

UNIVERSITY OF BIRMINGHAM

Research at Birmingham

DNA methylation in liver tumorigenesis in fish from the environment

Mirbahai, Leda; Yin, G; Bignell, JP; Li, N; Williams, Timothy; Chipman, James

DOI:

[10.4161/epi.6.11.17890](https://doi.org/10.4161/epi.6.11.17890)

License:

None: All rights reserved

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Mirbahai, L, Yin, G, Bignell, JP, Li, N, Williams, T & Chipman, J 2011, 'DNA methylation in liver tumorigenesis in fish from the environment' *Epigenetics* : official journal of the DNA Methylation Society, vol 6, no. 11, pp. 1319-1333. DOI: 10.4161/epi.6.11.17890

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

DNA methylation in liver tumorigenesis in fish from the environment

Leda Mirbahai,^{1,*} Guangliang Yin,² John P. Bignell,³ Ning Li,² Timothy D. Williams^{1,†} and James K. Chipman^{1,†}

¹School of Biosciences; University of Birmingham, Edgbaston; Birmingham, UK; ²BGI-Shenzhen; Shenzhen, China; ³Cefas; Weymouth Laboratory; Weymouth, Dorset UK

[†]These authors share senior authorship.

Key words: DNA methylation, dab, fish, tumor, estrogen

Abbreviations: BSP, bisulfite sequencing PCR; CSEMP, clean seas environmental monitoring programme; DEFRA, department for environment food and rural affairs; EAC, ecotoxicological assessment criteria; EDC, endocrine disrupting chemicals; ER, estrogen receptor; ERE, estrogen response element; HCA, hepatocellular adenoma; HCC, hepatocellular carcinoma; HPLC, high performance liquid chromatography; HTS, high-throughput sequencing; ICES, international council for the exploration of the sea; IPA, ingenuity pathway analysis; PBDE, polybrominated diphenyl ether; PCA, principal components analysis; PCB, polychlorinated biphenyl; SOAP package, short oligonucleotide analysis package; ST, tissue surrounding HCA tumors; TSS, transcription start site

The link between environment, alteration in DNA methylation and cancer has been well established in humans; yet, it is under-studied in unsequenced non-model organisms. The occurrence of liver tumors in the flatfish dab collected at certain UK sampling sites exceeds 20%, yet the causative agents and the molecular mechanisms of tumor formation are not known, especially regarding the balance between epigenetic and genetic factors. Methylated DNA Immunoprecipitation (MeDIP) combined with de novo high-throughput DNA sequencing were used to investigate DNA methylation changes in dab hepatocellular adenoma tumors for the first time in an unsequenced species. Novel custom-made dab gene expression arrays were designed and used to determine the relationship between DNA methylation and gene expression. In addition, the confirmatory techniques of bisulfite sequencing PCR (BSP) and RT-PCR were applied. Genes involved in pathways related to cancer, including apoptosis, wnt/β-catenin signaling and genomic and non-genomic estrogen responses, were altered both in methylation and transcription. Global methylation was statistically significantly 1.8-fold reduced in hepatocellular adenoma and non-cancerous surrounding tissues compared with liver from non-cancer bearing dab. Based on the identified changes and chemical exposure data, our study supports the epigenetic model of cancer. We hypothesize that chronic exposure to a mixture of environmental contaminants contributes to a global hypomethylation followed by further epigenetic and genomic changes. The findings suggest a link between environment, epigenetics and cancer in fish tumors in the wild and show the utility of this methodology for studies in non-model organisms.

Introduction

Cancer is a complex and multifactorial disease.¹ Traditionally, cancer was considered as a multistep genetic disease driven and initiated by mutations.^{2,3} However, data collected in the past decade demonstrate that both epigenetic and genetic changes interact and complement each other to enable cancer progression.^{4,5}

DNA methylation at CpG dinucleotides is one of the most intensively studied epigenetic modifications.^{6,7} Methylation at the promoter regions of genes inhibits access of transcription factors to their binding sites, resulting in gene silencing.⁸ However, the relationship between DNA methylation, chromatin modification and gene expression is complex. Several factors such as the presence of enhancer and suppressor elements and their

locations on the promoter can subsequently affect the outcome.^{8,9} Nevertheless, irrespective of the specific influences of DNA methylation upon gene expression, distortion of the methylation profile is a key event and a known hallmark of neoplastic cells with both global hypomethylation and gene specific hypermethylation (e.g., genes involved in apoptosis, metastasis, adhesion, cell cycle and DNA repair) reported in all cancers investigated.^{4,6,10-12} Evidence is emerging that epigenetic changes, including alterations in DNA methylation, occur at early stages of tumorigenesis, potentially preceding mutational changes.^{1,5,6,13,14}

The environment plays a key role in the development of complex diseases. Exposure to natural and artificial chemicals and physical agents has been recognized as the primary cause of human cancers (reviewed in ref. 15). Previously, adverse effects of

*Correspondence to: Leda Mirbahai; Email: lxm701@bham.ac.uk
Submitted: 08/02/11; Accepted: 08/26/11
DOI: 10.4161/epi.6.11.17890

chemicals were mainly associated with the ability of the chemicals to interact with DNA and induce mutations (genotoxic chemicals). However, non-genotoxic carcinogens have been identified as a second category, which induce tumorigenesis but are not directly mutagenic. These chemicals modulate cell growth and proliferation through alteration of signaling pathways and DNA methylation, resulting in changed gene expression.¹⁶ Growing amounts of experimental and epidemiological data demonstrate that accumulation of epigenetic modifications over time due to environmental insults (e.g., environmental stressors or sub-lethal concentrations of chemicals) increases the individual's susceptibility to disease and development of disorders such as cancer.^{13-15,17,18} This correlates with the time required for development of different types of tumors.^{14,18} For example, studies on exposure to metals (e.g., nickel, cadmium, chromium, arsenic), air pollutants and endocrine disruptors have clearly demonstrated a link between environment, epigenetics and cancer.¹⁸⁻²⁰

However, in contrast to the wealth of information available on environmental factors, epigenetic change and development of cancer in humans, this area is substantially understudied in aquatic biology. A few studies have investigated the effects of chemical exposures on methylation levels at either a global level or on selected genes in aquatic species. These are, hypomethylation of the *vitellogenin 1* promoter in adult zebrafish liver (*Danio rerio*) following exposure to 17 α -ethynodiol (EE2),²¹ global hypermethylation in male gonads of three-spine stickleback (*Gasterosteus aculeatus*) after exposure to 17 β -estradiol (E2),²² global hypermethylation in the liver of goldfish (*Carassius auratus*) following exposure to heavy metals,²³ change in methylation of aromatase and estrogen receptor (ER) in Japanese medaka (*Oryzias latipes*) following exposure to EE2,²⁴ genome-wide methylation profiling of adult zebrafish hepatocellular carcinoma tumors induced by 7, 12-dimethylbenz[α]anthracene,²⁵ global methylation studies in *Daphnia magna* following exposure to several chemicals²⁶⁻²⁸ and global methylation studies in the liver of bluegill sunfish (*Lepomis macrochirus*) and kelpfish (*Sebastiscus marmoratus*) following exposure to benzo[a]pyrene, tributyltin and triphenyltin.^{29,30} As a result, it is apparent that epigenetic mechanisms are affected by model environmental pollutants in aquatic species as well as mammals. This highlights the importance of investigating the influence of environmental factors upon the epigenome and determining its role in diseases of marine species. However, there have been no studies of marine species investigating epigenetic changes in relation to carcinogenesis so far, partly due to the difficulties of studying non-model organisms and the availability of tumor samples from wild species.

Common dab (*Limanda limanda*) is a flatfish caught from UK waters as part of the UK Clean Seas Environmental Monitoring Programme (CSEMP). Liver pathology in dab, including cancer and pre-neoplastic toxicopathic lesions, is used as an indicator of the biological effects of contaminants on the marine environment. Due to their living habits and close proximity to the ocean floor they are exposed to relatively high levels of sediment-associated chemicals, making them ideal species for biomonitoring and environmental carcinogenesis studies. An unusually high prevalence of liver tumors has been reported in dab, with some

UK sites exceeding 20%.³¹⁻³⁵ However, the causative factors of these tumors and the molecular mechanisms involved, especially the balance between epigenetic and genetic factors, is unclear. Previous studies in our laboratory indicated that mutation profiles of oncogenes and tumor-suppressor genes within tumors were different between fish species and humans.^{36,37} In our previous studies in reference 25, we demonstrated that pathways with differentially methylated genes in chemically induced zebrafish hepatocellular carcinoma (HCC) are similar to the pathways altered in human HCC. This highlighted the importance of DNA methylation alterations in the development of fish tumors and suggested involvement of epigenetic factors in the formation of dab tumors.

Therefore, the aims of this study were to identify differentially methylated regions of the genome as well as differentially expressed genes in dab hepatocellular adenoma (HCA) compared to healthy liver. In addition, using pathway analysis techniques we aimed to determine if methylation and expression of specific pathways were altered in tumors compared with healthy control samples.

As dab is an unsequenced species, gene-specific studies are challenging using traditional techniques. Hence to identify genome-wide methylation changes we combined methylated DNA immunoprecipitation with de novo high-throughput sequencing for the first time. Based upon these data, we designed the first dab-specific gene expression microarray. We further confirmed our data using bisulfite sequencing PCR (BSP) and real time polymerase chain reaction (RT-PCR). Finally, our data demonstrate the suitability and power of these methods for studying DNA methylation in unsequenced species.

Results

Global measurement of DNA methylation. To establish the global levels of cytosine methylation in the three tissue categories of interest—healthy dab liver, HCA and apparently healthy surrounding tissue of tumor containing dab liver (ST)—reverse phase high performance liquid chromatography (HPLC) was performed. Absence of RNA contamination was confirmed as previously described in reference 25. As shown in Figure 1 a statistically significant 1.8 fold hypomethylation was detected between healthy dab liver and tissue surrounding HCA tumors (ST) and between healthy dab liver and HCA (p value < 0.01). However, no significant changes were detected in overall cytosine methylation levels between ST and HCA. The level of cytosine methylation in healthy dab liver (Fig. 1) was 8%. This is in accord with previous published data demonstrating a 2-fold higher DNA methylation level in fish than mammals.³⁸

De novo high-throughput sequencing analysis of MeDIP DNA. Although no change was detected in overall cytosine methylation levels between ST and HCA, to further investigate the possibility of differentially methylated regions between HCA and ST at the gene level, we combined de novo high-throughput sequencing (HTS) with methylated DNA immunoprecipitation (MeDIP) for the first time. We performed genome-wide DNA methylation profiling of HCA and corresponding ST by first

generating MeDIP enriched DNA libraries. Immunoprecipitated fragments were subjected to high-throughput sequencing using an Illumina Genome Analyser II to obtain a comprehensive methylation map. The approximately 90 million and 88 million 44-base paired-end reads for HCA and ST respectively, were combined and assembled. This resulted in identification of 264,008 contigs with lengths greater than or equal to 200 bp. From the identified contigs, the ones that did not match the criteria for prediction of CpG islands (CGI) in vertebrates, ($\text{length} \geq 200 \text{ bp}$, $\text{O/E} \geq 0.6$, $\text{CG\%} \geq 50$) as described by Gardiner-Garden and Frommer,³⁹ were excluded (Table 1).

