RESEARCH

DNA methylation in human gastric epithelial cells defines regional identity without restricting lineage plasticity

Kristin Fritsche[1], Francesco Boccellato[1,2], Philipp Schlaermann[1], Max Koeppel[1], Christian Denecke[3], Alexander Link[4], Peter Malfertheiner[4], Ivo Gut[5], Thomas F. Meyer[1,6*] and Hilmar Berger[1,6*]

Abstract

Background: Epigenetic modifications in mammalian DNA are commonly manifested by DNA methylation. In the stomach, altered DNA methylation patterns have been observed following chronic *Helicobacter pylori* infections and in gastric cancer. In the context of epigenetic regulation, the regional nature of the stomach has been rarely considered in detail.

Results: Here, we establish gastric mucosa derived primary cell cultures as a reliable source of native human epithelium. We describe the DNA methylation landscape across the phenotypically different regions of the healthy human stomach, i.e., antrum, corpus, fundus together with the corresponding transcriptomes. We show that stable regional DNA methylation differences translate to a limited extent into regulation of the transcriptomic phenotype, indicating a largely permissive epigenetic regulation. We identify a small number of transcription factors with novel region-specific activity and likely epigenetic impact in the stomach, including GATA4, IRX5, IRX2, PDX1 and CDX2. Detailed analysis of the Wnt pathway reveals differential regulation along the craniocaudal axis, which involves non-canonical Wnt signaling in determining cell fate in the proximal stomach. By extending our analysis to pre-neoplastic lesions and gastric cancers, we conclude that epigenetic dysregulation characterizes intestinal metaplasia as a founding basis for functional changes in gastric cancer. We present insights into the dynamics of DNA methylation across anatomical regions of the healthy stomach and patterns of its change in disease. Finally, our study provides a well-defined resource of regional stomach transcription and epigenetics.

Keywords: Epigenetic regulation, Human stomach, Gastric epithelial development, Gastric cancer, Intestinal metaplasia

Background

Methylation of cytosines in the DNA represents the most studied heritable epigenetic modification that is highly important in embryonic development, tissue formation and cellular state maintenance [1]. DNA methylation patterns are established early in development after two waves of erasure in gametogenesis and early embryogenesis [2]. Methylation of promoter and enhancer regions during development often leads to stable repression of transcription factors (TFs) [3, 4] and is essential to stabilize cell lineages and phenotypes [5, 6]. Once established,

*Correspondence: tfm@mpiib-berlin.mpg.de; h.berger@ikmb.uni-kiel.de

[1] Department of Molecular Biology, Max Planck Institute for Infection Biology, Charitéplatz 1, 10117 Berlin, Germany
[6] Laboratory of Infection Oncology, Institute of Clinical Molecular Biology, Christian Albrecht University of Kiel and University Hospital Schleswig-Holstein – Campus Kiel, Rosalind-Franklin-Straße 12, 24105 Kiel, Germany
Full list of author information is available at the end of the article

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 2 of 19

these patterns remain stable throughout life. They can, however, undergo stochastic or deterministic changes, which mostly depend on the epigenetic context of a given site [7–10].

The human stomach consists of distinct regions along its cranial–caudal axis (gastric cardia, fundus, corpus or body and antrum), characterized by differences in mucosa phenotype, cellular composition and function [11]. These characteristics arise during development and are stable over time [12]. The role of DNA methylation in the establishment and homeostasis of these regional differences has not been elucidated in detail so far, despite some preliminary regional observations [13]. Aberrant DNA methylation has been linked to several pathological conditions, including cancer [14, 15]. In particular, it is also a hallmark of gastric adenocarcinoma that arises from a transformed healthy glandular epithelium [15]. Gastric cancer (GC) development is predominantly associated with chronic *Helicobacter pylori* infection [16]. Notably, changes in DNA methylation are already observed early in infected tissues and premalignant lesions, including intestinal metaplasia (IM) [13, 17, 18]. Although our understanding of DNA methylation dynamics increased drastically in the past years along with improved methods of its detection [19, 20], the mechanisms leading to aberrant DNA methylation in GC, apart from tumors with gastric CpG island methylator phenotype (CIMP) [21], remain largely unknown. Recent advances in primary cell culture techniques allow to maintain and expand pure epithelial cells derived from healthy or diseased tissue [22, 23]. Primary human gastric epithelial cells derived from the antrum and corpus can be cultivated in the presence of Wnt Family Member 3A (WNT3A) and R-Spondin 1 (RSPO1), where they maintain region-specific transcriptional programs [24]. Further, their differentiation to pit cells can be induced upon the withdrawal of the WNT3A and RSPO1 morphogens [22]. To understand the characteristics and dynamics of DNA methylation in human gastric epithelial cells, we cultivated primary cells, which were obtained from healthy antrum, corpus and fundus, as plane mucosoids

in the presence and absence of WNT3A and RSPO1 [22], followed by analyses of DNA methylation and gene expression patterns. We ran detailed bioinformatic analyses of epigenetic and transcriptional differences, comparing different stomach regions, differentiation states and pathological states. Our data were integrated with published methylomes of healthy gastrointestinal tissues, novel ex vivo-cultivated gastric primary cells of IM and public data sets of IM tissue and GC. We describe an epigenetic landscape that allows differentiation plasticity with only minor restrictions underlying the regional phenotypic differences in the stomach.
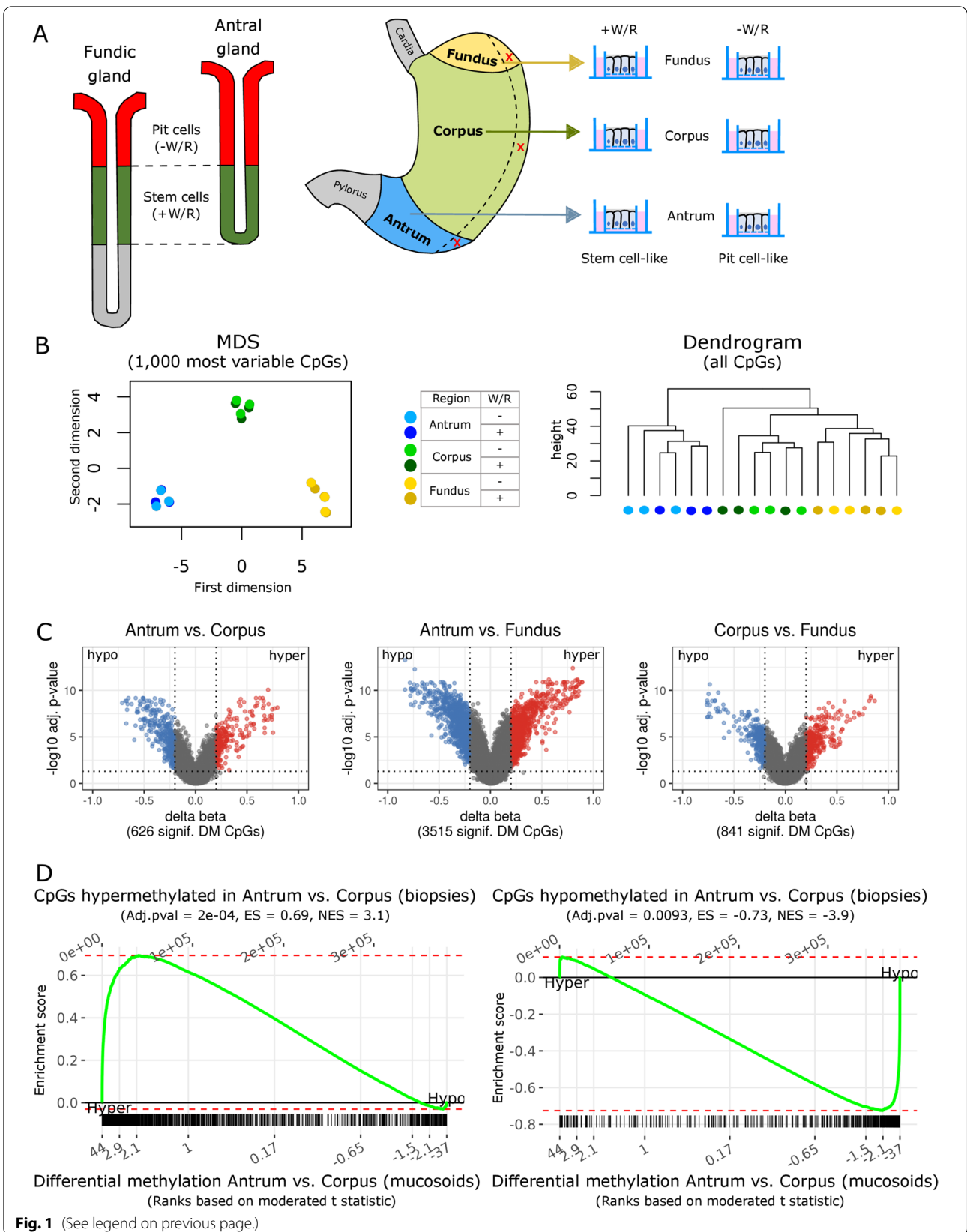
Results

In vitro model of primary human gastric epithelial cells recapitulates regional DNA methylation profiles

To investigate cell-type-specific DNA methylation patterns, we isolated human gastric epithelial cells of the antrum, corpus and fundus from sleeve resections and cultivated them as mucosoids [22]. Stem cell-enriched cell populations of the antrum, corpus and fundus were maintained in a medium containing WNT3A and RSPO1 (+W/R) or differentiated in vitro to pit cells after removal of both factors (−W/R, Fig. 1A). Genome-wide DNA methylation analysis of three biological replicates revealed that cells from different regions preserved a clear regional identity (Fig. 1B, C), while clusters distinguishing −W/R from +W/R cells were not observed. We sought to identify CpGs showing regional methylation differences irrespective of +W/R and -W/R treatment. Out of 738,115 CpGs, 3703 CpGs were found to be differentially methylated (DM) by comparing each two stomach regions (i.e., inter-regional comparison; delta beta > 0.2 or < − 0.2, FDR < 5%) (Fig. 1C, Additional files 1 and 7: Fig. S1A, Table S1). Inter-regional methylation differences determined in vitro correlated very well with those determined in in vivo samples from similar regions [13] (Fig. 1D, Additional file 1: Fig. S1B). Thus, the in vitro mucosoids maintain their region-specific DNA methylation profile. Because DNA methylation changes upon differentiation are known to be less prominent [25], we

(See figure on next page.)

Fig. 1 Global patterns of DNA methylation in the stomach. **A** Left: schematic representation of the human fundic and antral gland types and the respective location of pit cells and stem cells. Right: an experimental overview. Following sleeve resection, sample tissues (red X) from the antrum, corpus and fundus were cultivated as mucosoids using air–liquid interface cell culture inserts. The removal of WNT3A and RSPO1 (W/R) from the cell culture medium enriched for differentiated pit-like cells (−W/R) compared to stem cell-like cell populations obtained when cultivated in the presence of W/R (+W/R). **B** Left: multidimensional scaling (MDS) of methylation proportions (beta) based on the 1000 most variable CpGs in the normalized data set; samples from three biological replicates under the indicated conditions. Right: hierarchical cluster analysis (dendrogram) of all CpGs (n = 738,115) in the normalized data set. **C** Volcano plots representing stomach inter-regional comparisons (+W/R and −W/R combined). The dashed horizontal and vertical lines indicate cutoffs of FDR < 0.05 and delta beta between − 0.2 and 0.2, respectively. Red dots refer to hypermethylated CpGs (hyper) and blue dots to hypomethylated CpGs (hypo). **D** Enrichment of sets of CpGs hypermethylated (left) or hypomethylated (right) in antrum compared to corpus biopsies relative to differential methylation of CpGs between antrum and corpus mucosoids. Moderated t-scores were used for DM CpG ranking. ES—enrichment score; NES—normalized enrichment score. Hyper, Hypo indicate the direction of differential methylation in the antrum vs. corpus mucosoids comparison

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 3 of 19

Fig. 1 (See legend on previous page.)

