Integrated Omics Profiling Reveals Novel Patterns of Epigenetic Programming in Cancer-Associated Myofibroblasts

Hanna Najgebauer[1,*], Triantafillos Liloglou[2], Puthen V. Jithesh[2], Olivier T. Giger[1], Andrea Varro[1,3], Christopher M. Sanderson[1,*]

[1] Department of Cellular and Molecular Physiology, University of Liverpool, Crown Street, Liverpool L69 3BX, UK

[2] Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Crown Street, Liverpool L69 3BX, UK

[3] Department of Medicine, University of Szeged, Hungary

* Correspondence: Tel: +44 (0)151 794 4180; Fax: +44(0)151794 4434; Email: cmsand@liverpool.ac.uk

Correspondence may also be addressed to Hanna Najgebauer, Email: hanna.najgebauer@gmail.com

Present Addresses: Hanna Najgebauer, European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK; Puthen V. Jithesh, Biomedical Informatics Division, Sidra Medicine, PO Box 26999, Doha, Qatar; Olivier T. Giger, Division of Cellular and Molecular Histopathology, University of Cambridge, Hills Road, Cambridge, CB2 0QQ, United Kingdom.

ABSTRACT

There is increasing evidence that stromal myofibroblasts play a key role in tumour development, however the mechanisms by which they become reprogrammed to assist in cancer progression remain unclear. As cultured Cancer Associated Myofibroblasts (CAMs) retain an ability to enhance the proliferation and migration of cancer cells *in vitro*, it is possible that epigenetic reprogramming of CAMs within the tumour microenvironment may confer long-term pro-tumorigenic changes in gene expression. This study reports the first comparative multi-omics analysis of cancer-related changes in gene expression and DNA-methylation in primary myofibroblasts derived from gastric and oesophageal tumours. In addition, we identify novel CAM-specific DNA methylation signatures, which are not observed in patient-matched Adjacent Tissue-derived Myofibroblasts (ATMs), or corresponding Normal Tissue-derived Myofibroblasts (NTMs). Analysis of correlated changes in DNA methylation and gene expression show that different patterns of gene-specific DNA methylation have the potential to confer pro-tumourigenic changes in metabolism, cell signalling and differential responses to hypoxia. These molecular signatures provide new insights into potential mechanisms of stromal reprogramming in gastric and oesophageal cancer, while also providing a new resource to facilitate biomarker identification and future hypothesis driven studies into mechanisms of stromal reprogramming and tumour progression in solid tumours.

INTRODUCTION

Despite a recent decline in incidence [1] distal gastric cancer remains the third leading cause of cancer death worldwide [2]. As such, there remains a need to develop a better understanding of the molecular processes that contribute to the development and metastasis of gastric tumours.

There is strong evidence that Cancer Associated Myofibroblasts (CAMs), exhibit tumour-promoting properties, which are not observed in myofibroblasts derived from non-cancerous tissue [3-7]. Also, the presence of large numbers of CAMs within the tumour stroma are linked to poor prognosis [8, 9] and resistance to therapy [10]. Significantly, these tumour-promoting properties are retained in cultured gastric CAMs, when compared to patient matched Adjacent Tissue Myofibroblasts (ATMs) or Normal Tissue Myofibroblasts (NTMs) [11, 12], implying that some pro-tumourigenic properties may result from epigenetic reprogramming of CAMs within the tumour microenvironment. Interestingly, gastric and oesophageal CAMs both have distinct miRNA signatures compared to corresponding populations of ATMs and NTMs [13]. Also, gastric CAMs were reported to exhibit a global reduction in DNA methylation compared to patient matched ATMs [14]. However, previous studies were not performed at resolution, which allowed the mechanism, or functional consequences of CAM specific DNA methylation changes to be assessed.

This study, presents the first comparative genome-wide analysis of DNA-methylation patterns at individual CpG resolution in primary gastric CAMs, patient-matched ATMs and unrelated gastric NTMs. Significantly, a subset of these molecular signatures was also observed in oesophageal adenocarcinoma derived CAMs; suggesting that common mechanisms of stromal programming may operate in tumours derived from glandular cells in different tissues of the upper gastrointestinal tract. CAM specific methylation patterns also provide potential stromal biomarkers, which may improve stratification and prognosis of both gastric and oesophageal tumours. To confirm the robust nature of cancer related signatures identified in this study, comparative changes in the methylation status of a selection of genes with potential clinical relevance (SMAD3, SPON2, FOXF1, FENDRR) were validated in an independent set of gastric CAMs, patient-matched ATMs and unrelated NTMs by pyrosequencing and qPCR.

MATERIAL AND METHODS

Generation and culture of human primary myofibroblasts

Human primary myofibroblasts derived from resected gastric and oesophageal adenocarcinomas (CAM) and adjacent tissue (ATM) were obtained from patients undergoing gastric or oesophageal cancer surgery (Supplementary Table S1), as described previously [11, 12]. Normal gastric myofibroblasts (NTM) were generated from deceased transplant donors with normal morphology (Supplementary Table S2) as previously reported [15]. This work had been approved by the Ethics Committee of the University of Szeged, Hungary. Myofibroblasts were authenticated by qPCR and immunocytochemistry as described previously [11, 12]. The analysis showed positive expression of α-smooth muscle actin and vimentin (myofibroblast and mesenchymal markers), and lack of desmin (pericyte marker) and cytokeratin (epithelial cell marker) expression. Cells were tested prior usage in a new study or at least every six months by immunocytochemistry to ensure that their phenotype is maintained. Primary myofibroblast cultures were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS), 1% penicillin-streptomycin, 1% antibiotic-antimycotic and 1% non-essential amino acid solution. Medium was replaced routinely every 48–60 hours and cells were lysed at 80-90% confluence for DNA and RNA extraction. In all experiments myofibroblast cells were not used beyond passage 12.

Myofibroblast conditioned media preparation

To prepare CAM, ATM or NTM conditioned media (CM), 1.5 x 10^6 myofibroblast cells were seeded in 75cm^2 tissue culture flasks and left to attach for 24h. The next day the cells were washed 3 times in 1x PBS to get rid of any serum-derived factors. Then growth media was replaced with 15ml freshly prepared serum-free DMEM supplemented with 1% penicillin-streptomycin, 1% antibiotic-antimycotic and 1% non-essential amino acid solution and incubated for 24h at 37 °C in a humidified atmosphere with 5% CO2. The next day CM was collected and centrifuged at 800g for 7min to get rid of cell debris. The freshly prepared supernatants were immediately used for cancer cell migration and proliferation assays.

Cancer cell migration assay

The effects of myofibroblast conditioned media (CM) on gastric cancer cells migration was measured *in vitro* using transwell Boyden chamber assay (SLS; cat. no. 354578). Briefly, 1 x 10^4 AGS cells in 500µl serum-free DMEM were added to the 8µm pore chambers. The lower chambers contained either 750µl serum-free media, or myofibroblast CM to serve as a chemoattractant. Cells were incubated at 37°C and allowed to migrate overnight. Cells migrating through the membrane were fixed and detected on the lower surface using Reastain Quick-Diff Kit (Reagena; cat. no. 102164). The total cells in 15 fields per well were counted, and the mean of at least 3 independent membranes per experiment was taken.

Cancer cell proliferation assay

The effects of myofibroblast conditioned media (CM) on gastric cancer cell proliferation was assessed by incorporation of EdU (Salic and Mitchison 2008) and detected using the Click-iT EdU Alexa Fluor 488 Imaging Kit (Life Technologies; cat. no. C10337) according to the manufacturer's instruction.

DNA and RNA extraction

Genomic DNA was purified using a standard phenol/chloroform extraction method. Briefly, myofibroblast cells were lysed in Lysis Buffer (400mM Tris-HCl, pH 8, 10mM EDTA, 1% SDS, 150mM NaCl). DNA quantity was assessed using PicoGreen fluorimetry (Life Technologies, cat. no. Q-33130). DNA samples were analysed by Molecular Genetic Services (Gen-Probe Life Sciences, Manchester) using Illumina Infinium HumanMethylation450k BeadChip arrays for analysis of DNA methylation. Total RNA was purified using miRNeasy Mini Kit (Qiagen, cat. no. 217004) and sample degradation and purity was assessed using the Agilent 2100 Bioanalyzer and RNA 6000 Nano kit (Agilent Technologies, cat. no. 5067-1512). Samples were sent to Molecular Genetic Services (Gen-Probe Life Sciences, Manchester) for gene expression analysis using Illumina HumanHT-12 v4 arrays.

