

Clinical Epigenetics

Genome-wide analysis of DNA methylation and gene expression defines molecular characteristics of Crohn's disease-associated fibrosis --Manuscript Draft--

Manuscript Number:	CLEP-D-15-00158R1
Full Title:	Genome-wide analysis of DNA methylation and gene expression defines molecular characteristics of Crohn's disease-associated fibrosis
Article Type:	Research
Abstract:	<p>Background: Fibrosis of the intestine is a common and poorly understood complication of Crohn's disease (CD) characterized by excessive deposition of extracellular matrix and accompanied by narrowing and obstruction of the gut lumen. Defining the molecular characteristics of this fibrotic disorder is a vital step in the development of specific prediction, prevention and treatment strategies. Previous epigenetic studies indicate that alterations in DNA methylation could explain the mechanism by which mesenchymal cells adopt the requisite pro-fibrotic phenotype that promotes fibrosis progression. However, to date, genome-wide analysis of the DNA methylome of any type of human fibrosis is lacking. We employed an unbiased approach using deep sequencing to define the DNA methylome and transcriptome of purified fibrotic human intestinal fibroblasts (HIF) from the colons of patients with fibrostenotic CD.</p> <p>Results: When compared with normal fibroblasts, we found that the majority of differential DNA methylation was within introns, intergenic regions and not associated with CpG islands. Only a low percentage occurred in the promoters and exons of genes. Integration of the DNA methylome and transcriptome identified regions in three genes that inversely correlated with gene expression: WNT2B and two eicosanoid synthesis pathway enzymes (prostacyclin synthase and prostaglandin D2 synthase). These findings were independently validated by RT-PCR and bisulfite sequencing. Network analysis of the data also identified candidate molecular interactions relevant to fibrosis pathology.</p> <p>Conclusions: Our definition of a genome-wide fibrosis-specific DNA methylome provides new gene networks and epigenetic states by which to understand mechanisms of pathological gene expression that lead to fibrosis. Our data also provide a basis for development of new fibrosis-specific therapies, as genes dysregulated in fibrotic Crohn's disease, following functional validation, can serve as new therapeutic targets.</p>
Response to Reviewers:	<p>Dear Editor,</p> <p>On behalf of all the authors we would like to thank the reviewers for their insightful and helpful comments. Our revisions are included in red in the revised manuscript. In addition we have revised Figures 2, 3 and Additional file 1 and included a new additional file 3. A point-by-point rebuttal to the reviewers' comments is included below.</p> <p>Reviewer #2 Comment</p> <p>The study conducted an integrated Omics study to Crohn's disease-associated fibrosis. Genome-wide DNA methylation and transcriptome were explored with MBD-seq and RNA-seq in 3 pairs of Crohn's disease-associated fibrosis samples. Large number of potential differential methylation regions (DMR) and differential expressed gene (DMG) were identified. Several genes were validated with BSP and RT-PCR with solid result. The study was performed rigorously and the findings are interesting. What's more, the study were conducted timely which provided important insight of the contribution from DNA methylation to Crohn's disease-associated fibrosis. I only have several concerns as the following:</p> <p>Major Revisions</p> <p>1. The DMR analysis strategy should be revised. Since the sample size is too small,</p>

only DMR with large difference can be detected with high confidence. The present yes/no strategy would only detected some false positive signals. Meanwhile, the present study did not arrange technique negative control in the experiments, these false positive would be more serious. I found that similar study had ever been published in Clinical Epigenetics, such as Zhao et al. Clinical Epigenetics 2014, 6:18 or Zhao et al. Clinical Epigenetics 2014, 6:30. Maybe, this is also the reason why the difference of the BSP assay were not significant in Figure 3A. Actually, I preferred to take the right and left regions of the Figure 3A and 3B as the significant DMR. Could you provided the P-value for the right and left regions in Figure 3A and 3B? Whether these two regions were significantly different between cases and controls?

Response

We thank the reviewer for the insightful comments. False-discovery is an inevitable feature of any type of genome-wide analysis. However, our approach incorporates a number of strategies to minimize this issue. We note that there are two important aspects of our DMR analysis method. First, statistical approaches are employed both for determining the width of the DMRs when joining adjacent 50 bp windows (using a scan statistic) and for the differential magnitude of the DMR (using a linear mixed-effects model). Also, p-values are adjusted for false discovery control by using the Benjamini–Hochberg procedure. Second, the sharp yes/no determination adds additional requirements to each DMR that also reduce the false positive rate. Because a single outlier could cause significant DMRs to appear in the quantitative differential analysis above, we also require that 3 out of 3 samples in each group are consistently hypo- or hyper-methylated when binary methylation status is modeled with a Bayesian mixture of Poissons model. Thus, sharp yes/no DMRs are required to have 3 out of 3 samples methylated in one group and 3 out of 3 samples unmethylated in the other group. They are also required to show a statistically significant quantitative difference between the sequencing read counts under our linear mixed effects model.

For the negative control technique, we believe the reviewer is referring to MRE-seq used by Zhao et al (Clinical Epigenetics 2014; 6:18). This technique uses methylation sensitive restriction enzymes to provide a complementary unmethylated signal that can be integrated with the MBD-seq/MeDIP-seq methylation signal. We were unable to perform MRE-seq on our samples, as the amount of input DNA was not sufficient for both type of sequencing. Thus, we took the approach of performing targeted bisulfite sequencing to provide a gold-standard validation for the most biologically compelling candidates as described in the manuscript.

We note that the results from the targeted bisulfite sequencing assays shown in figure 3A and 3B do validate the hypermethylation in the Crohn's disease fibroblasts, as there is increased presence of hypermethylation in these samples and in each case these were statistically significant (see Additional table 1: PTGIS promoter region DMR, adjusted p value: 1.31×10^{-5} ; WNT2B DMR, adjusted p value: 3.12×10^{-3}). To address the reviewer's question, we have added the regions of the DMRs detected from our MiGS data to the figure (black box below browser screen capture in A and B) to show that the region validated by bisulfite sequencing overlaps with the DMR. All other flanking regions were not statistically significant in our MiGS analysis.

Minor Revisions,

Comment

1. Is there any miRNAs, ncRNAs in the DMR regions? The author should be check it again. Meanwhile, the signals or information of ENCODE in the main DMR regions should be provided in the manuscript or as the supplementary.

Response

We refer the Reviewer to Additional file 1 in which we had included annotations for DMRs detected across all types of annotated transcription units, which includes annotated miRNAs or ncRNAs. To address the reviewer's query, integration with ENCODE data indicating overlap between DNase I hypersensitive sites (HSS) and the DMRs we detected has been added to the table in Additional file 1.

Comment

2.The most important contribution of the study were that large number potential

epigenetic variations were found to be aberrant in Crohn's disease-associated fibrosis and several variations were validated by further experiments, therefore, a clear table should be provided to show the reader the most important 20 or 50 aberrant DMR and their expression.

Response

We have included a new table in the supplemental data (Additional file 3) to show the reader the genes for which there is an overlap between aberrant DMRs and their expression.

Comment

3.I do not quite understand the logic of the following sentence in the manuscript "Based on reported DNA methylome analyses in cancer and other fibrotic diseases, the number, size and genome-wide distribution of DMRs in fibrotic HIF strongly suggest that this epigenetic modification has an important role in intestinal fibrosis".

Response

We agree that this sentence is unclear and have revised this (Discussion, p.9) to state the following: The large number, size and genome-wide distribution of DMRs in fibrotic HIF suggest that this epigenetic modification has an important role in intestinal fibrosis.

Comment

4.Please check the characters in Line 33, page 11

Response

We have checked and corrected characters throughout the manuscript (shown in red).

Comment

5.In the line 25, page 14, the author mentioned "ENCODE", however, the result for the encode analysis were not shown in the whole manuscript?

Response

ENCODE data has been added to Additional file 1 (see response to point 1).

Comment

6.Additional file 1, 2 and so on which were involved with the table, should be provided as the Excel or txt files. Meanwhile, in the Additional file 1, the average enrichment level should be provided for cases and controls.

Response

All additional files have been uploaded as Excel files. Only one additional file that is an image file has been retained as a TIFF file. The average enrichment level for case and controls was already included in Additional file 1 under the heading "Ratio".

Comment

7.In the Figure 2, the font size for the gene label were too small. Something should be done here.

Response

The font size for the gene label in Figure 2 has been enlarged by splitting Fig 2A into New Fig. 2A and 2B and making 2B into New Fig. 2C. The New Figure 2 is included in the revised version of our manuscript.

Comment

8. The demographic information's for the samples should be included in the manuscript, such as gender, age and other important information.

Response

A table containing this information was already included (see Additional file 9).

Reviewer #3

Comment

The article submitted by Sadler et al identifies new gene networks and epigenetic states of genes within this network that were not previously linked to fibrosis in the

intestine. The study is well planned and the article is well written with logical flow of the research question being addressed. The results are presented properly and the inferences and discussion are appropriate. In this reviewer opinion's the article is good to be accepted for publication in Clinical Epigenetics.

Minor points

1. I will be good if read lengths and depth of reads is also provided for RNAseq as the study claims it performed deep sequencing which is evident in the DNA methylome profiling but not explicitly mentioned in the RNAseq methods.

Response

We did actually state in the RNA Seq methods section that “ Next-generation sequencing was performed on an Illumina HiSeq 2500” but did omit the number of read lengths and depth of reads which we have now added to this section (p.14) which reads as follows:

Read lengths of 100 bp were sequenced at an average depth of 53,390,000 reads (minimum: 39,470,000 reads, maximum: 62,550,000 reads).

We hope that the above revisions will make our manuscript acceptable for publication in Clinical Epigenetics and we look forward to hearing from you.

Sincerely,

Eleni Stylianou, Ph.D. – Corresponding author
Angela Ting, Ph.D. – Co-corresponding author

[Click here to view linked References](#)

Genome-wide analysis of DNA methylation and gene expression defines molecular characteristics of Crohn's disease-associated fibrosis

Tammy Sadler^{1#}, Jeffrey M. Bhasin^{2,3#}, Yaomin Xu⁴, Jill Barnholz-Sloan⁵, Yanwen Chen⁵, Angela H.

