OPEN

# Assessing DNA methylation in the developing human intestinal epithelium: potential link to inflammatory bowel disease

J Kraiczy[1], K Nayak[1], A Ross[1], T Raine[2], TN Mak[1], M Gasparetto[1], E Cario[3], V Rakyan[4], R Heuschkel[5] and M Zilbauer[1,5,6]

DNA methylation is one of the major epigenetic mechanisms implicated in regulating cellular development and cell-type-specific gene expression. Here we performed simultaneous genome-wide DNA methylation and gene expression analysis on purified intestinal epithelial cells derived from human fetal gut, healthy pediatric biopsies, and children newly diagnosed with inflammatory bowel disease (IBD). Results were validated using pyrosequencing, real-time PCR, and immunostaining. The functional impact of DNA methylation changes on gene expression was assessed by employing *in-vitro* assays in intestinal cell lines. DNA methylation analyses allowed identification of 214 genes for which expression is regulated via DNA methylation, i.e. regulatory differentially methylated regions (rDMRs). Pathway and functional analysis of rDMRs suggested a critical role for DNA methylation in regulating gene expression and functional development of the human intestinal epithelium. Moreover, analysis performed on intestinal epithelium of children newly diagnosed with IBD revealed alterations in DNA methylation within genomic loci, which were found to overlap significantly with those undergoing methylation changes during intestinal development. Our study provides novel insights into the physiological role of DNA methylation in regulating functional maturation of the human intestinal epithelium. Moreover, we provide data linking developmentally acquired alterations in the DNA methylation profile to changes seen in pediatric IBD.

## INTRODUCTION

The intestinal mucosa represents the largest area of the human body in direct contact with the exterior environment. In addition to its involvement in digestion and nutrient absorption, the intestinal epithelium has a key role in regulating barrier function and immune homeostasis in the gastrointestinal (GI) tract.[1] In mammals, development of a fully differentiated and functioning intestinal epithelium is a complex process, beginning *in utero* with formation of a stratified epithelial cell layer, derived from the visceral endoderm.[2] Although shortly before birth the final crypt-villus architecture and all major cell subsets (e.g., absorptive enterocytes, Paneth cells, goblet cells, and enteroendocrine cells) are present, the epithelium remains functionally immature.[3] At birth, a dramatic change occurs as the epithelium is colonized by the microbiota. Indeed, bacterial colonization, combined with exposure to an increasing variety of food antigens, is required for the normal postnatal functional development of the intestinal epithelium.[4] This early life interaction between the host epithelial cells and the environment has been shown to be particularly important in establishing mucosal barrier and immune functions such as the ability of the epithelium to sense microbial stimuli and mount an appropriate response.[5] Importantly, incomplete development or acquired impairment of intestinal epithelial cell/barrier function has been implicated in pathogenesis of

---

[1]University Department of Paediatrics, University of Cambridge, Cambridge, UK. [2]Department of Medicine, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK. [3]Division of Gastroenterology and Hepatology, Department of Experimental Gastroenterology, University Hospital of Essen, Essen, Germany. [4]Centre for Genomics and Child Health, Barts and The London School of Medicine and Dentistry, London, UK. [5]Department of Paediatric Gastroenterology, Hepatology and Nutrition, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK and [6]Witten/Herdecke University, Witten, Germany. Correspondence: M Zilbauer (mz304@medschl.cam.ac.uk)

several intestinal diseases, including necrotizing enterocolitis and inflammatory bowel disease (IBD).[5–7] Although the phenotypic and functional changes in the intestinal epithelium during human embryonic and early postnatal development have been well described, the underlying regulatory molecular mechanisms remain incompletely understood.

Epigenetic mechanisms are known to regulate gene expression and cellular function in the absence of changes to the underlying DNA sequence. DNA methylation is among the most extensively studied epigenetic modifications and occurs at the 5′ position of the pyrimidine ring of cytosines, primarily in the context of a CpG dinucleotide sequence.[8] CpG methylation is thought to regulate gene expression through its effect on chromatin state, as well as accessibility of transcription factor-binding sites.[9] In principle, increased methylation of CpGs (hypermethylation), in particular when located within promoter regions or in close proximity of the transcription start site, is associated with silencing of the respective gene, whereas hypomethylation has the opposite effect.[10]

DNA methylation has been shown to have a critical role in regulating lineage commitment of embryonic stem cells, cellular differentiation, as well as cell-type-specific gene expression of fully differentiated cell subsets.[11,12,13] However, to date, only limited information is available on the potential role of DNA methylation in regulating gene expression and cellular function in the human intestinal epithelium during embryonic and early life development. In addition, it has been proposed that epigenetic modifications may provide the mechanistic link between development, environmental change and altered gene expression.[14] Hence, they potentially represent a key biological mechanism mediating the rapid increase in complex conditions, including IBD, over the last century, in the absence of major changes to the human genome.[15,16]

Here we studied DNA methylation changes in the human intestinal epithelium during embryonic and postnatal development, with an aim to elucidate their functional role in regulating gene expression as well as their potential implication for IBD pathogenesis.

## RESULTS
### DNA methylation profiles of purified fetal and pediatric intestinal epithelium
First we performed genome-wide DNA methylation analysis on a discovery sample set ($n = 12$) of purified intestinal epithelial cells obtained from human fetal gut (i.e., proximal and distal) and pediatric biopsies derived from terminal ileum and two colonic segments (ascending colon and sigmoid colon). We performed multidimensional scaling analysis (MDS), which visualizes the level of similarity of individual samples, taking all of the ~470,000 data points present on the array into consideration. As shown in **Figure 1a**, the first dimension (i.e., $x$ axis), separates samples according to their developmental age (i.e., fetal vs. pediatric), whereas the second dimension (i.e., $y$ axis) divides samples by gut segment. These findings were further confirmed by performing unsupervised hierarchical clustering analyses, demonstrating major methylation

differences between fetal and pediatric intestinal epithelium (**Figure 1b**).

Next we categorized each CpG position into either methylated, partially methylated, or unmethylated, according to their respective methylation value. Differences in the distribution of probes were most obvious by developmental age and, to a lesser extent, by gut location, but were particularly pronounced in "unmethylated" and "partially methylated" categories (**Figure 1c**). Moreover, as demonstrated in **Figure 1c**, unmethylated probes were primarily located within CpG islands and enriched in promoter regions (**Figure 1c** and **Supplementary Figure S1** online).

### Identification of rDMRs
In order to investigate the role of DNA methylation in regulating gene expression in the intestinal epithelium, we performed simultaneous genome-wide expression profiling on purified cell samples. Similar to the observed differences in DNA methylation profiles, MDS analysis of gene expression data separated epithelial samples according to developmental age and gut location (**Figure 2a**).

Given this similar clustering pattern, we next aimed to identify genes for which expression might be primarily regulated through DNA methylation during the process of development from the immature fetal to the fully differentiated pediatric epithelium. We therefore performed comparative DNA methylation and gene expression analysis between the combined fetal (i.e., proximal and distal gut) and pediatric (i.e., terminal ileum, ascending colon, and sigmoid colon) sample set. As summarized in **Figure 2b**, we filtered out highly significant differentially methylated positions (DMPs), which were then grouped into differentially methylated regions (DMRs) by combining probes located in close proximity (see Methods). DMRs found to be associated with one unique gene, which also displayed significant differences in gene expression, were classed as potential regulatory DMRs (rDMRs). A very strict significant cut-off as well as log-fold change threshold was chosen in order to identify genes, for which DNA methylation changes are most likely to have a role in regulating gene expression. In total, we identified 259 rDMRs, which were associated with 214 genes (**Supplementary Table S1**). Methylation and gene expression profiles of selected rDMRs are displayed in **Figure 2c**, illustrating the robust absolute DNA methylation differences ranging up to 60% and associated changes in gene expression. Moreover, compared with all identified DMRs (i.e., including those in genes not associated with changed expression), rDMRs were more frequently found to overlap with annotated transcription start site, further highlighting the likely functional impact of DNA methylation changes on gene expression in identified genomic regions (**Figure 2d**).