Pyrosequencing analysis

DNA was extracted from 7 patient-matched CAM and ATM samples and 4 unrelated NTM samples using Wizard SV Genomic DNA Purification kits (Promega, cat. no. A2360). In each case, 1µg of genomic DNA was treated with sodium bisulphite using the EZ DNA Methylation-Gold Kit (ZymoResearch, cat. no. D5005). A full list of assays, primer sequences and annealing temperatures is shown in Supplementary Table S3. Pyrosequencing templates were prepared by PCR amplification using HotStarTaq Master Mix Kit (Qiagen, cat. no. 203603), 5µM biotinylated primer, 5-10µM non-biotinylated primer (corresponding to 1:1 or 1:2 ratio in Supplementary Table S3), 5mM dNTPs (Qiagen, cat. no. 201900) and 3 µl (~60 ng) bisulfite-treated DNA. The PCR thermal profile consisted of initial denaturation at 95°C for 5 min, followed by 40 cycles including 95°C for 30 sec, annealing temperature (Supplementary Table S3) for 30 sec, 72°C for 30 sec. A final extension step of 72°C for 10 min was also included. Purified biotinylated PCR products were made single stranded to act as a template in a pyrosequencing reaction run. PCR products were bound to streptavidin-coated Sepharose beads (GE Healthcare, cat. no. 17-5113-01), before washing, and denaturising in 0.2M NaOH. 0.5uM. Pyrosequencing primers were annealed to the purified single-stranded PCR products, and pyrosequencing was carried out using the Pyromark 96ID System (Qiagen). The methylation index for the analysed genomic region was calculated as the mean value of mC/(mC + C) for all examined CpG sites in the interrogated genomic region.

qPCR (TaqMan) expression analysis

TaqMan gene expression assays were used to quantify mRNA levels of target genes in stromal myofibroblasts. RNA samples were extracted from 6 patient-matched gastric CAM and ATM samples. Total RNA was purified using miRNeasy Mini Kit (Qiagen, cat. no. 217004) and QuantiTect Reverse Transcription Kits (Qiagen, cat. no. 205311) were used for cDNA synthesis and qPCR assays performed on the StepOne system (Applied Biosystems). Amplification mixture contained 7.5µl of 2x

TaqMan Universal Master Mix II (Life Technologies, cat. no. 4440042), 0.75µl of 20x TaqMan probe and primers, 1.25µl 10x ACTB, 2µl of the cDNA and 3.5µl ddH_2O, giving a final volume of 15µl. TaqMan assays were either designed using Oligo 7.0 software (Molecular Biology Insights Inc, USA) and synthesized by Eurofins MWG (Germany), or purchased as predesigned assays from Life Technologies (UK). A list of assays, nucleotide sequences and PCR product sizes are shown in Supplementary Table S4. Amplification mixtures were processed using standard conditions (50°C for 2 minutes and 95°C for 10 minutes followed by 45-50 cycles at 95°C for 15 seconds and 60°C or 61°C for 1 minute). β-actin was used as the endogenous control. The comparative ΔΔCt method was used to compute relative levels of target gene expression, subtracting Ct values of the endogenous control (β-actin) before comparing values to a calibrator sample, where the calibrator sample = 1.0 and other samples were expressed as n-fold relative to the calibrator.

Methods for data processing and bioinformatics analysis are provided in (Supplementary File S1).

RESULTS

Gastric CAMs retain a pro-tumorigenic phenotype *in vitro* and exhibit a global reduction in DNA methylation

Previous studies have shown that primary CAMs retain pro-tumorigenic properties following isolation and culture [16], including the ability to enhance cancer-cell migration and proliferation [11, 12, 17]. This phenotype was confirmed for all primary gastric myofibroblast populations used in this study (Supplementary Figure S1 and Figure S2). Illumina 450k probes that passed stringent filtering criteria were used to compute mean β-values as an indication of the global DNA methylation status of gastric myofibroblasts. In agreement with reported trends [14], gastric CAMs used in this study all exhibited a global reduction in DNA methylation compared to patient-matched ATMs (Wilcoxon test, $p < 2.2 \times 10^{-16}$) (Supplementary Figure S3A).

Genome-wide DNA methylation profiling of gastric myofibroblasts purified from different tissue microenvironments

To provide new insight into epigenetic changes that may be linked to the tumour-promoting properties of gastric CAMs, a comparative genome-wide DNA methylation analysis was performed on sets of gastric CAMs, patient-matched ATMs and unrelated NTMs, using the Illumina Infinium HumanMethylation450 BeadChip. This analysis identified numerous CpG sites that show consistent differences in DNA methylation in CAMs compared to either ATMs or NTMs. In total, 5688 differentially methylated CpG sites were identified in CAMs compared to ATMs, including 3404 hypomethylated and 2284 hypermethylated CpG sites. These loci were more frequently located in CpG shores than CpG islands. Comparison of the overall distribution of differentially methylated loci relative to RefSeq genes showed that hypomethylated CpG loci were overrepresented in promoters, gene bodies and intergenic regions (Figure 1, Figure 2 outer tract). In the CAM vs NTM comparison, a total of 8104 differentially methylated CpG loci were identified, including 4147 and 3957 loci that were respectively hypo- or hyper-methylated in CAMs. These genome-wide and gene-specific methylation patterns provide important signatures that facilitate differentiation between tumour-derived myofibroblasts (CAMs) and non–tumour derived myofibroblasts (ATMs or NTMs). Equally, identification of CAM-specific DNA methylation changes may aid biomarker identification for improved diagnosis/prognosis, or tumour/patient stratification. Therefore, a comparative analysis was performed to identify CpG loci that distinguish CAMs from non-tumour derived myofibroblasts, following CAM vs ATM and CAM vs NTM differential methylation analysis. The resulting overlap of 2006 CpG loci from these comparisons was then selected. Multiple CpG loci were found to be hypomethylated in CAMs but hypermethylated in both ATMs and NTMs and vice versa (Supplementary Figure S4; Figure 2 heatmap). As these genomic regions show distinct DNA methylation patterns in myofibroblast populations, it is possible that they facilitate distinction between different types of gastric myofibroblasts and different stages, or degrees of tumour reprogramming.

Technical validation of novel cancer related changes in DNA methylation

Pyrosequencing assays were performed to validate the methylation level of CpG sites identified in gastric CAMs by Illumina 450k arrays. Supplementary Figure S5 shows correlations between β-values and methylation index assessed by pyrosequencing in DNA samples that were used in the initial array experiments. Correlations between the two types of DNA methylation assays for 12 CpG loci interrogated were high (R^2 = 0.8177 - 0.9921, p-value = 7.1 x 10^{-3} – 1.41 x 10^{-7}), thus increasing confidence in the reliability of comparative differential DNA methylation trends identified in this study.

***In silico* molecular enrichment analysis**

To investigate the potential functional consequences of differentially methylated CpG loci identified in gastric CAMs, Gene Ontology (GO) enrichment and Ingenuity Pathway Analysis (IPA) were performed on subsets of genes with associated changes in methylation status. GO enrichment analysis revealed that aberrant DNA methylation has the potential to affect biological processes that are directly associated with tumour growth and progression, including: the regulation of cell development and differentiation, cell adhesion, chemotaxis, transmembrane transport, regulation of cholesterol biosynthesis and extracellular matrix organization (Supplementary File S2). Equally, Ingenuity Analysis identified processes involving aberrantly methylated genes, including gene expression, cellular movement, cell growth and proliferation, cellular development and morphology, cell signalling, energy production and lipid metabolism. Interestingly, TGF-β signalling was identified as one of the most overrepresented pathways that may be affected by cancer-imposed changes in CAM DNA methylation (Supplementary File S3).

Gene expression profiling of gastric myofibroblast purified from different tissue microenvironments

A comparative genome–wide gene expression analysis was performed on populations of gastric CAMs, ATMs and NTMs using the Illumina HumanHT-12v4 Expression BeadChip arrays. The analysis revealed 13381 consistently expressed genes.

Pairwise differential gene expression analyses performed between CAM vs ATM, CAM vs NTM and ATM vs NTM identified: 1215 genes (574 upregulated and 641 downregulated in CAMs) that were differentially expressed between CAMs and patient-matched ATMs; 987 genes that were differentially expressed between CAMs and unrelated NTMs (508 upregulated and 479 downregulated in CAMs) and 713 genes that were differentially regulated in ATMs compared to NTMs (407 upregulated and 306 downregulated in ATMs) (Figure 3A-C).

Technical validation of CAM-specific changes in gene expression

To validate the results from the differential gene expression analysis, candidate genes from CAM vs ATM and CAM vs NTM comparisons were selected and analysed by TaqMan qPCR (Figure 3D). Triplicate reactions were conducted on RNA samples that were used in the initial array experiment, and data were analysed using the comparative ΔΔCt method. Quantitative analysis of candidate genes confirmed the expression patterns observed by Illumina HT-12 array analysis, thus increasing confidence in the identified comparative differential gene expression trends.