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 4 of 19

performed a sensitivity analysis lowering the thresholds of delta beta (i.e., the difference in DM) and *p* value in the comparisons between−W/R and+W/R conditions for each stomach region (Additional file 1: Fig. S1C). In agreement with previous findings, no stem cell- or differentiation-specific gene was found to be consistently affected by DM across regions (Additional file 1: Fig. S1D). Instead, we detected only single CpGs with weak differences below the confidence threshold. It should be mentioned, though, that in the corpus, the suggested stem cell marker *TNFRSF19* (TNF Receptor Superfamily Member 19, also known as TROY) was hypermethylated in a fraction of differentiated cells compared to undifferentiated cells.
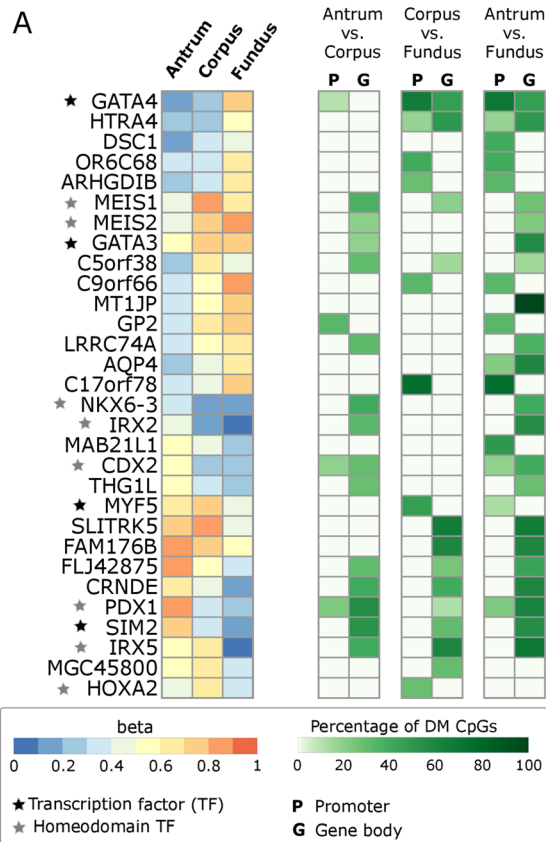
Regional DNA methylation differences in the gastric epithelium suggest stem cell plasticity

In total, 170 and 159 genes were found to be affected by hypermethylation or hypomethylation, respectively, comparing any two stomach regions (Additional files 2 and 8: Fig S2A, Table S2). The top 15 most affected genes between all regions, ordered by their percentage of DM CpGs out of all interrogated CpGs per gene, include several well-known developmental (homeodomain-) TFs (Fig. 2A). A gene set enrichment analysis of all DM genes revealed that almost half (147 out of 315 unique genes) of these genes are indeed involved in developmental processes like tissue development and pattern specification (Fig. 2B, Additional file 9: Table S3). Strikingly, only 11 genes of all DM genes demonstrated gene expression differences (Additional file 2: Fig. S2B). Of these, only four genes were found where promoter methylation negatively or positively correlated with gene expression (Fig. 2C). Interestingly, among genes most affected by differential methylation, we identified the pan-stomach developmental regulator GATA Binding Protein 4 (*GATA4*), Pancreatic And Duodenal Homeobox 1 (*PDX1*), which is a transcription factor essential for the development of the pancreas and gastroduodenal junction and expressed in gastric antrum [26–28], as well as

Iroquois Homeobox 2 and 5 (*IRX2* and *IRX5*), which have been implicated in the specification of stomach fundus in mice [29] (Fig. 2A). Furthermore, we detected the following inter-regionally DM genes: the Meis Homeobox family proteins *MEIS1* and *MEIS2*, which cooperatively bind DNA with several other homeodomain-containing TFs, like PDX1 or GATA TFs [30], the intestinal master regulator Caudal Type Homeobox 2 (*CDX2*) [31], and the transcription factor SIM BHLH Transcription Factor 2 (*SIM2*). Using additional in vivo and in vitro data sets (see Methods, Additional file 13: Table S7), we validated differential methylation of these genes at their respective promoter and enhancer sites (Additional file 10: Table S4) and evaluated their gene expression (Fig. 2D, Additional file 3: Fig. S3). Our analysis revealed that all of these DM sites bind Enhancer Of Zeste 2 Polycomb Repressive Complex 2 Subunit (EZH2), and most of them also SUZ12 Polycomb Repressive Complex 2 Subunit (SUZ12). EZH2 is the functional enzymatic subunit of the polycomb repressive complex 2 (PRC2), which catalyzes the methylation of lysine 27 of histone 3 (H3K27), leading to transcriptional repression [32]. PRC2 proteins, such as EZH2 and SUZ12, play a critical role in stem cell maintenance, lineage specification and differentiation [33, 34]. We performed locus overlap analysis [35] to identify known TF binding sites (TFBS) across discovered stomach-associated DM CpGs corresponding to published embryonic stem cell data sets [36, 37]. In line with the enrichment at binding sites for EZH2 and SUZ12 (Fig. 2E), we also detected a regional enrichment of DM CpGs at sites marked with trimethylation at lysine 9 and 27 of histone 3 (H3K9me3 and H3K27me3). In contrast, these genomic features were not enriched by comparing DM CpGs between the stomach and adjacent tissues (Fig. 2E). Instead, H3K4me1, a histone modification characteristic of enhancers [38, 39], was enriched. In contrast to the inter-regional differences, we further detected that differences with other tissues preferentially occur at open sea CpGs (Additional file 2: Fig. S2C), which might be related to cell-type-specific differences
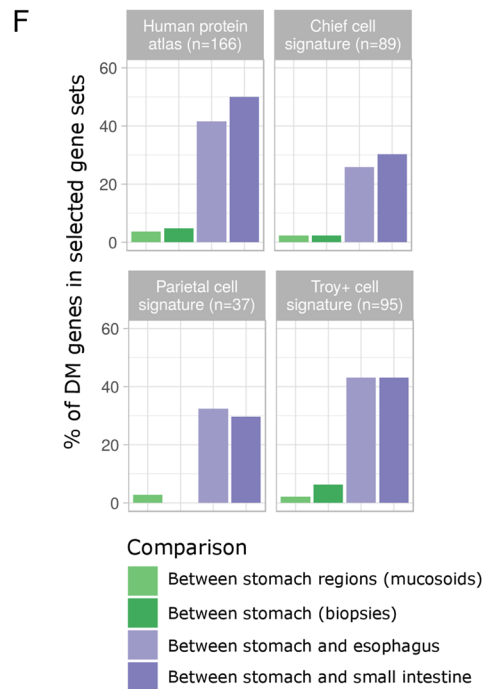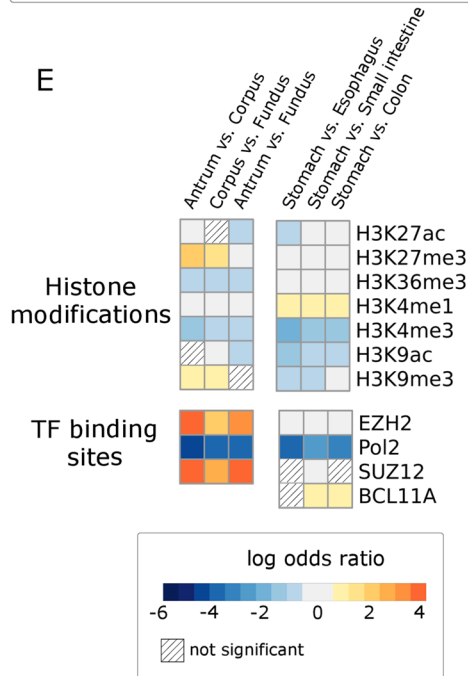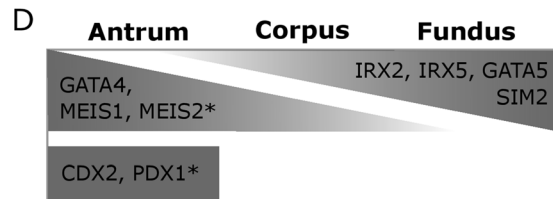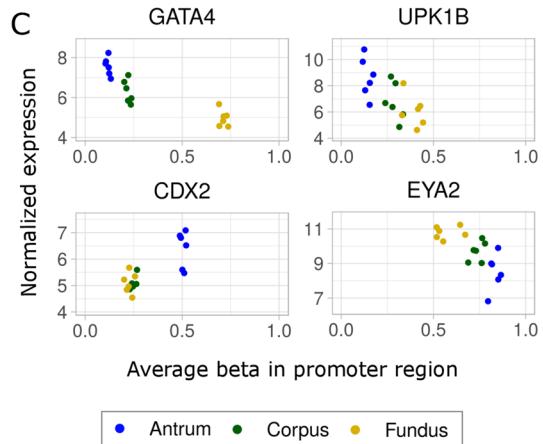
(See figure on next page.)

Fig. 2 Differential methylation within the stomach and between the stomach and adjacent tissues. **A** Displayed are the top 15 DM genes in all comparisons, ranked by the highest number of DM CpGs normalized to total CpGs/gene. Left: a heatmap of average DNA methylation (beta) values per gene. Color code ranges from 0 (unmethylated, blue) to 1 (methylated, red). Stars refer to (homeodomain) transcription factors (TFs). Right: the percentage of significantly DM CpGs out of the total number of interrogated CpGs per gene; color code ranges from white to dark green refer to percentage. **B** Table of selected gene ontology (GO) terms enriched in antrum vs. fundus, determined with StringDB [85]. **C** Average promoter methylation is plotted against the normalized gene expression values. The selected genes showed negative (GATA4, UPK1B, EYA2) and positive (CDX2) correlations between promoter methylation and gene expression. **D** Schematic representation of expression patterns of several DM TFs in different stomach regions, based on mucosoid and biopsy data sets (Additional files 3, 13: Fig. S3, Table S7). Asterisks indicate genes not differentially expressed in the mucosoid data set. **E** Enrichment and depletion of DM CpGs according to different genomic features—heatmap comparisons between stomach regions and between the stomach and adjacent tissues. Shown are significantly enriched (log odds ratio > 0.6) and depleted (log odds ratio < − 0.6) histone modifications and TF binding sites (FDR < 0.05). Color code ranges from blue (depleted) to red (enriched); light gray refers to a log odds ratio between − 0.6 and 0.6; shaded squares display nonsignificant results; Pol2—Polymerase 2 subunit. **F** Stomach- and stomach cell-type-specific gene sets—percentages of DM genes between stomach regions (green) and between the stomach and adjacent tissues (purple)

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 5 of 19

Fig. 2 (See legend on previous page.)