In order to classify each segment as hyper-, hypo-, and non-differentially methylated, the number of MeDIP-reads comprising each contig in tumor were divided by the number of MeDIP-reads that comprised the same contig in corresponding ST. We arbitrarily defined “hypo- and hyper-methylated” in HCA as 1.5 fold or greater decrease and increase in the calculated ratio, respectively. As dab is an unsequenced species, the identified candidate hypo- and hypermethylated contigs were annotated against Genbank nr protein database (E-value cut-off $< 10^{-6}$). This resulted in identification of 1,693 differentially methylated sequences containing CGI regions in tumor compared to control [Figure 2]. A complete list of identified contigs (1,693), annotations and fold changes are shown in Table S1]. To identify differentially methylated CGI-containing promoter regions in dab HCA compared to ST, the identified contigs (69,046) were mapped onto pufferfish (*Takifugu rubripes*) genome, available from UCSC vfr2 “refGene” file. Pufferfish was selected for contig mapping due to its higher sequence homology to dab compared with other fully sequenced fish species (zebrafish and stickleback). This resulted in identification of 60 hypo- and 51 hypermethylated (fold change >1.5 , Fig. 2) CGIs that could be aligned with the pufferfish genome between 1.5 kb upstream and 1 kb downstream of predicted transcription start sites (TSS) (Fig. 2).

Ingenuity Pathway Analysis (IPA) was performed to characterize the functional relationships between genes with altered methylation in HCA compared to ST. The most significant networks and functions included cell-to-cell signaling, cell cycle, DNA replication and cellular assembly and organization as well as canonical pathways including Wnt/ β -catenin signaling, growth hormone signaling and apoptosis signaling. Figure 3 represents the IPA biological functions associated with genes with altered methylation levels in HCA (FDR $< 5\%$). Table 2 illustrates a sub-section of the genes (hypo- and hypermethylated) possessing biological functions linked to carcinogenesis.

Confirmation of the MeDIP de novo high-throughput sequencing data using BSP and comparability of the data to additional individuals. To validate the data achieved from MeDIP de novo HTS, five differentially methylated genes were selected with a fold change greater than 2-fold, including both hypomethylated and hypermethylated groups of genes and a non-coding sequence. Direct bisulfite sequencing PCR (BSP) was carried out on the original samples used for MeDIP de novo HTS.²⁵ As an example, the data for *protocadherin Igamma22* (*pcdh1g22*) gene is shown in Figure 4. (The data for the remaining four genes, *microtubule-associated protein Iaa* (*map1aa*), *tubulin*

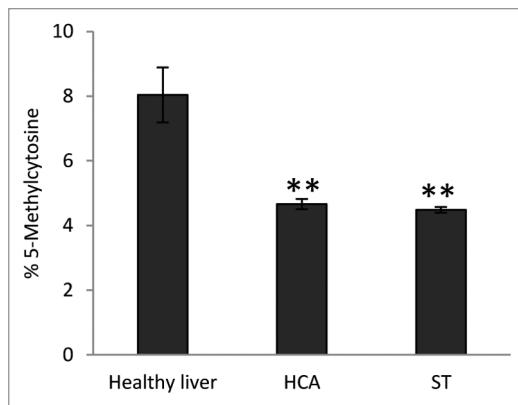


Figure 1. Measurement of global percentage of methylated cytosine. Overall methylation levels were measured in healthy dab liver, hepatocellular adenoma (HCA) and tissue surrounding HCA tumors in dab liver (ST) using HPLC. Data are shown as mean \pm SEM of five independent samples. **Significantly different from healthy liver (p value < 0.01).

tyrosine ligase-like member 7 (*ttll7*), *nidogen 1* (*nid1*) and a non-coding protein region, are shown in Fig. S4). For all the genes investigated the BSP data were in qualitative agreement with the data from MeDIP high-throughput sequencing.

In addition, we investigated the levels of variation and comparability of the methylation data derived from MeDIP de novo HTS of the one fish described above, to HCA and corresponding ST from 10 individual dab collected from five different sampling sites in the Irish Sea and Bristol Channel (sampling details are shown in Table S5). BSP was carried out for the same five genes on 10 HCA, 10 corresponding ST and 12 healthy tissues (Table S5). The data for *pcdh1g22* is shown in Figure 5 (the data for remaining genes are shown in Figs. S6–9). The same trends in methylation levels were observed between HCA and surrounding tissue for most genes and fish investigated. However, inter-individual variation was observed, especially between HCA samples of different fish. The lowest level of variation in methylation profile was observed between different healthy dab livers. The observed variation was not surprising as these samples are isolated from wild fish tumors, that are likely heterogeneous in terms of progression.

Design of 8 x 15k gene expression microarray and gene expression analysis of dab hepatocellular adenoma. Dab DNA sequences derived from the MeDIP-HTS were annotated against the Genbank nr protein database for identification of coding sequences. The annotated sequences were used to design an Agilent custom-made dab specific 8 x 15k 60-mer oligonucleotide gene expression array.

RNA extracted from the 3 groups of interest used in the direct BSP experiment—HCA, ST and healthy dab liver—were labeled with fluorescent Cy5-dCTP and hybridized to the dab custom-made gene expression array against a Cy3-dCTP labeled RNA reference pool made from all samples. Following quality checks and normalization, genes differentially expressed between the 3 groups investigated were identified (fold change >1.5 , FDR $<15\%$). One HCA sample did not pass quality checks and with its corresponding ST was removed from the analysis. As expected,

Table 1. CpG island discovery

Contigs Criteria	Total number	Total length	Average length	Maximum length	Minimum length
Length ≥ 200 bp	264,008	78,492,996	297.31	2,283	200
Length ≥ 200 bp, O/E ≥ 0.6, CG% ≥ 50	69,046	22,317,135	323.22	2,283	200

The reads obtained from MeDIP high-throughput sequencing of HCA and corresponding surrounding tissue were combined. The criteria described by Gardiner-Garden and Frommer³⁹ were used to obtain 69,046 CGI-containing contigs for further analysis and identification of differentially methylated regions between HCA and surrounding tissue.

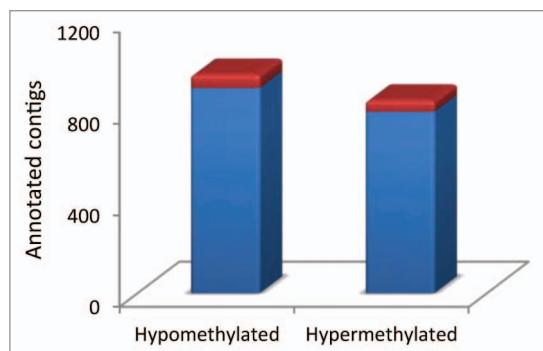


Figure 2. Differentially methylated contigs in dab HCA compared to surrounding tissue. Blue: 899 Hypomethylated and 794 hypermethylated CGIs in dab HCA as compared to ST could be annotated against Genbank nr protein database (fold change >1.5, E-value cut-off <10⁻⁶). Red: 60 Hypomethylated and 51 hypermethylated CGIs identified from dab that aligned 1.5 kb upstream to 1 kb downstream of predicted transcription start sites (TSS) in pufferfish genome (fold change >1.5, The list of 111 hypo- and hyper-methylated regions in TSS are shown in Table S2).

due to inter-individual variation, most differentially expressed genes were observed between HCA and ST with 520 genes statistically significantly altered in expression (208 genes repressed and 312 genes induced in HCA compared to ST, paired t-test, fold change >1.5, FDR <15%). Therefore, the focus has been on

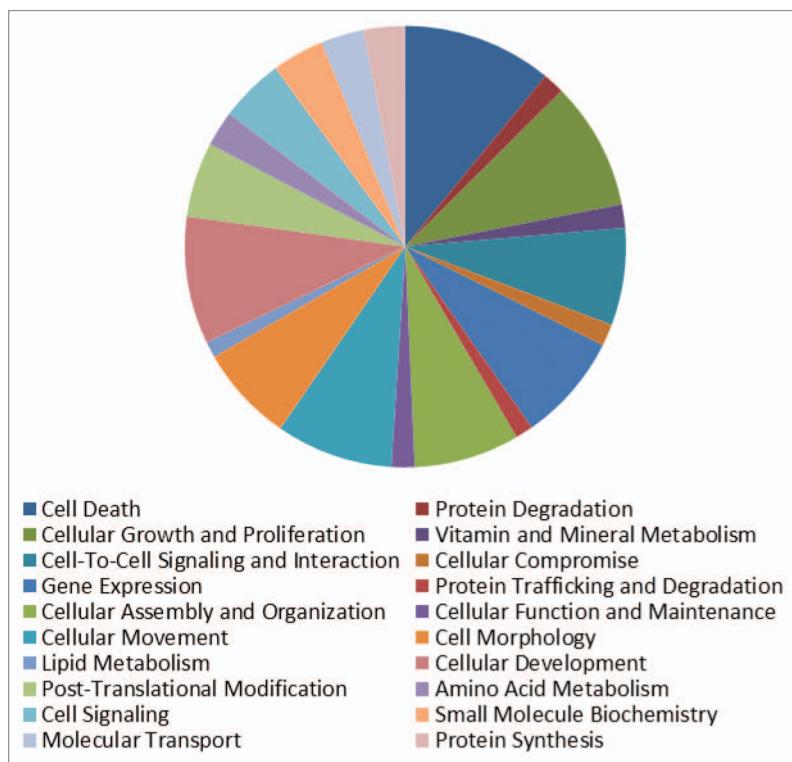


Figure 3. Biological functions enriched among genes with altered methylation levels (hypo- and hyper-methylated) in HCA. IPA was used to group the genes with altered methylation levels based on biological functions. Each section of the pie chart corresponds to the number of the genes associated with the mentioned biological function with altered methylation levels in HCA compared to ST (fold change >1.5, FDR < 5%).

the observed changes between HCA and ST. We found 91 genes to be statistically significantly altered in expression between HCA and healthy tissue with 73 and 18 induced and repressed in HCA compared to healthy dab liver, respectively (fold change >1.5, FDR <15%, unpaired t-test, Table S10). Furthermore, there were apparent gene expression changes between ST and healthy dab liver but due to inter-individual variation they did not pass the FDR-cut-off. Using principal components analysis (PCA) score plots of differentially expressed transcripts, the two groups of HCA and ST were separated based on treatment along the PC1 and PC2 axes (Fig. 6). In addition, within each condition, samples were clustered based on sampling sites. As shown in Figure 6, Cardigan Bay samples were distinctly separated from other sampling sites based on principal component 1. This indicates that a subset of these tumor-associated genes is expressed in a site-specific manner, implying an influence of environmental factors upon development of these tumors.