***In silico* enrichment analysis**

To assess processes that may be altered by induced changes in gene expression in gastric CAMs, a combination of Gene Ontology (GO) and Gene Set Enrichment (GSEA) analysis was performed (Supplementary Figure S6, File S4 and File S5). Significantly, 20% of genes differentially expressed between CAMs and patient-matched ATMs are known components of secretory exosomes [18]. While these data provide a new resource to drive future studies into the mechanisms by which stromal myofibroblasts promote tumour growth, it is important to establish which of the differential gene expression signatures may result from CAM-specific changes in DNA methylation.

Integration of CAM-specific DNA methylation and gene expression profiles

Although the relationship between DNA methylation and gene expression is complex, in some cases, the degree of promoter methylation is inversely correlated with gene expression, while methylation in gene bodies often shows a positive correlation with gene expression [19, 20]. This study identified a subset of 419 genes that show CAM-specific changes in both DNA methylation and gene expression; 230 of these genes are differentially methylated upstream of their annotated transcriptional start site, while 254 genes show altered DNA methylation patterns within gene bodies, with 65 genes exhibiting changes in both regions (Supplementary Figure S7).

Of the 124 genes with coordinated changes in promoter methylation and gene expression, 55 were hypermethylated and down-regulated while 69 were hypomethylated and up-regulated in comparison to corresponding ATMs (Supplementary Figure S7B). This data suggests that hypo- and hyper-methylated loci encode functionally distinct genes. In particular, genes that were hypermethylated and down-regulated in CAMs were over represented in highly relevant functional classes, including: gastrointestinal adenocarcinoma, gastrointestinal tract cancer, gastro-oesophageal carcinoma and gastric cancer. In contrast, CAM-specific genes that were hypomethylated and up-regulated were not associated with these processes. However, they did show enrichment in processes linked to extracellular vesicular exosomes, transport of amino acids and secretion of molecules (Supplementary Table S5).

With respect to the 152 genes that showed correlated changes in gene body methylation and gene expression (Supplementary Figure S7B), 79 genes that were hypomethylated and repressed were primarily associated with invasion, proliferation, transformation, transport of molecule and migration, while the 73 genes that were hypermethylated and induced show greater association with metabolic processes such as metabolism of amino acids and metabolism of heparin sulphate proteoglycans (Supplementary Table S6).

Verification of CAM-specific DNA methylation and gene expression signatures

To investigate the impact of cancer induced changes in DNA methylation on gene expression, pyrosequencing methylation and qPCR expression analyses were performed on an additional set of independent gastric CAMs, patient-matched ATMs and unrelated gastric NTMs, which were not

included in initial genome-wide profiling studies. Targeted pyrosequencing assays were performed on genomic regions associated with the regulation of a subset of candidate genes, including SMAD3, SPON2, FOXF1 and FENDRR, all of which have previously been implicated in cancer and tumour-stroma communication [18, 21-23], and were shown to be differentially methylated in this study.

Promoter hypermethylation represses SMAD3 expression in gastric CAMs

Two pyrosequencing assays were designed for SMAD3 covering a 224bp region spanning 10 CpG sites, including 2 CpGs identified in Illumina 450k arrays. These assays confirmed that the SMAD3 promoter is hypermethylated in CAMs (Figure 4A-B). Notably, DNA methylation levels in the SMAD3 promoter region were found to be very similar in ATMs and NTMs (Figure 4C), while qPCR expression analysis confirmed that SMAD3 is significantly downregulated in CAMs (Figure 4D). Therefore this genomic region may provide a proxy/biomarker for gastric CAM identification. Interestingly, SMAD3 hypermethylation in tumours is associated with poorer overall survival (Supplementary Figure S9). Collectively, these data provide a strong indication that SMAD3 expression may be repressed by cancer-induced reprogramming, resulting in SMAD3 promoter hypermethylation in gastric CAMs.

Promoter hypomethylation induces SPON2 expression in gastric CAMs

To investigate cancer-induced changes in SPON2 expression, a pyrosequencing assay was designed for SPON2 covering 117bp spanning 7 CpG sites, including 1 CpG site identified by the Illumina 450k array (Figure 4E-F). The pyrosequencing assay confirmed SPON2 promoter hypomethylation in CAMs, while qPCR assay confirmed that SPON2 expression is upregulated in CAMs. Interestingly, pyrosequencing data show that the extent of SPON2 promoter DNA methylation gradually changes in gastric stromal myofibroblasts, with low levels in CAMs, intermediate levels in patient-matched ATMs and high levels in NTMs (Figure 4G). Significantly, these trends show a good negative correlation with SPON2 expression patterns (Figure 4H, Figure 3D) and protein levels (Supplementary Figure S10) observed in secretome of these cells (unpublished data). Taken together, these data provide strong

evidence that SPON2 expression may be regulated by cancer-induced differential promoter DNA methylation in gastric CAMs, ATMs and NTMs.

DNA hypermethylation represses the expression of FOXF1 and FENDRR in gastric CAMs

A region on chromosome 16 spanning 526,426bp was identified as one of the largest differentially methylated regions in gastric CAM vs ATM and CAM vs NTM comparisons (Figure 5A). Notably, a smaller part of this region (ch16: 86,528,753 – 86,538,425) spanning 9,673bp was also identified as differentially methylated in oesophageal CAM vs ATM comparison (Supplementary Figure S11 and Figure 6). Differential DNA methylation within this region may regulate the expression of FOXF1 and several long non-coding RNAs (lncRNAs), including FENDRR.

To assess whether FOXF1 and FENDRR expression is regulated by DNA methylation in gastric CAMs, a pyrosequencing assay was designed to investigate the FOXF1 promoter region, which also overlaps with the FENDRR. The assay covers 104bp spanning 8 CpG sites (Figure 5B). This analysis confirms that the FOXF1 promoter region is consistently hypermethylated in CAMs when compared to either ATMs or NTMs. Notably, the FOXF1 promoter shows a gradual change in DNA methylation levels CAMs>ATMs>NTMs (Figure 5C). Interestingly, this genomic region is commonly hypermethylated in gastric cancer (Supplementary Figure S12) suggesting that CAMs might acquire some of the cancer-like DNA methylation patterns. In addition, qPCR analysis shows that FOXF1 and FENDRR expression are both downregulated in CAMs compared to ATMs (Figure 5D). Collectively, these data provide a strong indication that FOXF1 and FENDRR expression may be reprogrammed by DNA methylation within this region in both gastric CAMs and ATMs.

Data from these extended pyrosequencing and qPCR studies provide further evidence that CAM-specific promoter DNA methylation patterns may regulate the expression of associated genes. In addition, pyrosequencing analysis confirmed the existence of DNA methylation changes within border genomic regions, spanning several neighbouring CpG sites, between gastric CAMs and non-tumour derived myofibroblasts (ATMs and NTMs).

Comparison of DNA methylation patterns in CAMs derived from different tumour types

Although, the primary aim of this study was to identify genome-wide DNA methylation patterns in CAMs derived from gastric tumours, it was pertinent to question if CAMs that are reprogrammed in different adenocarcinomas from the upper GI tract show common changes in DNA methylation. To address this question, genome-wide DNA methylation profiles of primary patient-matched CAMs and ATMs, derived from oesophageal tumours were compared to signatures observed in gastric CAMs and ATMs.

The genome-wide analysis identified widespread DNA methylation alterations and confirmed that oesophageal CAMs also exhibit a global loss of DNA methylation compared to corresponding patient-matched ATMs (Supplementary Figure S3B-C). To identify common differentially methylated genes in gastric and oesophageal CAMs, differentially methylated CpG loci identified in both sets of CAMs were assigned to genes. Differentially methylated CpG loci from gastric CAMs were assigned to 5918 genes, whereas differentially methylated oesophageal CAM CpG loci were assigned to 4105 genes. Comparison of these gene lists identified 2223 common genes, which evidence of cancer-induced changes in DNA methylation in both gastric and oesophageal CAMs (Supplementary Figure S8). Notably, comparison analysis of differentially methylated CpG sites found 230 CpGs that are differentially methylated in both gastric and oesophageal CAMs when compared to corresponding patient-matched ATMs (Figure 2 inner most track; Figure 6). These conserved differentially methylated loci are distributed throughout the genome (Figure 2 inner most track). Further analysis of these loci shows that 65.22% are associated with 2 genes and 33.5% are associated with only one gene.