Ting^{2,3**} Eleni Stylianou^{1,6*}

¹Department of Pathobiology, Lerner Research Institute, Cleveland Clinic, Cleveland, OH,

²Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine at Case Western Reserve University, Cleveland, OH, ³Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, ⁴Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, TN, ⁵Institute for Computational Biology, Case Western Reserve University, Cleveland, OH, ⁶Department of Gastroenterology & Hepatology, Digestive Diseases Institute, Cleveland Clinic, Cleveland, OH

These authors contributed equally to this work

****To whom correspondence should be addressed:***

Dr. Eleni Stylianou, Ph.D.
Department of Pathobiology
Lerner Research Institute
Cleveland Clinic
9500 Euclid Avenue/NC-22
Cleveland, OH 44195
Tel: 216-445-7156
Fax: 216-636-0104
E-mail: styliae@ccf.org

***** Co-corresponding author:***

Dr. Angela Ting, Ph.D.
Genomic Medicine Institute
Lerner Research Institute
Cleveland Clinic
9500 Euclid Avenue/NC-22
Cleveland, OH 44195
E-mail: tinga@ccf.org

ABSTRACT

Background: Fibrosis of the intestine is a common and poorly understood complication of Crohn's disease (CD) characterized by excessive deposition of extracellular matrix and accompanied by narrowing and obstruction of the gut lumen. Defining the molecular characteristics of this fibrotic disorder is a vital step in the development of specific prediction, prevention and treatment strategies. Previous epigenetic studies indicate that alterations in DNA methylation could explain the mechanism by which mesenchymal cells adopt the requisite pro-fibrotic phenotype that promotes fibrosis progression. However, to date, genome-wide analysis of the DNA methylome of any type of human fibrosis is lacking. We employed an unbiased approach using deep sequencing to define the DNA methylome and transcriptome of purified fibrotic human intestinal fibroblasts (HIF) from the colons of patients with fibrostenotic CD.

Results: When compared with normal fibroblasts, we found that the majority of differential DNA methylation was within introns, intergenic regions and not associated with CpG islands. Only a low percentage occurred in the promoters and exons of genes. Integration of the DNA methylome and transcriptome identified regions in three genes that inversely correlated with gene expression: *WNT2B* and two eicosanoid synthesis pathway enzymes (prostacyclin synthase and prostaglandin D2 synthase). These findings were independently validated by RT-PCR and bisulfite sequencing. Network analysis of the data also identified candidate molecular interactions relevant to fibrosis pathology.

Conclusions: Our definition of a genome-wide fibrosis-specific DNA methylome provides new gene networks and epigenetic states by which to understand mechanisms of pathological gene expression that lead to fibrosis. Our data also provide a basis for development of new fibrosis-specific therapies, as genes dysregulated in fibrotic Crohn's disease, following functional validation, can serve as new therapeutic targets.

KEYWORDS

DNA methylome, transcriptome, intestinal fibrosis, next generation sequencing, RNA seq, omics, Crohn's disease, Inflammatory bowel disease.

BACKGROUND

Intestinal fibrosis is a devastating complication of Crohn's disease (CD), a major type of inflammatory bowel disease (IBD) [1]. Characterized by a chronic transmural inflammation of the intestine, CD is disabling, incurable and of unknown etiology. The associated fibrosis comprises prolonged abnormal wound repair and tissue remodeling leading to excessive deposition of a collagen-containing extracellular matrix (ECM). Hypertrophy of the submucosa and muscularis are major contributors to the increased rigidity and thickness of the bowel wall. These changes typically lead to stricture formation and fibrostenosis, a major cause of serious complications and surgical procedures in CD patients. In the absence of fibrosis-specific drugs, anti-inflammatory therapies have not prevented or reduced the incidence of fibrosis. For the CD patients that succumb to this complication, surgical intestinal resection is currently the only treatment option and provides temporary symptomatic relief without cure or alteration of disease progression [2]. In this context, the ability to predict the patients that develop fibrosis remains an important challenge that would significantly improve the clinical management of IBD.

A number of factors have been proposed to have a role in the etiology of CD [1] [3, 4]. These include genetic susceptibility, defects in innate immunity, undefined environmental factors, and alterations in the microbiome. Recent genome-wide association studies appear to explain only a minority of the risk associated with development of CD [5, 6]. The high rate of discordance among monozygotic twins and

the increased prevalence of CD over recent decades suggest environmental factors may be at play. As epigenetic changes are dynamically responsive to the environment, they are likely to play a key role in the pathogenesis of fibrosis and to offer a molecular explanation for how the intestine becomes pro-fibrotic.

DNA methylation is the addition of a methyl group to the 5-position of the DNA base cytosine. Genome-wide changes in DNA methylation have been shown to be major contributors to cancer, mammalian development, gene transcription and phenotype in a range of diseases [7, 8], including a variety of autoimmune and inflammatory conditions. Recent epigenetic profiling of intestinal tissue from IBD patients have comprised DNA methylome signatures for both major types of Inflammatory Bowel disease: Ulcerative Colitis (UC) and CD [9-12]. Of the few previous tissue-based studies, two of these correlated changes in DNA methylation with gene expression [11, 12] and none analyzed purified disease-relevant cell types. Furthermore, three previous studies have analyzed epigenetic changes in intestinal fibrogenesis but none of these defined changes in DNA methylation. One, from our lab, showed that chromatin modifications are linked with activation of type I collagen gene expression in endothelial to mesenchymal transition [13], a feature of intestinal fibrosis *in vivo*. The two other labs focused on specific miRNAs in the fibrotic intestine [14, 15]. Moreover, all published studies to date of the DNA methylation in fibrotic diseases have been limited to studying restricted subsets of genes or to the use of microarrays that lack genome-wide coverage [16-22]. We report here the use of an unbiased, genome-wide approach to define the DNA methylome and the transcriptome of fibroblasts isolated from colons of control and CD patients. This approach avoids the issues of heterogeneity of tissue/biopsy samples and employs next generation sequencing using the methyl-CpG binding domain (MBD) of MBD2, called MiGS (MBD-isolated genome sequencing). MiGS is based on the capacity of MBD2 to bind with high affinity and specificity to DNA containing densely methylated cytosines [23]. A sequencing library heavily enriched for these methylated sequences within sheared genomic DNA is coupled to next generation sequencing so that the location of DNA methylation at specific genomic loci can be quantified and compared. We integrated this information with RNA-seq data to identify key molecular interactions that lead to fibrosis pathology. Our identification of key differentially methylated regions (DMRs) in

1
2
3
4 intestinal fibrosis provides new molecular characteristics for fibrostenotic CD and a resource for studying
5
6 epigenetic mechanisms that could help classify different stages of fibrosis and identify patients
7
8 predisposed to developing this major complication of IBD.
9

10 11 12 13 14 15 **RESULTS** 16 17 18 19

20 **Genome-wide changes in DNA methylation in fibrotic human intestinal fibroblasts** 21

22 Genome-wide DNA methylation profiles were generated from HIF isolated from colon resection
23
24 specimens that were either normal or from patients with fibrostenotic CD. We identified statistically
25
26 significant regions of differential DNA methylation between the two groups at a false discovery rate
27
28 (FDR) threshold of $< 5\%$. Both quantitative and qualitative differences in DNA methylation were
29
30 detected (Additional file 1). Qualitative differences were defined as sharp yes/no DNA methylation with a
31
32 clear presence or absence of DNA methylation between the normal and fibrotic groups. In sharp yes/no
33
34 DMRs, one group has zero or statistically near-zero sequencing reads, indicating a lack of methylation
35
36 whereas the other group shows a strong enrichment of reads, indicating the presence of methylation.
37
38 Heatmap visualization and hierarchical clustering of the sharp yes/no DMRs showed a striking difference
39
40 between the differentially methylated regions (DMRs) in normal and CD samples (Fig. 1A). These sharp
41
42 yes/no DMRs were detected throughout the genome on all 22 autosomes (Fig. 1B). Of the sharp yes/no
43
44 DMRs, 1,180 DMRs represent hypermethylation in the fibrotic samples as compared to the normal
45
46 samples, and 802 represent hypomethylation. We focused further analysis on qualitative differences in
47
48 methylation because the quantitative differences (where both conditions are statistically different and
49
50 have high numbers of reads) are likely methylated in all conditions [24].
51
52
53
54
55
56
57

58 **Genomic and CpG island context of differential DNA methylation in intestinal fibrosis** 59

60 Evidence from recent studies suggests that the genomic location of DNA methylation is a major
61
62
63
64
65

contributor to the type of function performed by this epigenetic modification [7, 8]. Based on the global distribution of the hyper- and hypo- DNA methylated loci, we determined the frequency of loci in promoter, inter- and intra-genic regions (Fig. 1C & Additional file 1). Promoter regions were defined as +1000 bp to -500 bp relative to transcription start sites, 3' end gene regions as +1000 bp to -1000 bp of transcription termination sites and intergenic regions as the remaining regions of the genome. The NCBI reference sequence database (RefSeq) was used to define the transcription units for all genes. In the cases where DMRs were located in multiple gene regions the overlaps were prioritized as follows: promoter > 3' end > exon > intron > intergenic. Only a minor percentage of the mapped differential DNA methylation occurred in promoters (5.0% and 2.7% for hypermethylated and hypomethylated DMRs, respectively), and a large percentage occurred in intergenic regions (40.8% and 48.4%). DNA methylation within gene bodies has also been reported to have an important role in transcriptional control [25]. While only a small percentage of DMRs overlapped with annotated exons (3.0% and 2.9%), DMRs were abundant within introns (48.6% and 43.1%).

We also asked whether methylated promoter CpG islands (CGIs), associated with repression of gene transcription [26], are a feature of intestinal fibrosis (Fig. 1D). The vast majority of differential DNA methylation was found outside CGIs. For hypermethylated and hypomethylated DMRs, respectively, only 5.9% and 1.9% were in CGIs, 9.7% and 6.4% in shores (2000 bp flanking CpG islands) and 4.0% and 4.1% in shelves (2000 bp regions flanking shores). Of all the DMRs, the open sea regions (loci greater than 4 kb from CpG islands) contained most of the differential non-CpG island methylation (86.3% and 87.7%). Published data indicate that alterations in DNA methylation outside CGIs in shelves/shores can play a role in gene transcription and may be cell type specific [27].

Genome-wide transcriptome analysis of fibrotic and normal intestinal fibroblasts

RNA-seq analysis identified the fibrosis-associated changes in the HIF transcriptome associated with changes in DNA methylation. Using established criteria for analyzing RNA-seq data from normal and fibrotic RNA, we found 72 genes that were differentially expressed (Benjamin-Hochberg adjusted p-value <0.05, FDR < 5%, Fig. 2A & Additional file 2). The consistency of gene expression within each group is evident from the hierarchical clustering dendrogram. Of the 72 differentially expressed genes, 31 were downregulated and 41 upregulated. Only 3 genes: toll like receptor 4 (*TLR4*) [28], interleukin 33, (*IL-33*) [29] and insulin like growth factor -1 (*IGF-1*), [30] have been previously reported or hypothesized to have functions in intestinal inflammation, IBD or other fibrotic diseases [31, 32].