IPA was performed to characterize the functional relationships between genes with altered expression in HCA compared to healthy dab liver as well as HCA and ST. IPA identified significant networks and top functions that are commonly associated with cancer in comparisons of both of HCA to ST and HCA to healthy dab liver. Networks and functions including cell death, cell cycle, cellular growth and proliferation, cell morphology, cell signaling, amino acid synthesis and gene expression and canonical pathways including molecular mechanisms of cancer, Wnt/β-catenin signaling, aryl hydrocarbon receptor signaling and regulation of eIF4 and eIF2 signaling were associated with the differentially expressed genes for each comparison analyzed (FDR <5%). In both sets of analysis—HCA and ST, HCA and healthy dab liver—cancer was highlighted as the top predicted disorder based on categories of the genes with altered transcription (FDR <5%, fold change >1.5). As an example Table 3 shows a sub-section of the genes

Table 2. Subsection of the genes with altered methylation associated with development of tumors (FDR <5%) in HCA compared to ST

Biological function	Methylation level in HCA compared to ST
Cell death (cell death, apoptosis, survival of cells)	<i>caspase 6 (casp6,\uparrow), cell death-inducing dffa-like effector b (cideb,\downarrow), dedicator of cytokinesis 1 (dock1,\uparrow), apoptosis-associated tyrosine kinase (aatk,\downarrow), angiopoietin-related protein 4 precursor (angptl4,\downarrow), synovial apoptosis inhibitor synoviolin (syvn,\downarrow), signal transducer and activator of transcription 5 (stat5a,\downarrow), jagged 1 (jag1,\downarrow), jagged 2 (jag2,\downarrow), rho-associated protein kinase 1 (rock1,\uparrow), cyclin-dependent kinase 6 (cdk6,\uparrow)</i>
Cell growth and proliferation	<i>fibroblast growth factor receptor 3 (fgfr3,\downarrow), insulin-like growth factor 1 receptor (igf1r,\downarrow), jagged 1 (jag1,\downarrow), jagged 2 (jag2,\downarrow), transcription factor e2f7 (e2f7,\downarrow), DNA (cytosine-5-)methyltransferase 4 (dnmt4,\downarrow), signal transducer and activator of transcription 5 (stat5a,\downarrow), glucose-6-phosphate 1-dehydrogenase (g6pd,\downarrow), discs, large homolog 1 (dlg1,\downarrow), peroxisome proliferator-activated receptor beta (pparb,\downarrow)</i>
Cell-to-cell signaling and interaction	<i>jagged 1 (jag1,\downarrow), jagged 2 (jag2,\downarrow), signal transducer and activator of transcription 5 (stat5a,\downarrow), cadherin 22 (cdh22,\downarrow), integrin, alpha 3a (itga3,\downarrow), integrin beta 1 subunit (itgb1,\downarrow), integrin alpha 9 (itga9,\downarrow), protocadherin 1gamma22 (pcdh1g22,\downarrow)</i>
Gene expression	<i>cAMP responsive element binding protein 1 (crem,\downarrow), notch homolog 2 (notch2,\downarrow), aryl hydrocarbon receptor nuclear translocator-like protein 1 (arntl,\uparrow), hematopoietic transcription factor gata-1 (gata1,\uparrow), mitogen-activated protein kinase kinase kinase 4-like isoform 3 (mapk4k3,\uparrow), mitogen-activated protein kinase 12 (mapk12,\downarrow), mitogen-activated protein kinase 8 interacting protein 3 (mapk8ip3,\uparrow), mitogen-activated protein kinase 11 (mapk11,\downarrow), mitogen-activated protein kinase 1 (mapk1,\uparrow)</i>

\uparrow hypermethylated and \downarrow hypomethylated in HCA compared to ST. The complete list of the biological functions and associated genes are presented in Table S3.

(under and overexpressed) associated with cancer derived from comparison of HCA to ST. The list of all the genes with altered transcription levels and their associated biological functions (FDR <5%) in both groups is shown in Table S11. Figure 7 illustrates genes with altered expression levels in dab HCA compared to ST (fold change >1.5, FDR <5%) related to the canonical pathway “molecular mechanisms of cancer” in human including *v-akt murine thymoma viral oncogene homolog 1 (akt1)*, *cyclin d2 (ccnd2)* and *mitogen-activated protein kinase 10 (mapk10)*.

Confirmation of the gene expression data using RT-PCR. Based on microarray data, five genes with statistically significant altered transcription levels were selected for RT-PCR analysis (Table 4). A significant positive correlation was observed between array data and RT-PCR (p value < 0.05, Pearson’s correlation coefficient = 0.97).

Discussion

Global DNA methylation. In addition to the observed changes between HCA and surrounding liver tissue in methylation levels of specific genes, one of the most important findings of this study was the difference in global methylation levels. A significant 1.8-fold global hypomethylation was detected in HCA compared to liver from non-cancer bearing dab. This pattern was also identified when comparing “apparently healthy” liver (ST) surrounding HCA tumors with liver from non-cancer bearing dab (Fig. 1). However, there were no detectable differences in global methylation levels between ST and HCA dissected from the same liver.

There is evidence that global hypomethylation and gene specific hypermethylation can occur early during tumorigenesis, this has been reported in apparently normal tissues prior

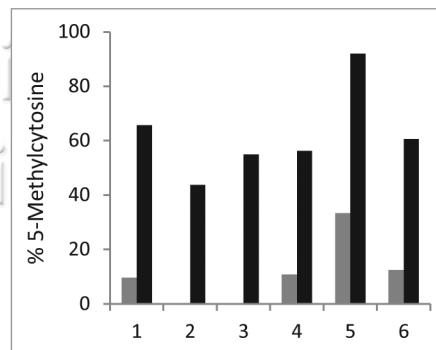


Figure 4. Direct bisulfite sequencing PCR data for *protocadherin 1gamma22 (pcdh1g22)* gene. To validate the MeDIP high-throughput sequencing data the same HCA and corresponding ST were used for direct BSP. MeDIP HTS illustrated a 2.08 fold hypomethylation of the investigated region of *pcdh1g22* in HCA compared to ST which is in agreement with the direct BSP data achieved from the same region. (numbers on the x-axis: CpG sites 1–6).

to mutational changes in patients who later developed cancer (reviewed in refs. 1 and 14). It has been suggested, as part of the “progenitor model of tumorigenesis”, that epigenetic changes arise as early as progenitor cells. This leads to genome instability, chromosomal rearrangement, loss of imprinting and activation of transposons, thereby predisposing for mutations in gatekeeper genes and an increased risk of tumorigenesis.^{1,4,5,14} Mutations in gatekeeper genes of the pre-neoplastic cells and further epigenetic and genetic changes can ultimately lead to fully formed tumors.^{5,14} Hence hypomethylation is a striking feature and a common early mechanism in carcinogenesis.¹⁴ This is in agreement with our findings of hypomethylation in apparently healthy portions of tumor-bearing dab livers.

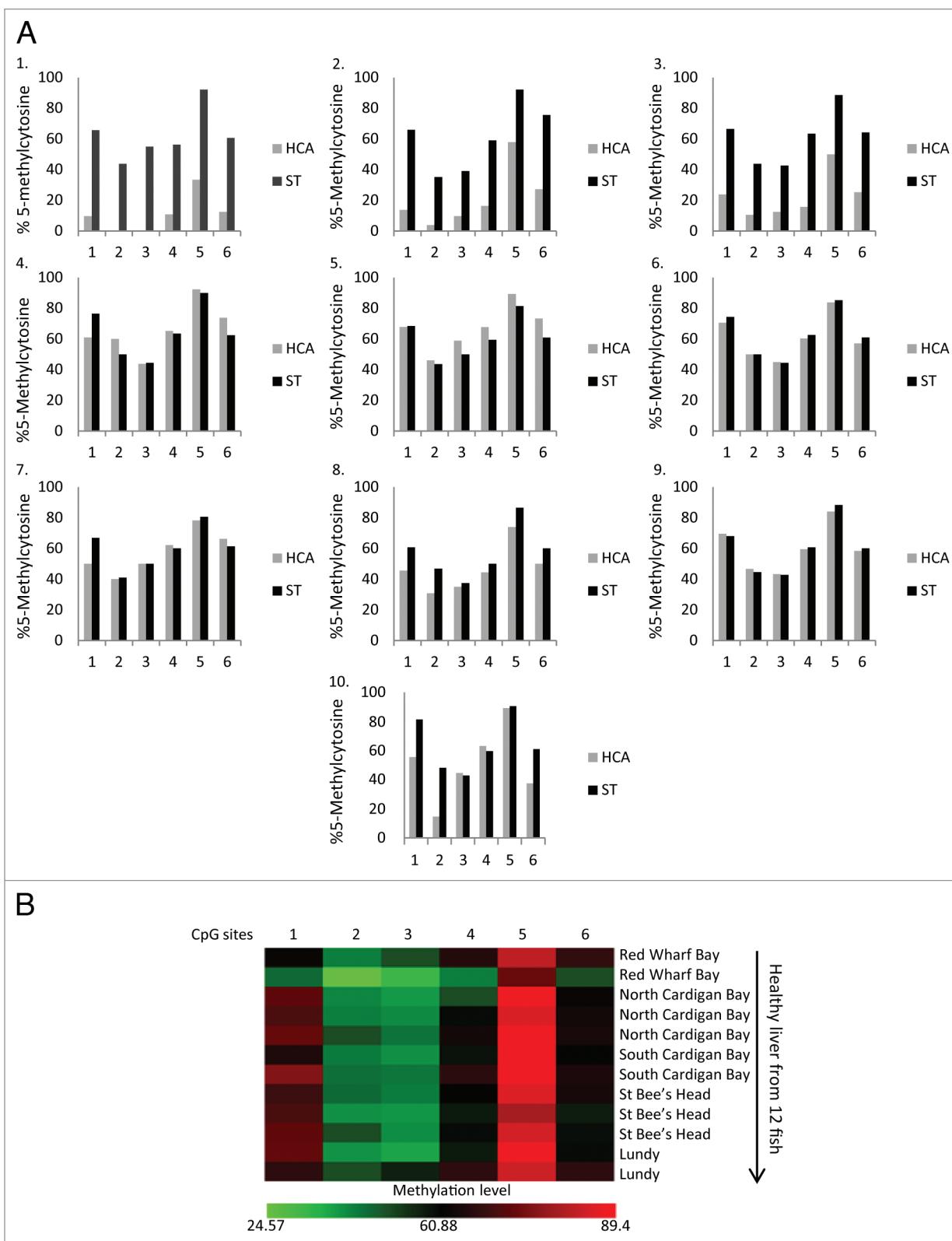


Figure 5. Comparing the methylation levels of *protocadherin 1gamma22* (*pcdh1g22*) in dab collected from five different sampling sites in Irish Sea via direct BSP. (A) Methylation levels of six CpG sites (x-axis) in 10 individual fish (10 HCA and 10 corresponding ST). Sampling sites: 1, 6, 9: South Cardigan Bay; 2: Inner Cardigan Bay; 3, 8: Red Wharf Bay; 4: North Cardigan Bay; 5, 10: St. Bee's Head. (B) Heat map of methylation levels of six CpG sites in 12 healthy dab livers. Each row represents one fish. Fish are labeled based on their sampling location.

Overview of the DNA methylation changes. Based on the epigenetic model of tumorigenesis, we anticipated that the “epigenetically primed” liver tissues could undergo further epigenetic and genetic changes, leading to formation of HCA with specific methylation changes at individual genes.^{5,14} Indeed, we have previously detected changes in methylation levels of specific genes in a preliminary study of dab tumors.⁴⁰ However, to achieve a comprehensive DNA methylation and transcriptomic profile for the unsequenced flatfish dab, we successfully combined methylation DNA immunoprecipitation with de novo HTS and designed a gene expression array based on the contigs achieved. As demonstrated for the first time by Down et al. MeDIP-HTS is a powerful, unbiased and quantitative technique for establishing whole genome methylation profiles in model organisms. However, this technique has never before been applied to study DNA methylation alterations in an unsequenced species.