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 6 of 19

in A-B compartments [40, 41]. Many stomach-specific genes are expressed in individual cell lineages, predominantly in selected areas of the stomach, for example, parietal cells in the corpus and fundus and G-cells in the antrum. We determined the percentage of DM among stomach-specific genes (Fig. 2F) and found that they are rarely affected inter-regionally in the stomach in contrast to the comparisons between different tissues. Together, the small number of stomach cell lineage-specific genes affected by differential methylation and the small number of regulated TFs indicate that the regulation of transcription programs within the stomach is rather permissive and does not include stringent specification of lineages.

Small but specific gene expression differences define the regional identity of stem cells in vitro

In contrast to methylation levels, the global transcriptomic landscape of +W/R and −W/R stomach regions appeared to be governed by changes between stem cells and pit cell-like cell populations (Fig. 3A, B, top). However, we also identified regional gene expression differences between the antrum and the corpus in vitro, which correlated well with those in vivo (Additional file 4: Fig. S4A). Of note, the expression of marker genes of dominant chief and parietal cells in the corpus and fundus was not induced in vitro (Additional file 4: Fig. S4B). The presence of antral endocrine G and D cells was confirmed by *GAST* (Gastrin) and *SST* (Somatostatin) expression, respectively (Additional file 4: Fig. S4B). Most proposed stem cell markers in the stomach, including the antral marker leucine-rich repeat containing G protein-coupled receptor 5 (*LGR5*), the corpus markers *TNFRSF19* (*TROY*), SRY-Box Transcription Factor 2 (*SOX2*) and RUNX Family Transcription Factor 1 (*RUNX1*), were expressed as reported [42–45] in the +W/R cultures (Fig. 3C). Despite these differences in stem cell gene expression, only a few other genes were differentially expressed inter-regionally (Fig. 3B, bottom). The antrum vs. fundus comparison revealed the relatively largest number of differences, including genes that were also DM (Fig. 3D, Additional file 2: Fig. S2B). A gene set enrichment analysis of the fundus compared to antrum further indicated a higher expression of genes associated with complement activation and adaptive and innate immune responses (Fig. 3E, Additional file 11:

Table S5). Interestingly, three of the leading-edge genes in adaptive and innate immune response gene sets, Beta-2-Microglobulin (*B2M*), Protein Tyrosine Kinase 2 Beta (*PTK2B*) and BCL6 Transcription Repressor (*BCL6*), were also hypermethylated in the antrum compared to fundus (Additional file 7: Table S1). Hypermethylation at *B2M* affected the gene body with annotated Fantom5 enhancer site [46], whereas *PTK2B* and *BCL6* hypermethylation occurred in the promoter region. Together, transcriptome analysis of stomach cells under +W/R and −W/R conditions in vitro present broadly similar profiles with only minor transcriptional differences.

Wnt responsiveness decreases from the antrum to the fundus

The withdrawal of the morphogens W/R induced differentiation to pit cells in all stomach regions (Fig. 4A). Differentiated pit cells characteristically expressed Gastrokine 1 and 2 (*GKN1* and *GKN2*) and Mucin 5AC (*MUC5AC*), while the stem cell-enriched population characteristically expressed Mucin 6 (*MUC6*) and *CD44*. Although canonical Wnt signaling was active in all three stomach regions (Fig. 4B), we detected regional differences in the expression of Wnt pathway genes, with antrum showing the highest number of up and down-regulated genes followed by corpus and fundus, and only a small number of differentially expressed (DE) genes shared by all three regions (Additional file 4: Fig. S4C). Thus, we determined the expression levels of crucial Wnt signaling genes upon induced differentiation. While the major Wnt target Axin 2 (*AXIN2*) was equally downregulated in differentiating mucosoids, indicating active Wnt signaling in cells from all regions, other genes demonstrated a region-dependent effect (Additional file 4: Fig. S4C, D). The differentiation effect on the expression of LGR5 and Transcription Factor 7 Like 1 (*TCF7L1*, one of the four TCF/LEF proteins that mediate Wnt signaling) weakened along the inferior–superior axis, while the remaining members of the family (TCF7, TCF7L2 and LEF1) displayed only minor or no change along this axis (Fig. 4C, Additional file 4: Fig. S4C, D). Interestingly, we detected an opposite trend in the expression of the non-canonical Wnt signaling genes Wnt Family Member 5A (*WNT5A*) and its putative receptor Receptor Tyrosine Kinase-Like Orphan Receptor 2 (*ROR2*) [47, 48], whose

---

(See figure on next page.)
Fig. 3 Differential gene expression in the stomach. **A** MDS of gene expression values in all samples. **B** Volcano plots of gene expression log2 fold changes between −W/R and +W/R (top) and between the stomach regions (+W/R and −W/R combined per region) (bottom). The dashed horizontal and vertical lines indicate cutoffs of adjusted *p* value (FDR) < 0.2 and a log2 fold change at −0.6 and 0.6, respectively. **C** Regional comparison of differentially regulated stem cell genes in undifferentiated samples (+W/R, *p* < 0.05). **D** Heatmap of the top 15 up- and downregulated genes (FDR < 0.2) in all regional comparisons. The normalized expression is represented by row-wise standardization (z-score). Color code ranges from blue (low expression) to red (high expression). **E** Selected GO terms of biological process based on the GSEA of antrum versus fundus. NES—normalized enrichment score

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 7 of 19
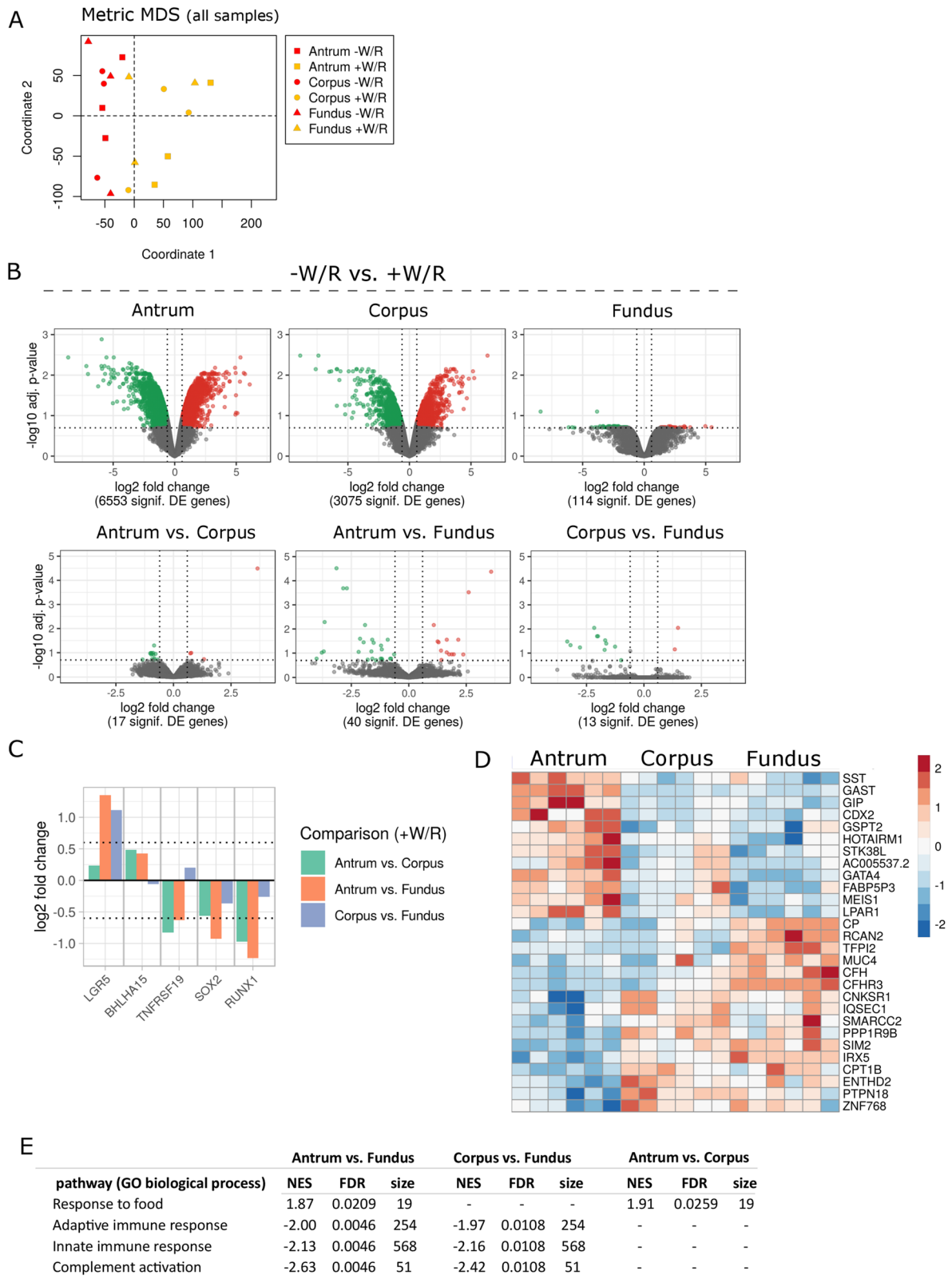
Fig. 3 (See legend on previous page.)

expression in the corpus and fundus was higher than the antrum, regardless of the differentiation state (Fig. 4C). Moreover, we detected DM of *WNT5A*, *TCF7L1* and *TCF7,* comparing the antrum and the fundus (Fig. 4D). In this comparison, *WNT5A* and *TCF7* were hypermethylated and hypomethylated, respectively, at one CpG, and *TCF7L1* was hypermethylated at two CpGs (Additional file 4: Fig. S4E). These CpGs are annotated by ENCODE as a region of an active promoter in embryonic stem cells, indicating a putative regulatory site for the expression of *WNT5A* and *TCF7* in gastric stem cells. In summary, transcriptomic and epigenomic data point toward the implication of the canonical and non-canonical Wnt pathways across stomach regions with differences in strength along an inferior–superior axis.