WNT2B, *PTGIS*, and *PTGDS* are subject to differential expression and differential DNA methylation

We next examined which of the sharp yes/no DMRs in promoter regions that overlapped with genes that were differentially expressed (Additional file 3). Promoter hypermethylation of two protein-coding genes, wingless type MMTV integration site family, member 2B (*WNT2B*) [33] and prostacyclin synthase (*PTGIS*) [34], were inversely correlated with alterations in their mRNA levels (Additional file 3). Validation by RT-PCR showed the changes in *PTGIS* mRNA and *WNT2B* mRNA were decreased in fibrotic compared with normal HIF (Fig.2B), consistent with the observed hypermethylation of each gene. In contrast, the mRNA level of prostaglandin D2 synthase (*PTGDS*), another member of the eicosanoid (prostaglandin) family [35], was increased in fibrotic HIF (Fig. 2B). Bisulfite sequencing was performed on the *WNT2B*, *PTGIS* and *PTGDS* genes (Figs. 3A, 3B & Additional file 4) to validate the DMRs at base-pair resolution. The DMRs in the *WNT2B* and *PTGIS* promoters were confirmed to be hypermethylated (Fig. 3A & B). Bisulfite sequencing confirmed that *PTGDS* was hypomethylated in a 550 bp region spanning its promoter and coding sequence (Additional file 4). These findings corroborate the corresponding changes in expression and DNA methylation levels for all 3 genes (Additional file 4, Fig. 2B, 3A & 3B).

Novel functional gene networks in Crohn's disease associated fibrosis

As a first step in determining how differential DNA methylation contributes to intestinal fibrosis pathogenesis, we used the GeneMANIA algorithm to predict network-based functional associations between genes differentially methylated and differentially expressed in fibrotic HIF (List of input genes in Additional file 5). The attributed functions fall into two main groups. The first is extracellular matrix structure and organization, which is highly relevant to fibrosis of the intestine and other fibrotic disorders [36, 37]. The second is guanine nucleotide exchange factor signaling, including the small GTP proteins RHO and RAC [38], previously described in gastrointestinal ulcer healing and in other fibrotic diseases but not in intestinal fibrosis [39, 40].

Sub-networks that include the three differentially methylated and expressed genes *WNT2B*, *PTGIS* and *PTGDS* (Fig.4) and the genes that interact with them were identified (Additional files 6-8). *PTGDS* and *PTGIS* were associated with fibulin 1 (*FBLN1*) and *IGF-1*, both linked with extracellular matrix structure and fibrosis [30, 41]. *PTGIS* is also linked with other genes involved in the organization and structure of the extracellular matrix, for example, ADAM metalloprotease with thrombospondin motifs 5, (*ADAMTS5*) and the alpha 2 chain of type I collagen (*COL1A2*) [41-43]. There are also other interactions not previously described in any fibrotic disease, for example, both *PTGDS* and *WNT2B* are linked with vasohibin 2 (*VASH2*) [44] and chromosome 7 open reading frame 69 (*C7orf69*).

DISCUSSION

The objective of our studies was to obtain novel insights into the molecular mechanisms that underlie intestinal fibrogenesis. Through definition and integration of the DNA methylome and transcriptome, we have revealed functional candidate gene networks including DMRs inversely correlated with gene

1
2
3
4 expression in three protein coding genes: *WNT2B*, *PTGIS* and *PTGDS*, none of which have previously
5
6 been described in CD or other fibrotic diseases. By employing the next generation sequencing-based
7
8 approach, MiGS [23] and RNA-Seq, we were able to achieve genome wide coverage and improved
9
10 accuracy over previous studies that have used microarrays to profile changes in DNA methylation and
11
12 gene expression in inflammatory bowel disease and other fibrotic disorders.
13
14

15
16
17 Our experiments were performed in HIF purified from fibrotic or normal colon tissue to minimize the
18
19 confounding effects of cell type heterogeneity of whole blood or tissue that have been widely used in
20
21 previous epigenetic studies [27]. The large number, size and genome-wide distribution of DMRs in
22
23 fibrotic HIF suggest that this epigenetic modification has an important role in intestinal fibrosis. We
24
25 observed sharp yes/no differences in DNA methylation throughout the genic and intergenic regions of the
26
27 fibrotic HIF genome. Only a comparatively low percentage of DMRs occurred in promoters of genes and
28
29 in CpG islands. The vast majority of DMRs were located within introns and intergenic regions. Intergenic
30
31 sequences contain enhancers and insulators that are associated with regulation of gene expression during
32
33 differentiation and organogenesis [7, 8]. Other distinct functions for DNA methylation have been
34
35 proposed within introns and CpG island shores [27, 45, 46]. For example, DNA methylation in introns
36
37 can modulate alternative exon splicing [45, 46] and in shores, differential DNA methylation is tissue-
38
39 specific and may regulate transcription from alternative start sites [27].
40
41
42
43
44
45

46
47 The precise mechanism by which DNA methylation regulates gene transcription remains poorly defined.
48
49 Methylated CpGs can prevent the binding of some transcription factors [7] and other evidence indicates
50
51 that methylation of DNA can alter nucleosome occupancy and alternative polyadenylation of mRNA
52
53 during transcription [47-49]. Whether *de novo* DNA methylation directly represses genes, or whether
54
55 gene silencing precedes or follows methylation, is still debated but the available evidence indicates that
56
57 this will be dependent in large part on the genomic location of the methylated DNA sequence [7, 8]. This
58
59 is a rapidly advancing area with new compelling evidence for causal associations between DNA
60
61
62
63
64
65

methylation changes and phenotype [50, 51]. Our detection of DMRs provides a resource of candidate loci for future work to understand the role of DNA methylation in different genomic locations at different stages of pro-fibrotic gene expression.

We identified 3 genes: *WNT2B*, *PTGIS* and *PTGDS* in which differential methylation of the DNA in each promoter region was inversely correlated with gene expression in fibrotic HIF. Complex roles for the WNT signaling pathway in the fibrosis of a number of organs including the lung and kidney and in controlling tissue homeostasis, cell proliferation, migration, differentiation, apoptosis and organogenesis have been reported [52, 53], however, this family of genes has not previously been associated with intestinal fibrosis. The decreased expression of hypermethylated *WNT2B* is a novel finding both in the context of fibrosis and CD, and increased expression of *WNT2B* has been detected in the intestinal mucosa of UC patients [54] suggesting that *WNT2B* may perform distinct functions in CD and UC.

In the case of the prostaglandin family, central functions are well established in homeostasis, inflammation and other fibrosis disorders [55]. Profibrogenic functions for prostaglandin F₂ α , and anti-fibrotic effects of prostaglandin E₂ have been documented in lung fibrosis [56, 57]. In addition, the product of *PTGIS* catalysis, PGI₂ has been shown to be anti-fibrotic in the lung [55] and also in a prostacyclin receptor null mouse model of cardiac fibrosis [58]. We found that *PTGIS* was hypermethylated and repressed in fibrotic HIF. Hypermethylation of the *PTGIS* promoter is a feature of colorectal cancer but decreased expression of *PTGIS* has not previously been linked with IBD [59]. A further finding of our studies was that the lipocalin type *PTGDS* gene was hypomethylated and its mRNA levels increased in fibrotic HIF. Our observations are supported by studies showing i) increased levels of lipocalin type *PTGDS* in colitis, ii) increased expression in UC patients in parallel with disease activity, iii) decreased DSS-induced colitis in L-*PTGDS* knockout mice [60]. Conflicting data in other studies indicate that in the kidney, PGD₂, the product of *PGDS* catalytic activity, promotes the development of tubulointerstitial fibrosis [61] whereas a protective function of PGD₂ has been reported in TNBS colitis

[62]. These observations may reflect tissue specific differences or methodological discrepancies between the aforementioned studies.

Based on previous studies including a recent study of fibrotic liver disease, the average age difference of 20 years between our fibrotic Crohn's disease and normal patient groups is highly unlikely to be a confounding factor in the differential DNA methylation we identified [22]. Comparison of the DNA methylomes of dermal fibroblasts from individuals with a much greater disparity in age (<23 years old versus > 60 years old) showed a change in DNA methylation levels of >15% at only 75 CpG sites, none of which were highly significant ($p < 0.001$) [63]. More importantly, none of these sites were differentially methylated in fibrotic HIF.

Studies in tissues such as the heart, and in diseases such as cancer, indicate that the same epigenetic pathways that are beneficial in organ development, are perturbed in the fully developed or diseased tissue. The concept of antagonistic pleiotropy [64] suggests that key pathways are advantageous at early stages of development, e.g., in myofibroblast differentiation and wound healing, but in later life, these same pathways become detrimental as they promote myofibroblast activation and progressive fibrosis [64, 65]. Thus, the available evidence indicates that alterations in DNA methylation during development could serve as key determinants of the pro-fibrotic phenotype. Together with genetic and environmental factors, DNA methylation could be a major contributor to the progression of injury in CD and to the establishment of intestinal fibrosis. An important aim of future studies will be to determine changes in DNA methylation in patients with different stages of fibrostenotic CD. This could help stratify distinct subtypes of CD and identify those patients most likely to develop intestinal fibrosis

CONCLUSIONS

Our detailed map of the methylome integrated with the transcriptome has revealed fibrosis-associated regulation of *WNT2B*, *PTGIS* and *PTGDS*. The networks for these genes reveal interactions with other factors important in development and extracellular matrix synthesis. This suggests that key sites of differential DNA methylation can lead to the molecular aberrations that underlie IBD-associated fibrosis and provide potential targets for future development of prognostic strategies and specific anti-fibrotic therapies. Our data also provide a basis for defining the contribution of genome wide changes in DNA methylation to the pathogenesis of fibrotic disorders.

METHODS

Colon tissue and cell culture

Colon resection specimens were obtained from patients with CD-associated fibrosis and control colon specimens were obtained from histologically normal tissue from patients with Diverticulitis (Additional file 9) under the approval of the Cleveland Clinic institutional review board (IRB 06-050). The approved protocol included a waiver of informed consent for redundant tissues obtained through the Cleveland Clinic Human Tissue Procurement Facility in accordance with Cleveland Clinic IRB policy. Tissue specimens were de-identified and made available following authorization by a Cleveland Clinic surgical pathologist. Human intestinal fibroblasts (HIF) were cultured from colon specimens and their purity confirmed as previously reported by established methods [66].

DNA methylome profiling

Genome-wide DNA methylation was mapped using MBD-isolated genome sequencing (MiGS) [23], with the following differences. Genomic DNA (DNEasy kit, Qiagen) was extracted from normal and fibrotic

cultured human intestinal fibroblasts (HIF). DNA (10µg) was sheared on a Covaris S220 to an average size of 120 bp. Methylated DNA fragments were purified on PrepEase DNA columns (USB) and captured with the methyl-CpG binding protein MBD2 (MethylMiner Methylated DNA Enrichment kit #ME10025, Invitrogen). A sequencing library for each immunoprecipitated DNA sample was prepared using Illumina's ChIP-seq DNA Sample Prep Kit. The quality of the DNA was assessed on an Agilent Bioanalyser. Sequencing libraries of methylated DNA fragments were analyzed and next generation sequencing performed on an Illumina HiSeq 2000. Read lengths of 50 bp were sequenced at an average depth of 161,226,405 reads (minimum: 130,528,176 reads, maximum: 190,943,105 reads). Reads were aligned to assembly *hg19* of the human genome (Genome Reference Consortium Human Build 37) using the *bowtie2* short read alignment software (options: -N 1 -L 20 --phred33) [67].

