Supplementary Figure Legends

Figure S1. CD associated, stable loss of MHC-I related DNAm in patient derived IEOs.

(A) i) Correlation heat map of 33 co-methylated CpG modules in SC identified by WGCNA and CD diagnosis. The first principal component of Module 9 (ME9) demonstrates hypomethylation and strong associations with CD diagnosis (R = -0.5, p-value < 0.001). ii) Number of CpGs in top modules in TI (module 17), SC (module 9) and shared probes between the two gut segments**. (B)** Differential DNAm analyses comparing CD with control patient derived IEOs from the Terminal Ileum (TI, n = 127) and Sigmoid Colon (SC, n = 131). i) Number of significantly differentially methylated probes (DMPs) in TI, SC and shared probes between the two gut segments. MHC-I related genes containing significant DMPs are highlighted. ii) Venn diagram showing the number of significantly differentially methylated positions (DMPs) comparing patient-derived organoids within TI (left) and SC (right). MHC-I related genes containing at least one significant DMP are highlighted. iii) Gene set enrichment analysis performed on CpGs showing loss of DNAm in CD organoids compared with controls in both TI and SC. CpGs located in genes involved in the MHC-I pathway are highlighted in bold. **(C and D)** DNAm of significant DMPs located within MHC-I related genes **(C)** in TI and **(D)** in SC IEOs. (Cohort 1, n = 127, delta beta > 0.05 and * FDR < 0.05, ** FDR < 0.01, *** FDR < 0.001, ns = not significant). **(E)** Correlation of *NLRC5* promoter DNAm between early and later passage IEOs from the same individuals including patients diagnosed with CD (blue), UC (yellow) and controls (grey). N = 22 patients. **(F)** DNAm (beta values) of selected, representative DMPs in high passage (>7) IEOs (Cohort 1, n = 22). **(G)** Correlation of summary MHC-I DNAm score between early and late passage IEOs (Cohort 1, n = 22 Spearman's Rank Correlation).

Figure S2. Stable loss of MHC-I related DNAm in CD patient derived IEO following exposure to inflammatory cytokines (A) DNAm of IEOs +/- treatment with IFNγ and TNFα. Selected CpGs are located within promoter regions of MHC-I related genes (delta beta > 0.05). **(B)** Gene transcription of MHC-I genes from bulk RNA-seq of the same IEOs used in A (normalised counts, * FDR < 0.05, ns = not significant).

Figure S3. Intestinal epithelial MHC-I DNAm correlates with gene expression. (A) The correlation between gene expression and DNAm is shown for additional MHC-I genes in primary purified TI IEC (Cohort 2, n = 32). i) DNAm of selected CpGs split by diagnosis ii) DNAm correlation with expression of the corresponding gene in the same patient derived IEOs (Spearman's Rank Correlation, * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001). **(B)** MHC-I gene expression is more negatively correlated with DNAm than expected by chance. In TI and SC, the correlation between DNAm and gene expression is shown for MHC-I pathway related genes with the blue curve. The grey distribution represents the background correlation of CpG-gene pairs. **(C)** DNAm is consistent in patients between time-points at the full range of DNAm values. i) A representative Bland–Altman plot of the difference in DNAm between at diagnosis and reassessment biopsies against mean DNAm at one promoter region CpG (*NLRC5*, cg07839457). ii) The magnitude and direction of differential DNAm are shown to be consistent between time-points. Delta betas comparing CD with controls and UC in the TI are shown for IEC samples at diagnosis and reassessment biopsies for all significantly differential CpGs. **(D)** Correlation between *NLRC5* promoter DNAm and relative gene expression of *NLRC5* normalised to controls in patient derived IEOs (n=3 lines per condition, Pearson's Correlation).

Figure S4. *NLRC5* acts as a transcriptional transactivator of human intestinal epithelial MHC-I. (A) Gene expression (normalised RNAseq read counts) of MHC-I pathway genes in IEOs with or without induction of *NLRC5* overexpression (dox) and +/- exposure to IFNγ or TNFα for 24h (N=4 independent technical replicates per condition. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, ns = not significant). **(B)** Flow cytometry on NLRC5 inducible overexpression IEOs (Dox) +/- stimulation with IFNγ for 24 hours (N=3 replicates per condition. P values were calculated by two-way ANOVA with Bonferroni test for multiple comparisons. ** $p < 0.01$, *** $p < 0.001$, ns = not significant. GMFI, geometric mean fluorescence intensity). **(C)** Quantitative PCR showing relative expression of MHC-I genes in WT and corresponding *NLRC5*$^{-/-}$ deficient IEOs (N = 3 technical replicates, two-way ANOVA with Turkey's test for multiple comparisons. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, ns = not significant). **(D)** Gene set enrichment analysis for genes annotated with the Gene Ontology term "Antigen processing and presentation of peptide and antigen by MHC Class I" in IEOs overexpressing *NLRC5* (Dox) versus WT (BSA).