Ex vivo-cultivated organoids of intestinal metaplasia show specific methylation patterns that are maintained in gastric cancer

IM is the emergence of epithelial cells with an at least partial intestinal phenotype in the stomach, associated with chronic *Helicobacter pylori* infection. We cultured epithelial cells obtained from biopsies of antral IM sites and normal control sites ex vivo. As expected, ex vivo IM organoids showed transcriptional downregulation of stomach-specific genes and upregulation of intestinal genes (Additional file 5: Fig. S5A). These organoids were also clearly segregated from their healthy controls in terms of DNA methylation (Additional file 5: Fig. S5B). Additional samples that we cultivated from atrophic mucosa biopsies showed an intermediate methylation phenotype and were not further characterized. The differences detected in IM compared to normal ex vivo samples (delta beta > 0.2 or < − 0.2, *p* < 0.05) correlated strongly with differences in vivo (Additional file 5: Fig. S5C–E), demonstrating the stability of IM epigenomic patterns ex vivo. We observed, however, higher DM CpG numbers in the in vivo data set compared to the ex vivo data set, most likely due to the higher power to detect DM in larger cohorts. We also observed a strong bias of hypermethylation in the in vivo data set (14,023 hyper- and 1199 hypomethylated CpGs, Additional file 5: Fig. S5E, left). Even stronger overlaps with our data, including affected genes, promoters and enhancers at the CpG level, were found when selecting the subset of in vivo samples with the highest global methylation levels and
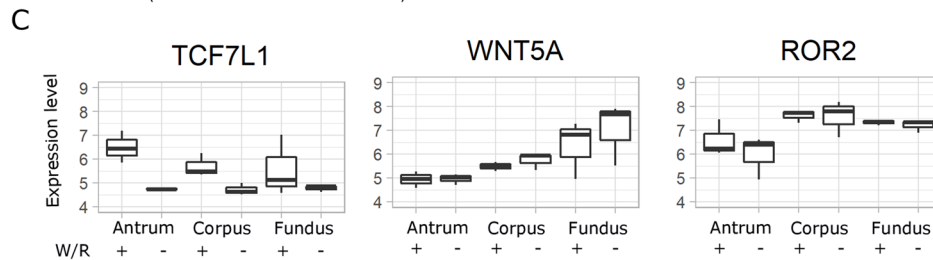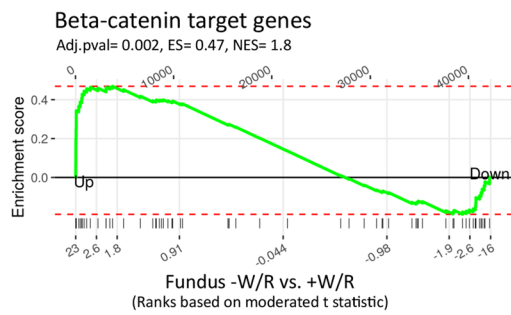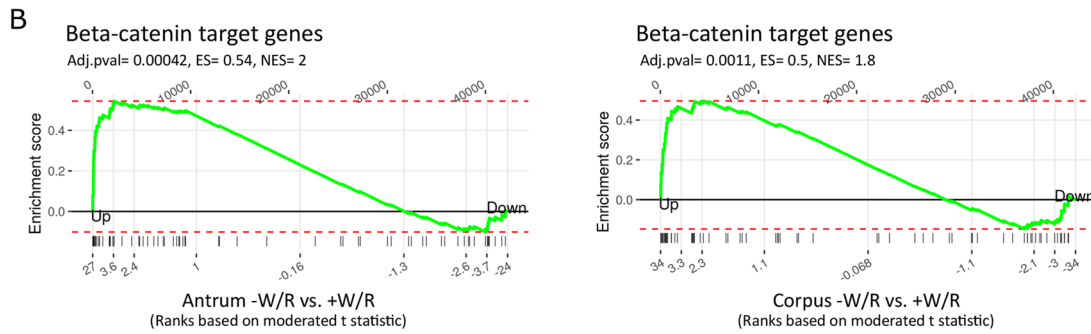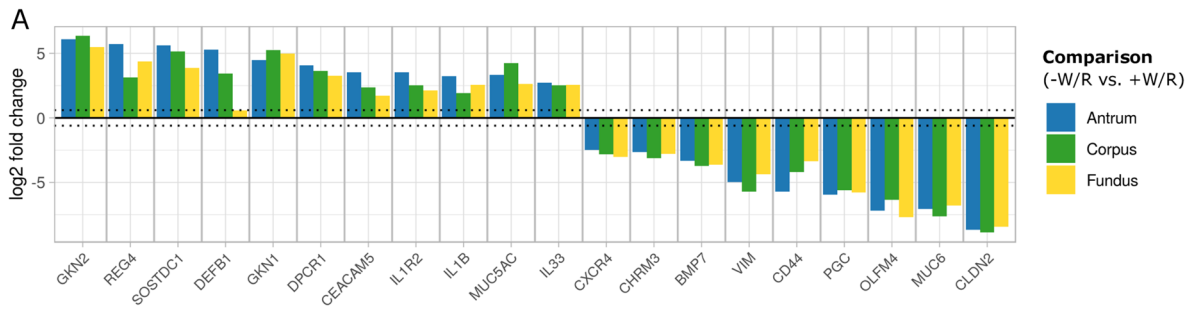
a higher purity [13] (Additional file 5: Fig. S5E, right). Higher levels of intermediate DNA methylation due to contaminating cells and proportionally more DM sites have been previously described when comparing results obtained with pure epithelial cells compared to biopsies [49]. Here, we show that the analysis of ex vivo cultured gastric cells results in lower admixture effects and improved calling of DM sites (Additional file 5: Fig. S5F).

The number of DM CpGs in IM ex vivo is similar to what we observed inter-regionally; however, these DM CpGs affected much higher numbers of genes and promoters (Additional files 2 and 5: Fig. S2A, S5E). Unexpectedly though, only a small number of genes with aberrant promoter hypermethylation or hypomethylation showed differences in gene expression (15/191 and 13/85, respectively; Tables 1, 2). Nevertheless, these differences were aligned with the general dogma of promoter hypomethylation, leading to increased transcription and vice versa. Hypomethylated and upregulated genes included the intestinal stem cell marker Achaete-Scute Family BHLH Transcription Factor 2 (*ASCL2*) as well as two other intestine-specific genes, Tripartite Motif Containing 15 (*TRIM15*) and Fucosyltransferase 6 (*FUT6*) [50]. To further support the regulation of these intestinal genes by promoter methylation, we treated healthy antral mucosoids with the demethylating agent 5-aza2'-deoxycytidine (5aza); this resulted in an increased expression (Table 2, Additional file 12: Table S6). We asked whether hypomethylated CpGs in ex vivo IM also played a possible role in gene deregulation in IM. Enrichment analysis of those DM CpGs revealed enriched binding sites for the intestinal TFs HNF1 Homeobox A (HNF1A), Interferon Regulatory Factor 1 (IRF1), PDX1 and Homeobox A5 (HOXA5) (Fig. 5A). *PDX1* and *HOXA5 gene* expression was increased in ex vivo-cultivated organoids of IM (log2 fold change = 1.06 and 0.54, and FDR = 0.174 and 0.025, respectively). CDX2 is the major TF driving intestinal transcriptional program and is known to be upregulated in IM [51, 52]. It has been reported that promoter methylation is not associated with *CDX2* expression in IM tissue [53]; the sites of regulation within *CDX2* that provoke its aberrant expression in IM remain unknown. In our IM analysis, we observed accompanying *CDX2* upregulation, promoter hypermethylation as well as a hypomethylation site in the *CDX2* gene body. This gene body site, which contains an EZH2 binding site (Additional file 10:

(See figure on next page.)

Fig. 4 Differentiation- and stem cell-specific gene expression. **A** Selected up- and downregulated pit cell- and stem cell-specific genes between −W/R and +W/R conditions in antrum-, corpus- and fundus-derived cell cultures (FDR < 0.05). **B** Gene set enrichment of beta-catenin target genes (n = 66) [90] compared to differential gene expression between −W/R vs. +W/R. The moderated t-score of the DE results was used for ranking the genes. ES—enrichment score; NES—normalized enrichment score. **C** Normalized gene expression boxplots of *TCF7L1* and the non-canonical Wnt signaling pathway genes *WNT5A* and *ROR2*. The boxplot displays the median, minimum and maximum normalized expression values of three biological replicates. **D** Table of DM CpGs within the genes displayed in C. n.s.—not significant

Fig. 4 (See legend on previous page.)

Fritsche *et al. Clinical Epigenetics* (2022) 14:193

Page 10 of 19

Table 1 Hypermethylated promoter regions in IM

Gene	Hypermethylated CpGs/ promoter region	Total CpGs/ promoter	logFC	*p* Value	FDR	Detected in
KDR	3	12	− 1.5	0.0000	0.0003	Ex vivo and in vivo
GCNT2	3	17	− 1.42	0.0000	0.0000	Ex vivo
CAPN13	3	6	− 1.39	0.0000	0.0000	Ex vivo
EYA4	7	51	− 1.37	0.0000	0.0013	Ex vivo and in vivo
ITGBL1	4	10	− 1.22	0.0001	0.0101	Ex vivo
SLC9A3	5	7	− 1.01	0.0001	0.0112	Ex vivo
CLIC6	5	11	− 0.89	0.0039	0.1244	Ex vivo
CHAD	7	8	− 0.69	0.0312	0.4136	Ex vivo
SORBS2	7	55	− 0.69	0.0003	0.0201	Ex vivo
TAS1R1	3	14	− 0.68	0.0293	0.4018	Ex vivo
PSCA	4	6	− 0.66	0.0266	0.3812	Ex vivo
CYP2E1	5	7	0.77	0.0063	0.1681	Ex vivo
TMEM176B	2	20	0.89	0.0022	0.0838	Ex vivo
GJB2	3	20	1.05	0.0000	0.0021	Ex vivo
CDX2	2	12	1.34	0.0000	0.0033	Ex vivo

Shown are the 15 genes with hypermethylated promoter regions in IM that also show altered gene expression. The in vivo samples represent antral samples with a high methylation cluster. logFC = log2 fold change; FDR = false discovery rate

Table 2 Hypomethylated promoter regions in IM

Gene	Hypomethylated CpGs/ promoter region	Total CpGs/ promoter	logFC	*p* Value	FDR	Detected in	5Aza logFC
TRIM15	9	25	2.11	0.0000	0.0000	Ex vivo and in vivo	1.66
ASCL2	6	45	1.27	0.0000	0.0003	Ex vivo and in vivo	0.79
HOXA2	15	28	1.24	0.0000	0.0053	Ex vivo	
HMGA1	3	17	1.09	0.0000	0.0000	Ex vivo	
TMEM154	3	7	0.92	0.0040	0.1257	Ex vivo	
KRT23	2	6	0.91	0.0045	0.1351	Ex vivo	4.25
FUT6	4	8	0.90	0.0033	0.1117	Ex vivo and in vivo	0.99
GAL3ST2	2	6	0.88	0.0006	0.0315	Ex vivo and in vivo	
GJC2	2	13	0.67	0.0370	0.4529	Ex vivo	2.89
SYNE1	2	21	0.66	0.0266	0.3812	Ex vivo	1.48
C8orf74	2	6	0.64	0.0469	0.5108	Ex vivo	
HOXA4	10	24	0.62	0.0262	0.3786	Ex vivo	
DCHS2	2	15	− 0.64	0.0093	0.2145	Ex vivo	

Shown are the 13 genes with hypomethylated promoter regions in IM that also show altered gene expression. The in vivo samples represent antral samples with a high methylation cluster. 5 aza = 5-aza-2′-deoxycytidine; logFC = log2 fold change; FDR = false discovery rate

Table S4), was also DM in a comparison between the healthy antrum and the corpus/fundus (Fig. 2A), indicating its putative regulatory role (Fig. 5B). Taking into account the in vivo data set, we noticed that the DM of this *CDX2* site might be obscured by the admixture effect described above (Additional file 5: Fig. S5F). As observed by Huang et al. [13], we observed an enrichment of DM CpGs in ex vivo IM at genes with a bivalent promoter state indicated by the enrichment of H3K4me1 and H3K27me3 [54] and TFBS of SUZ12 and EZH2 (Additional file 6: Fig. S6A). Additionally, in our data also hypomethylated CpGs show this enrichment, probably because we detected higher numbers of hypomethylated CpGs. Together, these data indicate that both key genes and TF binding sites might be affected by DM patterns in IM.