To investigate how DNA methylation changes in gastric and oesophageal CAMs may affect common pathways and processes, 2223 common differentially methylated genes were subjected to Ingenuity Pathway Analysis (IPA) and ConsensusPathDB [24] over-representation analysis (Enriched KEGG and Reactome pathways are reported in Supplementary File S6). IPA analysis showed that these commonly differentially methylated genes are involved in digestive organ tumour ($p = 1.67 \times 10^{-17}$), expression of RNA ($p = 3.46 \times 10^{-12}$), transcription ($p = 7.36 \times 10^{-12}$), gastroesophageal cancer ($p = 7.94 \times 10^{-11}$), tumorigenesis of the tissue ($p = 3.13 \times 10^{-11}$), cell movement ($p = 1.83 \times 10^{-10}$), migration

of cells ($p = 2.18 \times 10^{-9}$), proliferation of cells ($p = 1.28 \times 10^{-8}$), invasion of cells ($p = 3.43 \times 10^{-6}$), generation of fibroblasts ($p = 1.69 \times 10^{-5}$) and growth of tumour ($p = 4.97 \times 10^{-5}$).

In addition, the results from the gastric and oesophageal GO enrichment analysis of differentially methylated CpG loci (Supplementary File S2, Supplementary File S7) were compared to get further insight into common biological processes affected by DNA methylation changes; 32 unifying GO biological processes were identified (Supplementary Figure S8). Enriched biological processes linked to cancer related changes in DNA methylation in gastric and oesophageal CAMs include cell adhesion, cell differentiation and developmental processes, signalling, regulation of signal transduction and GTPase activity. As gastric stromal myofibroblasts have a neuroendocrine-like phenotype, which is associated with advanced cancer [15], it is interesting to note that synaptic transmission and regulation of calcium transport were one of the significantly enriched biological processes targeted by DNA methylation changes in both gastric and oesophageal CAMs (Supplementary File S2, Supplementary File S7). These conserved signatures represent an additional resource to inform future hypothesis driven studies into conserved mechanisms or functional consequences of tumour induced stromal reprogramming.

DISCUSSION

As the tumour microenvironment plays an important role in cancer progression, increasing efforts are being made to understand the molecular processes that drive pro-tumorigenic changes in stromal myofibroblasts.

While differential patterns of DNA methylation are well documented in cancer cells [25-27], there is relatively little information relating to induced epigenetic changes in stromal myofibroblasts derived from different tumours. Consequently, it is not yet clear to what extent common mechanisms of stromal reprogramming operate in different solid tumours. In this study an integrated multi-omics approach was used to provide the first evidence that myofibroblasts derived from the site of gastric and oesophageal adenocarcinomas (CAMs) are epigenetically reprogrammed to have distinct DNA methylation signatures, as compared to non-tumour derived patient matched myofibroblasts (ATMs), or corresponding tissue matched non-tumour associated myofibroblasts (NTMs).

Previous knowledge of genome-wide epigenetic changes within tumour stroma was largely based on the analysis of epithelial, myoepithelial and stromal fibroblasts, derived from either normal breast tissue, or *in situ* and invasive breast carcinomas [28]. These studies show that distinct epigenetic profiles were observed in tumour associated fibroblasts, revealing both stage and cell-type specific variations [28]. However, the link between imposed epigenetic changes and the molecular processes that contribute to reciprocal interactions between cancer and stromal myofibroblasts remains incomplete. CAMs derived from gastric [14] or non-small cell lung cancer [29] were found to exhibit reduced global DNA methylation, accompanied by a selective gain in focal DNA methylation. Significantly, all primary gastric and oesophageal CAMs used in this study also exhibited reduced global DNA methylation levels in comparison to patient-matched ATMs. Given the emerging consistency between CAMs derived from different tumours, a subtle yet significant reduction in global DNA methylation may be a useful common indicator of a functional transition from NTM or ATM status, to a cancer-promoting CAM phenotype.

Despite growing evidence that functional differences in myofibroblast populations are linked to conserved changes in DNA-methylation, previous studies did not investigated CAM-specific DNA methylation profiles at single CpG resolution [14, 28]. As such, the molecular mechanisms and functional consequences of cancer induced epigenetic programming remained unclear. In this context, the present study provides new insight into CAM-specific changes, identifying CpG loci with altered DNA methylation in both gastric and oesophageal CAMs compared to patient-matched ATMs or unrelated NTMs. Interestingly, unsupervised clustering, based on genome-wide DNA methylation profiles, shows that ATMs are more similar to NTMs then to corresponding patient matched CAMs. Also, loci that were consistently hypomethylated in CAMs were commonly hypermethylated in both ATMs and NTMs. Thereby providing evidence that CAMs are uniquely reprogrammed to have common gene/loci specific patterns of methylation. Therefore, novel conserved signatures may provide further insight into the molecular mechanism and functional consequences of cancer induced epigenetic reprogramming. Also, CAM-specific patterns of DNA methylation may facilitate the identification of proxy markers of regional stromal conversion.

Interestingly, altered patterns of promoter methylation observed in a number of genes (e.g. ZMIZ1, EYA4, SLC22A18AS, WIPF1, FAM49A, RUNX3, ESRRG) in both gastric and oesophageal CAMs

were also previously reported in CAMs derived from lung tumours [29], indicating that common mechanisms of epigenetic reprogramming may contribute to the aberrant expression of these genes in CAMs derived from different tumours.

Functional enrichment analysis of data from this study show that CAM-specific changes in DNA methylation have the potential to affect processes related to tumour development, including digestive organ tumours, tumourogenesis, cell movement, cell migration and proliferation, generation of fibroblasts, growth of tumours and cell adhesion. Loci showing CAM specific changes in methylation were also found to encode genes associated with Hedgehog, Wnt and Notch signalling pathways, several of which are differentially expressed in both gastric and oesophageal cancer [30-32]. Although deregulation of these pathways has been linked to tumour development, the spectrum of cell-types that contribute to the observed signals *in vivo* remain unclear. Loci encoding genes involved in glycosaminoglycan biosynthesis and metabolism, an important component of extracellular matrix involved in cell signalling, cell function and cancer progression [33], were also found to be differentially methylated in isolated gastric and oesophageal CAMs. In terms of reprogramming energy production in CAMs, it is interesting to note that CAM-specific DNA methylation signatures have the potential to affect fatty acid, triacylglycerol and ketone body metabolism, providing the first evidence that epigenetic reprogramming may contribute to a reverse Warburg phenotype in gastric and oesophageal CAMs. Interestingly, this phenotype would also be associated with a concomitant reduction in mitochondrial activity, which may in-turn contribute to a global reduction in DNA methylation [34-37]. In terms of paracrine communication within the tumour microenvironment, previous study in gastric has shown that myofibroblasts have neuroendocrine-like properties, which are lost in advanced stages of cancer [15]. Interestingly, both synaptic transmission and regulation of Ca^{2+} transport were identified as biological processes affected by CAM-specific changes in DNA methylation.

Further insight into the functional implications of CAM-specific changes in DNA methylation were provided by parallel genome-wide gene expression profiling of gastric CAM, ATM and NTM isolates. Previous studies in breast [38-41], non-small cell lung cancer [42], colon [43] and oral squamous cell carcinoma [44] reported CAM-specific changes in gene expression, compared to patient matched ATMs, or tissue matched NTMs. In each case, CAM-specific gene expression profile may provide proxy marker for diagnosis or prognostic predictions. As in the case of oral squamous cell carcinoma,

where gene expression studies identified two distinct CAM subtypes with differential tumour-promoting abilities [45] and showed that CAM expression profiles reflect the stage of tumour progression [44].

To assess the validity of correlated DNA methylation/gene expression signatures, a subsequent series of targeted pyrosequencing and qPCR studies were performed on additional set of gastric CAMs, ATMs and NTMs. This analysis confirmed initial CAM-specific patterns of DNA methylation/gene expression, while also revealing equivalent changes in DNA methylation in associated border genomic regions. These verification studies increase confidence in data derived from genome-wide studies, while also providing insight into the mechanisms by which imposed changes in CAM DNA methylation contribute to tumour progression.

As promoter hypermethylation of SMAD3 correlates with reduced gene expression in gastric CAMs, it is likely that these imposed changes may perturb TGF-β signalling responses. Interestingly, our data show that simultaneous promoter hypermethylation and gene body hypomethylation results in the transcriptional repression of TGFBR2 (type II TGF-β receptor). As selective ablation of TGFBR2 in mouse stromal fibroblasts induced neoplastic lesions and stromal expansion [46], it is possible that combined attenuation to TGF-β signalling in CAMs may play a key role in the development of gastric and oesophageal tumours. Notably, SMAD3 was also found to be down regulated in oesophageal CAMs and Vizoso et al. [29] also reported promoter hypermethylation-associated SMAD3 silencing in CAMs derived from lung tumours. In a related analysis of a panel of wound-related ECM genes, COL1A1, EDA-FN, LOX and SPARC were all found to be upregulated in CAMs compared to patient-matched control fibroblasts [29]. Significently, we also found that COL1A1, SPARC and LOX were upregulated in gastric CAMs compared to patient matched ATMs. Collectivly, these observations suggest that promoter silencing of SMAD3 expression may contribute to CAM aberrant phenotype and underlie tumour-promoting properties of CAMs derived from different tissues.