Figure S5. NLRC5 acts as a transcriptional transactivator of murine intestinal epithelial MHC-I. (A) Experimental design for the generation and treatment of murine IEOs with IFNγ. **(B)** Expression of H2K1 and B2M in murine IEOs generated from *NLRC5* WT and KO mice, +/- exposure to IFNγ for 24 hours. (**** $p < 0.0001$, ns = not significant). **(C)** Gene set enrichment analysis for genes annotated with the Gene Ontology term "MHC-I protein complex" in murine *NLRC5* WT (fl/fl) and KO (-/-) IEOs. **(D)** Heatmap showing transcript expression levels (RNAseq) for MHC-I pathway genes in the same IEOs +/- IFNγ. (N=4 replicates per condition). **(E)** Time course showing H2K^b cell surface expression by flow cytometry in *NLRC5* WT and KO murine IEOs stimulated with IFNγ. Cells are gated on Live EpCAM+ cells (N=4 replicates per group). MFI, mean fluorescence intensity. P values were

calculated using two-way ANOVA with Bonferroni's multiple comparisons test (** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, ns = not significant). Data is representative of 2 independent experiments.

Figure S6. ***NLRC5*** **gene expression correlates with MHC-I in primary purified small and large bowel IEC.** Correlation between *NLRC5* expression and selected MHC-I pathway genes in primary small and large bowel purified IECs obtained from children/adolescents diagnosed with CD, UC and matched healthy controls. Cohort 2; the correlation was calculated separately by gut segment TI (blue) and SC (green) (R=Spearman's Rank Correlation).

Figure S7. CD associated increased intestinal epithelial MHC-I and crypt-villus axis expression gradient validates in independent cohorts. (A) Overview of sampling sites of single-cell RNA-seq dataset from Cohort 4 (n=15 healthy adults) and UMAP overview of epithelial cells, coloured by gut segment (Small Bowel = TI, large bowel = SC and Rectum). **(B)** Distribution of crypt-villus scores split by cell type (top) and correlation between summary MHC-I score and crypt-villus score for all cells (bottom, Spearman's Rank Correlation) in **i**) Rectum, SC and TI. ii) Box plot showing MHC-I score split by intestinal region (*** $p < 0.001$). **(C)** i) Overview of sampling sites of the Cohort 5 scRNA-seq (n=46 CD patients, n=24 healthy controls). ii) Distribution of crypt-villus scores split by cell type in TI and SC respectively. Top panel: Violin plots showing crypt-villus scores of cells within each IEC sub-type (top left) and total cell number (top right). Bottom panel: Correlation between single cell MHC-I score and crypt-villus score for all cells (bottom left) and box plots of summary MHC-I expression score split by diagnosis (bottom right). **(D)** i) Schematic of scRNAseq datasets generated from TI organoids +/- 24h treatment with TNFα or IFNγ (n = 4 lines per condition). ii) UMAPs of cells coloured by cell type (bottom left) and by treatment (bottom right). **(E)** Top

panel: Violin plots showing crypt-villus scores of cells within each IEC sub-type (top left) and total cell number (top right). Bottom panel: Correlation between single cell MHC-I score and crypt-villus score for all cells split by treatment (bottom left) and box plots of summary MHC-I expression score split by treatment (bottom right). **(F)** Box plots showing average single cell MHC-I score split by treatment and cell sub-type.

Figure S8. NLRC5 acts as key a modulator of mucosal inflammation *in-vivo*. (A) Representative contour plots of $H2K^b$-SIINFEKL and pan-$H2K^b$ flow cytometry on live $EpCAM^+$ cells in murine WT and *NLRC5* KO (-/-) IEOs +/- IFNγ exposure for 48 hours and +/- pulse with OVA257–264 peptide (SIINFEKL) peptide. **(B to D)** Analyses of murine WT and *NLRC5* KO small bowel mucosa following DSS induced gut inflammation. **(B)** Representative contour plot of $H2K^b$ surface expression on $EpCAM^+$ cell populations within the lamina propria extractions of DSS treated mice. **(C)** Absolute cell counts by flow cytometry of indicated cell types in DSS treated mice. Neutrophils, $CD45^+$ $CD11b^+Gr1^+$; Monocytes, $CD45^+CD11b^+Ly6C^{hi}MHC\text{-}II^-$; MHC-II $(MHC\text{-}II)^+$ monocytes (monocytes maturing into macrophages), $CD45^+CD11b^+Ly6C^+MHC\text{-}II^+$; Macrophages, $CD45^+CD11b^+Ly6C^{lo}MHC\text{-}II^+$. **(D)** Flow cytometry of intracellular staining for pro-IL-1b in the indicated cell types. Significance for **(C)** and **(D)** was calculated using the non-parametric Mann-Witney U test (** $p \leq 0.01$). **(E)** i) Quantitative inflammation score of $Nlrc5^{fl/fl}$ and $Nlrc5^{-/-}$ murine small intestine and colon following exposure to DSS. Total score range is 0-8 (small intestine) and 0-6 (colon), consisting of combined ‘inflammatory cell infiltrate’ and ‘mucosal architecture’ scores (each 0-4 and 0-3 for small intestine and colon respectively). Data are shown as mean ± SEM (n= 3-4 each group), and statistical analysis was performed using the Mann-Whitney test (*$p < 0.05$). ii) absolute weight and length values of $Nlrc5^{fl/fl}$ and $Nlrc5^{-/-}$ murine colon following the DSS

treatment. Significance was calculated using the non-parametric Mann-Witney test (**p < 0.01, ***p < 0.001).