We were interested in understanding the mechanisms that drive aberrant methylation in IM and its relevance

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 11 of 19

to GC. Recently, a hypermethylation signature at specific CpG islands shared by most cancers and associated with extra-embryonic tissues was identified [55]. Interestingly, IM organoids displayed increased methylation levels at these sites compared to normal tissues. (Fig. 5C, Additional file 6: Fig. S6B, C). We observed the same pattern in the in vivo data set of IM [13] as well as in a data set of Barrett's esophagus (BE) samples [56], which is an IM that occurs in the lower esophagus. These results suggest that the dysregulation of putative responsible signaling pathways occurs already before the transformation. Finally, we asked whether any IM-related changes are preserved in GC. We determined the overlaps of hypermethylated and hypomethylated promoters in IM with those identified in molecular subtypes of GC, namely chromosomally instable (CIN), microsatellite instable (MSI) and genomically stable (GS) GCs. Most intestinal-like GCs (59%) are classified as belonging to the CIN subtype, whereas MSI and GS subtypes represent only 24 and 9% of intestinal-like GCs, respectively [14]. A recent study [13] has found a substantial overlap between hypermethylated regions in IM and those of CIN and MSI subtypes of GC. Our data is in good agreement with these findings. We found that 69% of the hypermethylated promoters and 19% of the hypomethylated promoters in IM overlapped with all three molecular GC subtypes (Fig. 5D). Since aberrant promoter methylation in IM was shared with all GC subtypes, it seems that these differences might not be limited to the intestinal phenotype alone. We further observed that while most genes with shared aberrant promoter methylation in IM and GC are not differentially expressed in IM, they show deregulation in cancer samples compared to healthy controls (Fig. 5D, lower row). This indicates that nascent aberrant methylation of these genes does not lead to considerable regulatory effects at the IM stage, but it does once the cells become malignant. As a prominent example, promoter hypermethylation of MutL Homolog 1 (*MLH1*) is present already in IM, but the hypermethylation does not affect the specific CpG island responsible for its silencing in cancer [57] (Additional file 6: Fig. S6D); *MLH1* was not differentially expressed (data not shown). These data support the hypothesis that aberrant methylation starting in IM continues in GC, eventually promoting deregulated expression.
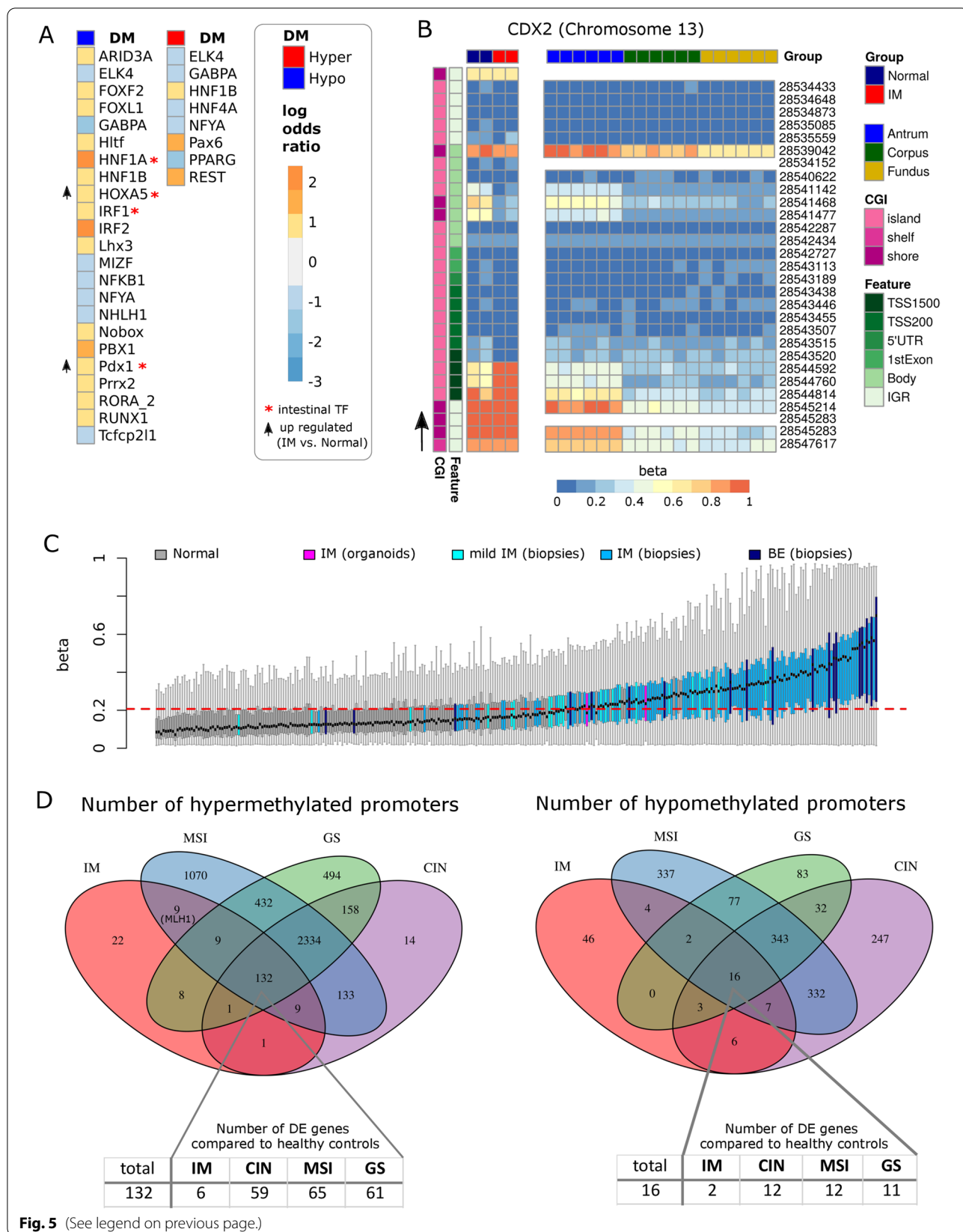
Discussion

The stomach, a major organ of the gastrointestinal tract, is divided into several functionally distinct domains along its longitudinal (craniocaudal) axis; these domains are characterized by phenotypically and functionally diverse mucosa with the predominance of distinct cell types and lineages. While some regulatory factors of the different zones have been identified [58], it is unclear what their role is in inter-regional stomach development and how these phenotypic differences are maintained.

By using a primary human gastric epithelial cell culture model [22], we show that global DNA methylation patterns define a regional identity that clearly and stably distinguishes the antrum, corpus and fundus. These results are supported by previous studies that have shown region-specific methylation profiles in individual tissues [59, 60] and different colon regions [49]. In addition, by comparing our in vitro data to published in vivo data from stomach biopsies we demonstrate that DNA methylation differences between regions are well preserved, indicating that methylation profiles are mainly stable during short-time culture. We present an extensive analysis across all three main stomach regions and differentiation states (stem and pit cells), thereby creating a whole map of DNA methylation and gene expression dynamics in the healthy human gastric epithelium. Interestingly, we find that many TFs DM between regions are no longer transcriptionally active in the adult stomach and that most region-specific lineage genes are not affected by DM. Together, these results infer plasticity of gastric epithelial stem cells and a specification of regional cellular phenotypes by only a few TFs and possibly cell-extrinsic

(See figure on next page.)

Fig. 5 Differential methylation in intestinal metaplasia. **A** Enrichment of hypermethylated and hypomethylated CpGs in TF binding sites (JASPAR collection) in organoids derived from intestinal metaplasia (IM) biopsies compared to organoids derived from normal gastric biopsies. Shown are log odds ratios of significantly enriched (log odds ratio > 0.6) or depleted (log odds ratio < −0.6) genomic features (FDR < 0.05). Color code ranges from blue (depleted) to red (enriched); gray refers to a log odds ratio between -0.6 and 0.6. Hyper—hypermethylated CpGs; Hypo—hypomethylated CpGs. **B** Heatmap of the *CDX2* DNA methylation pattern in the IM organoids vs. normal gastric organoids. The arrow indicates the transcriptional direction. CGI—CpG island, TSS—transcription start site, 5′UTR—5′ untranslated region, IGR—intergenic region. **C** ExE-hyper-CGI methylation of in healthy and precancerous samples. Shown are the mean ExE-hyper-CpG island methylation values per sample ranked by their means. Normal samples include all the samples from our mucosoid cultures (antrum, corpus and fundus) as well as all the healthy samples from the organoid and biopsy IM data sets [13] and the Barrett's esophagus (BE) biopsy data sets [56]. The dashed red line indicates the upper 95% confidence interval limit of all normal samples. Diseased samples include the respective in vivo mild IM and IM (biopsies), IM (organoids) and BE (biopsies). **D** Venn diagram of genes with hypermethylated or hypomethylated promoter regions in diseased samples as compared to healthy samples. Diseased samples: IM (organoids), microsatellite instable (MSI), genomically stable (GS) and chromosomal instable (CIN) molecular subtypes of gastric cancer [14]. Tables indicate the number of differentially expressed genes among those shared by all three cancer subtypes and IM

Fig. 5 (See legend on previous page.)

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 13 of 19

signals. Moreover, we found that significant changes in DNA methylation do not accompany the differentiation of stem cell-enriched gastric mucosoids to pit cells; this has been similarly observed in primary intestinal epithelial cells [61]. In addition to TFs known to be involved in stomach development and region specification, such as PDX1 and IRX5, we identified and validated several novel region-specific TFs, including SIM2, MEIS1, MEIS2 and CDX2. Mechanistically, most of those genes were DM at enhancers and promoters that bind EZH2 and SUZ12, pointing toward their possible regulation by PCR2. In different published in vivo data sets as well as in our in vitro data set, MEIS1 and MEIS2 showed higher expression in the antrum than the corpus, whereas SIM2 showed higher expression in the corpus than the antrum. The role of SIM2 in the stomach is yet to be investigated. Notably, in the small intestine, SIM2 has been shown to activate the transcription of the Wnt signaling mediator *TCF7L2* and to directly regulate the expression of several antimicrobial peptides in the small intestine [62].

Wnt and R-Spondin are major stemness regulators in the intestine and stomach [63, 64]. *Lgr5*, a main component of the Wnt pathway, is usually expressed in the antrum, while the mouse corpus usually lacks *Lgr5* [65], suggesting different roles of the Wnt pathway in different stomach regions. In our model,+W/R conditions led to comparable activation of beta-catenin signaling and maintained a stem-like phenotype in epithelial cells *ex vivo* independently of origin. Yet, the expression of the Wnt pathway mediator *TCF7L1* was attenuated in the proximal stomach, while major effectors of the antagonistic non-canonical Wnt signaling, *ROR2* and *WNT5A* [66], were enhanced compared to the inferior regions. These findings are supported by DM in the promoter and enhancer sites of *TCF7L1* and are similar to region-specific regulation of differentiation-associated genes observed in intestinal organoids [61]. In agreement with our findings, a supportive role for non-canonical *Wnt5a* signaling in the stem cell niche of the gastric corpus in contrast to that of the antrum has been suggested by Hayakawa et al. [67].

We extended our analysis of epithelial cell-specific DNA methylation in an attempt to characterize the differences in diseased primary epithelial cells. Comparing healthy samples to GC and its precancerous lesion IM suggested an early occurrence of GC-specific hypermethylation and hypomethylation patterns already at the IM stage. Hypermethylation in ex vivo-cultivated IM as well as in samples of IM in vivo and BE affected a set of particular CpG islands, related to extra-embryonic ectoderm-specific genes found in most cancer types, including GC, and thought to result from dysregulated fibroblast growth factor 2 (FGF2) signaling

[55]. Although we did not detect differential expression of FGF pathway genes in IM ex vivo (data not shown), the injection of the cytotoxin-associated gene A (CagA) by *Helicobacter pylori* is known to activate FGF signaling [68]. Similarly, dysregulated FGF signaling has been observed in BE [69, 70]. Most likely, due to our model of pure epithelial cells, we were also able to identify several hypomethylated promoters that are shared between IM and all GC subtypes, excluding the EBV subtype [14]. Further, shared genes with aberrant hypermethylated or hypomethylated promoters showed gene expression differences, compared to healthy controls, almost exclusively in GC samples but not in IM. This indicates that additional (epi-) genetic changes must occur during the progression of the pre-neoplastic lesions toward full-blown GC.