Extended verification studies confirmed a recurrent CAM specific signature in the genomic region encoding the FOX1 adjacent non-coding developmental regulatory RNA (FENDRR) and FOXF1 genes. FENDRR is a long non-coding RNA (lncRNA) transcribed bidirectionally on the opposite strand to the FOXF1 gene, which encodes a transcription factor involved in embryonic development and mesenchymal-epithelial interaction [21]. Also, there is evidence that FOXF1 acts as a tumour suppressor, as it is inactivated by DNA methylation in breast cancer [47] and its expression is reduced

in prostate cancer [48]. This study shows that cancer-imposed changes in DNA hypermethylation may lead to reduced expression of FENDRR and FOXF1 in gastric CAMs. Significantly, Xu et al showed that FENDRR regulates gastric cancer metastasis and is down-regulated in gastric cancer cells relative to cells derived from normal gastric epithelial [22]. Low levels of FENDRR were also found to correlate with poor patient prognosis and more aggressive tumour characteristics including greater invasion depth, higher tumour stage and lymphatic metastasis [22]. FOXF1 has also been shown to contribute to the tumour promoting properties of lung cancer associated fibroblasts (CAFs), including the production of hepatocyte growth factor (HGF) and fibroblast growth factor-2 (FGF-2), both of which promote tumour growth [49].

In conclusion, findings presented in this study provide new insight into the imposed molecular changes that contribute to epigenetic and functional reprogramming of gastric CAMs thereby improving our understanding of the complex range of reciprocal interactions that occur between developing tumours and the stromal microenvironment. Further studies are needed to establish the extent to which identified trends are also observed at protein level. As DNA methylation patterns are more robust and long lasting than mRNA or protein signatures, the signatures that differentiate CAM, ATM and NTM methylation profiles may also provide a useful resource to identify new markers to improve tumour stratification and the ability to define not just tumour boundaries but also the surrounding region of stromal reprogramming, which may be an important factor in defining optimal resection margins.

SUPPLEMENTARY MATERIAL

Supplementary information is available at http://carcin.oxfordjournals.org/

FUNDING

This work was supported by the Cancer Research UK non-clinical training award [C35628/A12779].

Funding for open access charge: University of Liverpool.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Bertuccio, P., et al., *Recent patterns in gastric cancer: A global overview.* International Journal of Cancer, 2009. **125**(3): p. 666-673.
2. Ferlay, J., et al., *Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012.* International Journal of Cancer, 2015. **136**(5): p. E359-E386.
3. Pietras, K. and A. Ostman, *Hallmarks of cancer: Interactions with the tumor stroma.* Experimental Cell Research, 2010. **316**(8): p. 1324-1331.
4. Polanska, U.M. and A. Orimo, *Carcinoma-associated fibroblasts: Non-neoplastic tumour-promoting mesenchymal cells.* Journal of Cellular Physiology, 2013. **228**(8): p. 1651-1657.
5. Ohlund, D., E. Elyada, and D. Tuveson, *Fibroblast heterogeneity in the cancer wound.* The Journal of experimental medicine, 2014. **211**(8): p. 1503-23.
6. Martinez-Outschoorn, U.E., et al., *Oxidative stress in cancer associated fibroblasts drives tumor-stroma co-evolution A new paradigm for understanding tumor metabolism, the field effect and genomic instability in cancer cells.* Cell Cycle, 2010. **9**(16): p. 3256-3276.
7. Martinez-Outschoorn, U.E., M.P. Lisanti, and F. Sotgia, *Catabolic cancer-associated fibroblasts transfer energy and biomass to anabolic cancer cells, fueling tumor growth.* Seminars in Cancer Biology, 2014. **25**: p. 47-60.
8. Yamashita, M., et al., *Role of stromal myofibroblasts in invasive breast cancer: stromal expression of alpha-smooth muscle actin correlates with worse clinical outcome.* Breast Cancer, 2012. **19**(2): p. 170-176.
9. Ha, S.Y., et al., *The Prognostic Significance of Cancer-Associated Fibroblasts in Esophageal Squamous Cell Carcinoma.* Plos One, 2014. **9**(6).
10. Li, X.Y., S.Q. Hu, and L. Xiao, *The cancer-associated fibroblasts and drug resistance.* European Review for Medical and Pharmacological Sciences, 2015. **19**(11): p. 2112-2119.
11. Holmberg, C., et al., *Release of TGFbetaig-h3 by gastric myofibroblasts slows tumor growth and is decreased with cancer progression.* Carcinogenesis, 2012. **33**(8): p. 1553-62.
12. Kumar, J.D., et al., *Increased expression of chemerin in squamous esophageal cancer myofibroblasts and role in recruitment of mesenchymal stromal cells.* PloS one, 2014. **9**(7).

13. Wang, L., et al., *Distinct miRNA profiles in normal and gastric cancer myofibroblasts and significance in Wnt signaling.* American Journal of Physiology - Gastrointestinal and Liver Physiology, 2016. **310**(9): p. G696-G704.
14. Jiang, L., et al., *Global hypomethylation of genomic DNA in cancer-associated myofibroblasts.* Cancer research, 2008. **68**(23): p. 9900-9908.
15. Balabanova, S., et al., *The neuroendocrine phenotype of gastric myofibroblasts and its loss with cancer progression.* Carcinogenesis, 2014. **35**(8): p. 1798-1806.
16. Orimo, A., et al., *Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion.* Cell, 2005. **121**(3): p. 335-348.
17. Hu, C., et al., *Effects of cancer-associated fibroblasts on the migration and invasion abilities of SGC-7901 gastric cancer cells.* Oncology letters, 2013. **5**(2): p. 609-612.
18. Mathivanan, S., et al., *ExoCarta 2012: database of exosomal proteins, RNA and lipids.* Nucleic Acids Research, 2012. **40**(D1): p. D1241-D1244.
19. Jones, P.A., *The DNA methylation paradox.* Trends in Genetics, 1999. **15**(1): p. 34-37.
20. Jones, P.A., *Functions of DNA methylation: islands, start sites, gene bodies and beyond.* Nature Reviews Genetics, 2012. **13**(7): p. 484-492.
21. Pruitt, K.D., et al., *RefSeq: an update on mammalian reference sequences.* Nucleic Acids Research, 2014. **42**(D1): p. D756-D763.
22. Xu, T.-P.P., et al., *Decreased expression of the long non-coding RNA FENDRR is associated with poor prognosis in gastric cancer and FENDRR regulates gastric cancer cell metastasis by affecting fibronectin1 expression.* Journal of hematology & oncology, 2014. **7**: p. 63.
23. Smith, R.N., et al., *InterMine: a flexible data warehouse system for the integration and analysis of heterogeneous biological data.* Bioinformatics, 2012. **28**(23): p. 3163-3165.
24. Kamburov, A., et al., *ConsensusPathDB: toward a more complete picture of cell biology.* Nucleic acids research, 2011. **39**(Database issue): p. 7.
25. Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer.* Nature reviews. Genetics, 2002. **3**(6): p. 415-428.
26. Witte, T., C. Plass, and C. Gerhauser, *Pan-cancer patterns of DNA methylation.* Genome Medicine, 2014. **6**(8): p. 66.
27. Chen, Y., et al., *Tissue-independent and tissue-specific patterns of DNA methylation alteration in cancer.* Epigenetics & Chromatin, 2016. **9**(1): p. 10.
28. Hu, M., et al., *Distinct epigenetic changes in the stromal cells of breast cancers.* Nature Genetics, 2005. **37**(8): p. 899-905.