### Identified genes containing rDMRs are enriched for pathways associated with immune function, GI development, and intestinal disease
In order to gain insight into which biological functions might be epigenetically regulated in the intestinal epithelium, we
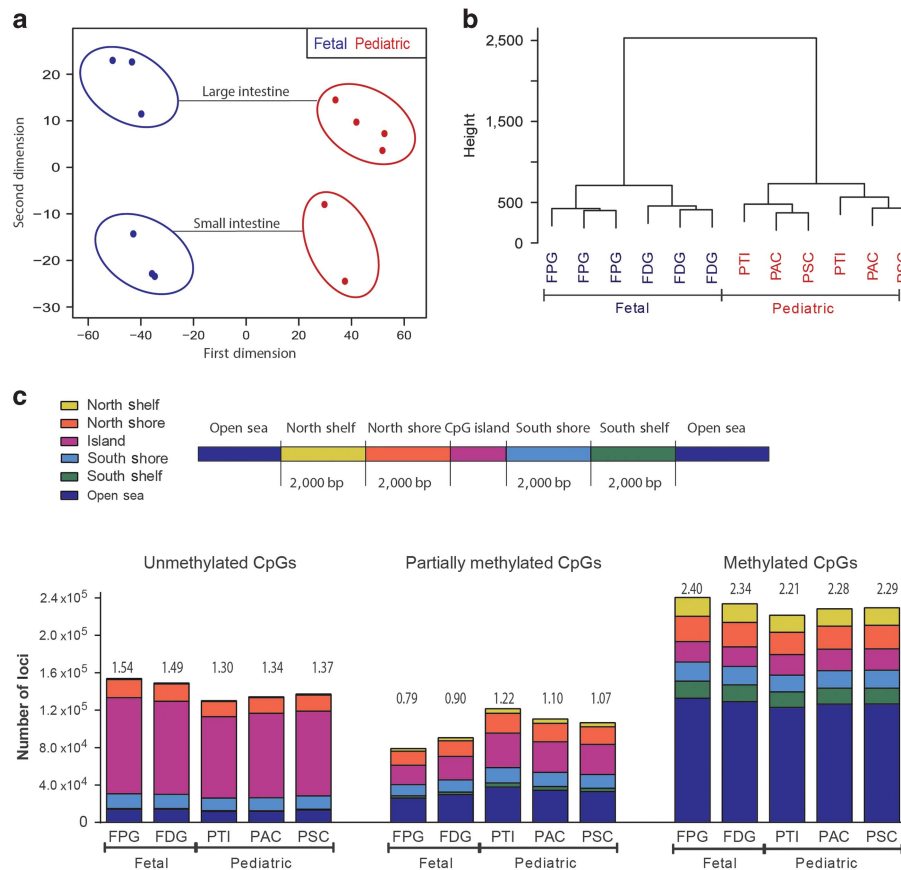
**Figure 1** Genome-wide DNA methylation profiles of purified human fetal ($n = 6$ from 3 donors) and pediatric intestinal epithelium ($n = 6$ from 2 donors). (**a**) Multidimensional scaling analysis (MDS) plot displaying the overall DNA methylation profiles considering all probes. The first dimension ($x$ axis) separates fetal from pediatric epithelial samples. The second dimension distinguishes epithelium according to gut location, separating proximal, small bowel from distal, large bowel. (**b**) Unsupervised hierarchical clustering confirms developmental age as the main factor separating samples into two groups, i.e., fetal and pediatric. (**c**) Distribution of loci according to their methylation status; "unmethylated," "partially methylated," and "fully methylated". The average number of loci in each group is plotted. Each bar is further sub-divided, indicating location of probes in relation to CpG islands. FDG, fetal distal gut ($n = 3$); FPG, fetal proximal gut ($n = 3$); PAC, pediatric ascending colon ($n = 2$); PSC, pediatric sigmoid colon ($n = 2$); PTI, pediatric terminal ileum ($n = 2$).

performed pathway analysis on all 214 genes containing rDMRs. We found that top gene networks and physiological functions include "Embryonic development," "Tissue morphology/development," and "Digestive system development and function" (**Figure 3a**). In addition, disease-related networks were significantly enriched for "Gastrointestinal disease" and "Immunological diseases." **Figure 3b** illustrates one of the major networks "Embryonic development, Tissue Morphology, Immunological Disease," which contained a number of key genes known to be involved in intestinal epithelial immune defence such as polymeric immunoglobulin receptor (*PIGR*) and trefoil factor 3 (*TFF3*), as well as molecules implicated in epithelial differentiation via the regulation of the Wnt and/or PI3K/Akt signaling pathway (e.g., *TCF7L1* and *WIF1*).[17]

**rDMRs validated in an independent sample cohort**
In order to validate findings derived from genome-wide array data, we used pyrosequencing and real-time PCR to examine DNA methylation and gene expression on a selection of

