nature aging

Supplementary information

https://doi.org/10.1038/s43587-025-01021-x

Iron homeostasis and cell clonality drive cancer-associated intestinal DNA methylation drift in aging

In the format provided by the authors and unedited

Supplementary Discussion

We found that the DNA hypermethylation is driven by an impaired iron homeostasis that affects the TET protein enzymatic activity. Epigenetic integration of changes in cellular metabolism is a common feature of aging human tissues [1,2]. Previous studies linked the *TET1* gene and colon cancer and demonstrated that *TET1* downregulation specifically synergizes with *BRAF* mutation in driving CIMP-mediated carcinogenesis [3,4,5]. We found that the TET2/3 enzymatic activity is affected in epigenetically drifted cells, suggesting that while TET1 may be primarily involved in CIMP+ tumors, our mechanism represents the ground stage for all the CRCs. Remarkably, the *TET* genes are not frequently found mutated nor transcriptionally regulated in CRCs, reinforcing the idea that cancer-associated DNAm alterations may be driven by their impaired enzymatic functionality. Iron metabolism is finely modulated by several both systemic and cellular processes, and it has been recently found dysregulated in many aging tissues [6]. Several factors could be the drivers of this dysregulated iron homeostasis in intestinal epithelial cells, for example, dietary and gut microbiome aging-associated changes, alterations of the cellular metabolism, impaired proteostasis, and inflammation [7]. Our data demonstrated that the Wnt pathway, key signaling to maintain the stemness and found impaired in the intestine during aging [8,9,10], can regulate the intracellular the iron level, the TET activity, and the ACCA drift, by inducing the expression of the transferrin receptor *Tfrc*.

Since a pro-inflammatory environment (especially interferon signaling) has been recently demonstrated to be a main cause of intestinal aging [11,12], it is worth thinking that inflammation can be the source of the observed aging-associated loss of intestinal iron homeostasis. Furthermore, inflammation has been shown to upregulate Lipocalin 2 (LCN2) in the gut, thus restricting iron availability by chelating siderophore-iron [13]. We showed that the interferon gamma treatment increased the protein level of the iron exporter FPN1, reducing the intracellular iron level as previously shown in immune cells [14,15], and the TET activity in intestinal organoids. Our data suggest that the preservation of functional and undifferentiated ISCs may be the key point in the maintenance of iron homeostasis. Indeed, intestinal cells upregulate ferroportin and downregulate the transferrin receptor during differentiation as a consequence of the diverse iron metabolism acting in ISCs or enterocytes [16]. It is tempting to speculate that a lack of proper Wnt signaling and/or a pro-inflammatory environment leads to a loss of stemness, and the ISCs start to modulate iron-related genes that can favor the ACCA drift initiation.

Supplementary Methods

QUANTIFICATION AND STATISTICAL ANALYSES

Bisulfite sequencing data analysis

Fastq files quality check was performed with FastQC v0.11.5. Only for WGBS, sequencing reads were adapter clipped with fastx_clipper by using the following parameters: -a GATCGGAAGAGCACACGTCTGAACTCCAGTCAC -Q33 -l 20 -n -v -M 10. Clipped reads were mapped to the mm9 genome using BSMAP v2.89. For RRBS, the SAM files were directly used for calculation of CpG methylation. The CpG methylation was calculated using methratio.py script in the BSMAP package with -z -g -x CG parameters (and --remove-duplicate only for WGBS). The methratio.py results were directly imported to R and all other downstream analyses were done using R v3.4.4. All the differentially methylated analyses