Bioinformatic Analysis of Differential DNA Methylation

Differential DNA methylation was detected using the *iDPT* (integrated signal Deconvolution, Pattern recognition, and differential Testing) framework [68] (<https://idpt.github.io/dptscan/>). Using this framework, reads were extended to the average fragment length. Average coverage (e.g., averages of the per-base coverage rounded to integers) was computed within 50 bp non-overlapping windows tiling the entire genome. Initial filtering removed all windows where no sample had a read count above a false discovery threshold established by a Poisson distribution that assumed an even coverage of reads across the genome. Signal deconvolution was performed using a Bayesian mixture of three Poissons model, producing a probability of methylation for each window in each sample. Pattern recognition used these probabilities and a scan statistic to produce regions of consistent methylation status, allowing for small gaps.

Finally, differential testing was performed using a linear mixed-effects model and an output table that contained fold-changes and multiple testing-adjusted p-values for each DMR was produced. The output differentiated between sites showing a sharp yes/no methylation difference (defined as all samples in one

group are predicted by the model to have no methylation and all samples in the other group are predicted to be methylated) versus sites that show a quantitative difference in reads (the model predicts all samples across all groups to be methylated, but the read counts show a statistically significant difference). Only the sharp sites were used for subsequent analysis because they are more likely to represent differential DNA methylation rather than other phenomena that can change read counts, such as copy number variation.

The R package *goldmine* was used to analyze the genomic context of the detected regions with respect to known genes and features (<http://jeffbhasin.github.io/goldmine>). The *goldmine* functions, *getCpgFeatures*, *getFeatures*, and *getGenes* were used to obtain CpG islands, ENCODE supertracks, and RefSeq gene models from the database of the UCSC Genome Browser [69, 70]. This data was used by the *goldmine* function to provide detailed annotation of how these datasets relate to the regions.

RNA Sequencing

For RNA Sequencing (RNA-seq), RNA was extracted from control and fibrotic HIF using the RNEasy mini kit (Qiagen). A gene sequencing library for each sample (1 µg RNA) was prepared using the TruSeq™ RNA Sample Prep Kit v2. Next-generation sequencing was performed on an Illumina HiSeq 2500. After quality assessment, reads were aligned to *hg19* using GSNAP [71]. Read lengths of 100 bp were sequenced at an average depth of 53,390,000 reads (minimum: 39,470,000 reads, maximum: 62,550,000 reads). Only reads that aligned to a single locus in the *hg19* genome assembly were retained for further analysis. Read counts for RefSeq genes were computed using HTSeq-count program [72]. Differential expression analysis was performed on this count table using the negative binomial test provided by the DESeq package [73] in the R statistical computing environment (<http://www.R-project.org/>). Transcripts with p-values less than 0.05 after adjustment by the Benjamini-Hochberg procedure were considered statistically significant.

Bisulfite sequencing

For validation of DMRs detected by the genome-wide DNA methylome analysis, genomic DNA (330ng) was prepared using the DNEasy kit (Qiagen). Bisulfite conversion of the DNA was performed with the EZ DNA Methylation-Lightning Kit #D5030 (Zymo). Bisulfite converted DNA (1μl) was amplified by PCR with iTaq DNA polymerase (BioRad) and primers designed using the MethPrimer website (see Additional file 10). DNA was cloned (Topo TA cloning kit, K4575-01, Invitrogen) and the cloned DNA (prepared using PureLink Miniprep kit, Invitrogen) was digested with EcoR1 to verify cloning of the inserted DNA. DNA from at least 10 different clones was then sequenced at the Genomics Core of Cleveland Clinic's Lerner Research Institute. DNA methylation of individual CpG sites was analyzed using Quma software (<http://quma.cdb.riken.jp/>).

Reverse Transcriptase-PCR (RT-PCR)

For validation of changes in mRNA levels of differentially DNA methylated genes, total RNA was extracted from fibrotic and normal HIF, reverse transcribed, and complementary DNA (cDNA) was amplified by RT-PCR. Five microliters of cDNA was amplified in the presence of 0.125 μmol/L each of the 5' and 3' primers (Biosynthesis, Lewisville, TX) and 1U of Taq DNA Polymerase (Roche, Mannheim, Germany). PCR was performed in a DNA thermal cycler using pre-optimized temperatures and times, and the primers (listed in Additional file 10) were used to quantify mRNA levels. Fifteen microliters of the PCR product were subjected to electrophoresis on 1.5% agarose gel and stained with 0.5μg/ml ethidium bromide, using 100 bp DNA ladder as a marker.

GeneMANIA network analysis

The GeneMANIA algorithm (version 3.2.1, <http://www.genemania.org>) employs functional interaction data to create a network containing putative functional links between genes. We used the GeneMANIA Cytoscape Plug-In [74] to generate a network for the genes identified in this study based on protein-

protein interactions, genetic interactions and co-expression profile databases. Data on shared protein domains and co-localization were eliminated from the analysis, to minimize the number of false positives. A sub-network was extracted by restricting the network to only nodes that interact with defined genes of interest, and was plotted using Gephi (<https://gephi.github.io/>)

Availability of supporting data

The DNA methylation and RNA-seq data sets (raw and normalized) supporting the results of this article are available in the NCBI Gene Expression Omnibus (GEO) repository: <http://www.ncbi.nih.gov/geo/>, accession number GSE67250.

ABBREVIATIONS

Crohn's disease (CD); human intestinal fibroblasts (HIF); extracellular matrix (ECM); inflammatory bowel disease (IBD); ulcerative colitis (UC); methyl CpG binding domain (MBD); methyl CpG binding domain-isolated genome sequencing (MiGS); differentially methylated regions (DMRs); false discovery rate (FDR); CpG islands (CGIs); wingless type mouse mammary tumor virus integration site family member 2B (WNT2B); prostacyclin synthase (PTGIS); prostaglandin D2 synthase (PTGDS); insulin-like growth factor 1 (IGF-1); fibulin-1 (FBLN1); vasohibin-2 (VASH2); retinoid-related orphan nuclear receptor beta (RORB).

COMPETING INTERESTS

The authors declare they have no competing interests.

AUTHORS' CONTRIBUTIONS

T.S performed the experiments, generated and analyzed data; J.M.B assembled, analyzed and interpreted data. Y.X. and A.H.T developed the DNA methylome algorithm. Y.X., J.B.-S. and Y.C. analyzed, and interpreted data and performed statistical analysis. T.S., J.M.B, A.H.T and E.S. drafted the manuscript. E.S devised the study concept; A.H.T and E.S. co-designed the study. E.S. obtained funding and supervised the study. All authors critically reviewed and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

Support for T.S. and E.S was from the Broad Foundation (IBD-032) and the National Institutes of Health (NIH) (DK093630 and DK050984). J.M.B. is a predoctoral student in the Molecular Medicine PhD Program of the Cleveland Clinic and Case Western Reserve University that is funded in part by the “Med into Grad” initiative of the Howard Hughes Medical Institute (HHMI). J.M.B. and A.H.T. are supported by the National Cancer Institute, NIH (R01CA154356 and F31CA195887). We are grateful to Eboni Ubinas and Tiffany Hollo and all staff of the Tissue Procurement Service of the Cleveland Clinic for help with the collection of intestinal tissue specimens. We also gratefully acknowledge the technical assistance with next generation sequencing of the McGill University and Génome Québec Innovation Center, Montreal, Canada, and the Genomics Core, Case Western Reserve University, Cleveland, Ohio.

REFERENCES

1. Kaser A, Zeissig S, Blumberg RS: Inflammatory bowel disease. *Annu Rev Immunol* 2010, 28:573-621.
2. Ahmed T, Rieder F, Fiocchi C, Achkar JP: Pathogenesis of postoperative recurrence in Crohn's disease. *Gut* 2011, 60(4):553-562.
3. Strober W: Impact of the gut microbiome on mucosal inflammation. *Trends in immunology* 2013, 34(9):423-430.
4. Scarpa M, Stylianou E: Epigenetics: Concepts and relevance to IBD pathogenesis. *Inflamm Bowel Dis* 2012, 18(10):1982-1996.
5. Ventham NT, Kennedy NA, Nimmo ER, Satsangi J: Beyond gene discovery in inflammatory bowel disease: the emerging role of epigenetics. *Gastroenterology* 2013, 145(2):293-308.
6. Stylianou E: Epigenetics: the fine-tuner in inflammatory bowel disease? *Curr Opin Gastroenterol* 2013, 29(4):370-377.
7. Jones PA: Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012, 13(7):484-492.
8. Schubeler D: Function and information content of DNA methylation. *Nature* 2015, 517(7534):321-326.
9. Nimmo ER, Prendergast JG, Aldhous MC, Kennedy NA, Henderson P, Drummond HE, Ramsahoye BH, Wilson DC, Semple CA, Satsangi J: Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. *Inflamm Bowel Dis* 2011.
10. Cooke J, Zhang H, Greger L, Silva AL, Massey D, Dawson C, Metz A, Ibrahim A, Parkes M: Mucosal genome-wide methylation changes in inflammatory bowel disease. *Inflamm Bowel Dis* 2012, 18(11):2128-2137.

11. Hasler R, Feng Z, Backdahl L, Spehlmann ME, Franke A, Teschendorff A, Rakyan VK, Down TA, Wilson GA, Feber A *et al*: A functional methylome map of ulcerative colitis. *Genome research* 2012, 22(11):2130-2137.
12. Adams AT, Kennedy NA, Hansen R, Ventham NT, O'Leary KR, Drummond HE, Noble CL, El-Omar E, Russell RK, Wilson DC *et al*: Two-stage genome-wide methylation profiling in childhood-onset Crohn's Disease implicates epigenetic alterations at the VMP1/MIR21 and HLA loci. *Inflamm Bowel Dis* 2014, 20(10):1784-1793.
13. Sadler T, Scarpa M, Rieder F, West G, Stylianou E: Cytokine-induced chromatin modifications of the type I collagen alpha 2 gene during intestinal endothelial-to-mesenchymal transition. *Inflamm Bowel Dis* 2013, 19(7):1354-1364.
14. Chen Y, Ge W, Xu L, Qu C, Zhu M, Zhang W, Xiao Y: miR-200b is involved in intestinal fibrosis of Crohn's disease. *Int J Mol Med* 2012, 29(4):601-606.
15. Nijhuis A, Biancheri P, Lewis A, Bishop CL, Giuffrida P, Chan C, Feakins R, Poulsom R, Di Sabatino A, Corazza GR *et al*: In Crohn's disease fibrosis-reduced expression of the miR-29 family enhances collagen expression in intestinal fibroblasts. *Clin Sci (Lond)* 2014, 127(5):341-350.
16. Wing MR, Devaney JM, Joffe MM, Xie D, Feldman HI, Dominic EA, Guzman NJ, Ramezani A, Susztak K, Herman JG *et al*: DNA methylation profile associated with rapid decline in kidney function: findings from the CRIC study. *Nephrol Dial Transplant* 2014, 29(4):864-872.
17. Ko YA, Mohtat D, Suzuki M, Park AS, Izquierdo MC, Han SY, Kang HM, Si H, Hostetter T, Pullman JM *et al*: Cytosine methylation changes in enhancer regions of core pro-fibrotic genes characterize kidney fibrosis development. *Genome Biol* 2013, 14(10):R108.
18. Komatsu Y, Waku T, Iwasaki N, Ono W, Yamaguchi C, Yanagisawa J: Global analysis of DNA methylation in early-stage liver fibrosis. *BMC Med Genomics* 2012, 5:5.