identified rDMRs in a second sample cohort ($n = 9$–12 per group). Given the association of genes containing rDMRs with immune function and related disease development, we selected genes based on their roles in the innate immune defence system (i.e., *MUC2*, *TLR3*, *PIGR*, and *IL6R*) or in epigenetic regulation (i.e., *TET1* (tet methylcytosine dioxygenase 1)). As shown in **Figure 4a**, results confirmed highly significant methylation differences ($>50\%$) between fetal and pediatric intestinal epithelium. Similarly, differences in gene expression were also confirmed as highly significant, reaching a log-fold scale in the majority of selected rDMR-associated genes (**Figure 4b**). An inverse relationship between gene expression and promoter DNA methylation was identified for all loci (**Figure 4c**). We also observed that with increasing age, hypomethylation of innate immune response genes occured, associated with upregulation in their expression. In contrast, the gene encoding TET1, an enzyme involved in DNA demethylation, gained methylation through development, which was associated with downregulation of gene expression. Previous reports have shown that unlike in upstream promoter regions, intragenic
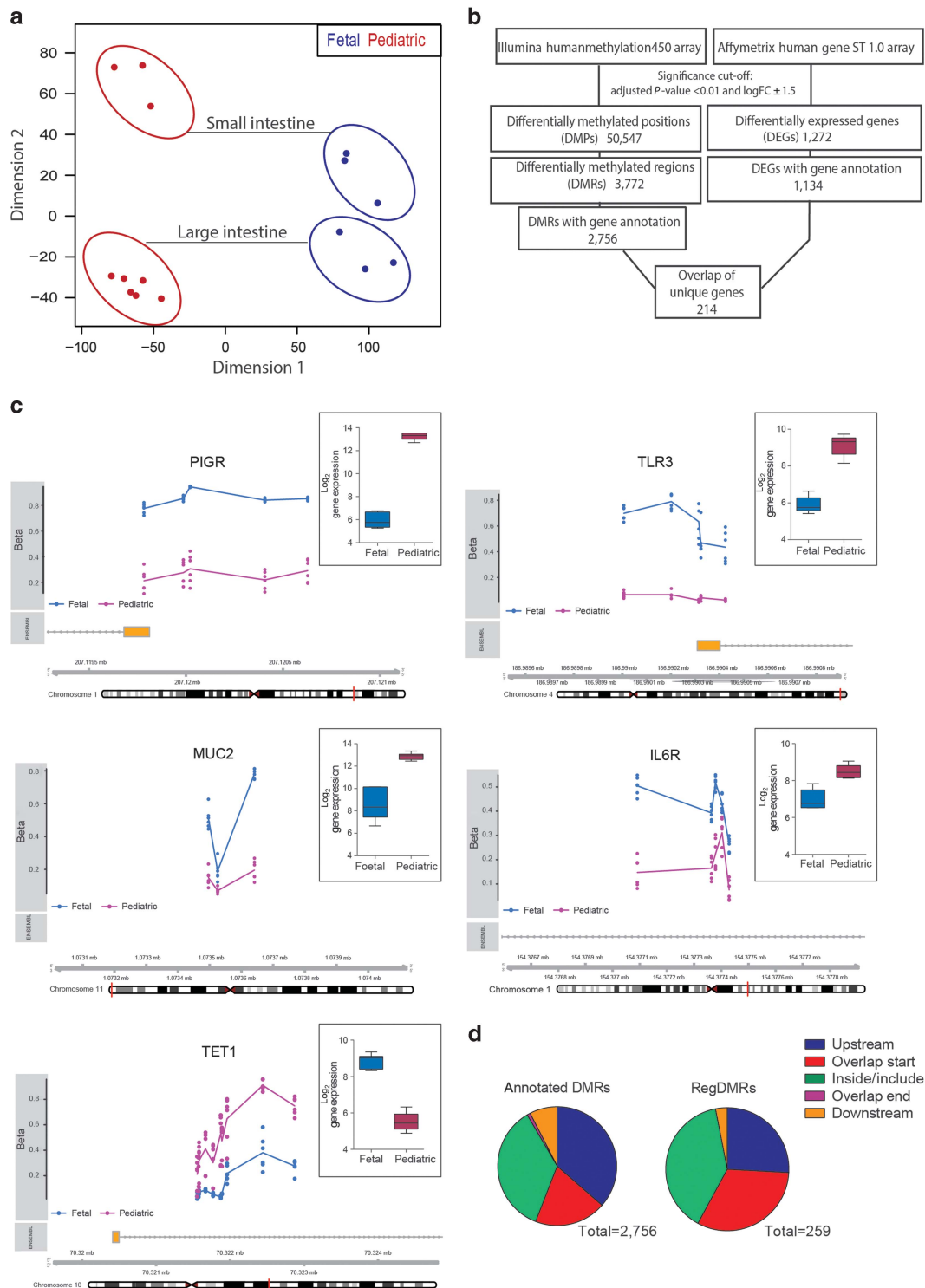
**Figure 2** 'Genome-wide gene expression analysis was performed on simultaneously extracted RNA from purified human intestinal epithelium. (**a**) Multidimensional scaling analysis (MDS) plot displaying clustering of samples according to developmental age in the first dimension, whereas the second dimension separates epithelial cells according to gut segment. (**b**) Overview of the algorithm used to identify regulatory differentially methylated regions (rDMRs) by integrating both genome-wide DNA and gene expression data. In total, we identified 259 rDMRs associated with 214 genes. (**c**) Summary plots of selected rDMRs displaying DNA methylation and gene expression of *PIGR* (polymeric immunoglobulin receptor), *TLR3*, *TET1* (tet methylcytosine dioxygenase 1), *MUC2*, and *IL6R*. The top panel indicates the level of DNA methylation across the DMR according to the methylation $\beta$-value ($y$ axis). Each point represents one sample and the line indicating the average methylation value. Additional panels indicate relation of DMR to the reference human genome. Inset displays corresponding gene expression plotted as $\log_2$ average expression of each sample. (**d**) Distribution of all DMRs (i.e., including DMRs for which associated genes do not display expression changes) vs. rDMRs according to their genomic location. rDMRs are found more frequently to be located either fully inside or overlapping with transcription start site (TSS). DMP, differentially methylated position; FDG, fetal distal gut ($n=3$); FPG, fetal proximal gut ($n=3$); PAC, pediatric ascending colon ($n=3$); PSC, pediatric sigmoid colon ($n=3$); PTI, pediatric terminal ileum ($n=3$).
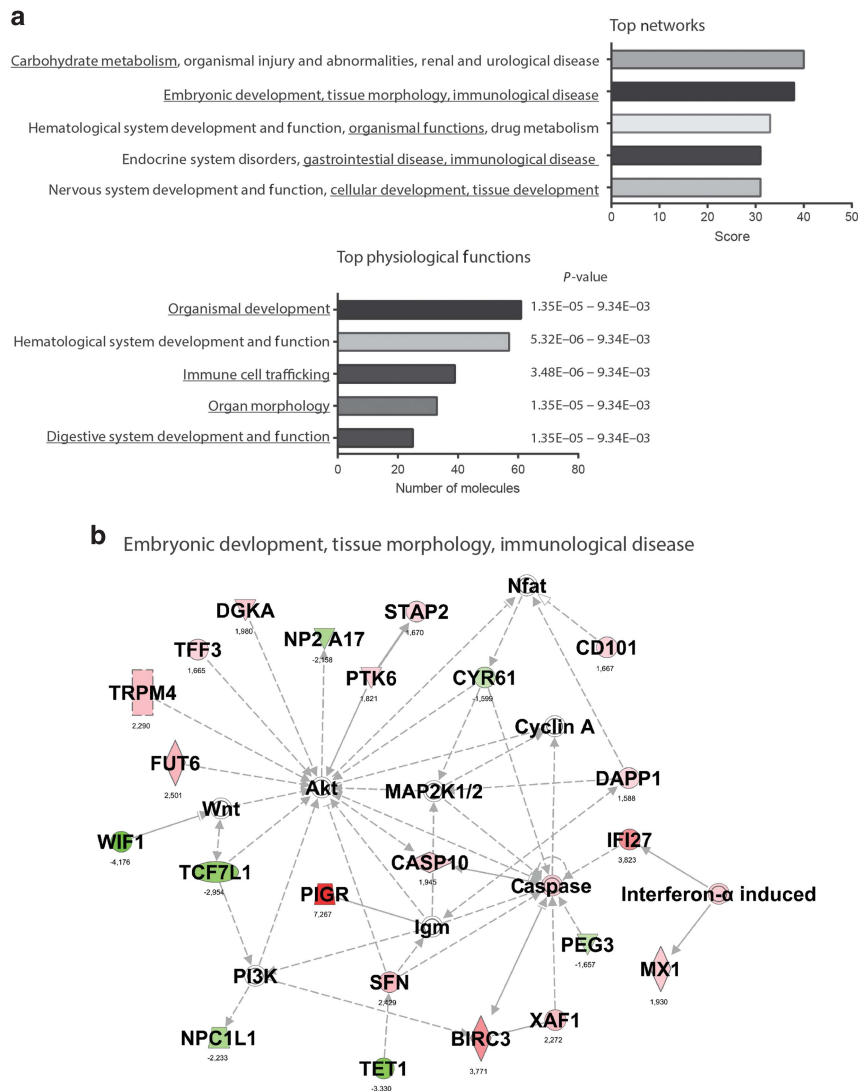
**Figure 3** Ingenuity pathway analysis (IPA) performed on identified regulatory differentially methylated region (rDMR)-containing genes. (**a**) Bar plot displaying scores of the top five networks and physiological functions, indicating the number of involved genes in each group and the enrichment *P*-value. (**b**) Major network "Embryonic development, tissue morphology, immunological disease". Graph shows gene symbols with color coding of differentially methylated and expressed genes. Red color indicates gene expression upregulated, green color represents gene expression downregulated in pediatric epithelium (compared with fetal). Numbers below symbols indicate the expression logFC (pediatric vs. fetal). Non-colored symbols indicate genes not in the input data but involved in the network shown. Molecule relationships were simplified using the "IPA Path Designer" tool.

hypermethylation is associated with an increase in gene expression.[18] Consistent with these reports, we found that in the gene body of *TLR3*, hypermethylation correlated with increased gene expression (**Supplementary Figure S2**).

Subsequent immunostaining for PIGR, MUC2, and TLR3 on fetal and pediatric samples suggests that real-time PCR data are reflective of protein levels. Specifically, in contrast to absent or low protein expression in the distal fetal gut epithelium, all proteins were highly expressed in the intestinal epithelium of pediatric colonic samples (**Figure 5**).