Combining MeDIP with high-throughput sequencing resulted in considerable numbers of overlapping reads for construction of contigs, 34% of which could be annotated based on homology between protein-coding regions and were apparently hypo- or hypermethylated (fold change >1.5) in tumor samples compared to control. This method is not without limitations. As the sequences were identified based on conserved regions of orthologous sequences, most of the identified regions corresponded to the CGIs located within exons, and only a fraction of the data focused on promoter regions. Therefore, the main outcome of this study was identification of genes and pathways with altered methylation in HCA compared to ST irrespective of the position of the changes observed. Nevertheless, this approach resulted in identification of a substantial number of genes and pathways with altered methylation in HCA samples, as shown in Tables S1–3. Most importantly, the data confirmed our hypothesis that although similar at global levels, HCA samples differ in DNA methylation at the gene level compared with primed surrounding tissue. Our finding that methylation of exons was different in the tumor samples is in agreement with previous findings in humans.⁴² Furthermore, in tumorigenesis, identifying methylation alterations is more important than the absolute levels of methylation.⁴² In humans, aberrant methylation mediates changes in most of the key pathways associated with cancer, resulting in loss of cell cycle control, cell-to-cell signaling, apoptosis signaling, altered receptor and transcription factor function as well as genomic instability.⁴³ Our study identified expression and methylation changes in key genes of the same biological pathways in dab HCA. DNA methylation is therefore an important factor in tumorigenesis in the wild flatfish dab.

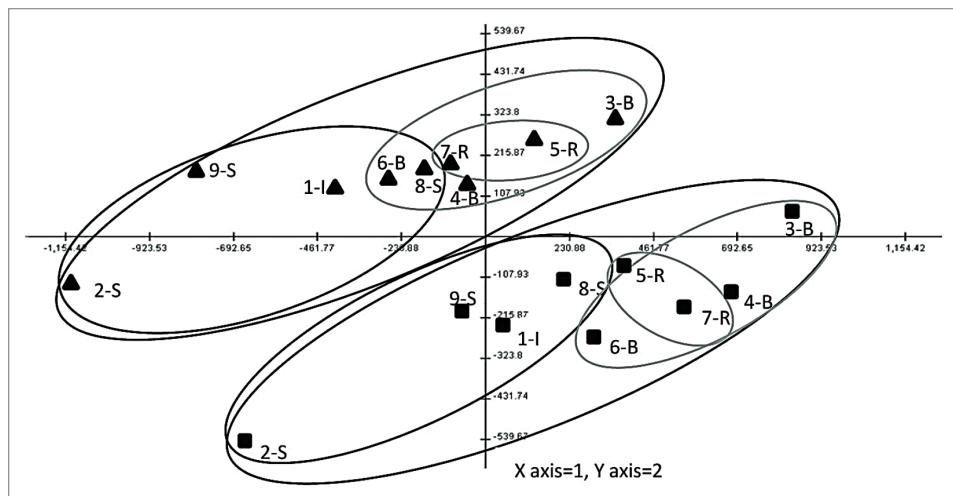


Figure 6. Principal component analysis (PCA) scores plot of differentially expressed transcripts for HCA and ST. Dab HCA samples (triangle) and dab ST (square) were separated based on treatment along the PC1 and PC2 axes. Numbers 1–9 represent nine individual fish. Sampling sites: I, Inner Cardigan Bay; S, South Cardigan Bay; B, St. Bee’s Head; R, Red Wharf Bay.

Methylation and transcriptional changes associated with tumorigenesis in dab HCA. IPA identified changes in methylation of genes involved in the Wnt/β-catenin signaling pathway. Activation of this pathway has been associated with development of cancer mediated by stabilization of β-catenin and inappropriate activation of β-catenin target genes such as *c-MYC* and *c-JUN* in humans.^{44,45} Our data showed that methylation of genes such as *dishevelled*, *dsh homolog 2* (*dsh*), acting as a positive mediator of Wnt signaling, *frizzled* gene, responsible for encoding Wnt receptors, tumor suppressor gene *adenomatous polyposis coli* (*apc*) and *c-myc* were altered in HCA compared to ST. In addition, transcription of *c-myc* and *frizzled homolog 7a* were induced and *frizzled homologs 10* and *7b* were suppressed in HCA samples.

Changes in methylation of genes involved in the apoptosis signaling pathway, related to release of mitochondrial cytochrome *C* and activation of a range of caspases, were detected in HCA samples. These genes included *Diablo*, which is associated with neutralization of members of inhibitory proteins (IPA), and release of cytochrome *C*,^{46,47} caspase activator *rock1*^{48,49} and pro- and anti-apoptotic *nfκb*.⁵⁰

Based on transcriptional changes, the canonical pathway “molecular mechanism of cancer” and “cancer” as a disease annotation term were enriched among differentially expressed genes in HCA compared to ST and healthy liver. Genes involved in different aspects of tumorigenesis were significantly enriched in HCA. For example, genes involved in cell cycle, proliferation and cell death [e.g., *cyclin d2* (*ccnd2*), *cyclin d1 binding protein* (*ccndbp1*), *ras* and map kinases (*mapk3k9*, *mapk10*)] were significantly overexpressed in tumors, while genes related to repair mechanisms such as an *excision repair* (*ercc6*) and *suppression of tumorigenesis 14* (*st14*) were significantly downregulated. Genes associated with energy production and ribosomal protein synthesis were significantly overexpressed in tumors. This funding is not unexpected as both functions are required for highly proliferating tumor cells.³³

Table 3. Subsection of the genes with altered gene expression in HCA compared to ST

Biological function	Gene expression level in HCA compared to ST
Cell death	mitogen-activated protein kinase kinase kinase 10 (<i>map3k10</i> ,↑), ras-like protein 1 (<i>ras</i> ,↑), histone de-acetylase 1 (<i>hdac1</i> ,↓), retinoid x receptor beta (<i>rxb</i> ,↑), cyclin D2 (<i>ccnd2</i> ,↑), v-myc myelocytomatosis viral oncogene-like protein (<i>myc</i> ,↑), dnaJ (<i>hsp40</i>) homolog, subfamily B, member 1 (<i>dnajb1</i> ,↑)
Cellular growth and proliferation	s-adenosylhomocysteine hydrolase (<i>ahcy</i> ,↑), retinoid x receptor beta (<i>rxb</i> ,↑), histone de-acetylase 1 (<i>hdac1</i> ,↓), ras-like protein 1 (<i>ras</i> ,↑), myc-associated zinc finger protein (<i>maz</i> ,↑), cyclin d2 (<i>ccnd2</i> ,↑), 26s protease regulatory subunit 8 (<i>psmc5</i> ,↑), DnaJ (<i>hsp40</i>) homolog, subfamily B, member 1 (<i>dnajb1</i> ,↑), suppression of tumorigenicity 14 (<i>st14</i> ,↓)
Cell cycle	cell cycle regulator mat89bb homolog (<i>m89bb</i> ,↓), mediator of RNA polymerase ii transcription subunit 29 (<i>med29</i> ,↑), retinoid x receptor beta (<i>rxb</i> ,↑), histone de-acetylase 1 (<i>hdac1</i> ,↓), cyclin d2 (<i>ccnd2</i> ,↑)
Protein and lipid synthesis and reproduction	vitellogenin b (<i>vtgb</i> ,↑), vitellogenin a (<i>vtga</i> ,↑), 60s ribosomal protein l9 (<i>rpl9</i> ,↑), 60s ribosomal protein l19 (<i>rpl19</i> ,↑), 60s ribosomal protein l28 (<i>rpl28</i> ,↑), aldehyde dehydrogenase 8 member a1 (<i>aldh8a1</i> ,↑), Similar to human cytochrome P450, family 2, subfamily J, polypeptide 2 (<i>cyp2j2</i> ,↑), eukaryotic translation initiation factor subunit 3 isoform cra_d (<i>eif3</i> ,↑), methionine aminopeptidase 2 (<i>metap2</i> ,↑)

These genes are associated with the most relevant biological functions altered in tumors (FDR <5%, fold change >1.5). ↑ and ↓ over- and under-expressed in HCA compared to ST, respectively. A complete list of the biological functions and associated genes is presented in **Table S11**.

A link between environmental contaminants, changes in DNA methylation, transcription and dab liver tumors. A statistically significant induction in expression of the egg yolk precursor proteins (*vitellogenins vtga*, *vtgb* and *vtgc*) was observed in HCA compared to healthy liver. As this induction was also observed when comparing HCA to ST of the same fish, inter-individual variability, such as the female reproductive cycle, was not responsible. During oocytegenesis in most oviparous species, vitellogenin proteins are transferred from the liver to the growing oocyte where they are cleaved to form lipovitellin and phosphovitin, which serve as food reserves for the growing embryo.^{51,52} In contrast to high levels of vitellogenin in female fish during seasonal reproduction, very little vitellogenin is observed in male fish and juvenile female fish.⁵² Expression of vitellogenin in the liver is controlled mainly through binding of estrogen receptors to estrogen response elements. Therefore exposure to estrogen mimicking chemicals (EDC-endocrine disrupting chemicals) can elevate vitellogenin expression in fish, which can indicate feminization of male fish. Hence induction of vitellogenin in male fish and signs of intersex are used as biomarkers of exposure to EDC.⁵³ Our findings agree with previous studies demonstrating induction of vtg in female and male dab and intersex in male dab.^{33,53,54}

Different types of PCBs, PBDE and heavy metals such as cadmium (Cd) and lead (Pb) were detected in dab tissue. The concentration of these chemicals, as shown in **Table S12**, varied at different sampling sites with Cardigan Bay, and in particular South Cardigan Bay, identified as the most polluted site in regards to tissue concentrations of PCBs and PBDEs. For example, from the seven non-planar PCBs recommended for monitoring by International Council for the Exploration of the Sea (ICES), PCB118 concentration in dab liver at South Cardigan Bay was higher than the recommended hazardous ecotoxicological assessment criteria (EAC). This indicates the possibility of chronic exposure to this chemical, which is in agreement with the data reported by the Department for Environment Food and

Rural Affairs (DEFRA).⁵⁵ The differences in chemical contamination of sampling sites were also reflected in the results of PCA. As shown in the **Figure 6**, samples in each disease category were further clustered based on sampling sites. It is therefore possible that environmental contaminants are influencing transcriptional profiles. In addition, it has been reported that vitellogenin levels in male dab collected at different sampling sites varies, with fish collected from Cardigan Bay having the highest levels of vitellogenin compared to fish collected from other sampling sites investigated in this study (St. Bee's Head, Red Wharf Bay and Lundy).⁵³ Furthermore, a study by Lyons et al.³⁵ reported that dab collected from Cardigan Bay, and specifically South Cardigan Bay, have higher levels of tumor prevalence compared to dab collected from Red Wharf Bay. Cancer prevalence in dab increases with age; however, the age of onset of cancer in dab collected at geographically distinct sites varies, ranging from 4–7 years. This phenomenon could be linked to different susceptibilities of the fish collected from various sampling sites.⁵⁶ As the levels of environmental contaminants vary between sampling sites it is plausible that environmental driven epigenetic changes are responsible for different susceptibilities of these fish. Therefore, it is plausible that a mixture of weak, but persistent, lipophilic estrogenic chemicals such as polychlorinated biphenyls (PCB), polybrominated diphenyl ethers (PBDE) and some of their metabolites⁵⁷ are responsible for the observed increased expression of vitellogenin and formation of these tumors.