29. Vizoso, M., et al., *Aberrant DNA methylation in non-small cell lung cancer-associated fibroblasts.* Carcinogenesis, 2015. **36**(12): p. 1453-1463.
30. Kang, H., et al., *Notch3 and Jagged2 contribute to gastric cancer development and to glandular differentiation associated with MUC2 and MUC5AC expression.* Histopathology, 2012. **61**(4): p. 576-586.
31. Wang, D.H., et al., *Aberrant Epithelial–Mesenchymal Hedgehog Signaling Characterizes Barrett's Metaplasia.* Gastroenterology. **138**(5): p. 1810-1822.e2.
32. Kurayoshi, M., et al., *Expression of Wnt-5a Is Correlated with Aggressiveness of Gastric Cancer by Stimulating Cell Migration and Invasion.* Cancer Research, 2006. **66**(21): p. 10439-10448.
33. Afratis, N., et al., *Glycosaminoglycans: key players in cancer cell biology and treatment.* Febs Journal, 2012. **279**(7): p. 1177-1197.
34. Pavlides, S., et al., *The reverse Warburg effect Aerobic glycolysis in cancer associated fibroblasts and the tumor stroma.* Cell Cycle, 2009. **8**(23): p. 3984-4001.
35. Martinez-Outschoorn, U., F. Sotgia, and M.P. Lisanti, *Tumor Microenvironment and Metabolic Synergy in Breast Cancers: Critical Importance of Mitochondrial Fuels and Function.* Seminars in Oncology, 2014. **41**(2): p. 195-216.
36. Afanas'ev, I., *Mechanisms of Superoxide Signaling in Epigenetic Processes: Relation to Aging and Cancer.* Aging and Disease, 2015. **6**(3): p. 216-227.
37. Smiraglia, D.J., et al., *A novel role for mitochondria in regulating epigenetic modification in the nucleus.* Cancer Biology & Therapy, 2008. **7**(8): p. 1182-1190.
38. Bauer, M., et al., *Heterogeneity of gene expression in stromal fibroblasts of human breast carcinomas and normal breast.* Oncogene, 2010. **29**(12): p. 1732-1740.
39. Al-Rakan, M.A., et al., *Breast stromal fibroblasts from histologically normal surgical margins are pro-carcinogenic.* Journal of Pathology, 2013. **231**(4): p. 457-465.
40. Peng, Q., et al., *Biological Characteristics and Genetic Heterogeneity between Carcinoma-Associated Fibroblasts and Their Paired Normal Fibroblasts in Human Breast Cancer.* Plos One, 2013. **8**(4).
41. Singer, C.F., et al., *Differential gene expression profile in breast cancer-derived stromal fibroblasts.* Breast Cancer Research and Treatment, 2008. **110**(2): p. 273-281.
42. Navab, R., et al., *Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(17): p. 7160-7165.

43. Herrera, M., et al., *Functional Heterogeneity of Cancer-Associated Fibroblasts from Human Colon Tumors Shows Specific Prognostic Gene Expression Signature.* Clinical Cancer Research, 2013. **19**(21): p. 5914-5926.
44. Lim, K.P., et al., *Fibroblast gene expression profile reflects the stage of tumour progression in oral squamous cell carcinoma.* Journal of Pathology, 2011. **223**(4): p. 459-469.
45. Costea, D.E., et al., *Identification of Two Distinct Carcinoma-Associated Fibroblast Subtypes with Differential Tumor-Promoting Abilities in Oral Squamous Cell Carcinoma.* Cancer Research, 2013. **73**(13): p. 3888-3901.
46. Bhowmick, N.A., et al., *TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia.* Science, 2004. **303**(5659): p. 848-851.
47. Lo, P.-K., et al., *Epigenetic Inactivation of the Potential Tumor Suppressor Gene FOXF1 in Breast Cancer.* Cancer Research, 2010. **70**(14): p. 6047-6058.
48. Watson, J.E.V., et al., *Integration of high-resolution array comparative genomic hybridization analysis of chromosome 16q with expression array data refines common regions of loss at16q23-qter and identifies underlying candidate tumor suppressor genes in prostate cancer.* Oncogene, 2004. **23**(19): p. 3487-3494.
49. Saito, R.-A., et al., *Forkhead Box F1 Regulates Tumor-Promoting Properties of Cancer-Associated Fibroblasts in Lung Cancer.* Cancer Research, 2010. **70**(7): p. 2644-2654.

FIGURE LEGENDS

Figure 1. DNA methylation profiling of primary myofibroblasts purified from different tissue microenvironments **A.** Unsupervised clustering of 5688 CpG loci with marked differential methylation in CAMs and patient-matched ATMs with projection including related values observed in NTMs. **B.** Unsupervised clustering of 8104 CpG loci with marked differential methylation in CAMs and unrelated NTMs. Heatmaps represent differentially methylated CpG loci identified in respective comparisons |Δβ|>0.2, p-value <0.05. **C, D.** Distribution of differentially methylated CpG loci identified in CAMs vs ATMs. **C.** in CpG islands, shores, shelves and sea regions or **D.** relative to RefSeq gene promoters, gene bodies and intergenic regions.

Figure 2. Circular plot of genome-wide DNA methylation changes in stromal myofibroblasts. The outer ring represents human ideograms. The first tract shows differentially methylated CpG sites between gastric CAMs and patient-matched ATMs. The heatmap represents CpG loci that distinguish CAMs from non-tumour derived myofibroblasts (ATMs and NTMs). These loci may serve as proxies for the identification of gastric CAMs. The innermost track represents conserved CAM susceptible loci identified in both gastric and oesophageal CAMs compared to patient-matched ATMs; *red - hypermethylated loci in CAMs, blue - hypomethylated loci in CAMs; |Δβ|>0.2, p-value <0.05.*

Figure 3. Differential gene expression signatures in gastric myofibroblasts purified from different tissue microenvironments. **A.** CAM vs ATM **B.** CAM vs NTM **C.** ATM vs NTM. Heatmaps represent differentially expressed genes in respective comparisons *FDR p-value < 0.05;* volcano plots represent differentially expressed genes *p-value <0.05;* dashed lines 1.6 fold change. **D.** Quantitative PCR validations of genes identified as differentially expressed in CAM vs ATM and CAM vs NTM comparisons. Each TaqMan assay was done in triplicates for CAM (n=3), ATM (n=3) and NTM (n=3) samples. The comparative ΔΔCt method was used and samples were normalized to calibrator. Error bars represent SEM; CAM vs ATM *t-test* ****p<0.0001*; CAM vs NTM *t-test* ****p<0.0001, **p<0.01, *p<0.05*.

Figure 4. Changes in DNA methylation in the SMAD3 and SPON2 promoter regions regulates gene expression in gastric CAMs and ATMs. **A.** Differentially methylated CpG loci identified by Illumina 450k array in the SMAD3 promoter region. Mean β values (n=3) for probes identified as differentially

methylated in CAM vs ATM and CAM vs NTM comparisons. X-axis indicates distance of Illumina 450k probes from SMAD3 transcription start site. Positions highlighted in magenta (-1344) and blue (-1240) are within the genomic region examined by corresponding pyrosequencing assays. **B.** Pyrosequencing analysis of the SMAD3 promoter region was performed on patient-matched CAM (n=7) and ATM (n=7) samples. Methylation means for 10 individual CpG sites in the SMAD3 promoter region are plotted. X-axis indicates the chromosomal position of examined CpG sites. Positions marked with * correspond to the Illumina 450k probes. **C.** The overall methylation level of the SMAD3 promoter region interrogated by pyrosequencing analysis. Boxplots represent methylation distribution and mean for 10 CpG sites in CAM (n=7), ATM (n=7) and NTM (n=4) samples. **D.** Quantitative PCR analysis of SMAD3 gene expression in CAM (n=6) and ATM (n=6) samples; *t-test* ****$p<0.0001$. Error bars represent SEM. DNA methylation in the SPON2 promoter region correlates with SPON2 gene expression in gastric CAMs and ATMs. **E.** Differentially methylated CpG sites identified by Illumina 450k array in the SPON2 promoter region. Mean β values (n=3) for probes found to be differentially methylated in CAM vs ATM and CAM vs NTM comparisons. X-axis indicates distance of Illumina 450k probes from the SPON2 transcription start site. Position highlighted in magenta (-29185) is within the genomic region examined by pyrosequencing. **F.** Pyrosequencing analysis of the SPON2 promoter region in patient-matched CAM (n=7) and ATM (n=7) samples. Methylation means for 7 individual CpG sites in the interrogated promoter region are plotted. The X-axis indicates the chromosomal position of examined CpG sites. Position marked with * corresponds to the Illumina 450k probe highlighted in magenta (-29185) in Panel E. **G.** The overall methylation level of the SPON2 promoter region interrogated by pyrosequencing. Boxplots represent methylation distribution and mean for 7 CpG sites in CAM (n=7), ATM (n=7) and NTM (n=4) samples. **H.** Quantitative PCR analysis of SPON2 gene expression in CAM (n=6) and ATM (n=6) samples; *t-test* ****$p<0.0001$. Error bars represent SEM.