### Functional relevance of DNA methylation changes on gene transcription

Next, we aimed to demonstrate a direct impact of changes in DNA methylation on gene expression by employing two frequently used *in-vitro* models. First, we treated intestinal epithelial cell line (Caco-2) with a DNA methyltransferase inhibitor for 24 to 48 h. As demonstrated in **Figure 6a**, DNA methylation of candidate genes displayed a significant reduction, accompanied by a corresponding increase in gene expression. Next, we tested the functional relevance of DNA methylation on gene transcription in a locus-specific manner by cloning promoter regions of selected genes into a luciferase vector containing a CpG-free backbone.[19] As shown in **Figure 6b**, *in-vitro* methylation of the *PIGR* and *TLR3* promoter-containing plasmid led to a significant reduction of luciferase signal compared with the unmethylated construct. Basic luciferase activity of the unmethylated *MUC2* promoter plasmid was found to be very low and no further reduction was observed following *in-vitro* methylation (**Figure 6b**).
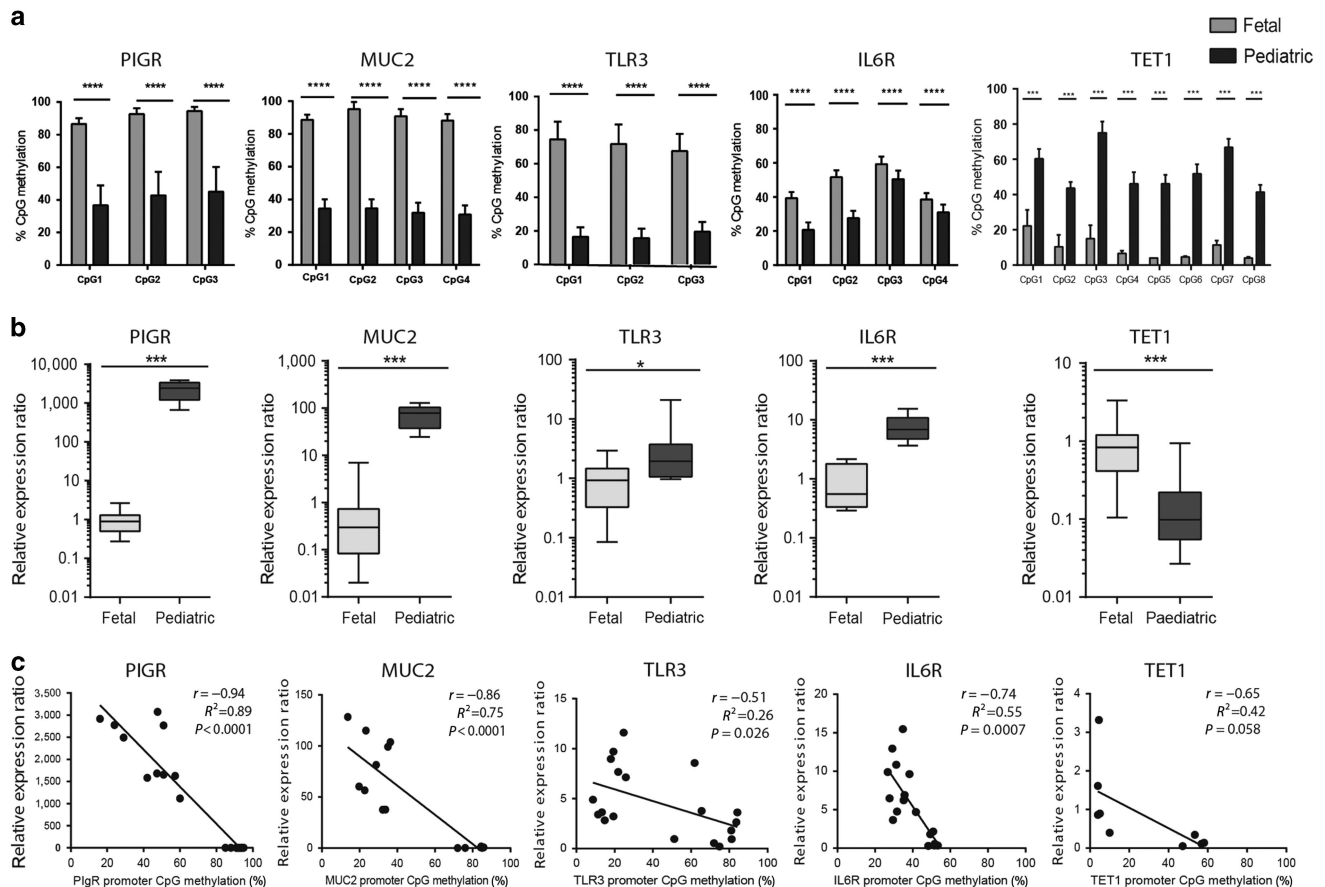
**Figure 4** Validation of DNA methylation and gene expression differences between fetal and pediatric epithelial cells in a second sample cohort using pyrosequencing and reverse transcription–PCR. (**a**) CpG methylation in promoter regions of *PIGR* (polymeric immunoglobulin receptor), *MUC2*, *TLR3*, *IL6R*, and *TET1* (tet methylcytosine dioxygenase 1). Data are presented as the average and s.d. of absolute DNA methylation values in %. Two-way analysis of variance, \*$P < 0.05$, \*\*\*$P < 0.005$, \*\*\*\*$P < 0.001$, $n = 6$–16 (minimum of four donors in each group). (**b**) Gene expression was assessed in parallel displaying significant differences between fetal and pediatric epithelial samples. Mann–Whitney test, \*$P < 0.05$, \*\*\* $P < 0.005$, $n = 7$–12. (**c**) Correlation between DNA methylation across the assessed region (average of three to eight CpGs) and respective mRNA expression was assessed by calculating the Pearson's correlation coefficient (*r*). Each dot represents one sample.

## Altered DNA methylation in the intestinal epithelium of children with IBD

The above data suggest that methylation changes at key immune regulators controls their expression during the development of healthy intestinal epithelium and alterations may be involved in GI and immunological diseases pathogenesis. Therefore, we next investigated whether aberrant epigenetic modifications are found in children with a confirmed diagnosis of IBD. We performed genome-wide DNA methylation profiling of purified colonic intestinal epithelium obtained from an additional cohort of children newly diagnosed with IBD (treatment naive), as well as matched healthy controls. As shown in **Figure 7a**, unsupervised hierarchical clustering analysis of DNA methylation profiles revealed clear separation into two groups, one containing all healthy control samples and five IBD samples, with the second cluster containing only samples obtained from children with IBD. Importantly, subclustering observed within IBD patients could not be solely attributed to the presence of intestinal inflammation, (**Figure 7a**). MDS analyses confirmed these findings, with

the first dimension separating control samples from IBD samples (**Figure 7b**), whereas the second dimension divides samples according to their gut location (i.e., sigmoid vs. ascending colon).

## IBD-specific intestinal epithelial methylation profiles significantly overlap with probes undergoing methylation changes during GI development

Next, we aimed to investigate whether the observed IBD-specific alterations in DNA methylation could primarily occur in loci that are dynamically methylated during intestinal epithelial development. We therefore performed differential methylation analysis comparing IBD samples with control samples. Using the same stringent significance cut-off as before (i.e., adjusted $P < .01$ and log-fold change $> \pm 1.5$), we identified a total of 233 DMPs (**Supplementary Table S2**). Interestingly, a statistically highly significant proportion of these IBD-specific DMPs overlapped with DMPs from the fetal vs. pediatric sample data set comparison (**Figures 7c**, $P < 8.5e - 32$). This was also the case when lowering the
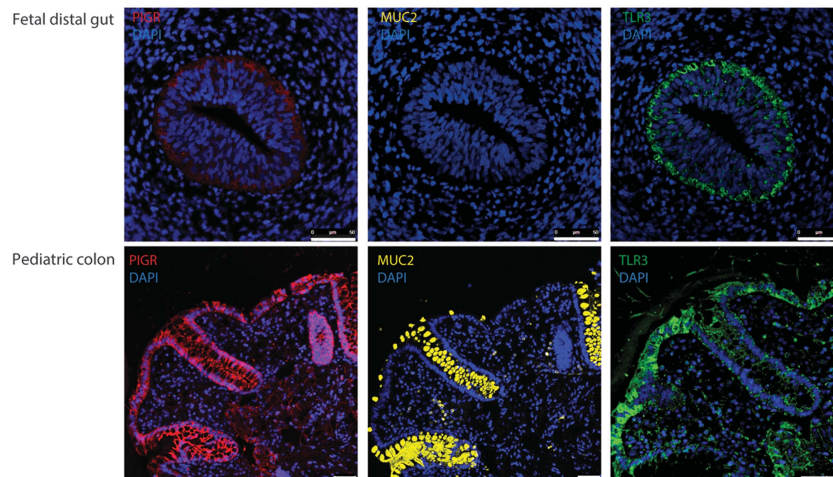
**Figure 5** Protein expression of polymeric immunoglobulin receptor (PIGR), MUC2, and TLR3 was assessed by immunofluorescent staining in fetal distal gut (upper panel) and pediatric ascending colon (lower panel). Expression of all three proteins was found to be either absent or low in fetal gut samples and highly abundant in pediatric colonic epithelium. Fetal gut sections display the presence of a stratified epithelial layer with a complete lack of crypt villus axis. Protein expression pattern corresponds with mRNA levels of the respective genes and is confined to the intestinal epithelium. Staining was performed on tissue samples obtained from three donors in each group, representative images are shown. White bar represents 50 μm. Blue, 4',6-diamidino-2-phenylindole-counterstained nuclei; red, PIgR; yellow, Muc2; green, Tlr3.