(single CpG, CpG islands and promoter) were done using methylKit v1.4.1 R package. To increase the coverage, replicates were merged before calculating statistics. For single CpGs, we used those that have at least 10 times coverage in both groups (for statistical analysis after merging the replicates), or in all the replicates (for PCA analysis). Definition of promoters was based on the Ensemble mm9 annotation with ±1kb from the transcription start site. Only promoters that are covered at least 50 times by minimum 5 CpG sites were used for further analysis. For CpG islands (reference from the UCSC Genome Browser), we considered just those that have more than 100 times coverage by minimum 10 CpGs. Differentially methylated single CpGs, CpG islands and promoters were defined based on adjusted p-value (q-value) and methylation difference (q-value<0.05, absolute number of Δ methylation >= 10). For defining differentially methylated genes, we merged the data of promoters and CpG islands (the latter less than ±20kb to the transcription start site), with a priority for the promoters (if the promoter was not differentially regulated then we looked for the differentially methylated CpG islands). We assigned only one TSS to each CpG island (the nearest TSS). The GO analysis was done using IPA software.

Analysis of TCGA data on human colon samples

The DNA methylation data (beta values) for Colorectal cancer tumors and healthy tissue for adult donors was downloaded from TCGA data base (The Cancer Genome Atlas). Only the data created using Illumina 450K methylation array was used. All the tumor samples are from stages 0-1 (https://www.cancer.net/cancer-types/colorectal-cancer/stages) of colorectal cancer development. Supplementary Table S6 shows the clinical data for the collected samples from the TCGA database. The Illumina 450K methylation array data for the healthy samples from young patients (<15 years old) were downloaded from the previous study [75]. The raw methylation array data (.idat) was normalized using champ.norm function from ChAMP R package (v2.32.0) [76] using default parameters. To calculate the methylation on promoters, the average methylation (beta value) of single CpGs at less than 1kb distance from the transcription start site of each gene was calculated (hg19 genome version). Only promoters with at least 3 covered single CpGs were considered for the downstream analysis. For principal component analysis (PCA) the promoters with an interquartile range (IQR) of at least 0.1 (10% difference in beta value) were used. PCA was calculated using prcomp R function with center = TRUE,scale. = TRUE parameters. To calculate the differentially methylated promoters, t-test were done comparing the age groups <15 yo with >60 yo from the normal samples for aging comparison, and >60 yo healthy with >60 yo Tumor for cancer comparison using t.test function in R with alternative = "two.sided", paired = FALSE, var.equal = FALSE parameters. Adjusted p-value (correction for multiple comparisons) is calculated using p.adjust function in R with method=”holm”. The promoters with adjusted p-value <0.05 and difference in methylation >10% (comparing normal tissue young vs old, or normal vs tumor) were considered as differentially methylated promoters.
The DNA methylation data (beta values) for Colon Adenocarcinoma and the other nine cancer types (Esophageal Carcinoma (ESCA), Liver Hepatocellular Carcinoma (LIHC), Head and Neck Squamous Cell Carcinoma (HNSC), Bladder Urothelial Carcinoma (BLCA), Kidney Renal Clear Cell Carcinoma (KIRC), Breast Invasive Carcinoma (BRCA), Lung Adenocarcinoma (LUAD), Glioblastoma Multiforme (GBM), Skin Cutaneous Melanoma (SKCM)) were downloaded from TCGA database (The Cancer Genome Atlas) using the TCGAbiolinks R package. Only the data created using Illumina 450K methylation array was selected. Promoters were defined as the ±1kb region surrounding the transcription start site. Only promoters with at least 3 covered single CpGs were considered for the downstream analysis, and the data were normalized on promoters of two housekeeping genes (ACTB and