Figure 5. DNA methylation pattern in the genomic region associated with regulation of FOXF1 and FENDRR expression in gastric CAMs and ATMs. **A.** Differentially methylated CpG sites identified by Illumina 450k array in the region downstream of the FOXF1 transcription start site. Mean β values (n=3) for probes identified as differentially methylated in CAM vs ATM and CAM vs NTM comparisons are plotted. The X-axis indicates distance of Illumina 450k probes to FOXF1 transcription start site. **B.**

Pyrosequencing analysis of the FOXF1 promoter region in patient-matched CAM (n=7) and ATM (n=7) samples. Methylation means for 8 individual CpG sites in the interrogated promoter region are plotted. The X-axis indicates chromosomal position of examined CpG sites. **C.** The overall methylation level of the FOXF1 promoter region interrogated by pyrosequencing assay. Boxplots represent methylation distribution and mean for 8 CpG sites in CAM (n=7), ATM (n=7) and NTM (n=4) samples. **D.** Quantitative PCR analysis of FOXF1 and FENDRR gene expression in CAM and ATM samples; t-test FOXF1 (n=5) ****p=0.0008*; FENDRR v1 (splice variant 1; n=4) *****p<0.0001*; FENDRR v2 (splice variant 2; n=4) ***p=0.0028*. Error bars represent SEM.

Figure 6. Representative conserved DNA methylation patterns in gastric and oesophageal patient-matched CAM and ATM samples. Probes highlighted in magenta are also identified as proxies for gastric CAMs. Numbers in brackets indicate the distance to transcription start site (TSS) of a given gene (indicated at the top of each plot); *magenta – CAMs, purple – ATMs; |Δβ|>0.2, p-value <0.05.*