significance cut-off by including all DMPs with an adjusted $P < .01$ regardless of log-fold change (**Figure 7c**). Moreover, a number of the previously highlighted genes (e.g., *PIGR* and *MUC2*) containing identified rDMRs were found to display distinct methylation differences in IBD patients compared with controls, both in array data of the identified rDMR, as well as using locus-specific pyrosequencing (**Figure 7d,e**).

## DISCUSSION

Epigenetic mechanisms are emerging as key factors in regulating the fundamental biological processes driving cellular differentiation as well as cell-type-specific function.[8,20] Evidence is also accumulating that epigenetic mechanisms mediate the effects of a changing environment on mammalian embryonic and early-life development, something that has been linked to the increase in incidence of complex, multifactorial diseases including IBD.[14,21] In this study we investigated the role of DNA methylation in regulating intestinal epithelial function during GI development and assessed whether alterations may have an impact on IBD pathogenesis.

We first generated and analyzed genome-wide DNA methylation profiles derived from healthy fetal and pediatric epithelium. Results revealed distinct differences between fetal and pediatric samples, providing a first indication that dynamic changes in DNA methylation are involved in the maturation of the developing gut epithelium. We then went on to identify genomic regions for which an association was observed between DNA methylation and gene expression (rDMRs). Using a stringent significance cut-off, we identified a total of 214 genes containing one or more rDMRs. Given the relatively small sample size of our discovery cohort, we further validated a selected subset of rDMRs in a second sample cohort, confirming highly significant absolute changes in methylation

of up to 60%. Furthermore, these methylation changes are observed consistently across multiple CpG sites. Together with corresponding changes in gene expression, as well as the strong inverse correlation between methylation and gene expression, we suggest that expression of those genes is highly likely to be regulated by DNA methylation.

Despite a large body of existing evidence supporting a role for DNA methylation in regulating gene transcription, an association between changes within a given gene does not prove these changes are functionally related.[8] However, demonstrating a causative relationship between an epigenetic mark and a change in gene expression remains one of the major challenges in the field. We were able to demonstrate that treatment of an intestinal epithelial cell line with DNA methyltransferase inhibitors led to reduced levels of DNA methylation and an associated increase in gene expression of all tested genes. Moreover, *in-vitro* methylation of promoter regions in a CpG-free luciferase reporter plasmid led to a significant reduction in promoter activity of *TLR3* and *PIGR*. These results strongly suggest that DNA methylation has a functional role in regulating gene expression at our identified rDMRs.

Based on our findings so far, we hypothesized that these genomic regions might represent important epigenetic regulators in gut development and postnatal functional maturation of the intestinal epithelium. Specifically, we speculated that network analysis on identified rDMRs could provide further insight into what aspects of the intestinal epithelial functional development might be, at least in part, epigenetically regulated. Indeed, pathway analysis performed on all rDMRs revealed a significant enrichment for networks involved in embryonic, tissue, and GI development, suggesting an essential role for DNA methylation in regulating these fundamental biological processes. In fact, performing a detailed
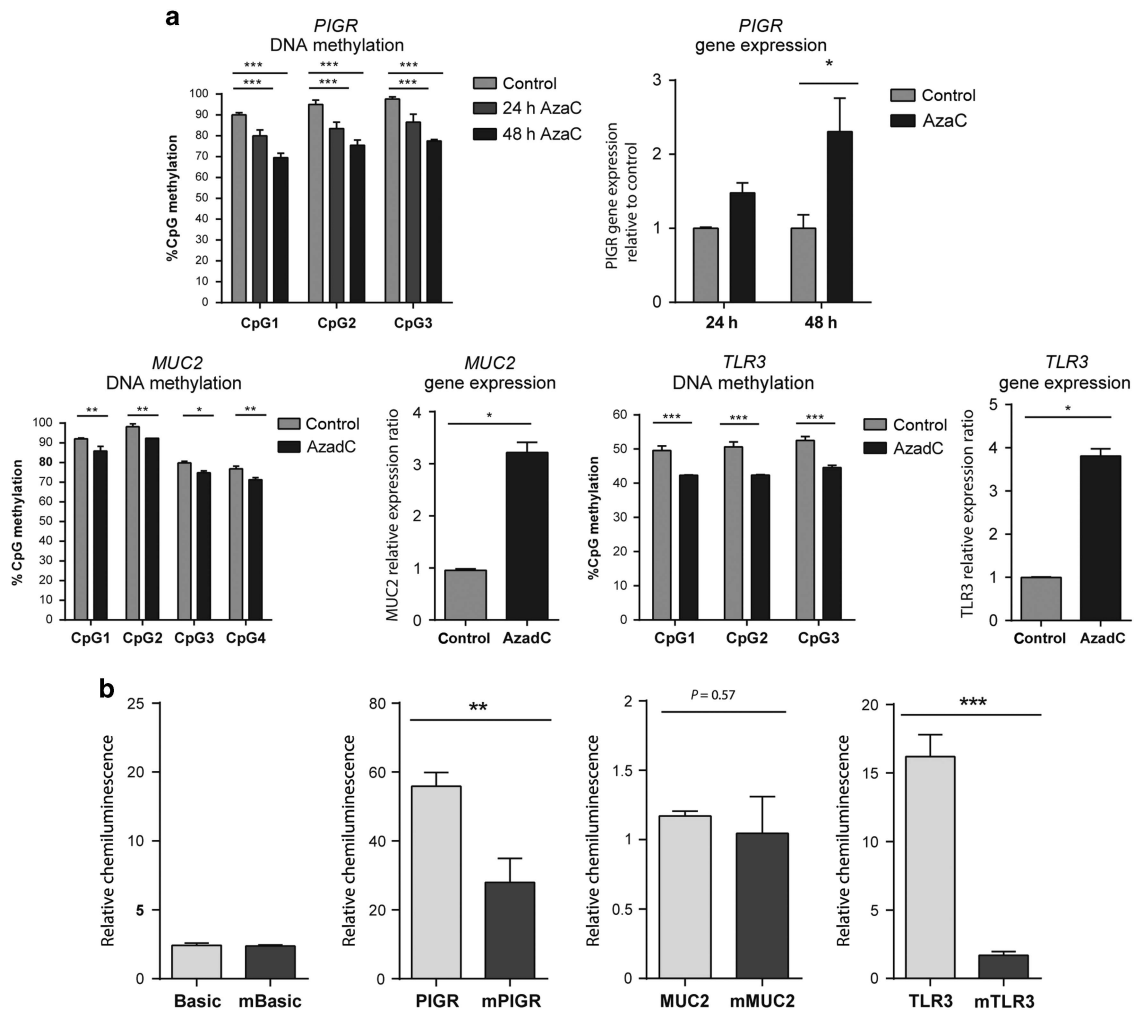
**Figure 6** *In-vitro assays* demonstrate the direct impact of DNA methylation on transcription of regulatory differentially methylated region (rDMR)-containing genes. (**a**) Caco-2 cells were treated with DNA methyltransferase (DNMT) inhibitor (10 μM 5-Azacytidine (AzaC) for 24 h and for 48 h or 5 mM 5-Azadeoxycytidine (AzadC) for 24 h. DNA methylation of the polymeric immunoglobulin receptor (*PIGR*), *MUC2*, and *TLR3* promoter was assessed by pyrosequencing. Reduced levels of CpG methylation with a corresponding increase in mRNA expression was observed. Data are displayed as mean + s.d. of three independent experiments performed in duplicates. Two-way analysis of variance, post test; Sidak, *$P < 0.05$, **$P < 0.005$, ***$P < 0.001$ vs. dimethylsulfoxide (DMSO) vehicle control. (**b**) Promoter regions of *MUC2*, *TLR3*, and *PIGR* were cloned into a CpG-free luciferase reporter plasmid. *In-vitro* methylation of *TLR3* and *PIGR* promoter plasmids led to a significant reduction in luciferase signal compared with mock-treated unmethylated plasmids. Basic luciferase activity of MUC2 promoter containing plasmid was found to be low in Caco-2 cells and no further reduction was observed following *in-vitro* methylation. Basic, CpG-free plasmid with no promoter. Renilla-luciferase served as control to correct for transfection efficiency. Mean + s.d. of three independent experiments performed in duplicates; Student's unpaired *t*-test, two tailed, $n = 3$ per condition. ***$P < 0.0001$, **$P < 0.005$, $P < 0.57$ (NS, not significant).