A link between exposure to chemicals such as EDCs (PCB, EE2, E2) and heavy metals (e.g., As, Cd), changes in DNA methylation (reviewed in ref. 58) and tumorigenesis in rodents and humans has been established.^{18–20,59,60} Although recent studies have demonstrated that EDCs can cause changes in DNA methylation in aquatic species,^{21,22} with separate studies suggesting that EDCs can cause tumors in fish liver,⁶¹ this connection between exposure to EDCs, change in DNA methylation and tumorigenesis has not been established in fish.

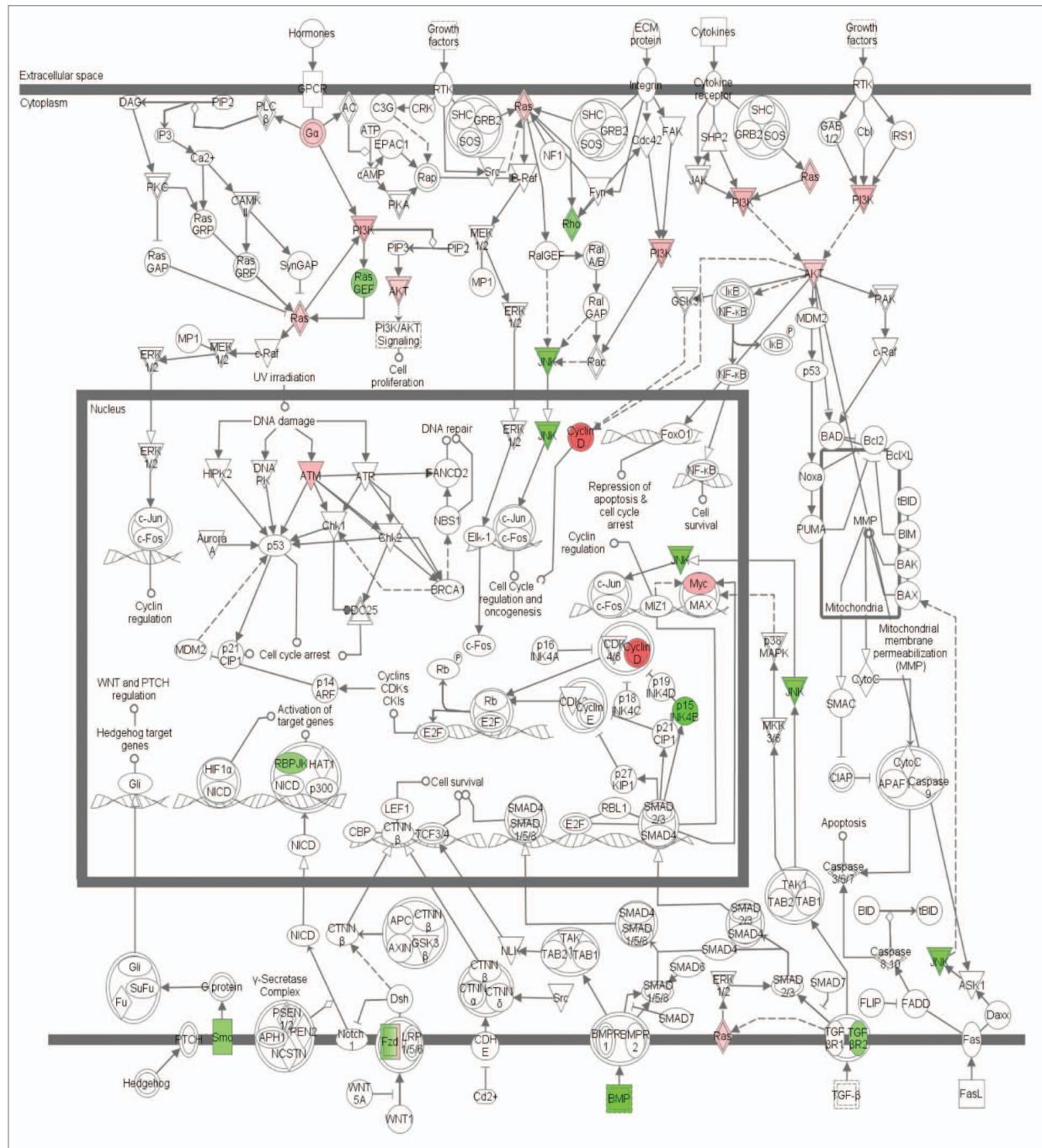


Figure 7. Biological network of genes linked to the canonical pathway “molecular mechanisms of cancer” with altered gene expression levels in dab HCA compared to ST (fold change >1.5, FDR < 5%). Green, underexpressed; red, overexpressed in HCA compared to ST.

Our data show that methylation and/or expression of several genes in both non-genomic and genomic (estrogen response element (ERE)-dependent and ERE-independent) estrogenic pathways are altered in HCA samples compared to healthy liver and ST. These genes are involved in biological functions such as proliferation and growth, differentiation, cell cycle, apoptosis and oncogenesis.⁶²⁻⁶⁴ In our study methylation and/or expression of several genes controlled through the ERE-dependent pathway that are commonly implicated in cancers were altered.

For example, induced transcription of *c-myc* oncogene, *vtg a* and *vtg b*, reduced transcription of protease *cathepsin (cts)* with change in methylation of *c-myc* and *cts* were identified in HCA compared to ST.^{62,63} E-ER complex can also modulate transcription of a range of genes without directly binding to DNA. This ERE-independent mechanism occurs in the nucleus and through protein-protein interaction with other DNA binding transcription factors (TF) and subsequent binding of TFs to their recognition sites in the promoter regions.⁶² We found that methylation/

Table 4. Comparison of fold change in gene expression for the selected genes by RT-PCR and microarray analysis

Gene	Grouping	QPCR fold change	Microarray fold change (FDR <15%, p value < 0.05, fold change > 1.5)
<i>vitellogenin b</i>	HCA and H	3.17**	2.13
<i>vitellogenin c</i>	HCA and ST	7.89*	8.5
<i>s-adenosylhomocysteine hydrolase</i>	HCA and ST	2.3*	1.78
<i>udp-glucose 6-dehydrogenase</i>	HCA and ST	2.51**	3.24

The same mRNA was used for both RT-PCR and microarray analysis. *syntrphin*, beta2 (*sntβ2*) gene was used as a reference gene. Five biological and three technical replicates were used. (H: healthy dab liver, **p value < 0.01, *p value < 0.05).

expression of several transcription factors that bind to ER and their target genes were altered in HCA compared to ST, such as *specificity protein 1 (Sp1)*, involved in crucial biological processes (such as differentiation, proliferation, apoptosis) and commonly overexpressed in tumors,⁶⁴ *signal transducer and activator of transcription (stat) 5, activating transcription factor 2 (atf-2)* and *nuclear factor nf-κb* associated with inflammation and oncogenesis⁶⁵ and their target genes such as *c-myc*, *vtgs* and *cyclin d1*.⁶² In addition, to the changes observed in ER-dependent and independent pathways, methylation and/or expression of several genes involved in the non-genomic estrogenic mechanism of action and development of cancer were also changed. For example, *insulin growth factor-1 receptor (igf1r)*, known to be affected in terms of methylation and expression in various cancers,²⁵ several different MAP kinases and the oncogene *Akt kinase*, involved in cell proliferation and survival of tumor cells,⁶⁶ were altered in HCA samples. In the non-genomic pathway of estrogenic compounds, estrogen binds to the ER located at the cytoplasmic membrane and in combination with several other signaling molecules activates MAP kinases. Activated kinases regulate transcription factors associated with ER (such as SP1, ATF2, STAT5 and NFκB) as well as the nuclear ER through phosphorylation.⁶² Figure 8, adapted from Bjornstrom and Sjoberg,⁶² illustrates the 3 main estrogen receptor pathways and the changes detected in dab HCA samples.

Hypothesis. Based on the evidence discussed above (chemical data and vitellogenin induction), one possible hypothesis is that chronic exposure to a mixture of contaminants such as weak endocrine disrupting chemicals (e.g., PCB, PBDE) or metals (e.g., Cd) cause the initial changes in overall methylation levels seen in the entire liver of tumor bearing dab. Indeed, inhibition of the activity of DNA methyltransferases is a reported effect of Cd.¹⁸ Exposure to EDCs and metals, especially during critical stages such as embryogenesis and gametogenesis, as well as increasing the individual's susceptibility to disease, can cause heritable epigenetic changes.¹⁸ It is plausible that these chemicals further affect the ERE-dependent and independent genomic and non-genomic estrogen response pathways. Changes in the methylation of these pathways can alter the expression of genes involved

in biological functions such as proliferation and cell cycle (Fig. 8). It is postulated that additional epigenetic and genetic changes in the primed pre-neoplastic cells and genome instability caused by global hypomethylation could lead to development of these tumors (Fig. 9 and adapted from Feinberg et al.).

Although the combined data on gene expression, DNA methylation, chemical exposure and evidence available in the literature are in accord with this interpretation, it is also possible that altered DNA methylation and disturbance of estrogen signaling within ST is a secondary change due to proximity of this tissue to HCA in the organ. However, there is no evidence in relation to fish tumors to support this possible explanation. Further support for a direct environmental influence on methylation changes might come from analysis of additional tissues in animals with modified methylation in the liver. Our data highlighted that DNA methyltransferase 3b was hypomethylated at its promoter region, indicating a possible increase in transcription of this gene in HCA samples compared to ST. As this gene is involved in DNA methylation pathway and is commonly overexpressed in cancers, we are currently investigating changes in the transcripts and metabolites involved in this pathway.

Material and Methods

All chemicals were obtained from Sigma-Aldrich (Poole, Dorset, UK) unless otherwise stated.

Collection of fish liver. As part of the United Kingdom Clean Seas Environmental Monitoring Program (CSEMP) several hundred dab flatfish (*Limanda limanda*) were captured from sampling sites in Irish Sea and Bristol Channel (Table S5) during June and July 2007, 2008 and 2009 using 30 minute tows of a standard Granton trawl by the Centre for Environment, Fisheries and Aquaculture Science (Cefas, Weymouth, UK). Dab were immediately removed from the catch and placed into flow-through tanks containing aerated seawater. Fish were assessed for external diseases, sacrificed and livers were visually assessed for the presence of macroscopic lesions (nodules). Methods described by Feist et al.⁶⁷ were used for dissecting the nodules and the healthy liver tissues. Sections of the collected samples were immediately snap frozen in liquid nitrogen for use in methylation and transcriptomic studies and the remainder of the samples were fixed in 10% Neutral Buffered Formalin (NBF) for 48 h.