Figure 1

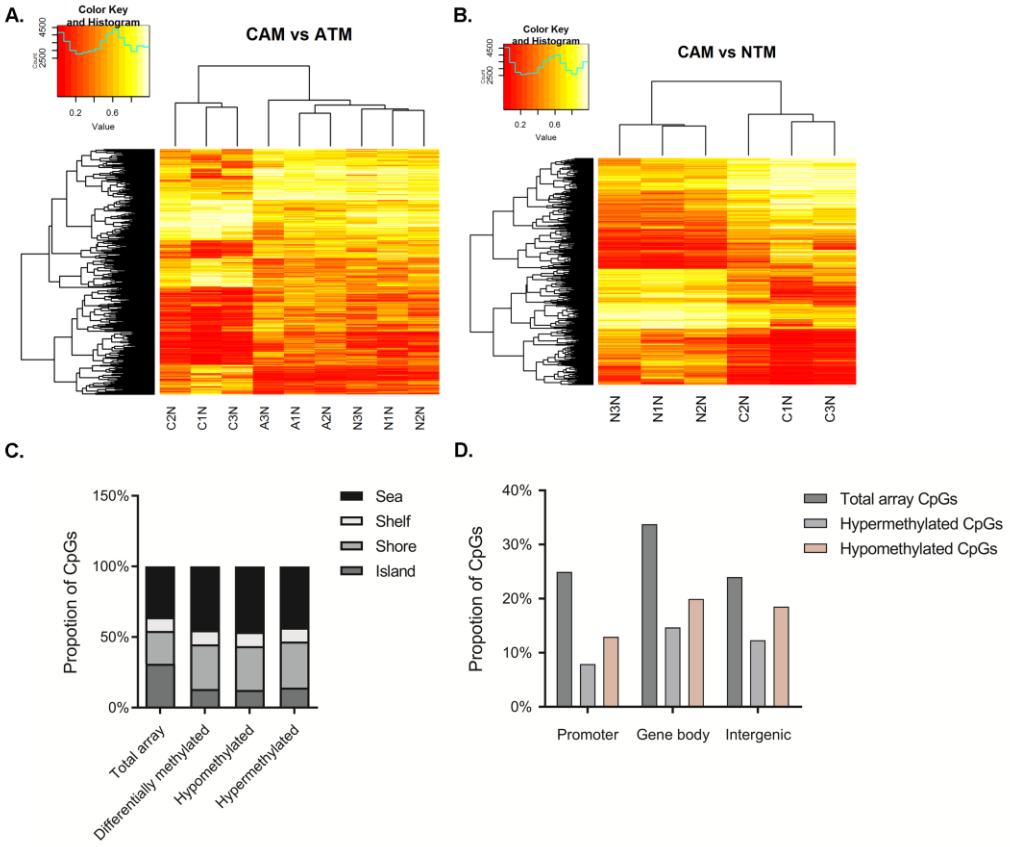

Figure 2

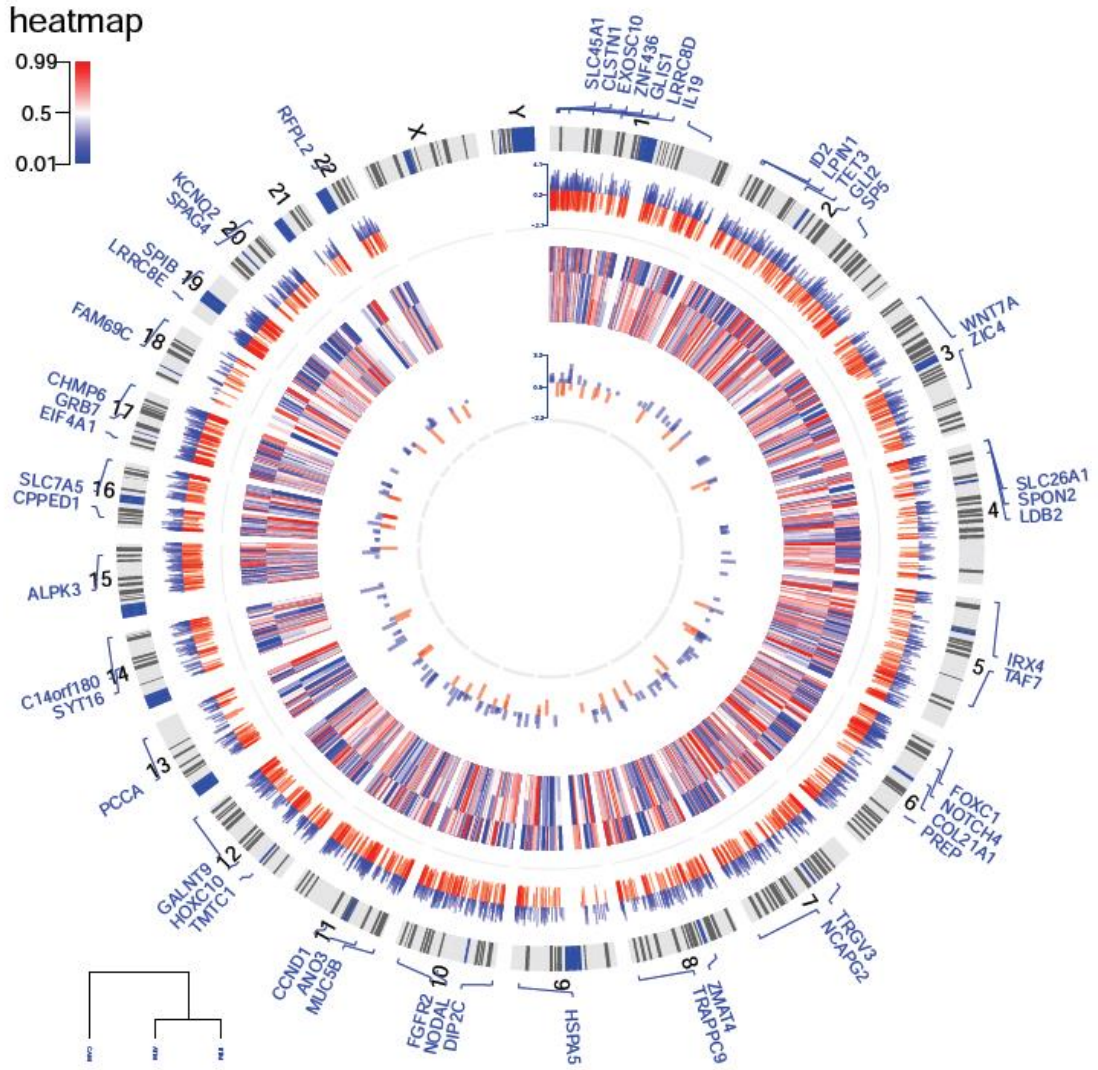

Figure 3

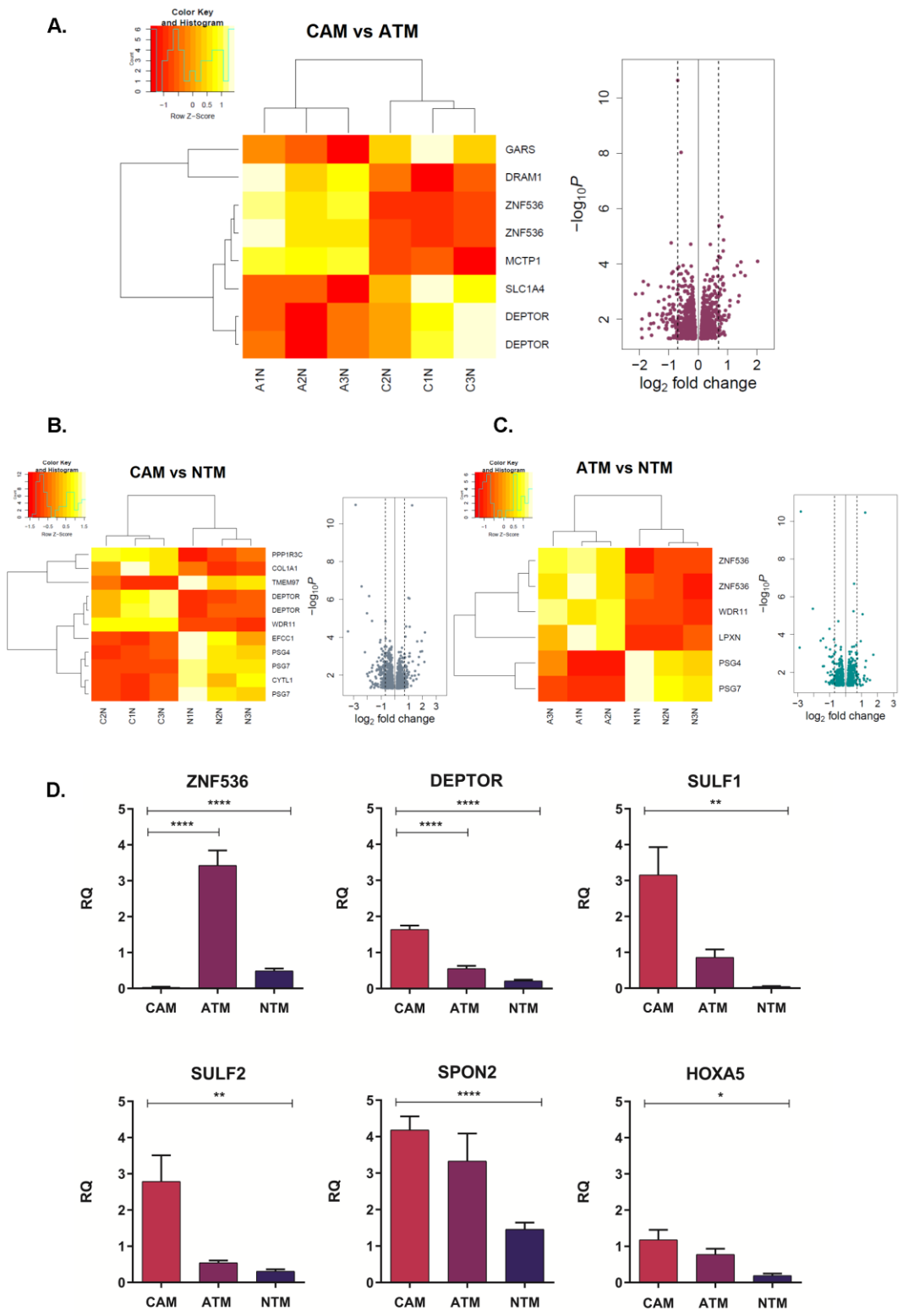

Figure 4

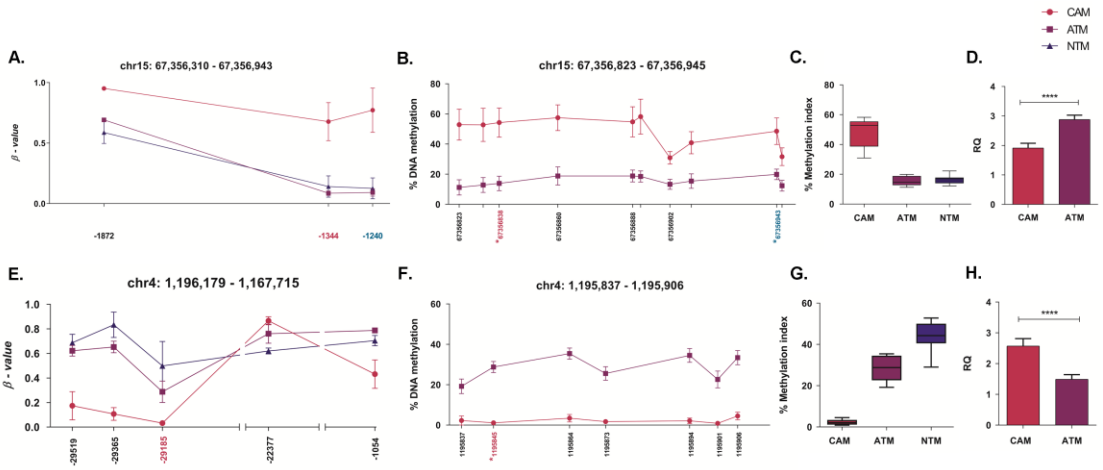

Figure 5

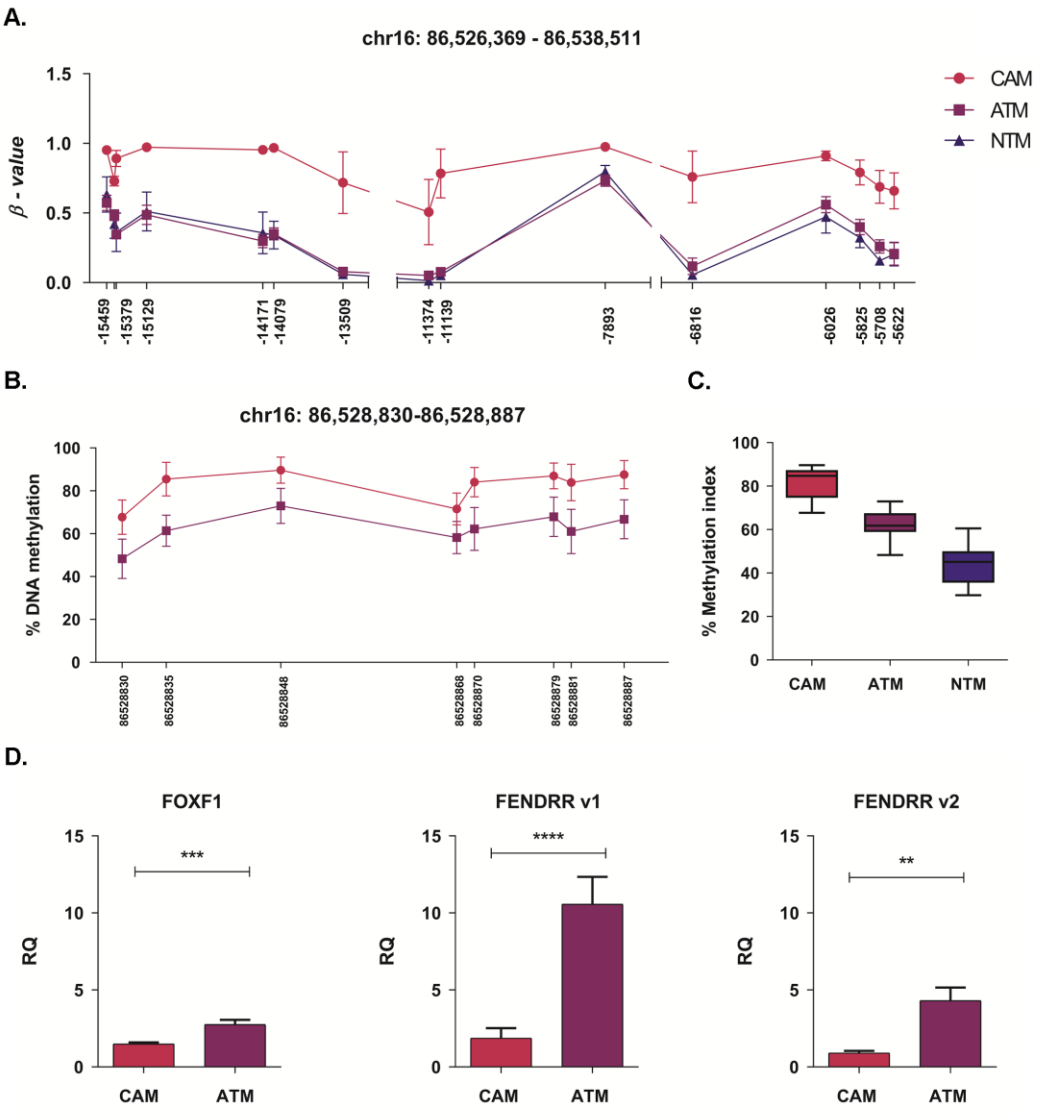

Figure 6