Changes in ex vivo-cultivated IM organoids correlated well with those differences found in vivo, while the lack of cell admixture in cultured cells [49] allowed more sensitive detection of hypomethylated CpGs sites in IM as compared to studies involving biopsies [13]. Thus, we found hypomethylation of the promoter region and upregulation of the putative stem cell marker of the small intestine, *ASCL2*, essential for maintaining adult intestinal stem cells [71]. In addition to the induction of intestinal genes, we also found that hypomethylated CpGs were enriched at binding sites of intestinal TFs, such as HNF1A, HOXA5, IRF2 and PDX1. These intestinal TFs were partially upregulated in IM ex vivo, compatible with the altered epigenetic regulation of transcriptional networks in IM. Both the binding of TFs at hypomethylated sites [72] and TF binding-induced hypomethylation [73] have been previously observed. Further, enhancer hypomethylation in the healthy intestine has been shown to be associated with inappropriate TF binding [25]. Therefore, our results reveal several potential epigenetic mechanisms and target genes that could contribute to molecular changes that promote carcinogenesis in gastric epithelial cells.

Conclusions

In summary, we present an extensive characterization of epigenetic and transcriptional landscapes that vary across stomach regions. Our study reveals new candidate regulators for region-specific phenotypes and illuminates crucial mechanisms of cellular transformation.

Methods
Human material
Gastric sleeve: Human gastric tissue samples were obtained from the Center for Bariatric and Metabolic Surgery at the Charité University Medicine, Berlin, Germany. Patient samples, negative for *Helicobacter pylori*,

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 14 of 19

were used for isolation of gastric glands from the antrum, corpus and fundus.

Gastric biopsy: Human gastric biopsies of antral normal, atrophic and metaplastic gastric tissue were provided by the Department of Gastroenterology, Hepatology and Infectious Diseases, Otto-von-Guericke-University of Magdeburg, Magdeburg, Germany. Adjacent biopsies were assessed macro- and microscopically by a pathologist, and samples, where macro- and microscopic assessment matched, were used for the isolation of epithelial cells.

Primary cell culture

Organoids and mucosoids were cultivated as previously described [22, 24]. Primary epithelial cells from the corpus and fundus were cultivated under the same conditions as the antral mucosoids. Differentiation of the mucosoids was achieved by replacing Wnt and R-spondin in the cell culture medium with advanced Dulbecco's modified Eagle medium/F12 (Invitrogen) for seven days. Samples of undifferentiated antrum and corpus have also been used for gene expression analysis by Wölffling et al. [74].

5-Aza-2'-deoxycytidine treatment of mucosoid cultures

Freshly seeded healthy antral undifferentiated mucosoid cultures were treated every 24 h with 4 µM 5aza (in 0.02% acetic acid; Sigma-Aldrich). A global decrease in DNA methylation, measured with the 5mC ELISA kit (Zymo Research), to 12% of the total human 5-methylated cytosine (5mC) content was achieved after 10 days, after which cells were harvested for RNA isolation.

Nucleic acid isolation

DNA and RNA were extracted from mucosoid cultures using the QIAamp DNA Mini Kit (Qiagen) and Trizol (Thermo Fisher Scientific) methods, respectively. To extract DNA and RNA from ex vivo-cultivated organoids and 5aza-treated mucosoids, we used the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen). RNA integrity and quantity of RNA were assessed using the Agilent 2100 Bioanalyzer (Agilent). DNA concentration was measured using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific), and integrity was assessed by gel electrophoresis (0.8% agarose gel, 120 V, 1 h).

DNA methylation array (450 k and EPIC)

The Illumina Infinium® human 450 k (Illumina, WG-314-10031) and EPIC methylation (Illumina, SWG-317-1001) bead chips were applied for ex vivo-cultivated organoids and healthy mucosoids, respectively, according to the manufacturer's protocol at the Life&Brain research center (Bonn, Germany) and in collaboration with Per

Hoffmann from the Institute of Human Genetics, University of Bonn.

Gene expression microarray

Single color Agilent SurePrint G3 Custom Gene Expression Microarray $8 \times 60K$ (Agilent, Agilent-048908; GEO Platform GPL21272) was used for gene expression analysis according to the manufacturer's protocol. Note: a part of the gene expression results obtained with the samples of $+W/R$ antrum and $+/W/R$ corpus has already been published [74]. Here, we used these samples combined with the respective $-W/R$ samples and extended the data set to $+W/R$ and $-W/R$ fundus.

Bioinformatic methods

Data analysis was performed in the R/Bioconductor environment [75]. Public data sets used in this work were downloaded from GEO (Additional file 13: Table S7), TCGA (STAD) [14], GTEx Portal [76] and the Human Protein Atlas [50].

DNA methylation

Raw intensity files (idats) were processed using the ChAMP package [77, 78] using default settings. When data sets of the same platform were combined, all idat files were loaded and processed together. Only when combining 450k and EPIC data sets, the minfi function combineArrays was used to virtually combine the already loaded and filtered 450k and EPIC objects [79, 80]. Subset quantile within array normalization (SWAN) was used to create data sets in this study. For combined data sets, functional normalization [81] was applied. Patient and slide effects were adjusted in the healthy in vitro stomach data set using the champ.SVD function. Differential methylation analysis was performed with limma [82], applying a moderated t-test for the mucosoids and a paired moderated t-test for the ex vivo data set using M-values. Since beta-values, ranging from 0 (unmethylated) to 1 (methylated), are more intuitive, DM in this manuscript is reported as delta beta. For the regional comparisons, $+W/R$ and $-W/R$ samples of the same region were combined and analyzed as six biological replicates. Genes were considered DM when at least two CpGs/gene were affected by significant differential methylation (FDR < 0.05, delta beta > 20%). The threshold was lowered to 1 CpG/gene when Wnt pathway genes were determined as DM. Promoters were considered DM when at least one CpG/promoter region, defined as TSS1500, TSS200, 5'UTR or 1st Exon, was significantly differentially methylated. Reference for genes of the canonical and non-canonical Wnt pathways was taken from the GO Wnt and GO-non-canonical Wnt gene lists. Stomach-specific genes were taken from the

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 15 of 19

Human Protein Atlas as those that show fivefold higher mRNA levels compared to average levels in all other tissues [50]. Chief and parietal cell-specific genes were determined from re-analyzing the data set of Ramsey and others [83], who profiled murine gastric epithelial cells, as compared to the other cell types. Symbols of mouse genes whose expression was logFC > 1 were translated into human gene symbols and used as a gene set. Genes of the Troy + cell signature were taken from the Additional file 7: Table S1 [45]. DM genes were classified as TFs according to [84]. Protein–protein interactions and enrichment in GO biological processes of DM genes were determined using the platform StringDB [85]. The segmentation classes for specific gene annotations were taken from the UCSC table browser bed files of Broad Hmm in H1hesc for table annotation and the average segmentation combined H1hesc for visualization. CpG sets corresponding to hypermethylated or hypomethylated CpGs were created from DM results. Only CpGs with delta beta > 20% were included to reduce the maximum size of the CpG set to < 3000; 5000 permutations were applied. The distribution of DM CpGs along genomic features was determined using the UCSC RefGene Group from the Illumina annotation of the EPIC or 450k array by calculating Pearson residuals of DM CpGs for each genomic feature and CpG island relation. Enrichment of DM CpGs was determined using Locus overlap analysis (LOLA) [35], the ENCODE Segmentation classes and TFBSs [37], the UCSC features [86], the Roadmap epigenomics histone marks [36] and the JASPAR motifs collection [87]. For all core databases, only the data of untreated embryonic cell lines were included. The Roadmap epigenomics collection was subsetted for tissues corresponding to the respective comparisons. For the inter-regional comparison in the stomach, histone marks determined in the fetal stomach, stomach mucosa, gastric and stomach smooth muscle are included. For the comparison between tissues, we additionally included data from the fetal small and large intestine, small intestine, sigmoid colon, colonic mucosa, duodenum mucosa and esophagus. The function fisher.test within the runLOLA function was modified in order to allow a two-sided Fisher's exact test. We determined the high methylation cluster in samples of IM in vivo [13] by performing hierarchical clustering of the 108 IM samples and selecting the 39 antral samples with the high methylation cluster. Extra-embryonic ectoderm (ExE)-hypermethylated CpG islands [55] were orthologously mapped to the human hg19 genome using the UCSC Genome Browser liftOver tool [86]. Bed files of the genomic regions were mapped to the closest human CpG island, and the mean methylation/CpG island was determined. The mucosoid, the in vivo IM samples and the BE data set included 483 of the 489 mouse CpG islands, whereas in the ex vivo IM data set 487 CpG islands were represented. Significance between diseased and normal samples was determined by a Wilcoxon rank-sum test with continuity correction between the mean methylation of all ExE-hyper-CpG islands in diseased data sets compared to all normal means. Healthy samples from the Roadmap genome collection that were included in the Smith publication were included as well. The distribution of average methylation of all other CpG island promoters, which were interrogated on the methylation arrays, and of the ExE-hyper-CpG islands was assigned as a control. Hypermethylated sites of *MLH1* in cancer were taken from [57], using also the UCSC genome browser function lift over tool to visualize the hg18 annotated sites in hg19 with custom tracks of DM CpGs ex vivo [86].

Gene expression

Background correction was performed using the normexp method [88], and inter-array normalization was performed with the quantile method of [89]. Differential gene expression was determined using limma by applying a moderated paired t-test. Gene set enrichments included gene sets with sizes between 15 and 2000 genes; 5000 permutations were applied. The beta-catenin gene set was taken from [90]. Gene set enrichment analysis was performed using the fgsea package and gene sets of hallmark, pathway, motif, GO_BP, oncogenic and immunologic of the Molecular Signature DataBase [91]. RNA-Seq-based gene expression values (RSEM FPKM) from normal stomach samples of the GTEX project and stomach adenocarcinoma samples of the TCGA project were obtained from the TOIL project website (https://xenabrowser.net/datapages/?cohort=TCGA%20TARGET%20GTEx) and used to identify gene expression changes in DM genes in GC [92].

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13148-022-01406-4.

Additional file 1: Figure S1. **A)** Venn diagram of differentially methylated (DM) CpGs by regional comparisons of the stomach. **B)** Enrichment of sets of CpGs hypermethylated (left) or hypomethylated (right) between antrum and cardia biopsies compared to differential methylation of CpGs between antrum and fundus mucosoids. Moderated t-scores were used for DM CpG ranking. ES—enrichment score; NES—normalized enrichment score. Hyper, Hypo indicate the direction of differential methylation in the antrum vs. fundus mucosoids comparison. **C)** Volcano plots representing comparisons between differentiated (− W/R) vs. undifferentiated (+ W/R) states in each stomach region with a less stringent threshold. Delta beta is plotted against the negative logarithm of the raw p-values. The dashed horizontal and vertical lines indicate cutoffs of p < 0.001 and delta beta between − 0.05 and 0.05, respectively. Red dots refer to hypermethylated CpGs and blue dots to hypomethylated CpGs. **D)** Venn diagrams showing

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 16 of 19

overlaps of DM CpGs upon differentiation in the antrum, corpus and fundus. The tables provide further details on the overlapping DM CpGs, including stem cell-specific genes (*TNFRSF19*=Troy).