- 1
2
3
4 19. Huang SK, Scruggs AM, McEachin RC, White ES, Peters-Golden M: Lung fibroblasts from
5 patients with idiopathic pulmonary fibrosis exhibit genome-wide differences in DNA methylation
6 compared to fibroblasts from nonfibrotic lung. *PLoS One* 2014, 9(9):e107055.
7
8
9
10
11 20. Rabinovich EI, Kapetanaki MG, Steinfeld I, Gibson KF, Pandit KV, Yu G, Yakhini Z, Kaminski
12 N: Global methylation patterns in idiopathic pulmonary fibrosis. *PLoS One* 2012, 7(4):e33770.
13
14
15 21. Sanders YY, Ambalavanan N, Halloran B, Zhang X, Liu H, Crossman DK, Bray M, Zhang K,
16 Thannickal VJ, Hagood JS: Altered DNA methylation profile in idiopathic pulmonary fibrosis.
17 *Am J Respir Crit Care Med* 2012, 186(6):525-535.
18
19
20
21 22. Zeybel M, Hardy T, Robinson SM, Fox C, Anstee QM, Ness T, Masson S, Mathers JC, French J,
22 White S *et al*: Differential DNA methylation of genes involved in fibrosis progression in non-
23 alcoholic fatty liver disease and alcoholic liver disease. *Clinical epigenetics* 2015, 7(1):25.
24
25
26
27 23. Serre D, Lee BH, Ting AH: MBD-isolated Genome Sequencing provides a high-throughput and
28 comprehensive survey of DNA methylation in the human genome. *Nucleic acids research* 2010,
29 38(2):391-399.
30
31
32
33 24. Robinson MD, Strbenac D, Stirzaker C, Statham AL, Song J, Speed TP, Clark SJ: Copy-number-
34 aware differential analysis of quantitative DNA sequencing data. *Genome research* 2012,
35 22(12):2489-2496.
36
37
38
39 25. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G: Gene body methylation can alter
40 gene expression and is a therapeutic target in cancer. *Cancer Cell* 2014, 26(4):577-590.
41
42
43
44 26. Feinberg AP: Phenotypic plasticity and the epigenetics of human disease. *Nature* 2007,
45 447(7143):433-440.
46
47
48
49 27. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione
50 M, Webster M *et al*: The human colon cancer methylome shows similar hypo- and
51 hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009, 41(2):178-186.
52
53
54
55 28. Franchimont D, Vermeire S, El Housni H, Pierik M, Van Steen K, Gustot T, Quertinmont E,
56 Abramowicz M, Van Gossum A, Deviere J *et al*: Deficient host-bacteria interactions in
57
58
59
60
61
62
63
64
65

- inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis. *Gut* 2004, 53(7):987-992.
29. Pastorelli L, De Salvo C, Cominelli MA, Vecchi M, Pizarro TT: Novel cytokine signaling pathways in inflammatory bowel disease: insight into the dichotomous functions of IL-33 during chronic intestinal inflammation. *Therap Adv Gastroenterol* 2011, 4(5):311-323.
 30. Simmons JG, Pucilowska JB, Keku TO, Lund PK: IGF-I and TGF-beta1 have distinct effects on phenotype and proliferation of intestinal fibroblasts. *Am J Physiol Gastrointest Liver Physiol* 2002, 283(3):G809-818.
 31. McHedlidze T, Waldner M, Zopf S, Walker J, Rankin AL, Schuchmann M, Voehringer D, McKenzie AN, Neurath MF, Pflanz S *et al*: Interleukin-33-dependent innate lymphoid cells mediate hepatic fibrosis. *Immunity* 2013, 39(2):357-371.
 32. Seki E, De Minicis S, Osterreicher CH, Kluwe J, Osawa Y, Brenner DA, Schwabe RF: TLR4 enhances TGF-beta signaling and hepatic fibrosis. *Nat Med* 2007, 13(11):1324-1332.
 33. Poulain M, Ober EA: Interplay between Wnt2 and Wnt2bb controls multiple steps of early foregut-derived organ development. *Development* 2011, 138(16):3557-3568.
 34. Yokoyama C, Yabuki T, Inoue H, Tone Y, Hara S, Hatae T, Nagata M, Takahashi EI, Tanabe T: Human gene encoding prostacyclin synthase (PTGIS): genomic organization, chromosomal localization, and promoter activity. *Genomics* 1996, 36(2):296-304.
 35. Joo M, Sadikot RT: PGD synthase and PGD2 in immune resposne. *Mediators Inflamm* 2012, 2012:503128.
 36. Rieder F, Fiocchi C: Intestinal fibrosis in IBD--a dynamic, multifactorial process. *Nat Rev Gastroenterol Hepatol* 2009, 6(4):228-235.
 37. Spagnolo P, Sverzellati N, Rossi G, Cavazza A, Tzouvelekis A, Crestani B, Vancheri C: Idiopathic pulmonary fibrosis: An update. *Ann Med* 2015:1-13.
 38. Narumiya S: The small GTPase Rho: cellular functions and signal transduction. *J Biochem* 1996, 120(2):215-228.

- 1
2
3
4 39. Choi SS, Sicklick JK, Ma Q, Yang L, Huang J, Qi Y, Chen W, Li YX, Goldschmidt-Clermont PJ,
5
6 Diehl AM: Sustained activation of Rac1 in hepatic stellate cells promotes liver injury and fibrosis
7
8 in mice. *Hepatology* 2006, 44(5):1267-1277.
9
- 10
11 40. Tarnawski AS: Cellular and molecular mechanisms of gastrointestinal ulcer healing. *Dig Dis Sci*
12
13 2005, 50 Suppl 1:S24-33.
14
- 15 41. Jaffar J, Unger S, Corte TJ, Keller M, Wolters PJ, Richeldi L, Cerri S, Prele CM, Hansbro PM,
16
17 Argraves WS *et al*: Fibulin-1 predicts disease progression in patients with idiopathic pulmonary
18
19 fibrosis. *Chest* 2014, 146(4):1055-1063.
20
- 21
22 42. Hattori N, Carrino DA, Lauer ME, Vasanji A, Wylie JD, Nelson CM, Apte SS: Pericellular
23
24 versican regulates the fibroblast-myofibroblast transition: a role for ADAMTS5 protease-
25
26 mediated proteolysis. *J Biol Chem* 2011, 286(39):34298-34310.
27
- 28
29 43. Schnapp LM, Hatch N, Ramos DM, Klimanskaya IV, Sheppard D, Pytela R: The human integrin
30
31 alpha 8 beta 1 functions as a receptor for tenascin, fibronectin, and vitronectin. *J Biol Chem* 1995,
32
33 270(39):23196-23202.
34
- 35 44. Wang X, Zhu H, Yang X, Bi Y, Cui S: Vasohibin attenuates bleomycin induced pulmonary
36
37 fibrosis via inhibition of angiogenesis in mice. *Pathology* 2010, 42(5):457-462.
38
- 39
40 45. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE,
41
42 Hong C, Nielsen C, Zhao Y *et al*: Conserved role of intragenic DNA methylation in regulating
43
44 alternative promoters. *Nature* 2010, 466(7303):253-257.
45
- 46 46. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P,
47
48 Sandberg R, Oberdoerffer S: CTCF-promoted RNA polymerase II pausing links DNA
49
50 methylation to splicing. *Nature* 2011, 479(7371):74-79.
51
- 52
53 47. Choy JS, Wei S, Lee JY, Tan S, Chu S, Lee TH: DNA methylation increases nucleosome
54
55 compaction and rigidity. *Journal of the American Chemical Society* 2010, 132(6):1782-1783.
56
57
58
59
60
61
62
63
64
65

48. Wood AJ, Schulz R, Woodfine K, Koltowska K, Beechey CV, Peters J, Bourc'his D, Oakey RJ: Regulation of alternative polyadenylation by genomic imprinting. *Genes & development* 2008, 22(9):1141-1146.
49. Lay FD, Triche TJ, Jr., Tsai YC, Su SF, Martin SE, Daneshmand S, Skinner EC, Liang G, Chihara Y, Jones PA: Reprogramming of the human intestinal epigenome by surgical tissue transposition. *Genome research* 2014, 24(4):545-553.
50. Yu DH, Waterland RA, Zhang P, Schady D, Chen MH, Guan Y, Gadkari M, Shen L: Targeted p16(Ink4a) epimutation causes tumorigenesis and reduces survival in mice. *The Journal of clinical investigation* 2014, 124(9):3708-3712.
51. Chaikind B, Ostermeier M: Directed evolution of improved zinc finger methyltransferases. *PLoS One* 2014, 9(5):e96931.
52. Kahn M: Can we safely target the WNT pathway? *Nature reviews Drug discovery* 2014, 13(7):513-532.
53. Verzi MP, Shivdasani RA: Wnt signaling in gut organogenesis. *Organogenesis* 2008, 4(2):87-91.
54. You J, Nguyen AV, Albers CG, Lin F, Holcombe RF: Wnt pathway-related gene expression in inflammatory bowel disease. *Dig Dis Sci* 2008, 53(4):1013-1019.
55. Castellino FV: Lipids and eicosanoids in fibrosis: emerging targets for therapy. *Curr Opin Rheumatol* 2012, 24(6):649-655.
56. Wilborn J, Crofford LJ, Burdick MD, Kunkel SL, Strieter RM, Peters-Golden M: Cultured lung fibroblasts isolated from patients with idiopathic pulmonary fibrosis have a diminished capacity to synthesize prostaglandin E2 and to express cyclooxygenase-2. *The Journal of clinical investigation* 1995, 95(4):1861-1868.
57. Oga T, Matsuoka T, Yao C, Nonomura K, Kitaoka S, Sakata D, Kita Y, Tanizawa K, Taguchi Y, Chin K *et al*: Prostaglandin F(2alpha) receptor signaling facilitates bleomycin-induced pulmonary fibrosis independently of transforming growth factor-beta. *Nat Med* 2009, 15(12):1426-1430.