literature research revealed substantial evidence that supports the role of a large proportion of rDMR-containing genes in intestinal epithelial-specific functions (**Supplementary Table S3**). Examples include genes involved in innate immune/antimicrobial defence such as mucin (*MUC2*), pattern recognition receptors (*TLR3* and *TLR4*) fucosyltransferases (*FUT3* and *FUT6*), *PIGR*, and *TFF3*, as well as genes involved in antigen presentation such as HLA class II genes (e.g., *HLA-C*, *HLA-DMA*, *HLA-DMB*, *HLA-DPB1*, and *HLA-DRA*). Importantly, many of these genes have been implicated in the pathogenesis of IBD including *TLR3* and *TLR4*,[22] *MUC2*,[23] *PIGR*,[24] *TFF3*,[25] and *HLA-DRA*.[26] Moreover, the majority of genes contributing to intestinal epithelial defence were found to

be hypermethylated, with expression either absent or at low levels in fetal samples. During the transition to a fully developed pediatric epithelium, these genes became hypomethylated and hence actively expressed.

In addition to genes associated with immune function, rDMRs were also enriched within genes known to have key roles in intestinal development. DNA methylation has been shown to regulate cellular differentiation through the activation of lineage-specific genes at the appropriate time during cellular development.[20] We identified several rDMR-containing genes known to be involved in driving cellular development and differentiation. These include *TCF7L1* and *WIF1*, both recognized regulators of the Wnt signaling pathway as well as
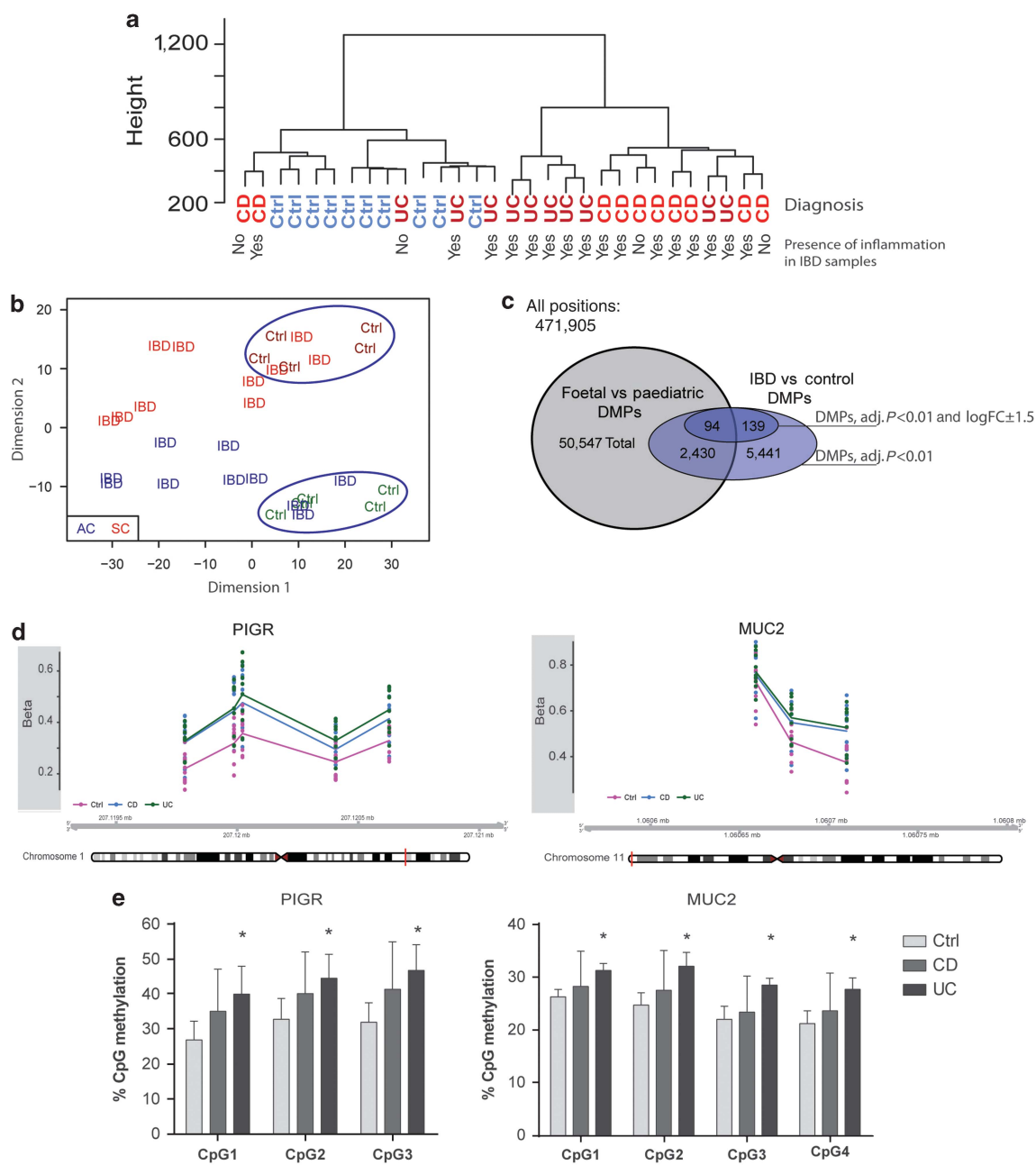
**Figure 7** Alteration of DNA methylation in colonic epithelial cells of children diagnosed with inflammatory bowel disease (IBD) occurs at sites undergoing dynamic methylation changes during gastrointestinal (GI) development. (**a**) Hierarchical clustering of genome-wide DNA methylation profiles of colonic intestinal epithelial cells obtained from children newly diagnosed with IBD (treatment naive) and matched healthy controls ($n = 5$ patients per group and 2 biopsy samples per patient). The presence of mucosal inflammation is indicated. (**b**) Multidimensional scaling analysis (MDS) analysis confirms methylation differences present in the majority of IBD samples compared with healthy controls. (**c**) Venn diagram showing the overlap of significant differentially methylated positions (DMPs) derived from the comparison of fetal vs. pediatric samples in the discovery cohort and newly diagnosed IBD epithelium vs. healthy controls. Up to 40% of CpGs displaying significant alterations in IBD epithelium were also found to undergo significant DNA methylation changes during the transition from fetal to pediatric epithelium. Overlaps between DMPs were found to be highly significant (hypergeometric enrichment test $P < 8.5 \times E - 32$). (**d**) Alterations of DNA methylation in the intestinal epithelium of children newly diagnosed with IBD were found to be present in regulatory differentially methylated regions (rDMRs) of *MUC2* and polymeric immunoglobulin receptor (*PIGR*). Methylation $\beta$-values derived from array data (IBD sample cohort) are plotted in genomic context. (**e**) DNA methylation of respective DMRs assessed using pyrosequencing. Mean + s.d., $n = 5$ individuals per group (i.e., 5 healthy controls, 5 CD, and 5 UC). Analyses were performed in duplicates on two biopsy samples per patient, multiple *t*-test, post test: Holm–Sidak, *$P < 0.05$ vs. control. CD, Crohn's Disease; Ctrl, Control; UC, ulcerative colititis.

olfactomedin4 (*OLFM4*).[27,28] Although the experimental design of our study does not allow differentiating between epigenetic modification as a cause or consequence to changes in gene expression, our findings support and further extend existing evidence that DNA methylation plays a critical role in the development of the human intestinal epithelium.