Histopathology. Following fixation, samples were processed to wax in a vacuum infiltration processor using standard histological protocols.⁶⁷ Using a rotary microtome, 3–5 µm tissue sections were prepared and mounted onto glass slides for haematoxylin and eosin staining. Stained sections were analyzed by light microscopy (Eclipse E800, Nikon, UK). The type of liver tumor was established based on guidelines described by Feist et al.⁶⁷ In this study 12 non-cancerous liver samples from 12 individual female fish, 10 hepatocellular adenoma and 10 corresponding surrounding tissue samples from 10 female fish were used. Lesions were classified as HCA based on a combination of distinct morphological criteria including presence of lesion with relative lack of macrophage aggregates, compression of

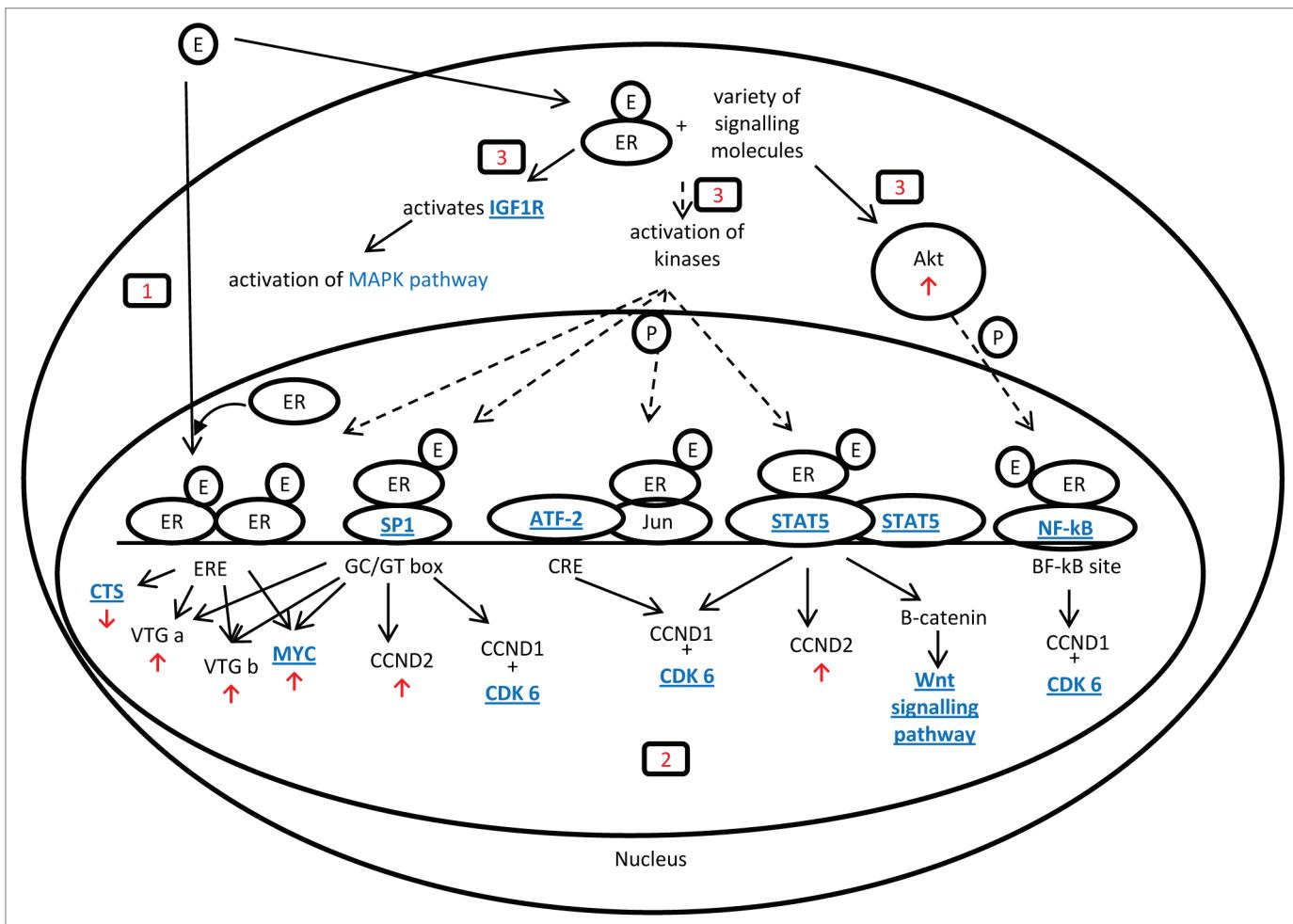


Figure 8. Methylation and transcription changes to estrogenic pathways observed in dab samples. Blue, genes with altered methylation in HCA compared to ST; red arrows, genes with altered transcription (induced, repressed) in HCA compared to ST or healthy dab liver; P, phosphorylation; E, estrogen or estrogen mimicking compounds. (1) Classic genomic ERE-dependent pathway. (2) Genomic ERE-independent pathway. (3) Non-genomic pathway. Adapted from original figure of Bjornstrom and Sjoberg.⁶¹

surrounding parenchyma, thickened trabecular structure and the relative absence of atypical nuclear and cellular profiles (histopathology images of HCA and surrounding tissue are presented in Fig. S13).⁶⁷ Both hepatocellular adenoma and carcinoma are observed within the liver of dab flatfish. However as the prevalence of HCA is much higher compared to HCC, HCA samples were used in this experiment.

Sample preparation. Dab liver tissues provided by Cefas were kept at -80°C until extraction. Each liver was homogenized in 10 µl/mg (wet tissue mass) water using a Precellys-24 ceramic bead-based homogenizer (Stretton Scientific Ltd.). Aliquots of the homogenized tissues (no less than 10 mg) were immediately added to ice cold RLT (600 µl) and β-mercaptoethanol (6 µl) mixture. Following centrifugation and removal of lipid layer, RNA and DNA were extracted from the same homogenates using the AllPrep DNA/RNA kit (Qiagen Ltd., West Sussex, UK) according to manufacturers' guidelines. The RNA and DNA purity and concentration were determined using a NanoDrop ND-1000 UV-VIS Spectrometer version 3.2.1 (Nanodrop, USA). Purity was assessed using the ratio A_{260}/A_{280} and for further RNA

quality assessments the 2100 Bioanalyzer (Agilent Technologies, Berkshire, UK) was used with a Eukaryote Total RNA Nano assay chip.

Measurement of genome-wide DNA methylation. Overall cytosine methylation in DNA samples extracted from five ST, HCA and non-cancerous liver were measured as described by Ramsahoye.⁶⁸ Briefly, following removal of RNA contamination from DNA samples, DNA was enzymatically digested to single nucleotides. An AKTA Explorer 10 with P900 pump, automatic UV detector (Amersham Biosciences), APEX ODS C18 column, 250 x 4.6 mm i.d., 5 µm particle size (Waters HPLC Ltd., UK, Phenomenex, UK), and grade column (Phenomenex, UK) were used. The retention time of each peak was established using nucleotide standards. In addition, a standard for uracil was used to monitor the purity of the DNA samples.

MeDIP de novo high-throughput sequencing. The MeDIP de novo high-throughput sequencing was carried out at Beijing Genomics Institute (BGI, China) based on methods described by Li et al.⁶⁹ Briefly, genomic DNA was extracted from one HCA sample and corresponding surrounding tissue (5 µg) and

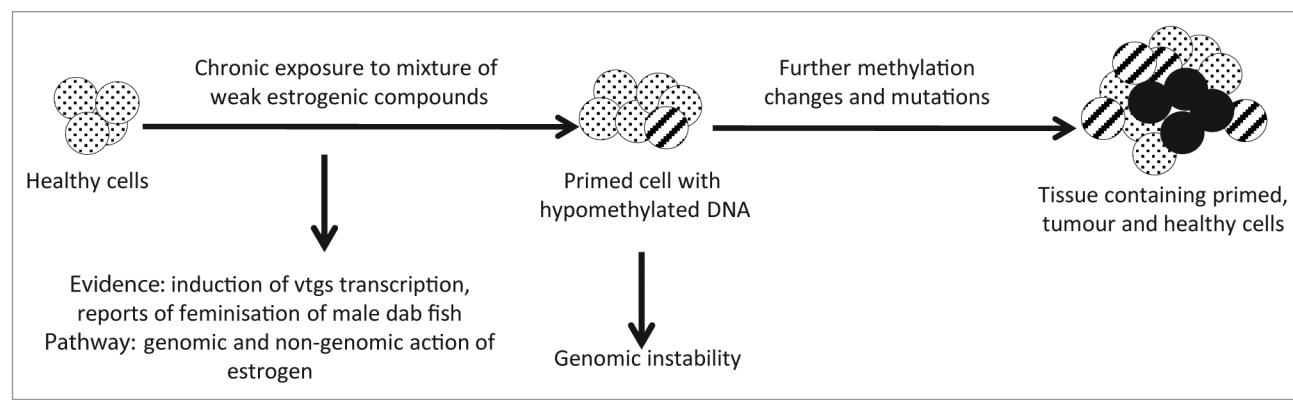


Figure 9. Hypothetical mechanism of tumorigenesis in dab. Chronic exposure to a mixture of contaminants such as weak estrogenic chemicals causes the initial global hypomethylation. ●, Healthy cell; ○, "Primed cell"; ●, tumor cell.

sonicated. Fragmentation was carried out using a Covaris sonication system (16 cycles of duty cycle: 10%, intensity: 5, cycles/burst: 200, time: 960 s; Covaris, Inc., USA), which resulted in 100–500 bp rich DNA fragments. The Illumina Paired-end library preparation kit (Illumina, Inc., USA) and the protocol provided were used to generate adaptor-attached DNA fragments for sequencing. During this process the overhangs from fragmentation were converted into blunt ends using T4 DNA polymerase and Klenow polymerase. “A” base was added to the 3' end of the blunt end phosphorylated DNA fragments, which provided a recognition site for a single “T” base overhang containing adaptors.

Methylated adaptor ligated DNA fragments were separated from the un-methylated fragments using MagMeDIP kit (Diagenode, Belgium) according to the manufacturer's instructions. The positive and negative control primers provided with the kit and the SYBR green mastermix (Applied Biosystems) with 40 cycles of 95°C for 15 s and 60°C for 1 min were used for validation of the immunoprecipitation procedure.

The methylated fractions were purified and eluted in EB buffer (27 µl) using ZYMO DNA Clean & Concentrator-5 column following the manufacturer's instructions. The eluted DNA was amplified by adaptor-mediated PCR in a final reaction volume of 50 µl consisting of purified DNA (23 µl), Phusion DNA polymerase mix (25 µl) and PCR primers (2 µl). Cycling parameters consisted of 94°C 30 s, 16 cycles of 94°C 30 s, 60°C 30 s, 72°C 30 s with a final step of 1 min at 72°C. Products were sequenced using an Illumina Genome Analyzer II (Illumina, Inc., USA). Raw data for both samples, hepatocellular adenoma HCA-18-RA09065-830 and healthy surrounding tissue ST-19-RA09065-830, were submitted to NCBI database with accession numbers GSE31124 and GSM770685, respectively.

Identification and annotation of differentially methylated regions. HTS of MeDIP genomic DNA from the HCA sample and ST produced 89,925,735 and 87,826,470 paired-end reads, respectively. The SOAP de novo software (a member of Short Oligonucleotide Analysis Package family) was used to assemble the reads with a 25 k-mer overlap to achieve sets of contiguous sequences (contigs) for both samples. The assembled contigs later contributed to microarray design.

