-0.80	-0.33	0.96	0.98	0.98	0.98
CB170	0.97	0.97	1.00	0.98	0.98	0.53	-0.38	-0.88	-0.44	0.99	0.94	0.92	0.98
CB180	0.97	0.98	0.98	1.00	0.98	0.42	-0.34	-0.87	-0.46	0.97	0.96	0.96	1.00
CB187	0.98	0.93	0.98	0.98	1.00	0.42	-0.47	-0.85	-0.43	0.95	0.91	0.89	0.96
CB194	0.42	0.40	0.53	0.42	0.42	1.00	0.08	-0.69	-0.64	0.61	0.41	0.30	0.46
CB28	-0.34	-0.33	-0.38	-0.34	-0.47	0.08	1.00	0.00	-0.32	-0.28	-0.37	-0.20	-0.32
CB31	-0.89	-0.80	-0.88	-0.87	-0.85	-0.69	0.00	1.00	0.77	-0.92	-0.76	-0.80	-0.87
CB44	-0.46	-0.33	-0.44	-0.46	-0.43	-0.64	-0.32	0.77	1.00	-0.53	-0.35	-0.37	-0.47
CB47	0.94	0.96	0.99	0.97	0.95	0.61	-0.28	-0.92	-0.53	1.00	0.94	0.92	0.98
CB49	0.87	0.98	0.94	0.96	0.91	0.41	-0.37	-0.76	-0.35	0.94	1.00	0.95	0.97
CB52	0.91	0.98	0.92	0.96	0.89	0.30	-0.20	-0.80	-0.37	0.92	0.95	1.00	0.97
CB66	0.96	0.98	0.98	1.00	0.96	0.46	-0.32	-0.87	-0.47	0.98	0.97	0.97	1.00

891

892

893

Table S6. Levels of hydroxypyrene (ng/g, mean ± SD) measured in the bile of fish sampled from Rye Bay (*n* = 15), Newhaven (*n* = 13), North East Dogger (*n* = 19), North Dogger (*n* = 17), West Dogger (*n* = 15) and Central Dogger (*n* = 15) and EROD (pmol/min/mg protein, mean ± SD) activities in the liver of fish from Rye Bay (*n* = 20), Newhaven (*n* = 20), North East Dogger (*n* = 20), North Dogger (*n* =19), West Dogger (*n* = 20) and Central Dogger (*n* = 20). The significance of EROD and PYR1OH at the different sites was assessed using the contrast method (Hastie & Pregibon 1992) with the Esticon function in R (DoBy package, Hojsgaard 2004). Values with different letters are significantly different (GLIM, contrasts method, *P*-value < 0.05).

	Rye Bay	Newhaven	North East Dogger	North Dogger	West Dogger	Central Dogger
PYR1OH	203 ± 63^{c}	231 ± 57^{c}	167 ± 67^{b}	124 ± 52^{a}	165 ± 53^{b}	166 ± 57^{b}
EROD	138 ± 122^{c}	39 ± 21^{a}	84 ± 74^{b}	83 ± 58^{b}	287 ± 213^{d}	61 ± 36^{ab}

Table S6. Supporting Statistics Method detail:

GLIM A: Effects of the site and sex on the PYR1OH levels.

```
> glmA=glm.nb(PYR1OH~SITE+SEX, data=myd)

Analysis of Deviance Table
Model: Negative Binomial(10.2374), link: log
Response: PYR1OH
```

Table showing model estimates for GLIM A.

```
     Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
NULL                    93    128.010
SITE  5   32.410        88     95.599 4.928e-06 ***
SEX   1    0.154        87     95.445    0.6946
---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

GLIM B: Effects of the site and sex on the EROD levels.

```
> glmB=glm.nb(EROD~SITE+SEX, data=myd)

Analysis of Deviance Table
Model: Negative Binomial(1.7933), link: log
Response: EROD
```

Table showing model estimates for GLIM B.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			**118**	**222.55**	
SITE	**5**	**90.332**	**113**	**132.22**	**<2e-16 *****
SEX	**1**	**0.803**	**112**	**131.42**	**0.3701**

Signif. codes: 0 ‘*’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1**