Increasing evidence suggests that epigenetic mechanisms including DNA methylation might have a crucial role in IBD pathogenesis.[15,16] Although limited studies have been performed in human mucosal tissue, results have been supportive.[29–31] However, given the fact that DNA methylation profiles are highly cell-type specific, analysis on mixed cell tissue samples (e.g., mucosal biopsies), in particular in the context of an inflammatory condition, carry a high risk of yielding results confounded by cellular composition.[32] Here we generated genome-wide DNA methylation profiles of highly purified intestinal epithelium obtained from a treatment-naive pediatric patient cohort, as well as matched healthy controls. Unsupervised clustering and MDS analysis of these methylation profiles clearly separated out a subset of IBD samples with all controls grouping closely together. Importantly, this clustering pattern appeared not to be entirely due to the presence or absence of inflammation, implying the possibility of an IBD-specific intestinal epithelial DNA methylation signature in these patients.

A developmental origin of health and disease has been proposed by many as a plausible framework in which epigenetic mechanisms mediate environmental effects during vulnerable stages of development, ultimately causing disease when exposed to appropriate later triggers (second/multiple hit theory).[33] We hypothesized that genomic areas displaying differential methylation between fetal and pediatric samples may represent potential susceptibility regions, in which alteration of DNA methylation status might predispose to disease. We therefore tested for an overlap between highly significant DMPs identified between fetal and pediatric samples, and DMPs derived from comparison of children with IBD and healthy controls. We observed a highly significant overlap, with up to 40% of CpGs differentially methylated in children with IBD, among those undergoing significant changes in DNA methylation between the first trimester of pregnancy and childhood. These findings suggest that epigenetic alterations during critical periods of human fetal and postnatal development, potentially mediated via exposure to specific environmental factors, could predispose and ultimately cause development of chronic intestinal inflammation.

In summary, our study provides novel insight into the role of DNA methylation in regulating human intestinal epithelial development and function. The striking overlap between genomic loci undergoing significant DNA methylation changes during GI development with those observed in children diagnosed with IBD supports a developmental origin of IBD hypothesis, mediated through epigenetic mechanisms.

## METHODS

**Human fetal gut and intestinal biopsy samples**. Colonoscopy was performed by experienced pediatric gastroenterologists, who collected three additional mucosal biopsies each from the terminal ileum, ascending colon, and sigmoid colon. Biopsies were taken within 2–3 cm of those taken for routine histological assessment, which was performed by experienced consultant GI histopathologists. Diagnosis of Crohn's disease and ulcerative colitis was based on standard criteria using clinical, radiological, endoscopic, and histopathological findings in accordance with established criteria. Children with macroscopically and histologically normal mucosa, with no subsequent evidence of any underlying GI pathology, served as a healthy control group. All studies were conducted with informed patient and/or carer consent as appropriate, and with full ethical approval (REC-12/EE/0482). Fetal intestine was obtained with full ethical approval (REC-96/085) following informed consent following elective terminations at 8–12 week gestation, carried out for reasons other than maternal physical health or fetal disease. Fetal gut was dissected and divided into the proximal and distal sections at the reliably identifiable ileocecal junction. Sample and patient details are provided as **Supplementary Material**.

**Purification of intestinal epithelial cells**. Purification of intestinal epithelium from pediatric mucosal biopsies and fetal gut samples was performed using enzyme digestion, followed by magnetic bead sorting as described previously.[32] This protocol yields highly purified epithelial cell fractions, which was routinely confirmed by assessing lymphocyte contamination using flow cytometry. Details and a representative analysis are provided as **Supplementary Material** (**Supplementary Figure S3**).

**Genome-wide DNA methylation and gene expression arrays**. DNA and RNA were extracted simultaneously from the same sample using the AllPrep DNA/RNA mini kit (Qiagen, Hilden, Germany) following the manufacturer's instructions.[11] Genome-wide DNA methylation was assessed using bisulfite-converted DNA on the Illumina Infinium HumanMethylation450 BeadChip platform (Illumina, Cambridge, UK). Gene expression was analyzed using Affymetrix Human Gene ST 1.0 Array (Affymetrix UK Ltd, High Wycombe, UK). Both methods have been described in more detail previously.[11] Data have been deposited in the ArrayExpress database under the following accession numbers: E-MTAB-3703 (Illumina K450 DNA methylation arrays discovery cohort), -E-MTAB-3704 (Affymetrix gene expression arrays discovery cohort), and E-MTAB-3709 (Illumina K450 DNA methylation arrays paediatric IBD and control cohort).

**RT and real-time PCR**. One hundred to 500 ng of RNA were reverse transcribed using QuantiTect Reverse Transcription Kit (Qiagen). Complementary DNA corresponding to 5 ng RNA was used for real-time PCR using QuantiFast SYBR Green PCR Master Mix (Qiagen) on the 7500 Fast real-time PCR system, 7500 software v2.0.6 (Applied Biosciences by Thermo Fisher Scientific, Paisley, UK). Relative quantitation was calculated using the $\Delta\Delta C_t$ method following normalization against GAPDH.[34]

**Bisulfite-conversion and pyrosequencing**. Extracted DNA was bisulfite treated using EZ DNA methylation Gold kit (Zymo Research, Irvine, CA) according to the manufacturer's instructions. Locus-specific PCR amplification and quantification of DNA methylation levels were performed on bisulfite-converted DNA using Pyromark Gold reagents on a Pyromark Q24 (Qiagen) pyrosequencing system as described previously.[32] Further details are provided as **Supplementary Material**.

**Immunohistochemistry**. Tissue specimens were fixed in 4% formaldehyde in phosphate-buffered saline, stepwise dehydrated, and embedded in paraffin. Paraffin sections (4 μM) were rehydrated by subsequent incubation with xylene and descending concentrations

of ethanol. Heat-induced antigen retrieval was performed for 20 min in a conventional steamer. Nonspecific binding of antibodies was prevented by blocking with 10% normal goat serum and 1% bovi-neserum albumin in phosphate-buffered saline. Primary antibodies were incubated overnight at 4 °C. After washing, secondary antibodies were incubated for 1 h at room temperature before washing and mounting. Details of antibodies are provided as **Supplementary Material**. Images were taken using a SP5 Leica confocal microscope (Leica Microsystems, Milton Keynes, UK).

**DNA methyltransferase inhibitor treatment**. Caco-2 cells were treated with $10 \mu M$ 5-Azacytidine or $5 \mu M$ 5-Azadeoxycytidine for 24–48 h, whereas dimethylsulfoxide served as vehicle control. Following incubation periods, cells were lysed before simultaneous DNA and RNA extraction.