In order to identify differentially methylated regions that contained CGIs in the HCA sample compared with ST, a third assembly was generated by combining the reads from both individual samples (HCA and ST). Contigs were generated using the list of combined reads from both samples and SOAP de novo software. N50 contigs with length <200 bp were discarded and the remaining N50 contigs were defined as “methylated DNA fragments”. (Note: N50 contig size was defined as the length of the smallest contig S in the sorted list of all contigs where the cumulative length from the largest contig to contig S was at least 50% of the total length). Reads from HCA and ST were aligned to the “methylated DNA fragments” assembled by SOAPaligner/soap2. The sequences that were uniquely mapped with ≤2 bp mismatches were retained. This assembly resulted in 264,008 contigs >200 bases, of which 69,046 passed the CGI criteria ($O/E > 0.6$ and $CG\% \geq 0.5$), with an average length of 323,22 bases. As the numbers of short reads contributing from each sample to each contig was known, the ratio of enrichment between HCA and ST was calculated to identify candidate differentially methylated CGIs. The identified candidate CGIs were annotated using BLASTx versus Genbank nr protein database employing Blast2GO at E-value <1E-06.^{70,71}

DNA methylation analysis using bisulfite sequencing PCR. The MethPrimer database was used to design BSP primers (primer sequences and annealing temperatures used are listed in Table S14).⁷² The EZ DNA Methylation kit (Cambridge Biosciences, UK) was used for bisulfite conversion according to manufacturer's protocol. Briefly, for each sample, 500 ng of genomic DNA was bisulfite treated and amplified using Zymo Taq DNA polymerase (Cambridge Biosciences, UK). The PCR products were analyzed via DNA gel electrophoresis followed by sequencing using an ABI3730 DNA analyzer.

Design of flatfish dab specific 8 x 15k oligo microarray based on contigs achieved from MeDIP de novo high-throughput sequencing. The oligonucleotide microarray was designed via Agilent eArray (Agilent technologies, Berkshire, UK). The array was designed using the 1,693 differentially methylated contigs with known protein-coding sequences. These contigs were supplemented with additional sequences to produce a more comprehensive resource for dab transcriptomics. 84 of the most highly

differentially methylated unidentifiable CGIs were included, necessitating design of both forward and reverse array probes. Additional putative protein coding sequences were derived from the CGI contigs that displayed no differential methylation (7,416), the HCA and ST assemblies >60 bases that aligned with flatfish ESTs (1,599) and that matched with other protein coding sequences (3,864) and 5 gene sequences from other flatfish species. These were manually curated to reduce duplication of probes for sequences aligning with the same orthologs. In total 14,951 sequences were uploaded to Agilent eArray, resulting in successful design of 14,919 60-mer probes. The microarrays were printed in 8 x 15k format with each sub-array including 14,919 experimental spots, 290 empty positions and 77 negative and 459 positive Agilent control spots (Agilent-031032_8 x 15k_BhamDab). The array details and experimental design (Transcription profiling of dab hepatocellular adenoma liver) are available from ArrayExpress under accession A-MEXP-2084 and A-MEXP-2084, respectively.

Sample labeling and hybridization. The Agilent Two Color Low Input Linear Amplification Kit with Spike-ins (Agilent technologies, Berkshire, UK) and the protocol provided were used for labeling. The total RNA from both HCA and ST was converted to cRNA and labeled with fluorophore Cy5. A reference sample was generated by combining the RNA from all samples. The total RNA from reference sample was labeled with fluorophore Cy3. Any uncoupled Cy3 and Cy5 were removed using the RNeasy mini kit (Qiagen Ltd., West Sussex, UK) following manufacturer's instructions. Labeling efficiency was determined by NanoDrop ND-1000 UV-VIS Spectrophotometer, version 3.2.1 (Nanodrop, USA). A yield of 0.825 µg cRNA, and a specific activity (Cy3 and Cy5 dye incorporation) of >6 pmol Cy3/µg cRNA was considered sufficient.

Microarray hybridization was performed using an Agilent Gene Expression Hybridisation kit according to the manufacturer's protocol. Briefly, the hybridization mixes were loaded onto 8 x 15K format slides (Agilent-031032_8 x 15k_BhamDab) hybridized overnight (65°C), washed, stabilized and dried. The dried slides were scanned using an Agilent G2565CA microarray scanner system (Agilent technologies, Berkshire, UK).

Statistical analyses. MIAME-compliant raw microarray data were submitted to ArrayExpress at EMBL-EBI and can be found under accession A-MEXP-2084. GeneSpring vGX11 (Agilent, UK) and MultiExperimental Viewer v4.7⁷³ were used for analyzing the data. Agilent feature extraction software (FE) was used for quality control, both at sample and gene level. The samples that did not pass the 8 feature extraction values were removed from further analysis. In addition, saturated or below background level entities were also removed from analysis. Data flagged as present or marginal in at least 3 out of 32 samples were used for analysis. Following background correction, log intensity and log ratio calculation, data were normalized by Lowess. Data were standardized by calculating the mean and SD for each data point (x) in all samples (standardized value = x - mean/SD). Statistically differentially expressed genes were identified using Welch paired t-tests between matching HCA and ST, one way ANOVA between the three groups (HCA, ST and healthy liver),

and un-paired t-test between HCA and healthy liver (p value < 0.05) with a fold change of greater than 1.5 and Benjamini-Hochberg (B-H) correction.⁷⁴ Welch paired t-test was used for comparing HCA and ST samples as these samples could be readily paired coming from the same fish.

The "Core Analysis" function of Ingenuity Pathway Analysis (IPA)⁷⁵ was used to aid interpretation of the hypo- and hypermethylated genes as well as under and overexpressed genes (fold change >1.5). A Benjamini and Hochberg multiple testing correction was used to determine significant enrichment of annotation with biological functions and canonical pathways amongst these gene sets.

Real time PCR. Real time PCR was conducted to validate the microarray data for selected probes which significantly varied between HCA and ST or HCA and healthy dab liver. Primers were designed using Primer3⁷⁶ and synthesized by AltaBioscience (Birmingham, UK). Primer sequences are presented in Table S15. Primers were validated using cDNA (80 ng) and DreamTaq DNA polymerase (Fermentas Life Sciences, UK) as recommended by manufacturer's guidelines. PCR products were sequenced on a capillary sequencer ABI3730 (Functional Genomics and Proteomics Unit, Birmingham, UK) and compared to the sequences obtained from HTS.

RT-PCR was conducted on an ABI Prism 7000 (Applied Biosystems, USA) using SYBR Green SensiMix (Quantace, Watford, UK). Five biological replicates, with three technical replicates of each were run, containing 80 ng cDNA with cycling parameters of 95°C for 30 s (denaturing step) and 60°C for 30 s (combined annealing and extension). Melting curves were plotted using the ABI Prism 7000 SDS software to ensure only a single product was amplified and no primer dimers were formed. Absolute fluorescence values were used to calculate PCR efficiencies for each well.⁷⁷ Threshold cycle (CT) values were recorded in the linear phase of amplification and the data were analyzed using the delta-delta CT method of relative quantification.⁷⁸ The internal reference gene (*syntrophin, beta2*) was selected as unvarying during the microarray experiment and was used to further normalize the data.

Conclusions

In this study we demonstrated the suitability of MeDIP-HT de novo DNA sequencing for studying DNA methylation in unsequenced species. We also showed for the first time, a link between methylation abnormalities and liver tumors in wild fish. In addition, we proposed a hypothetical mechanism behind formation of dab liver tumors that arise in wild fish chronically-exposed to contaminants such as EDCs. It was shown that methylation and transcription of cancer related pathways were altered in HCA compared to ST with inter-individual variation. In addition, the information gathered on variation levels observed between groups of fish sampled from different sampling sites in both gene expression and BSP experiments will aid power studies to determine the sample size required in the design of any future experiments. Our data support the epigenetic model of tumorigenesis in which methylation changes initiate tumorigenesis and

therefore have the potential to be used for identification of tumor bearing livers, possibly even prior to manifestation of microscopic tumors. In particular, hypomethylation and the methylation status of DNA methyltransferases and the genes associated with the non-genomic and genomic estrogen pathways have the potential to be used as early biomarkers of environmentally-induced epigenetic change, thus acting as indicators for identification of potential tumor inducing, hazardous environments for marine species. However, further studies are required to investigate if the observed transcription and methylation changes are consequences or causes of these tumors. Nevertheless, it appears that there is a link between environment and dab tumorigenesis as the prevalence of these tumors differed between sampling sites. Overall, our study highlights the importance of incorporating and considering epigenetic mechanisms, especially in chronic