- 1
- 2
- 3
- 4 58. Francois H, Athirakul K, Howell D, Dash R, Mao L, Kim HS, Rockman HA, Fitzgerald GA,
- 5
- 6 Koller BH, Coffman TM: Prostacyclin protects against elevated blood pressure and cardiac
- 7
- 8 fibrosis. *Cell Metab* 2005, 2(3):201-207.
- 9
- 10
- 11 59. Frigola J, Munoz M, Clark SJ, Moreno V, Capella G, Peinado MA: Hypermethylation of the
- 12
- 13 prostacyclin synthase (PTGIS) promoter is a frequent event in colorectal cancer and associated
- 14
- 15 with aneuploidy. *Oncogene* 2005, 24(49):7320-7326.
- 16
- 17 60. Hokari R, Kurihara C, Nagata N, Aritake K, Okada Y, Watanabe C, Komoto S, Nakamura M,
- 18
- 19 Kawaguchi A, Nagao S *et al*: Increased expression of lipocalin-type-prostaglandin D synthase in
- 20
- 21 ulcerative colitis and exacerbating role in murine colitis. *Am J Physiol Gastrointest Liver Physiol*
- 22
- 23 2011, 300(3):G401-408.
- 24
- 25
- 26 61. Ito H, Yan X, Nagata N, Aritake K, Katsumata Y, Matsushashi T, Nakamura M, Hirai H, Urade Y,
- 27
- 28 Asano K *et al*: PGD2-CRTH2 pathway promotes tubulointerstitial fibrosis. *J Am Soc Nephrol*
- 29
- 30 2012, 23(11):1797-1809.
- 31
- 32
- 33 62. Ajuebor MN, Singh A, Wallace JL: Cyclooxygenase-2-derived prostaglandin D(2) is an early
- 34
- 35 anti-inflammatory signal in experimental colitis. *Am J Physiol Gastrointest Liver Physiol* 2000,
- 36
- 37 279(1):G238-244.
- 38
- 39
- 40 63. Koch CM, Suschek CV, Lin Q, Bork S, Goergens M, Joussen S, Pallua N, Ho AD, Zenke M,
- 41
- 42 Wagner W: Specific age-associated DNA methylation changes in human dermal fibroblasts.
- 43
- 44 *PLoS One* 2011, 6(2):e16679.
- 45
- 46 64. Thannickal VJ: Aging, antagonistic pleiotropy and fibrotic disease. *Int J Biochem Cell Biol* 2010,
- 47
- 48 42(9):1398-1400.
- 49
- 50
- 51 65. Hecker L, Vittal R, Jones T, Jagirdar R, Luckhardt TR, Horowitz JC, Pennathur S, Martinez FJ,
- 52
- 53 Thannickal VJ: NADPH oxidase-4 mediates myofibroblast activation and fibrogenic responses to
- 54
- 55 lung injury. *Nat Med* 2009, 15(9):1077-1081.
- 56
- 57
- 58
- 59
- 60
- 61
- 62
- 63
- 64
- 65

- 1
2
3
4 66. Strong SA, Pizarro TT, Klein JS, Cominelli F, Fiocchi C: Proinflammatory cytokines
5 differentially modulate their own expression in human intestinal mucosal mesenchymal cells.
6 *Gastroenterology* 1998, 114(6):1244-1256.
7
8
9
10
11 67. Langmead B, Salzberg SL: Fast gapped-read alignment with Bowtie 2. *Nature methods* 2012,
12 9(4):357-359.
13
14
15 68. Xu Y, Hu B, Choi AJ, Gopalan B, Lee BH, Kalady MF, Church JM, Ting AH: Unique DNA
16 methylome profiles in CpG island methylator phenotype colon cancers. *Genome research* 2012,
17 22(2):283-291.
18
19
20
21
22 69. Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, Dreszer TR, Fujita PA,
23 Guruvadoo L, Haeussler M *et al*: The UCSC Genome Browser database: 2014 update. *Nucleic*
24 *acids research* 2014, 42(Database issue):D764-770.
25
26
27
28
29 70. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, Wong MC,
30 Maddren M, Fang R, Heitner SG *et al*: ENCODE data in the UCSC Genome Browser: year 5
31 update. *Nucleic acids research* 2013, 41(Database issue):D56-63.
32
33
34
35 71. Wu TD, Nacu S: Fast and SNP-tolerant detection of complex variants and splicing in short reads.
36 *Bioinformatics* 2010, 26(7):873-881.
37
38
39
40 72. Anders S, Pyl PT, Huber W: HTSeq--a Python framework to work with high-throughput
41 sequencing data. *Bioinformatics* 2015, 31(2):166-169.
42
43
44 73. Anders S, Huber W: Differential expression analysis for sequence count data. *Genome Biol* 2010,
45 11(10):R106.
46
47
48
49 74. Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, Morris Q, Bader GD:
50 GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics*
51 2010, 26(22):2927-2928.
52
53
54
55
56
57
58
59
60
61
62
63
64
65

FIGURE LEGENDS

Figure 1. Genome-wide differentially methylated regions (DMRs) in fibrotic human intestinal fibroblasts (HIF).

(A) Heatmap of read counts at all sharp yes/no DMRs. Square root transformed read counts were plotted for each detected DMR. Higher read counts (red) indicate stronger evidence for the presence of methylation. Lower and zero read counts (blue) indicate absence of methylation. Each row represents one DMR. Read counts have been normalized to the number of 50 bp windows in each DMR. (B) Karyogram showing genome-wide coverage of differentially methylated regions (DMRs). Hypermethylated (red) and hypomethylated (blue) regions in Crohn's disease HIF when compared to normal controls are shown. Black lines above each chromosome represent regions covered by sequencing reads to show genome-wide coverage. (C) The proportion of DMRs that overlap with promoters, gene 3' ends, exons, introns, and intergenic regions is shown. RefSeq genes were used to define transcription units. (D) The number of sharp yes/no DNA methylated loci in CpG islands versus CpG shores (2000 bp flanking CpG islands), shelves (2000 bp regions flanking shores) and open sea regions (loci greater than 4 kb from CpG islands) is shown.

Figure 2. Differential gene expression profiles of HIF isolated from normal colons and Crohn's Disease (CD) fibrotic colons.

(A) Heatmap and hierarchical clustering dendrogram of transcript abundance from RNA-seq performed on HIF RNA from three normal colons and three CD fibrotic colons depicting differentially expressed genes with reduced expression in CD fibrotic colon compared to normal colon. Red represents up-regulation of the gene's expression, and blue down-regulation. (B) Heatmap depicting differentially expressed genes with increased expression in CD fibrotic colon compared to normal colon. (C) Validation

by RT-PCR of the fold change in *PTGDS*, *PTGIS*, and *WNT2B* mRNA levels in CD fibrotic versus control fibroblasts is shown (n=7; *PTGDS* p=0.0074, *PTGIS* p = 0.021 and *WNT2B* p = 0.015).

Figure. 3. Bisulfite sequencing validation of differentially methylated *WNT2B* and *PTGIS* in fibrotic versus control HIF.

(A) On left, UCSC genome browser capture showing the MiGS read depth for three CD fibrotic (CD) and three control colon (NL) samples that shows the *WNT2B* promoter DMR and adjacent flanking regions. The range for the methylation counts (y axis) for all samples in the UCSC genome browser was set at 0 to 150. The solid black box underneath the browser screen capture shows the location of the yes/no DMR region called from the MiGS data. On right, targeted bisulfite sequencing validation of a region overlapping the DMR as indicated by the rectangle over the genome browser tracks. Dark circles indicate methylated and open circles unmethylated cytosines. Each row consists of a single sequenced clone. (B) MiGS data and targeted bisulfite sequencing illustrated as in (A) for *PTGIS*. The y-axis scale for all samples in the browser capture was set at 0 to 75 reads.

Figure. 4. Interaction network for genes differentially expressed and differentially DNA methylated in fibrotic HIF.

The blue circles represent the genes entered into the Cytoscape plug in for GeneMANIA. While the network was built for all differentially expressed genes, only the interactions from the subset that show both differential expression and DMRs are shown here. The gray circles are additional genes closely associated with the input genes. The size of the circle (node) is the number of neighbors each gene connects to. The edges are indicated by associations found through previously published co-expression, co-localization, genetic and physical interactions. The genes/nodes with edges connected to *PTGDS* are colored green, the edges that connect to *PTGIS* are colored dark blue and those to *WNT2B* are colored red.

ADDITIONAL FILES

Additional file 1 (Excel file)

Supplementary Table 1

Differentially methylated regions (DMRs) between normal and fibrotic HIF.

Each row represents one DMR. For each DMR, a unique ID, coordinates in the form of chr, start, end (hg19), and the DMR width are provided. Each DMR is categorized by a pattern code: "01" represents hypermethylation in fibrotic samples, and "10" represents hypomethylation in fibrotic samples. The "ratio" column represents the magnitude of the difference in sequencing signal as a log2 fold change, "p.value" is the unadjusted p-value, and "q.value" is the Benjamini-Hochberg adjusted p-value. "Genomic_context" specifies if the DMR falls within a portion of an annotated gene model. Multiple overlaps are resolved using the priority: promoter > exon > intron > 3' end > intergenic. Overlap with CpG shores, islands, and shelves is provided, along with the distance to the nearest gene and the names of the nearest genes.

Additional file 2 (Excel file)

Supplementary Table 2

RNA-seq analysis of differentially expressed genes in control and fibrotic HIF.

Total mRNAs from 3 different normal (2,3,4) and fibrotic fibroblasts (6,7,8) were subjected to RNA-seq analysis. Normalized FPKM values were averaged and the log base 2 fold-change for fibrotic versus normal HIF calculated. FDR adjusted p-values are shown. Genes were considered significant if the adjusted p-value was <0.05 and the log fold change was at least 1.5 fold. The differentially expressed genes are highlighted.

Additional file 3 (Excel file)

Supplementary Table 3: Genes differentially DNA methylated and expressed in fibrotic HIF

The genes that were found to be both differentially DNA methylated in their promoter regions and inversely correlated with gene expression are shown. Hyper/Hypo column denotes whether genes are yes/no hyper- (01) or hypo- (10) methylated in fibrotic HIF. Italicised *01* and *10* represent genes that were quantitatively differentially DNA methylated. The location of the DMR, the fold change in mRNA expression detected by RNA-Seq and the adjusted p value for this change are shown. * PTGDS fold change and p-values are from RT-PCR validation.

Additional file 4 (TIFF file)

Supplementary Figure 1

Bisulfite sequencing validation of differentially methylated *PTGDS* in fibrotic versus control HIF.

UCSC genome browser capture (left panel) of regions of differential DNA methylation for *PTGDS* from MiGS are shown next to the corresponding region of the gene validated by bisulfite sequencing (right panel). Three CD fibrotic (top panel) and three control (NL, lower panel) samples are shown. Dark circles indicate methylated and open circles unmethylated cytosines. Each row consists of a single sequenced clone. The range for the methylation counts (y axis) in the UCSC genome browser was set at 1 to 128.