**DNA methylation-specific Luciferase reporter gene assays**. Genomic promoter regions of interest were cloned from human DNA into a CpG-free luciferase reporter vector as described previously.[19] Following, plasmids were *in vitro* methylated by incubation with recombinant CpG methyltransferase *M.SssI* (New England Biolabs, Ipswich, MA) in the provided buffer and supplemented with 3 mM *S*-adenosylmethionine at 37 °C for 4 h. Control plasmids were processed in the same way without addition of *M.SssI*. Following enzyme heat inactivation at 65 °C for 20 min, plasmid DNA was isolated using a PCR purification kit (Qiagen) and quantified by spectrophotometer. Caco-2 were transfected at 60% confluence for 24 h using Lipo-fectamine with Plus Reagent (Life Technologies by Thermo Fisher Scientific) and 100 ng reporter vector per well (96-well plate). Differences in transfection efficiency and cell density were corrected for by co-transfection with 3 ng Renilla luciferase reporter vector pRL-CMV E226V (Promega, Madison, WI). Cells were given 6 h of recovery in normal medium before addition of passive lysis buffer. Reporter gene activity was measured with dual luciferase assay reagents (Promega) in a GloMax 96 Microplate Luminomer (Promega). Relative chemiluminescence is expressed as ratio of raw reading of Firefly luciferase/Renilla luciferase, multiplied by factor $10^3$, to improve legibility. pCpGL-Basic (no promoter) served as negative control and pCpGL-CMV served as positive control.

**Statistical analysis**. DNA methylation analysis were performed using R software for statistical analysis and the Bioconductor packages *minfi* and *limma* were used for normalization and the majority of downstream analysis including generation of MDS plots and hier-archical clustering analysis.[35,36] Affymetrix gene expression array analysis was performed according to established standard protocols as published previously.[11] Further details can be found in the **Supplementary Methods**.

Data were further analyzed through the use of Qiagen's Ingenuity Pathway Analysis (www.qiagen.com/ingenuity).

Statistical tests for data other than genome-wide arrays were performed in GraphPadPrism 6 software (GraphPad Software, La Jolla, CA). Two groups were compared by Student unpaired, two-tailed *t*-test, in case of unequal variances with Welch correction, or Mann–Whitney rank test. Two groups in more than two conditions were compared by two-way analysis of variance with Sidak's correction for multiple testing. Multiple ($>2$) groups were compared by multiple *t*-test with Holm–Sidak post test.

**SUPPLEMENTARY MATERIAL** is linked to the online version of the paper at http://www.nature.com/mi

**DISCLOSURE**
The authors declare no conflict of interest.

**REFERENCES**
1. Peterson, L.W. & Artis, D. Intestinal epithelial cells: regulators of barrier function and immune homeostasis. *Nat. Rev. Immunol.* **14**, 141–153 (2014).
2. Noah, T.K., Donahue, B. & Shroyer, N.F. Intestinal development and differentiation. *Exp. Cell Res.* **317**, 2702–2710 (2011).
3. de Santa Barbara, P., van den Brink, G.R. & Roberts, D.J. Development and differentiation of the intestinal epithelium. *Cell. Mol. Life Sci.* **60**, 1322–1332 (2003).
4. Renz, H., Brandtzaeg, P. & Hornef, M. The impact of perinatal immune development on mucosal homeostasis and chronic inflammation. *Nat. Rev. Immunol.* **12**, 9–23 (2012).
5. Dupaul-Chicoine, J., Dagenais, M. & Saleh, M. Crosstalk between the intestinal microbiota and the innate immune system in intestinal homeostasis and inflammatory bowel disease. *Inflamm. Bowel Dis.* **19**, 2227–2237 (2013).
6. Cilieborg, M.S., Boye, M. & Sangild, P.T. Bacterial colonization and gut development in preterm neonates. *Early Hum. Dev.* **88**, S41–S49 (2012).
7. Salim, S.Y. & Soderholm, J.D. Importance of disrupted intestinal barrier in inflammatory bowel diseases. *Inflamm. Bowel Dis.* **17**, 362–381 (2011).
8. Jones, P.A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492 (2012).
9. Cedar, H., Bergman, Y. & Linking, D.N.A. methylation and histone modifi-cation: patterns and paradigms. *Nat. Rev. Genet.* **10**, 295–304 (2009).
10. Deaton, A.M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–1022 (2011).
11. Zilbauer, M. *et al.* Genome-wide methylation analyses of primary human leukocyte subsets identifies functionally important cell-type-specific hypomethylated regions. *Blood* **122**, e52–e60 (2013).
12. Smith, Z.D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220 (2013).
13. Ziller, M.J. *et al.* Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477–481 (2013).
14. Feil, R. & Fraga, M.F. Epigenetics and the environment: emerging patterns and implications. *Nat. Rev. Genet.* **13**, 97–109 (2011).
15. Jenke, A.C. & Zilbauer, M. Epigenetics in inflammatory bowel disease. *Curr. Opin. Gastroenterol.* **28**, 577–584 (2012).
16. Ventham, N.T., Kennedy, N.A., Nimmo, E.R. & Satsangi, J. Beyond gene discovery in inflammatory bowel disease: the emerging role of epigenetics. *Gastroenterology* **145**, 293–308 (2013).
17. Anderson, E.C. & Wong, M.H. Caught in the Akt: regulation of Wnt signaling in the intestine. *Gastroenterology* **139**, 718–722 (2010).
18. Deaton, A.M. *et al.* Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res.* **21**, 1074–1086 (2011).
19. Klug, M. & Rehli, M. Functional analysis of promoter CpG methylation using a CpG-free ND SC. *Epigenetics* **1**, 127–130 (2006).
20. Lipka, D.B. *et al.* Identification of DNA methylation changes at cis-regulatory elements during early steps of HSC differentiation using tagmentation-based whole genome bisulfite sequencing. *Cell Cycle* **13**, 3476–3487 (2014).
21. Manzel, A., Muller, D.N., Hafler, D.A., Erdman, S.E., Linker, R.A. & Kleinewietfeld, M. Role of "Western diet" in inflammatory autoimmune diseases. *Curr. Allergy Asthma Rep.* **14**, 404 (2014).
22. Cario, E. Toll-like receptors in inflammatory bowel diseases: a decade later. *Inflamm. Bowel Dis.* **16**, 1583–1597 (2010).

23. Shaoul, R., Okada, Y., Cutz, E. & Marcon, M.A. Colonic expression of MUC2, MUC5AC, and TFF1 in inflammatory bowel disease in children. *J. Pediatr. Gastroenterol. Nutr.* **38**, 488–493 (2004).

24. Bruno, M.E., Frantz, A.L., Rogier, E.W., Johansen, F.E. & Kaetzel, C.S. Regulation of the polymeric immunoglobulin receptor by the classical and alternative NF-kappaB pathways in intestinal epithelial cells. *Mucosal Immunol.* **4**, 468–478 (2011).

25. Podolsky, D.K., Gerken, G., Eyking, A. & Cario, E. Colitis-associated variant of TLR2 causes impaired mucosal repair because of TFF3 deficiency. *Gastroenterology* **137**, 209–220 (2009).

26. Yap, L.M., Ahmad, T. & Jewell, D.P. The contribution of HLA genes to IBD susceptibility and phenotype. *Best Pract. Res. Clin. Gastroenterol.* **18**, 577–596 (2004).

27. Clevers, H., Loh, K.M. & Nusse, R. Stem cell signaling. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science* **346**, 1248012 (2014).

28. Sheaffer, K.L. *et al.* DNA methylation is required for the control of stem cell differentiation in the small intestine. *Genes Dev.* **28**, 652–664 (2014).

29. Hasler, R. *et al.* A functional methylome map of ulcerative colitis. *Genome Res.* **22**, 2130–2137 (2012).

30. Harris, R.A. *et al.* DNA methylation-associated colonic mucosal immune and defense responses in treatment-naive pediatric ulcerative colitis. *Epigenetics* **9**, 1131–1137 (2014).

31. Cooke, J. *et al.* Mucosal genome-wide methylation changes in inflammatory bowel disease. *Inflamm. Bowel Dis.* **18**, 2128–2137 (2012).

32. Jenke, A.C. *et al.* DNA methylation analysis in the intestinal epithelium-effect of cell separation on gene expression and methylation profile. *PLoS One* **8**, e55636 (2013).

33. Kellermayer, R. Epigenetics and the developmental origins of inflammatory bowel diseases. *Can. J. Gastroenterol.* **26**, 909–915 (2012).

34. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).

35. Smyth, G.K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**: Article 3, 1–25 (2004).

36. Aryee, M.J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics (Oxford, England)* **30**, 1363–1369 (2014).