Figure S9. Increased MHC-I gene expression in the large bowel intestinal mucosa of CD patients across different cohorts and data sets. (A) Overview of sampling sites and data types of Cohorts 2 and 7. **(B)** Average/summary MHC-I score and expression of selected MHC-I pathway genes. Expression split by diagnosis in SC samples from Cohort 2 (n = 32) and Cohort 7 (n = 116). **(C)** Expression of *NLRC5* and *IFNγ* in colonic biopsies from Cohort 7 (n = 116), comparing CD versus UC patients based on the presence or absence of mucosal inflammation.

Figure S10. IEC MHC-I DNAm stratifies patients into distinct subgroups that correlate with phenotype and disease severity. (A) i) Unsupervised hierarchical clustering of average MHC-I DNAm in SC (n = 131) IEOs (Cohort 1). Average MHC-I DNAm score is shown for each sample (middle annotation bar) and grouped by cluster (bottom annotation bar). Distribution of IEOs based on the MHC-I DNAm score cluster is shown on the right. ii) Unsupervised hierarchical clustering of average MHC-I DNAm in primary purified SC (n = 72) IECs (Cohort 2). Average MHC-I DNAm is shown for each sample (middle annotation bar) and grouped by cluster (bottom annotation bar). Distribution of primary purified IECs based on the MHC-I DNAm score cluster split by diagnosis is shown on the right.

(B) *NLRC5* and selected MHC-I CpG DNAm in TI and SC IEOs comparing CD patients with severe versus mild/moderate disease outcomes. i) *NLRC5* promoter CpGs (cg07839457 and cg07862320); ii) *HLA-E* (cg20105257); and iii) *PSMB9* (cg00045690). Comparisons for panels C and D were calculated using Wilcox tests (* p < 0.05, ** p < 0.01). **(C)** Box plot of average MHC-I DNAm and selected 53 diagnostic CpGs in TI (n = 127) IEOs (Cohort 1), for CD and

non-CD (UC patient or healthy control) samples respectively. P values were calculated by two-way Welch's t-test (**** $p < 0.0001$). ii) ROC curves and AUC scores of logistic regression classifiers for CD diagnosis using TI (n = 127) IEOs (Cohort 1), based on selected 53 diagnostic CpGs, 628 MHC-I CpGs and median of random selections of 628 CpGs respectively. ROC and AUC are both estimated means through repeated cross validations. iii) Box plot of CD severity prognostic risk score based on methylation of selected 28 CpGs processed by logistic regression classifiers, for CD samples with severe and mild/moderate disease outcomes respectively. P values were calculated by two-way Welch's t-test (** $p < 0.01$).

Figure S11. MHC-I gene transcription of primary IEC and mucosal biopsies separates patients into distinct subgroups across multiple patient and sample cohorts. (A) Unsupervised hierarchical clustering of selected MHC-I pathway gene expression in IECs from Cohort 2 in TI (n = 70). Samples are labelled by diagnosis (top annotation bar). Average MHC-I pathway gene expression score is shown for each sample (middle annotation bar) and grouped by cluster (bottom annotation bar). Distribution of TI IECs based on average MHC-I pathway gene expression split by diagnosis is shown on the right. **(B)** Clustering of selected MHC-I pathway gene expression in IECs from Cohort 2 in SC (n = 72). Samples are labelled by diagnosis (top annotation bar). Average MHC-I pathway gene expression score is shown for each sample (middle annotation bar) and grouped by cluster (bottom annotation bar). Distribution of SC IECs based on average MHC-I pathway gene expression split by diagnosis is shown on the right. **(C)** Unsupervised hierarchical clustering of whole biopsy gene expression of selected MHC-I pathway genes from Cohort 6 (n = 322). Samples are labelled by diagnosis (top annotation bar), tissue inflammation (second row annotation bar) and presence or absence of deep ulcers (third row annotation bar). Average MHC-I pathway gene expression is shown for each sample (fourth row annotation bar) and grouped by cluster

(bottom annotation bar). **(D)** Unsupervised hierarchical clustering of whole biopsy gene expression of selected MHC-I pathway genes from Cohort 7 TI samples (n = 78). Samples are labelled by diagnosis (top annotation bar), tissue inflammation (second row annotation bar) and disease activity (third row annotation bar). Average MHC-I pathway gene expression is shown for each sample (forth row annotation bar) and grouped by cluster (bottom annotation bar). **(E)** Unsupervised hierarchical clustering of whole biopsy gene expression of selected MHC-I pathway genes from Cohort 7 colon samples (n = 116). Samples are labelled by diagnosis (top annotation bar), tissue inflammation (second row annotation bar) and disease activity (third row annotation bar). Average MHC-I expression is shown for each sample (fourth row annotation bar) and grouped by cluster (bottom annotation bar).