GAPDH). For cancer samples n=50, for healthy samples n=2-50 according to the availability of the healthy samples on the TCGA.
To assess the relationship between age and cancer-associated DNA methylation (DNAm) drift and specific Wnt signaling antagonists, we computed Pearson correlation coefficients in R (v4.1.2) between the mean DNAm levels of ACCA drift genes and the mean DNAm levels of either two (*DKK2* and *SFRP1*) or four (*DKK1/2*, *SFRP1/2*) selected Wnt antagonist genes. The DNAm values used were derived from normalized beta values across the promoter regions of each gene.
To evaluate the statistical significance of the observed correlations, we performed a permutation test using custom R scripts. Specifically, for each test (2-gene and 4-gene sets), 100 random gene sets of matching size were sampled from the pool of all genes identified in the experiment. Correlation coefficients were computed between each random gene set and the mean DNAm of ACCA drift genes, generating a null distribution against which the observed correlation was compared.
For functional and pathway analysis, IPA (Ingenuity Pathway Analysis v45868156) software has been used with differentially methylated promoters. For the co-expression analysis, for the donors that both expression and methylation data were provided by TCGA database, the gene expression quantification was downloaded and fpkm_uq_unstranded data were used for further analysis. For the hypermethylated genes in both aging and cancer, the correlation between methylation of each promoter and expression of all genes was calculated using cor.test R function with alternative ="two.sided", method ="pearson" parameters. The genes with a correlation p-value < 0.05 and a correlation coefficient >0.5 or <-0.5 are considered as significantly correlated genes (SCG). The SCGs were uploaded in IPA with the correlation coefficient as the direction of the expression (for activity z-score calculation). P-values are calculated by IPA by using Fisher's Exact Test with Benjamini-Hochberg correction.

RNAseq data analysis

Quality assessment of Fastq files was performed using FastQC v0.11.5. Subsequently, alignment to the mm9 genome was performed using TopHat v2.1.0 employing the parameters --bowtie1 --no-coverage-search -a 5. HTSeq-Count 0.11.2 was then employed to determine the read coverage for each gene, using the parameters -s no -a 0 -t exon -m intersection-nonempty. Following the removal of rRNA genes from the count data, DESeq2 R package v1.20.0 with default settings was utilized for differential expression analysis and normalization. To identify differentially expressed genes, normalized expression counts were compared between colorectal cancer patients and healthy controls using the t.test function in R. Log_2 fold change (log_2FC) was computed for each gene and multiple testing correction was applied using the Benjamini–Hochberg method (p.adjust function in R, method = "BH"). Genes with a false discovery rate (FDR) < 0.05 were considered significantly differentially expressed; genes with positive or negative log_2FC were classified as up- or downregulated, respectively. For the PCA plot, normalized counts were filtered based on the criteria: count>10 in at least 3 samples, and an IQR of the log-transformed count >= 1.5. Post filtration, log10 transformed counts were employed in the prcomp R function with parameters center = TRUE and scale = TRUE, with PC1 and PC2 used for plotting. For functional and pathway analysis as well as for identification of potential upstream regulators, DESeq2 differentially expression analysis (adjusted p-value<0.05) results were uploaded in IPA (Ingenuity Pathway Analysis v45868156). Only the highly predicted or experimentally validated predicted upstream regulators mRNAs were used for downstream analysis. For gene set enrichment analysis, normalized counts (for each ge3ne in all the samples) were scaled using scale function in R (with center = TRUE, scale = TRUE

parameters). The Z-score was calculated by multiplying the scaled counts by +1 or -1 that shows the direction of the expression in aging (+1 for up-regulated genes and -1 for down-regulated genes). The average of Z-scores were calculated for each group and used for drawing plot and downstream analysis. The Delta Z-score has been calculated in this way: average of the Zscores of old mice (mean of all the 8 old replicates) minus the average of the Zscores of the young mice (mean of the 4 young replicates).
For single-crypts RNA-seq only protein coding genes were used for downstream analysis. Principal component analysis (PCA) was performed on count data after filtering to retain genes with counts greater than 10 in at least 50% of samples and with an interquartile range (IQR) greater than 2 on log-transformed counts incorporating a pseudocount. The selected genes were centered and scaled prior to PCA computation using prcomp in R. Visualization was performed with the autoplot function, which projected samples onto the first two principal components and colored them according to their group annotation (Low 5mC vs High 5mC). Convex hulls were drawn around groups and no additional clustering method was applied. and Gene Ontology analysis for Biological Processes was performed by using DAVID online tool.