Additional file 2: Figure S2. A) Table of the numbers of hypermethylated and hypomethylated genes and promoters, comparing regions within the stomach and the stomach to adjacent tissues. Bottom: Venn diagrams showing overlaps of DM genes and promoters between the stomach regions. **B)** Differential gene expression of DM genes between the stomach regions. Shown are genes with log2 fold changes > 0.6 or < − 0.6, * p < 0.01, ** FDR < 0.05. **C)** DM affects genomic features: comparisons of mucosoid samples from different stomach regions (left) and between the stomach to adjacent tissues (right). The plots show the Pearson residuals of the Chi-square statistic of counts of significant DM CpGs located in in TSS1500, TSS200, 5′UTR, 1stExon, gene body, 3′UTR and IGRs with respect to CpGs islands, shores, shelves and opensea. The dashed line marks the threshold of 95% confidence interval. TSS1500 and TSS200: 200-1500, and 0-200 bases upstream of the transcription start site (TSS); 5′UTR—5′ untranslated region; Body—gene body; 1stExon—first exon; 3′UTR—3′ untranslated region; IGR—intergenomic region.

Additional file 3: Figure S3. Differential expression of DM TFs between the antrum and the corpus, based on several *in vitro* and *in vivo* data sets (Additional table 7). The expression level is relative to the antrum mean; the p-values of a student t-test between antrum and corpus are shown on each plot, except for cases with less than two samples in each group. For each study, all microarray probes/probe sets mapping to the respective gene are shown.

Additional file 4: Figure S4. A) Gene set enrichment analysis of genes overexpressed in antrum (left) or corpus (right) biopsies compared to differential gene expression between antrum and corpus mucosoids. +W/R and −W/R samples of each region were combined for the mucosoid analysis. Moderated t-scores were used for differential expression ranking. Gene sets of biopsies were determined from (Nookaew et al., 2013, Additional table 7). ES—enrichment score; NES—normalized enrichment score. **B)** Boxplots of biopsy (red) and mucosoid (green) log2 gene expression fold changes, comparing the corpus and the antrum of selected cell-type-specific genes. The dashed line marks the threshold of log2 fold change > 0.6 or < − 0.6. Biopsy samples were taken from Nookaew et al., 2013 (Additional table 7). **C)** Venn diagrams showing the overlaps of significant upregulated (top) and downregulated (bottom) genes, comparing -W/R (differentiated) to +W/R (undifferentiated) in the different stomach regions (FDR < 0.2; log2 fold change > 0.6 or < − 0.6). Numbers in brackets indicate the number of genes belonging to the GO Wnt signaling pathway gene set, and the corresponding genes are listed accordingly. **D)** Normalized expression boxplots of Wnt signaling pathway genes. The boxplots display the median, minimum and maximum normalized expression values of three biological replicates. **E)** Heatmap of *TCF7L1* DNA methylation. Displayed are the normalized methylation values (beta) ranging from 0 (blue, unmethylated) to 1 (red, methylated). Significantly DM CpGs between the antrum and the fundus are marked with red arrows. The black arrow refers to the transcriptional direction. CGI—CpG island, TSS—transcription start site, 5′UTR—5′ untranslated region, IGR—intergenomic region.

Additional file 5: Figure S5. A) Gene expression log2 fold changes of the stomach- and intestine-specific genes comparing IM and normal organoids (FDR < 0.05). The dashed line indicates a cutoff of log2 fold change > 0.6 or < − 0.6. **B)** DNA methylation MDS plot of healthy mucosoids from the antrum, corpus and fundus, or normal, atrophic and intestinal metaplasia (IM) organoids. **C)** Volcano plot of IM vs. normal organoids. Changes in methylation (delta beta) are plotted against the negative logarithm of p-values. The red horizontal and the dashed vertical lines indicate cutoffs of p < 0.05 and delta beta between -0.2 and 0.2, respectively. **D)** Enrichment of sets of CpGs hypermethylated (left) or hypomethylated (right) between IM biopsies and normal biopsies compared to differentially methylation of CpGs between IM and normal organoids. Moderated t-scores were used for ranking the CpGs. Hyper, Hypo indicate the direction of differential methylation in the IM vs. normal organoid comparison. **E)** Number of DM CpGs, genes, promoters and enhancers determined in the comparison of IM samples to healthy samples. Left:

organoid and biopsy antral IM samples, right: the corresponding samples of the high methylation clusters. **F)** Dot plots of methylation levels (beta values) of the genes *TRIM15* and CDX2: healthy organoid and biopsy samples (top) and IM organoid and biopsy samples (bottom). Squares and crosses denote organoid and biopsy data, respectively. Colors below the plot show DM calls of IM vs. healthy in both data sets.

Additional file 6: Figure S6. A) Comparison of organoid IM samples to normal organoid samples—enrichment or depletion of hypermethylated and hypomethylated CpGs in various genomic features. Genomic features were defined based on Encode segmentation classes and transcription factor binding sites (Siggens and Ekwall, 2014), and Roadmap histone modification in gastric tissues (Roadmap Epigenomics Consortium et al., 2015). Shown are log odds ratios of significantly (FDR < 0.05) enriched (log odds ratio > 0.6) or depleted (log odds ratio < − 0.6) DM CpGs. The color code ranges from blue (depleted) to red (enriched), gray refers to a log odds ratio < 0.6 or > − 0.6. Pol2—Polymerase 2 subunit; Hyper—hypermethylated CpGs; Hypo—hypomethylated CpGs. **B)** Boxplots of ExE CpG island (CGI) hypermethylation. Shown are the mean ExE CGI hypermethylation values per sample. Healthy samples include, in addition to those of the respective data sets, Roadmap epigenome healthy samples (n = 77), except for the extra-embryonic trophoblast cell line and placenta samples. Diseased samples include the respective mild IM and IM (biopsies), IM (organoids), and BE and GC (both biopsies). P values displayed refer to the Wilcoxon rank-sum test with continuity correction. **C)** Histograms of the methylation distribution in ExE hypermethylated CGI and all other CGI taken from the healthy mucosoid data set of the stomach regions and the data set of normal and IM samples from biopsies. The red line and number indicate the median methylation value. **D)** UCSC Genome Browser Image of the *MLH1* promoter region with custom tracks showing the hypermethylation sites in samples of *ex vivo* IM and MSI type of GC (Kent et al., 2002). The sites responsible for *MLH1* silencing in cancer (Morak et al., 2008) are marked in red. Additional tracks show UCSC CpG islands (green) and the UCSC layered H3K27Ac marks.

Additional file 7: Differentially methlyated CpGs between stomach regions and culture conditions.

Additional file 8: Differentially methylated genes between stomach regions.

Additional file 9: Gene set enrichment analysis of genes differentially methylated between stomach regions.

Additional file 10: Validation of differentially methylated CpGs between stomach regions in published data sets.

Additional file 11: Gene set enrichment analysis of genes differentially expressed between stomach regions.

Additional file 12: Differential gene expression after treatment of healthy antral primary epithelial cells with 5aza.

Additional file 13: DNA Methylation and gene expression data sets used in this manuscript.

Author contributions

KF, FB, HB and TFM conceived the project and designed experiments, KF performed air liquid culture experiments of stomach samples, KF performed methylation analysis, PS and MK performed organoid cultures of ex vivo stomach samples, CD, AL and PM selected and provided gastric samples, and KF and HB performed the bioinformatic analysis. IG generated and provided RNA-Seq data for ex vivo samples. KF, HB and TFM wrote the manuscript. All authors read and approved the final manuscript.

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 17 of 19

Declarations

Ethical approval and consent to participate
The Charité University Medicine ethics committee approved the scientific usage of gastric sleeve resection material (EA1/129/12) and pseudonymized samples from patients. The Otto-von-Guericke-University of Magdeburg ethics committee approved the scientific usage of human gastric biopsies (EA 80/11).

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Molecular Biology, Max Planck Institute for Infection Biology, Charitéplatz 1, 10117 Berlin, Germany. [2]Ludwig Institute for Cancer Research, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK. [3]Center for Bariatric and Metabolic Surgery, Center of Innovative Surgery (ZIC), Department of Surgery, Campus Virchow Klinikum and Campus Mitte, Charité-Universitätsmedizin Berlin, Berlin, Germany. [4]Department of Gastroenterology, Hepatology and Infectious Diseases, Otto-Von-Guericke University Hospital, Magdeburg, Germany. [5]Centro Nacional de Análisis Genómico (CNAG-CRG), Barcelona, Spain. [6]Laboratory of Infection Oncology, Institute of Clinical Molecular Biology, Christian Albrecht University of Kiel and University Hospital Schleswig-Holstein – Campus Kiel, Rosalind-Franklin-Straße 12, 24105 Kiel, Germany.