Additional file 5 (Excel file)

Supplementary Table 4

List of input genes for GeneMANIA

The list of genes used to construct an interaction network in GeneMANIA in order to obtain functional enrichment is provided. A subset of the identified interactions are shown in Fig. 6.

Additional file 6 (Excel file)

Supplementary Table 5

Network gene associations for *WNT2B*

List of genes in RNA-seq data set (bold italics) and in GeneMANIA database (italics) that interact with *WNT2B* are shown.

Additional file 7 (Excel file)

Supplementary Table 6

Network gene associations for *PTGIS*

List of genes in RNA-seq data set (bold italics) and in GeneMANIA database (italics) that interact with *PTGIS* are shown.

Additional file 8 (Excel file)

Supplementary Table 7

Network gene associations for *PTGDS*

List of genes in RNA-seq data set (bold italics) and in GeneMANIA database (italics) that interact with *PTGDS* are shown.

Additional file 9 (Excel file)

Supplementary Table 8

Patient characteristics

The characteristics of the patients with Crohn's disease and controls from whom colon specimens were obtained are shown.

Additional file 10 (Excel file)

Supplementary table 9

Primers for Gene Expression and Bisulfite sequencing

The bisulfite sequencing primers for specific chromosomal locations and the qRT-PCR primer sequences and accession numbers, for *WNT2B*, *PTGIS* and *PTGDS* are shown.

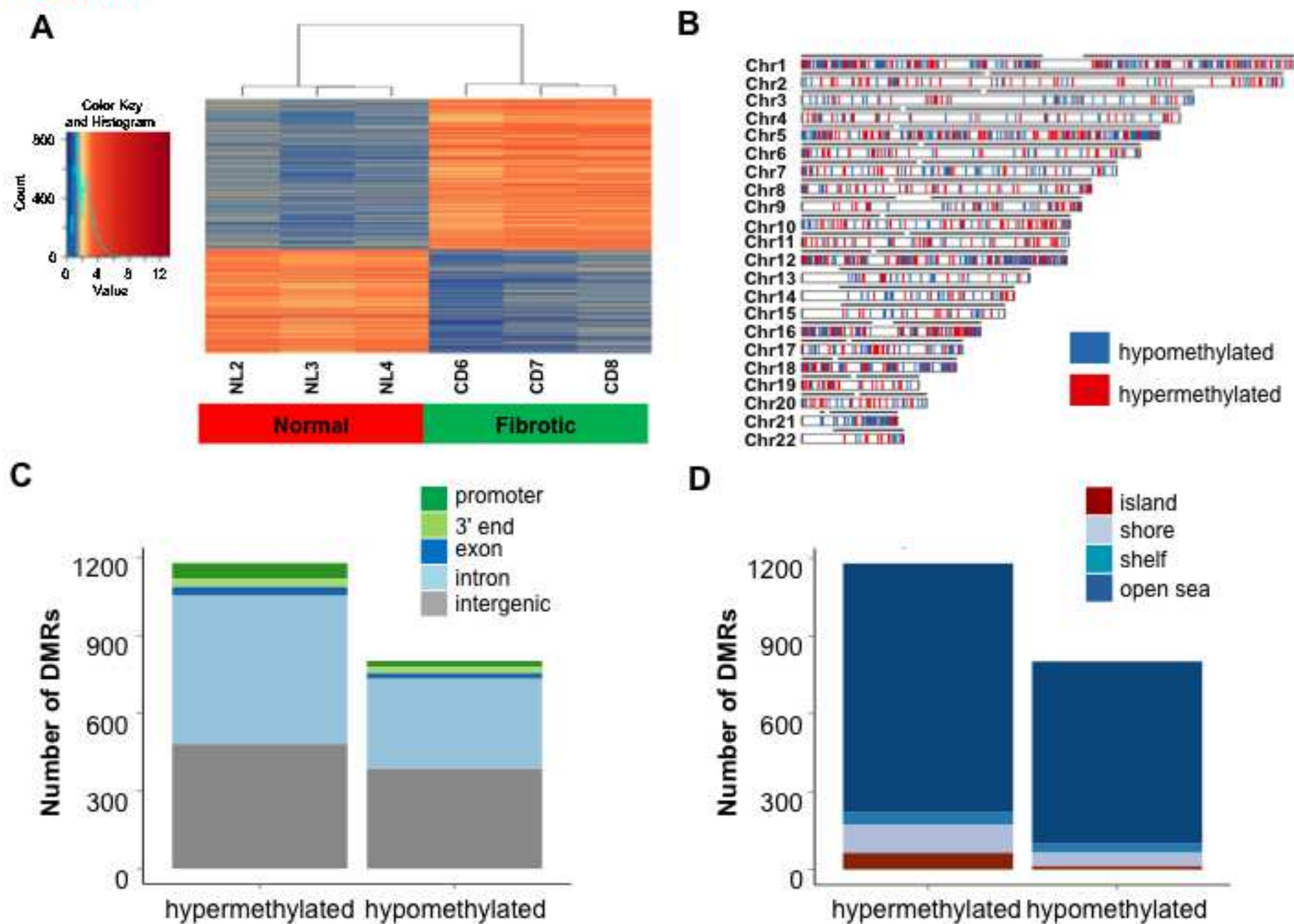
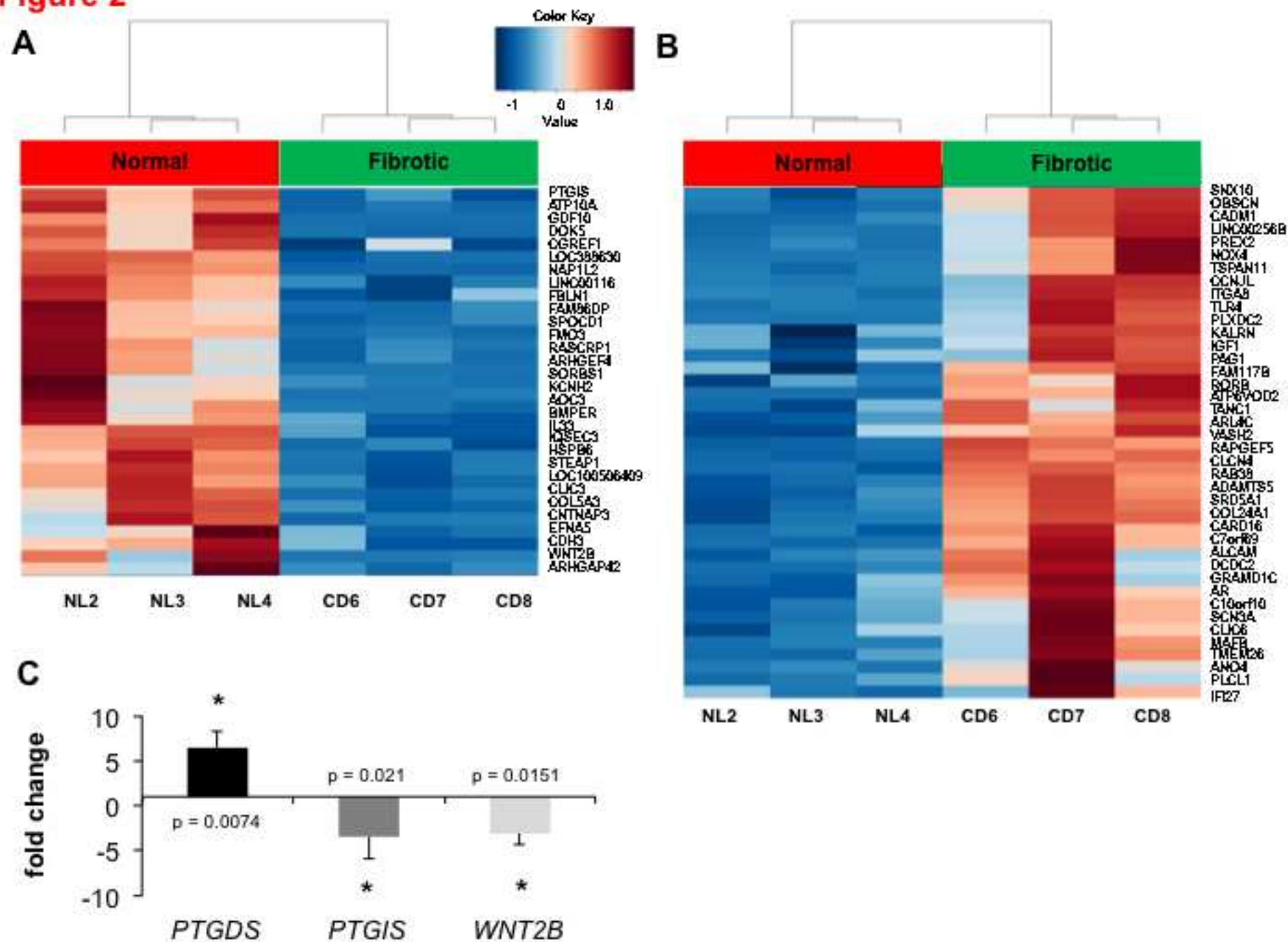
Figure 1