Single-crypt DNAm analysis

For Fig. 3d (DNAm correlation heatmap), Pearson correlation was performed using cor R function (method = "pearson") and the distance between samples (correlation coefficient values) was calculated using dist R function (method = "euclidean") for hierarchical clustering. For Fig. 3f, we calculated the median value of DNAm in all the crypts for each of the 4 genes analyzed and we defined each gene for each crypt as methylated when the value was higher than the median. We considered only the crypts having three genes analyzed. We counted the number of crypts having 0, 1, 2 or 3 genes methylated (number of events) and we subtracted to this number the number of expected crypts having that combination of methylated genes.

Supplementary References

1. Ermolaeva, M., Neri, F., Ori, A. & Rudolph, K. L. Cellular and epigenetic drivers of stem cell ageing. *Nature reviews. Molecular cell biology* 19, 594–610 (2018).

2. Brunet, A. & Rando, T. A. Interaction between epigenetic and metabolism in aging stem cells. *Current Opinion in Cell Biology* 45, 1–7 (2017).

3. Neri, F. *et al.* TET1 is a tumour suppressor that inhibits colon cancer growth by derepressing inhibitors of the WNT pathway. *Oncogene* 34, 4168–4176 (2015).

4. Tricarico, R. *et al.* TET1and TDGsuppress intestinal tumorigenesis by down-regulating the inflammatory and immune response pathways. *bioRxiv* 62, 4075 (2019).

5. Noreen, F. *et al.* DNA methylation instability by BRAF-mediated TET silencing and lifestyle-exposure divides colon cancer pathways. *Clin Epigenetics* 11, 196–17 (2019).

6. Zeidan, R. S., Han, S. M., Leeuwenburgh, C. & Xiao, R. Iron homeostasis and organismal aging. *Ageing Res. Rev.* 72, 101510 (2021).

7. Anderson, G. J. & Frazer, D. M. Current understanding of iron homeostasis. *Am. J. Clin. Nutr.* 106, 1559S-1566S (2017).

8. Nalapareddy, K. *et al.* Canonical Wnt Signaling Ameliorates Aging of Intestinal Stem Cells. *Cell reports* 18, 2608–2621 (2017).

9. Pentinmikko, N. *et al.* Notum produced by Paneth cells attenuates regeneration of aged intestinal epithelium. *Nature* 571, 398–402 (2019).

10. Annunziata, F. *et al.* Paneth Cells drive Intestinal Stem Cell Competition and Clonality in Aging and Calorie Restriction. *Eur J Cell Biol* 151282 (2022) doi:10.1016/j.ejcb.2022.151282.

11. Omrani, O. *et al.* IFNγ-Stat1 axis drives aging-associated loss of intestinal tissue homeostasis and regeneration. *Nat. Commun.* 14, 6109 (2023).

12. Funk, M. C. *et al.* Aged intestinal stem cells propagate cell-intrinsic sources of inflammaging in mice. *Dev. Cell* 58, 2914-2929.e7 (2023).

13. Moschen, A. R. *et al.* Lipocalin 2 Protects from Inflammation and Tumorigenesis Associated with Gut Microbiota Alterations. *Cell Host Microbe* 19, 455–469 (2016).

14. Abreu, R., Essler, L., Giri, P. & Quinn, F. Interferon-gamma promotes iron export in human macrophages to limit intracellular bacterial replication. *PLoS ONE* 15, e0240949 (2020).

15. Nairz, M. *et al.* Interferon-γ limits the availability of iron for intramacrophage Salmonella typhimurium. *Eur. J. Immunol.* 38, 1923–1936 (2008).

16. Harnik, Y. *et al.* A spatial expression atlas of the adult human proximal small intestine. *Nature* 632, 1101–1109 (2024).