References

- Pogribny IP. Epigenetic events in tumorigenesis: putting the pieces together. *Exp Oncol* 2010; 32:132-6.
- Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med* 2004; 10:789-99.
- Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000; 100:57-70.
- Gronbaek K, Hother C, Jones PA. Epigenetic changes in cancer. *APMIS* 2007; 115:1039-59.
- Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. *Carcinogenesis* 2010; 31:27-36.
- Lopez J, Perchard M, Coley HM, Webb A, Crook T. The context and potential of epigenetics in oncology. *Brit J Cancer* 2009; 100:571-7.
- Anway MD, Skinner MK. Epigenetic transgenerational actions of endocrine disruptors. *Endocrinology* 2006; 147:43-9.
- Siegfried Z, Simon I. DNA methylation and gene expression. *WIREs* 2010; 2:362-71.
- Jones PA. The DNA methylation paradox. *Trends Genet* 1999; 15:34-7.
- Tischoff I, Tannapfel A. DNA methylation in hepatocellular carcinoma. *World J Gastroentero* 2008; 14:1741-8.
- Esteller M, Heman JG. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumors. *J Pathol* 2002; 196:1-7.
- Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypermethylation profile of human cancer. *Cancer Res* 2001; 61:3225-9.
- Reamon-Buettrner SM, Mutschler V, Borlak J. The next innovation cycle in toxicogenomics: environmental epigenetics. *Mutat Res* 2008; 659:158-65.
- Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006; 7:21-33.
- Wild CP. Environmental exposure measurement in cancer epidemiology. *Mutagenesis* 2009; 24:117-25.
- Klaunig JE, Kamendulis LM, Xu Y. Epigenetic mechanisms of chemical carcinogenesis. *Hum Exp Toxicol* 2000; 19:543-55.
- Bollati V, Baccarelli A. Environmental epigenetics. *Heredity* 2010; 105:112.
- Baccarelli A, Bollati V. Epigenetics and environmental chemicals. *Curr Opin Pediatr* 2009; 21:243-51.
- Skinner MK, Manikkam M, Guererro-Bosagna C. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends Endocrinol Met* 2010; 21:214-22.
- Arita A, Costa M. Epigenetics in metal carcinogenesis: nickel, arsenic, chromium and cadmium. *Metallomics* 2009; 1:222-8.
- Stromqvist M, Tooke N, Brunstrom B. DNA methylation levels in the 5' flanking region of the vitellogenin I gene in liver and brain of adult zebrafish (*Danio rerio*)-sex and tissue differences and effects of 17 α -ethynodiol exposure. *Aquat Toxicol* 2010; 98:275-81.
- Aniagu SO, Williams TD, Allen Y, Katsiadaki I, Chipman JK. Global genomic methylation levels in the liver and gonads of the three-spine stickleback (*Gasterosteus aculeatus*) after exposure to hexabromocyclododecane and 17-[beta] oestradiol. *Environ Int* 2008; 34:310-7.
- Zhou XW, Zhu GN, Jilisa M, Sun JH. Influence of Cu, Zn, Pb, Cd and their heavy metal ion mixture on the DNA methylation level of the fish (*Carassius auratus*). *China Environ Sci* 2001; 21:549-52.
- Contrator RG, Foran CM, Li S, Willett KL. Evidence of gender- and tissue-specific promoter methylation and the potential for ethynodiol-induced changes in Japanese medaka (*Oryzias latipes*) estrogen receptor and aromatase genes. *J Toxicol Env Health* 2004; 67:1-22.
- Mirbahai L, Williams TD, Zhan H, Gong Z, Chipman JK. Comprehensive profiling of zebrafish hepatic proximal promoter CpG island methylation and its modification during chemical carcinogenesis. *BMC Genomics* 2011; 12:1-16.
- Vandegehuchte MB, Kyndt T, Vanholme B, Haegeman A, Ghysen G, Janssen CR. Occurrence of DNA methylation in *Daphnia magna* and influence of multigeneration Cd exposure. *Environ Int* 2009; 35:700-6.
- Vandegehuchte MB, Lemiere F, Janssen CR. Quantitative DNA methylation in *Daphnia magna* and effects of multigeneration Zn exposure. *Comp Biochem Physiol* 2009; 150:343-8.
- Vandegehuchte MB, Lemiere F, Vanhaecke L, Vanden Berghe W, Janssen CR. Direct and transgenerational impact on *Daphnia magna* of chemicals with a known effect on DNA methylation. *Comp Biochem Physiol* 2010; 151:278-85.
- Shugart LR. 5-methyl deoxycytidine content of DNA from bluegill sunfish (*Lepomis macrochirus*) exposed to benzo[a]pyrene. *Environ Toxicol Chem* 1990; 9:205-8.
- Wang Y, Wang C, Zhang J, Chen Y, Zuo Z. DNA hypomethylation induced by tributyltin, triphenyltin, and a mixture of these in *Sebastiscus marmoratus* liver. *Aquat Toxicol* 2009; 95:93-8.
- Anon, UK national marine monitoring programme. Second report. Marine Environment Monitoring group, Cefas, UK 2004.
- Stentiford GD, Bignell JP, Lyons BP, Feist SW. Site-specific disease profiles in fish and their use in environmental monitoring. *MEPS* 2009; 381:1-15.
- Small HJ, Williams TD, Sturve J, Chipman JK, Southam AD, Bean TP, et al. Gene expression analyses of hepatocellular adenoma and hepatocellular carcinoma from the marine flatfish *Limanda limanda*. *DAO* 2010; 88:127-41.
- Southam AD, Easton JM, Stentiford GD, Ludwig C, Arvanitis TN, Viant M. Metabolic changes in flatfish hepatic tumors revealed by NMR-based metabolomics and metabolic correlation networks. *J Proteome Res* 2008; 7:5277-85.
- Lyons BP, Stentiford GD, Bignell J, Goodsir F, Sivyer DB, Devlin MJ, et al. A biological effects monitoring survey of Cardigan Bay using flatfish histopathology, cellular biomarkers and sediment bioassays: findings of the Prince Madog prize 2003. *Mar Environ Res* 2006; 62:342-6.
- Rothcell JM, Lee J, Chipman JK, Ostrander GK. Structure, expression and activation of fish ras genes. *Aquat Toxicol* 2001; 55:1-21.
- Franklin TM, Lee J, Kohler A, Chipman JK. Analysis of mutations in the p53 tumor suppressor gene and ki- and ha-ras proto-oncogenes in hepatic tumors of European flounder (*Platichthys flesus*). *Mar Environ Res* 2000; 50:251-5.
- Jabbbari K, Bernardi G. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene* 2004; 333:143-9.
- Gardiner-Garden M, Frommer M. CpG islands in vertebrate genome. *J Mol Biol* 1987; 196:261-82.
- Mirbahai L, Aniagu SO, Williams TD, Chipman JK. Comparative quantification of genome-wide methylation in fish liver using reverse-phase high performance liquid chromatography (HPLC). PRIMO 15. Bordeaux, France. 17-20 May 2009.
- Down TA, Rakyan VK, Turner DJ, Flicek P, Li H, Kulesha E, et al. A Bayesian deconvolution strategy for immunoprecipitation based DNA methylome analysis. *Nat Biotechnol* 2008; 26:779-85.
- Liang G, Salem CE, Yu MC, Nguyen HD, Gonzales FE, Nguyen TT, et al. DNA methylation differences associated with tumor tissues identified by genome scanning analysis. *Genomics* 1998; 53:260-8.
- Baylin SB, Esteller M, Rountree MR, Bachman KE, Schuebel K, Herman JG. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum Mol Genet* 2001; 10:687-92.
- Polakis P. Wnt signaling and cancer. *Gene Dev* 2000; 14:1837-51.
- Mann B, Gelos M, Siedow A, Hanski ML, Gratchev A, Ilyas M, et al. Target genes of β -catenin-T cell-factor/lymphoid-enhancer-factor signaling in human colorectal carcinomas. *PNAS* 1999; 96:1603-8.

exposures, for establishing acceptable levels of contaminants in marine environments, which may subsequently affect the health of organisms including humans.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors would like to thank the Functional Genomics and Proteomics Unit, Birmingham, UK, especially Lorraine Wallace, and Centre for Environment, Fisheries and Aquaculture Science (Cefas, Weymouth, UK), specifically Dr. Brett Lyons and Dr. Grant Stentiford for their input throughout this project.

This work was funded by Natural Environment Research Council (NERC).

46. Adrian C, Creagh EM, Martin SJ. Apoptosis-associated release of Smac/Diablo from mitochondria requires active caspases and is blocked by bcl-2. *EMBO J* 2001; 20:6627-38.
47. Kohli M, Yu J, Seaman C, Bardelli A, Kinzler KW, Vogelstein B, et al. SMAC/Diablo-dependent apoptosis induced by nonsteroidal anti-inflammatory drugs (NSAIDs) in colon cancer cells. *PNAS* 2004; 101:16897-902.
48. Tsai N, Wei L. RhoA/Rock1 signaling regulates stress granules formation and apoptosis. *Cell Signal* 2010; 22:668-75.
49. Coleman ML, Sahai EA, Yeo M, Bosch M, Dewart A, Olson MF. Membrane blebbing during apoptosis results from caspase-mediated activation of ROCK 1. *Nat Cell Biol* 2001; 3:339-46.
50. Hu W, Johnson H, Shu H. Activation of NF κ B by FADD, casper and caspase-8. *J Biol Chem* 2000; 275:10838-44.
51. Zhang S, Wang S, Li H, Li L. Vitellogenin, a multivalent sensor and an antimicrobial effector. *Int J Biochem Cell Biol* 2011; 43:303-5.
52. Sumpter JP, Jobling S. Vitellogenesis as a biomarker for estrogenic contamination of the aquatic environment. *Environ Health Persp* 1995; 103:173-8.
53. Scott AP, Sanders M, Stentiford GD, Reese RA, Katsiadaki I. Evidence for estrogenic endocrine disruption in an offshore flatfish, the dab (*Limanda limanda*). *Mar Environ Res* 2007; 64:128-48.
54. Stentiford GD, Feist SW. First reported cases of intersex (ovotestis) in the flatfish species dab *Limanda limanda*: Dogger Bank, North Sea. *MEPS* 2005; 301:307-10.
55. Charting Progress 2 Feeder report: clean and safe seas. UKMMS, DEFRA, UK 2010.
56. Stentiford GD, Bignell JP, Lyons BP, Thain JE, Feist SW. Effect of age on liver pathology and other diseases in flatfish: implications for assessment of marine ecological health status. *MEPS* 2010; 411:215-30.
57. Siddiqi MA. Polybrominated diphenyl ethers (PBDEs): new pollutants-old diseases. *CMR* 2003; 4:281-90.
58. Vandegheuchte MB, Janssen CR. Epigenetics and its implications for ecotoxicology. *Ecotoxicology* 2011; 20:607-24.
59. Li S, Hursting SD, Davis BJ, McLachlan JA, Barrett JC. Environmental exposure, DNA methylation and gene regulation. *Ann NY Acad Sci* 2003; 983:161-9.
60. Lo P, Sukumar S. Epigenomics and breast cancer. *Pharmacogenomics* 2008; 9:1879-902.
61. Tilton SC, Givan SA, Pereira CB, Bailey GS, Williams DE. Toxicogenomics profiling of the hepatic tumor promoters indole-3-carbinol, 17 β -estradiol and β -naphthoflavone in rainbow trout. *Toxicol Sci* 2006; 90:61-72.
62. Bjornstrom L, Sjoberg M. Mechanisms of estrogen receptor signaling: Convergence of genomic and non-genomic actions on target genes. *Mol Endocrinol* 2005; 19:833-42.
63. Obrero M, Yu DV, Shapiro DJ. Estrogen receptor-dependent and estrogen receptor-independent pathways for tamoxifen and 4-hydroxytamoxifen-induced programmed cell death. *J Biol Chem* 2002; 277:45695-703.
64. Safe S, Kim K. Nonclassical genomic ER/Sp and ER/AP-1 signaling pathways. *J Mol Endocrinol* 2008; 41:263-75.
65. Dolcet X, Liobet D, Pallares J, Matias-Guiu X. NF κ B in development and progression of human cancer. *Virchows Arch* 2005; 446:475-82.
66. Rosen N, She QB. AKT and cancer-is it all mTOR? *Cancer Cell* 2006; 10:254-6.
67. Feist SW, Lang T, Stentiford GD, Koehler A. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. *ICES Tech Mar Environ Sci* 2004; 38:1-42.
68. Ramsahoye BH. Measurement of genome wide DNA methylation by reversed-phase high-performance liquid chromatography. *Methods* 2002; 27:156-61.
69. Li N, Ye M, Li Y, Yan Z, Butcher LM, Sun J, et al. Whole genome DNA methylation analysis based on high-throughput sequencing technology. *Methods* 2010; 52:203-13.
70. Gotz S, Garcia-Gomez J, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *NAR* 2008; 36:3420-35.
71. Conesa A, Gotz S, Garcia-Gomez J, Terol J, Talon M, Robles M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005; 21:3674-6.
72. Methprimer software web link [<http://www.urogene.org/methprimer/index1.html>]
73. MultiexpEriment Viewer software web link [<http://www.tm4.org/mev/>]
74. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 1995; 57:289-300.
75. Ingenuity Pathway Analysis software web link [<http://www.ingenuity.com/>]
76. Rozen S, Skaltsky HJ. Primer3 on the www for general users and for biologist programmers. *Methods Mol Biol* 2000; 365:86.
77. Ramakers C, Ruijter JM, Lekanne RH, Antoon FM. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* 2003; 339:62-6.
78. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods* 2001; 25:402-8.

© 2011 Landes Bioscience.
Do not distribute.