Figure 2

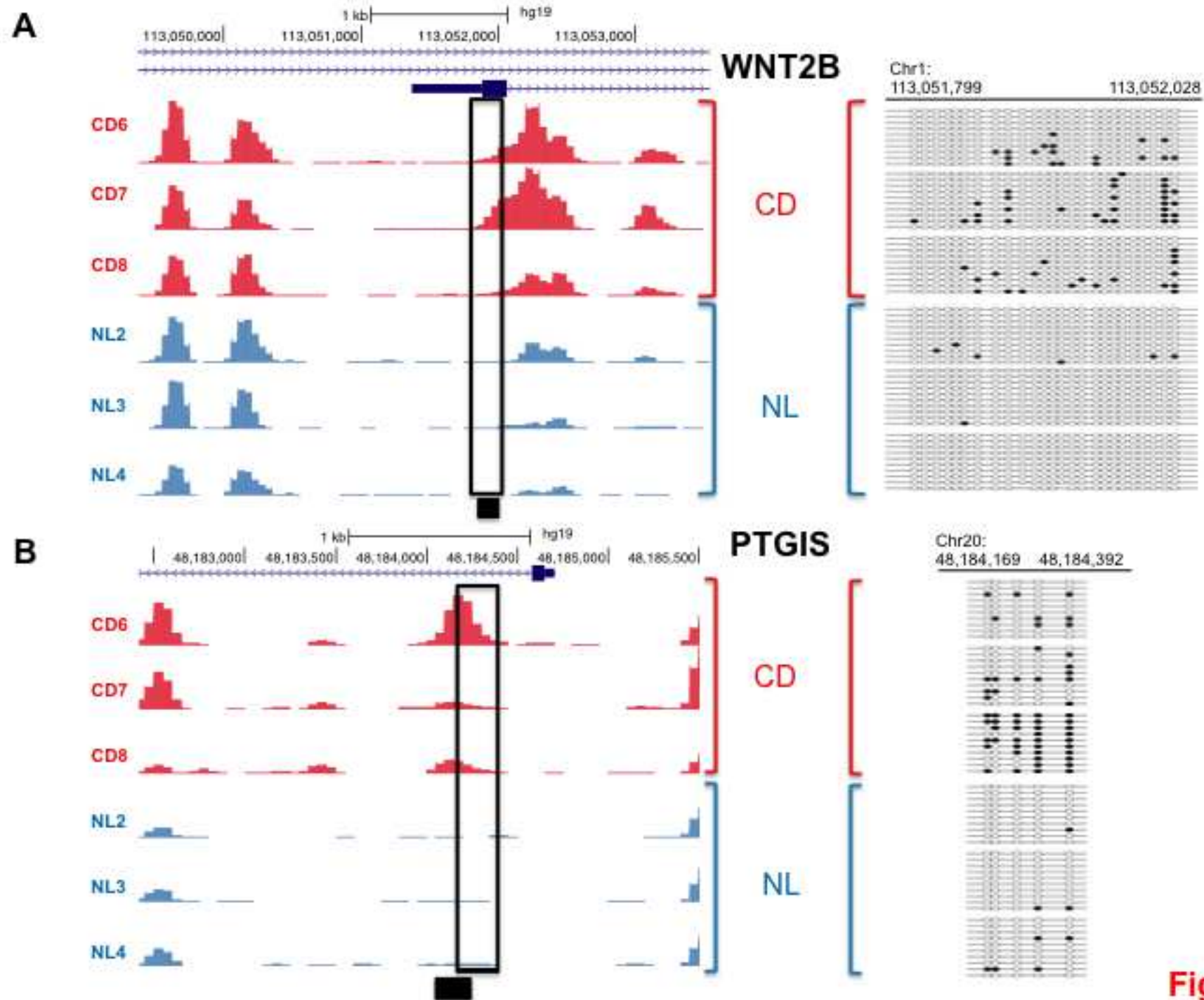
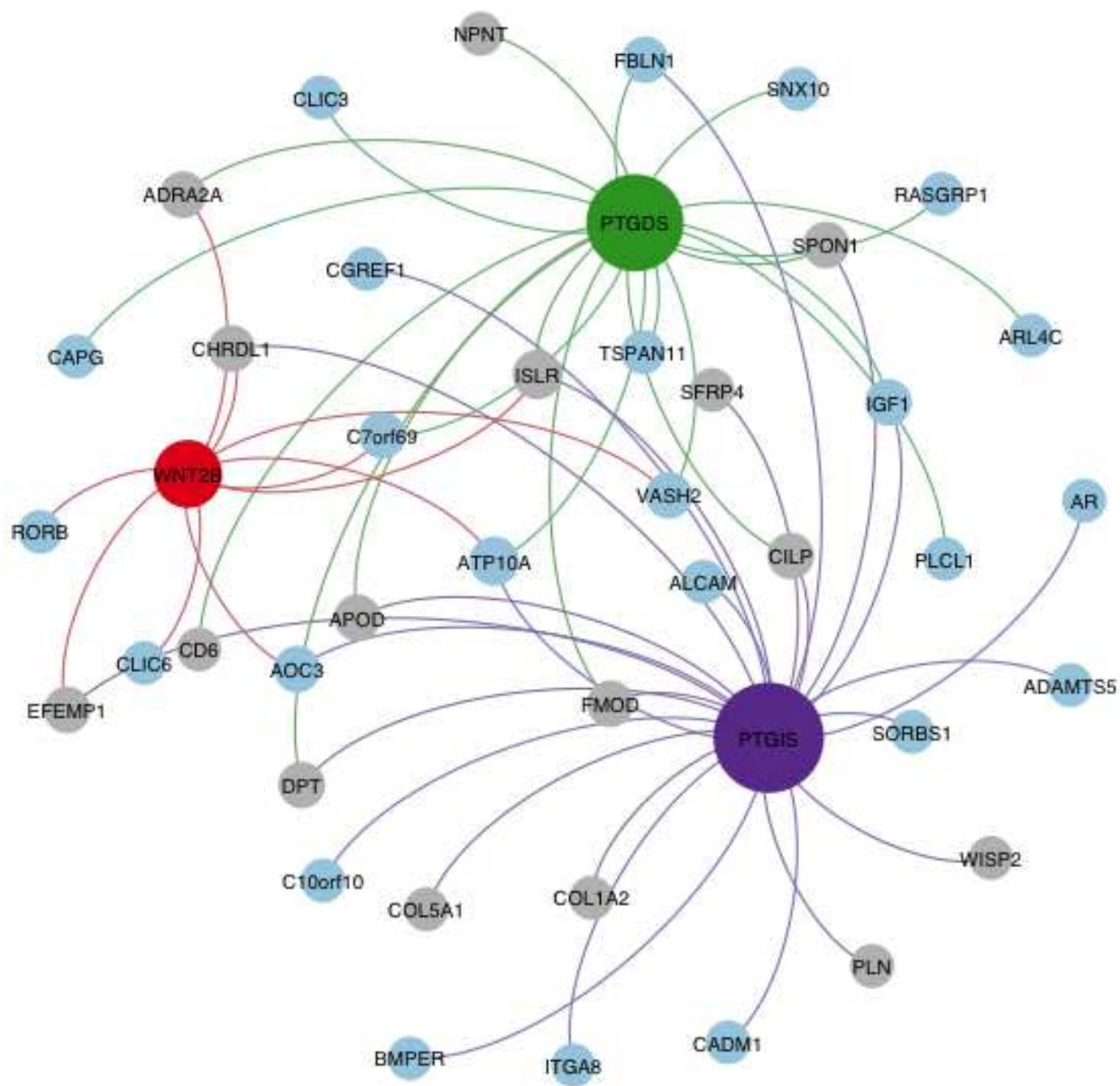
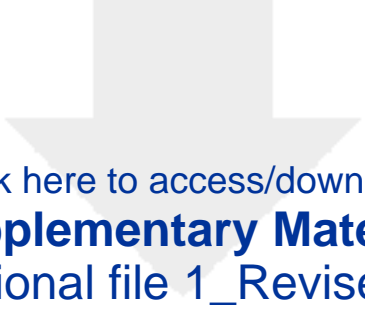
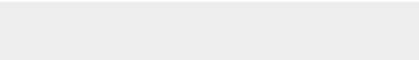
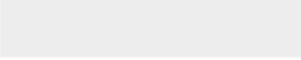
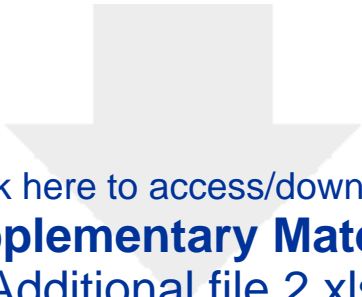
**Figure 3**

Figure 4

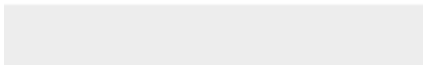
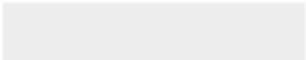


[Click here to access/download](#)
Supplementary Material
Additional file 1_Revised.xls





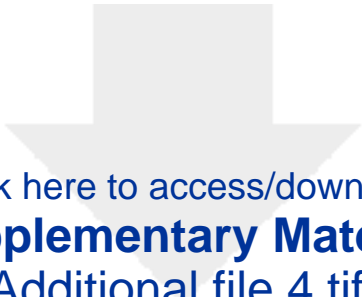
Click here to access/download
Supplementary Material
Additional file 2.xls



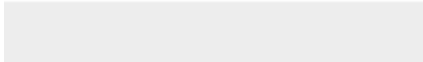


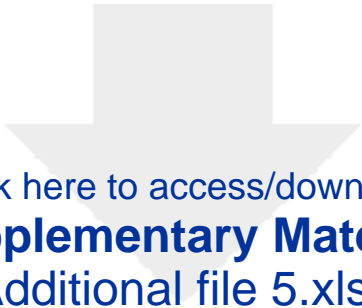
[Click here to access/download](#)
Supplementary Material
New Additional file 3.xlsx






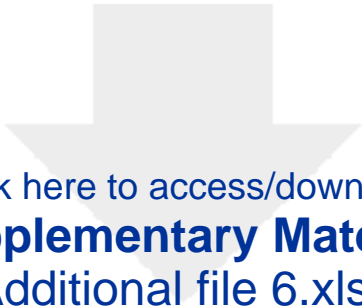
Click here to access/download
Supplementary Material
Additional file 4.tiff



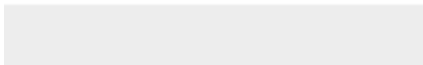



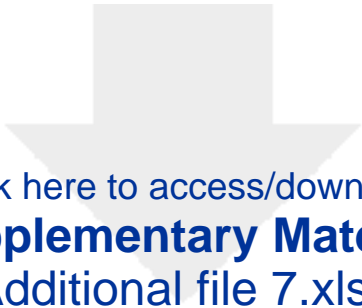
[Click here to access/download](#)
Supplementary Material
Additional file 5.xlsx



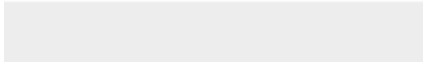



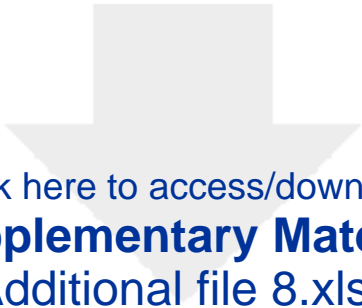
Click here to access/download
Supplementary Material
Additional file 6.xlsx



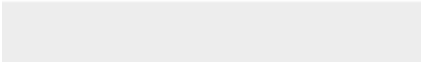



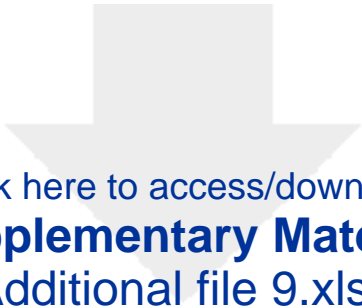
Click here to access/download
Supplementary Material
Additional file 7.xlsx



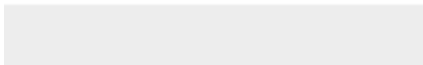
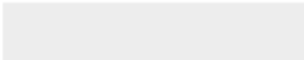


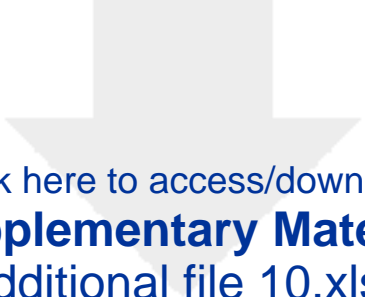
Click here to access/download
Supplementary Material
Additional file 8.xlsx





Click here to access/download
Supplementary Material
Additional file 9.xlsx





[Click here to access/download](#)
Supplementary Material
Additional file 10.xlsx