References
1. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol. 2019;20(10):590–607.
2. Monk D. Germline-derived DNA methylation and early embryo epigenetic reprogramming: the selected survival of imprints. Int J Biochem Cell Biol. 2015;67:128–38.
3. Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A, Van Neste L, et al. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell. 2008;2(2):160–9.
4. Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z, Reik W, et al. Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nat Cell Biol. 2008;10(11):1280–90.
5. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16(1):6–21.
6. Bock C, Beerman I, Lien WH, Smith ZD, Gu H, Boyle P, et al. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. Mol Cell. 2012;47(4):633–47.
7. Beerman I, Bock C, Garrison BS, Smith ZD, Gu H, Meissner A, et al. Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. Cell Stem Cell. 2013;12(4):413–25.
8. Marttila S, Kananen L, Häyrynen S, Jylhävä J, Nevalainen T, Hervonen A, et al. Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. BMC Genomics. 2015;14(16):179.
9. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet. 2007;39(2):232–6.
10. Zhou W, Dinh HQ, Ramjan Z, Weisenberger DJ, Nicolet CM, Shen H, et al. DNA methylation loss in late-replicating domains is linked to mitotic cell division. Nat Genet. 2018;50(4):591–602.
11. Hoffmann W. Current status on stem cells and cancers of the gastric epithelium. Int J Mol Sci. 2015;16(8):19153–69.
12. Kim TH, Shivdasani RA. Stomach development, stem cells and disease. Development. 2016;143(4):554–65.
13. Huang KK, Ramnarayanan K, Zhu F, Srivastava S, Xu C, Tan ALK, et al. Genomic and epigenomic profiling of high-risk intestinal metaplasia reveals molecular determinants of progression to gastric cancer. Cancer Cell. 2018;33(1):137–50.
14. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202–9.
15. Yoda Y, Takeshima H, Niwa T, Kim JG, Ando T, Kushima R, et al. Integrated analysis of cancer-related pathways affected by genetic and epigenetic alterations in gastric cancer. Gastric Cancer. 2015;18(1):65–76.
16. Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, et al. *Helicobacter pylori* infection and the development of gastric cancer. N Engl J Med. 2001;345(11):784–9.
17. Huang FY, Chan AOO, Rashid A, Wong DKH, Cho CH, Yuen MF. *Helicobacter pylori* induces promoter methylation of E-cadherin via interleukin-1β activation of nitric oxide production in gastric cancer cells. Cancer. 2012;118(20):4969–80.
18. Niwa T, Tsukamoto T, Toyoda T, Mori A, Tanaka H, Maekita T, et al. Inflammatory processes triggered by *Helicobacter pylori* infection cause aberrant DNA methylation in gastric epithelial cells. Cancer Res. 2010;70(4):1430–40.
19. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011;98(4):288–95.
20. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics. 2016;8(3):389–99.
21. Hughes LAE, Melotte V, de Schrijver J, de Maat M, Smit VTHBM, Bovée JVMG, et al. The CpG island methylator phenotype: What's in a name? Cancer Res. 2013;73(19):5858–68.
22. Boccellato F, Woelffling S, Imai-Matsushima A, Sanchez G, Goosmann C, Schmid M, et al. Polarised epithelial monolayers of the gastric mucosa reveal insights into mucosal homeostasis and defence against infection. Gut. 2018;68:400–13.
23. Nanki K, Toshimitsu K, Takano A, Fujii M, Shimokawa M, Ohta Y, et al. Divergent routes toward Wnt and R-spondin niche independency during human gastric carcinogenesis. Cell. 2018;174(4):856–69.
24. Schlaermann P, Toelle B, Berger H, Schmidt SC, Glanemann M, Ordemann J, et al. A novel human gastric primary cell culture system for modelling *Helicobacter pylori* infection in vitro. Gut. 2016;65(2):202–13.
25. Sheaffer KL, Kim R, Aoki R, Elliott EN, Schug J, Burger L, et al. DNA methylation is required for the control of stem cell differentiation in the small intestine. Genes Dev. 2014;28(6):652–64.
26. Holland AM, Garcia S, Naselli G, Macdonald RJ, Harrison LC. The Parahox gene Pdx1 is required to maintain positional identity in the adult foregut. Int J Dev Biol. 2013;57(5):391–8.
27. Offield MF, Jetton TL, Labosky PA, Ray M, Stein RW, Magnuson MA, et al. PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development. 1996;122(3):983–95.
28. Stoffers DA, Heller RS, Miller CP, Habener JF. Developmental expression of the homeodomain protein IDX-1 in mice transgenic for an IDX-1 promoter/lacZ transcriptional reporter. Endocrinology. 1999;140(11):5374–81.
29. McCracken KW, Aihara E, Martin B, Crawford CM, Broda T, Treguier J, et al. Wnt/β-catenin promotes gastric fundus specification in mice and humans. Nature. 2017;541(7636):182–7.
30. Schulte D, Geerts D. MEIS transcription factors in development and disease. Development. 2019;146(16):dev174706.
31. Stringer EJ, Duluc I, Saandi T, Davidson I, Bialecka M, Sato T, et al. Cdx2 determines the fate of postnatal intestinal endoderm. Development. 2012;139(3):465–74.

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 18 of 19

32. Laugesen A, Højfeldt JW, Helin K. Molecular mechanisms directing PRC2 recruitment and H3K27 methylation. Mol Cell. 2019;74(1):8–18.

33. Batool A, Jin C, Liu YX. Role of EZH2 in cell lineage determination and relative signaling pathways. Front Biosci (Landmark Ed). 2019;01(24):947–60.

34. Rajasekhar VK, Begemann M. Concise review: roles of polycomb group proteins in development and disease: a stem cell perspective. Stem Cells. 2007;25(10):2498–510.

35. Sheffield NC, Bock C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. Bioinformatics. 2016;32(4):587–9.

36. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317–30.

37. Siggens L, Ekwall K. Epigenetics, chromatin and genome organization: recent advances from the ENCODE project. J Intern Med. 2014;276(3):201–14.

38. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007;39(3):311–8.

39. Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, Holota H, et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. EMBO J. 2011;30(20):4198–210.

40. Fortin JP, Hansen KD. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. Genome Biol. 2015;28(16):180.

41. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326(5950):289–93.

42. Arnold K, Sarkar A, Yram MA, Polo JM, Bronson R, Sengupta S, et al. Sox2(+) adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell Stem Cell. 2011;9(4):317–29.

43. Barker N, Huch M, Kujala P, van de Wetering M, Snippert HJ, van Es JH, et al. Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell. 2010;6(1):25–36.

44. Matsuo J, Kimura S, Yamamura A, Koh CP, Hossain MZ, Heng DL, et al. Identification of stem cells in the epithelium of the stomach corpus and antrum of mice. Gastroenterology. 2017;152(1):218-231.e14.

45. Stange DE, Koo BK, Huch M, Sibbel G, Basak O, Lyubimova A, et al. Differentiated Troy+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium. Cell. 2013;155(2):357–68.

46. Lizio M, Harshbarger J, Shimoji H, Severin J, Kasukawa T, Sahin S, et al. Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol. 2015;5(16):22.

47. Oishi I, Suzuki H, Onishi N, Takada R, Kani S, Ohkawara B, et al. The receptor tyrosine kinase Ror2 is involved in non-canonical Wnt5a/JNK signalling pathway. Genes Cells. 2003;8(7):645–54.

48. Saldanha J, Singh J, Mahadevan D. Identification of a frizzled-like cysteine rich domain in the extracellular region of developmental receptor tyrosine kinases. Protein Sci. 1998;7(7):1632–5.

49. Barnicle A, Seoighe C, Golden A, Greally JM, Egan LJ. Differential DNA methylation patterns of homeobox genes in proximal and distal colon epithelial cells. Physiol Genomics. 2016;48(4):257–73.

50. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.

51. Barros R, Freund JN, David L, Almeida R. Gastric intestinal metaplasia revisited: function and regulation of CDX2. Trends Mol Med. 2012;18(9):555–63.

52. Boyd M, Hansen M, Jensen TGK, Perearnau A, Olsen AK, Bram LL, et al. Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). J Biol Chem. 2010;285(33):25115–25.

53. Pereira B, Oliveira C, David L, Almeida R. CDX2 promoter methylation is not associated with mRNA expression. Int J Cancer. 2009;125(7):1739–42.

54. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011;470(7333):279–83.

55. Smith ZD, Shi J, Gu H, Donaghey J, Clement K, Cacchiarelli D, et al. Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. Nature. 2017;549(7673):543–7.

56. Krause L, Nones K, Loffler KA, Nancarrow D, Oey H, Tang YH, et al. Identification of the CIMP-like subtype and aberrant methylation of members of the chromosomal segregation and spindle assembly pathways in esophageal adenocarcinoma. Carcinogenesis. 2016;37(4):356–65.

57. Morak M, Schackert HK, Rahner N, Betz B, Ebert M, Walldorf C, et al. Further evidence for heritability of an epimutation in one of 12 cases with MLH1 promoter methylation in blood cells clinically displaying HNPCC. Eur J Hum Genet. 2008;16(7):804–11.

58. McCracken KW, Catá EM, Crawford CM, Sinagoga KL, Schumacher M, Rockich BE, et al. Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. Nature. 2014;516(7531):400–4.

59. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet. 2009;41(2):178–86.

60. Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LTY, Kohlbacher O, et al. Charting a dynamic DNA methylation landscape of the human genome. Nature. 2013;500(7463):477–81.

61. Kraiczy J, Nayak KM, Howell KJ, Ross A, Forbester J, Salvestrini C, et al. DNA methylation defines regional identity of human intestinal epithelial organoids and undergoes dynamic changes during development. Gut. 2017;68:49–61.

62. Chen KJ, Lizaso A, Lee YH. SIM2 maintains innate host defense of the small intestine. Am J Physiol Gastrointest Liver Physiol. 2014;307(11):G1044-1056.

63. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M, et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature. 2007;449(7165):1003–7.

64. Fischer AS, Sigal M. The role of Wnt and R-spondin in the stomach during health and disease. Biomedicines. 2019;7(2):44.

65. Barker N, Huch M, Kujala P, van de Wetering M, Snippert HJ, van Es JH, et al. Lgr5+ve stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell. 2010;6(1):25–36.

66. Miyoshi H, Ajima R, Luo CT, Yamaguchi TP, Stappenbeck TS. Wnt5a potentiates TGF-β signaling to promote colonic crypt regeneration after tissue injury. Science. 2012;338(6103):108–13.

67. Hayakawa Y, Ariyama H, Stancikova J, Sakitani K, Asfaha S, Renz BW, et al. Mist1 expressing gastric stem cells maintain the normal and neoplastic gastric epithelium and are supported by a perivascular stem cell niche. Cancer Cell. 2015;28(6):800–14.

68. Katoh M, Katoh M. FGF signaling network in the gastrointestinal tract (review). Int J Oncol. 2006;29(1):163–8.

69. Barclay C, Li AW, Geldenhuys L, Baguma-Nibasheka M, Porter GA, Veugelers PJ, et al. Basic fibroblast growth factor (FGF-2) overexpression is a risk factor for esophageal cancer recurrence and reduced survival, which is ameliorated by coexpression of the FGF-2 antisense gene. Clin Cancer Res. 2005;11(21):7683–91.

70. Kresty LA, Clarke J, Ezell K, Exum A, Howell AB, Guettouche T. MicroRNA alterations in Barrett's esophagus, esophageal adenocarcinoma, and esophageal adenocarcinoma cell lines following cranberry extract treatment: Insights for chemoprevention. J Carcinog. 2011;10:34.

71. van der Flier LG, van Gijn ME, Hatzis P, Kujala P, Haegebarth A, Stange DE, et al. Transcription factor achaete scute-like 2 controls intestinal stem cell fate. Cell. 2009;136(5):903–12.

72. Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schübeler D. Competition between DNA methylation and transcription factors determines binding of NRF1. Nature. 2015;528(7583):575–9.

73. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature. 2011;480(7378):490–5.

74. Wölffling S, Daddi AA, Imai-Matsushima A, Fritsche K, Goosmann C, Traulsen J, et al. EGF and BMPs govern differentiation and patterning in human gastric glands. Gastroenterology. 2021;161:623–36.

75. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2017. https://www.R-project.org/.

76. GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund,

Fritsche *et al. Clinical Epigenetics*   (2022) 14:193

Page 19 of 19

NIH/NCI, et al. Genetic effects on gene expression across human tissues. Nature. 2017;550(7675):204–13.

77. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. ChAMP: 450k chip analysis methylation pipeline. Bioinformatics. 2014;30(3):428–30.

78. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics. 2017;33(24):3982–4.

79. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363–9.

80. Fortin JP, Triche TJ, Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinformatics. 2017;33(4):558–60.

81. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol. 2014;15(12):503.

82. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.

83. Ramsey VG, Doherty JM, Chen CC, Stappenbeck TS, Konieczny SF, Mills JC. The maturation of mucus-secreting gastric epithelial progenitors into digestive-enzyme secreting zymogenic cells requires Mist1. Development. 2007;134(1):211–22.

84. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. Nat Rev Genet. 2009;10(4):252–63.

85. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43(Database issue):D447–52.

86. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002;12(6):996–1006.

87. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. 2018;46(D1):D1284.

88. Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A, et al. A comparison of background correction methods for two-colour microarrays. Bioinformatics. 2007;23(20):2700–7.

89. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003;19(2):185–93.

90. Herbst A, Jurinovic V, Krebs S, Thieme SE, Blum H, Göke B, et al. Comprehensive analysis of β-catenin target genes in colorectal carcinoma cell lines with deregulated Wnt/β-catenin signaling. BMC Genomics. 2014;28(15):74.

91. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545–50.

92. Vivian J, Rao AA, Nothaft FA, Ketchum C, Armstrong J, Novak A, et al. Toil enables reproducible, open source, big biomedical data analyses. Nat Biotechnol. 2017;35(4):314–6.