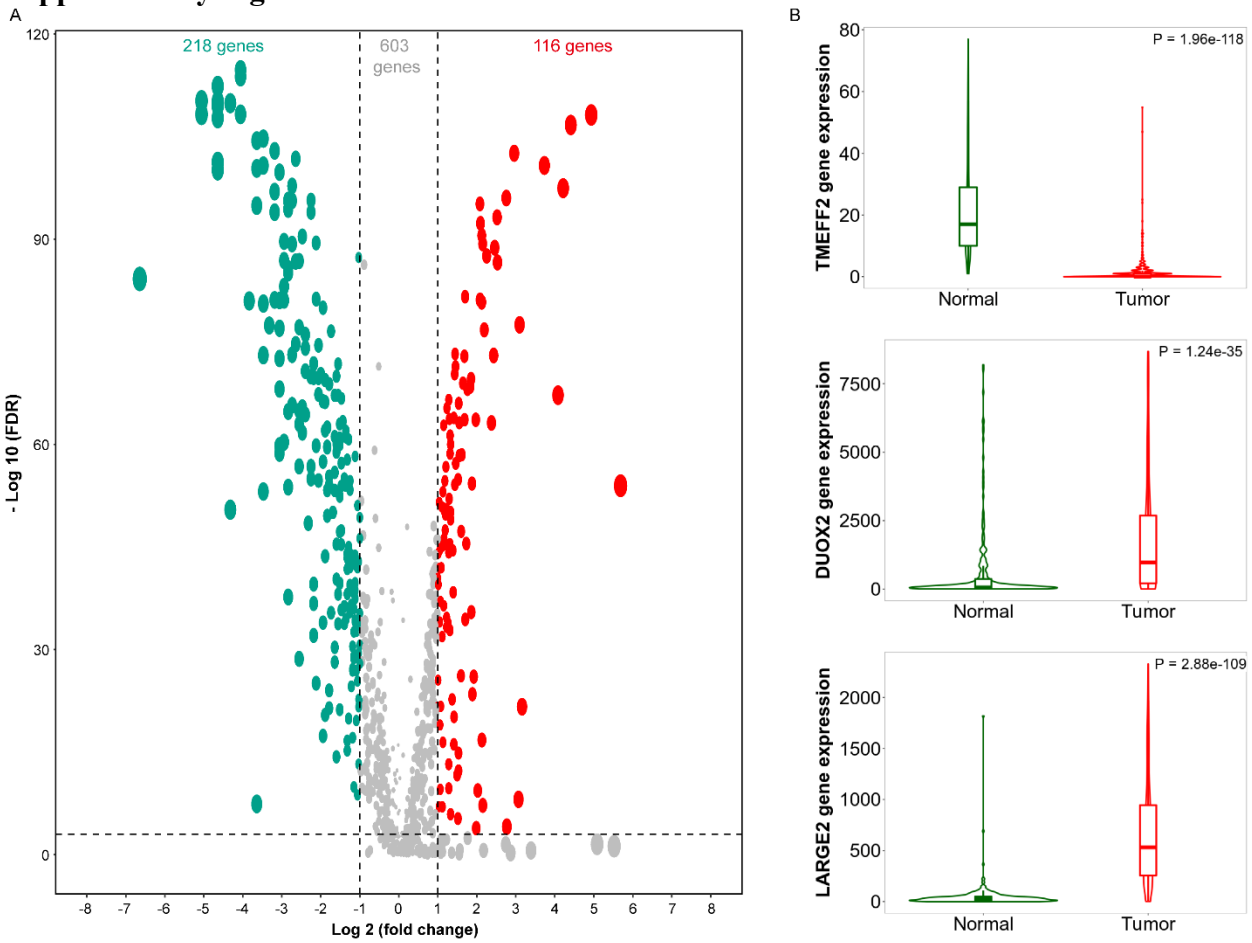
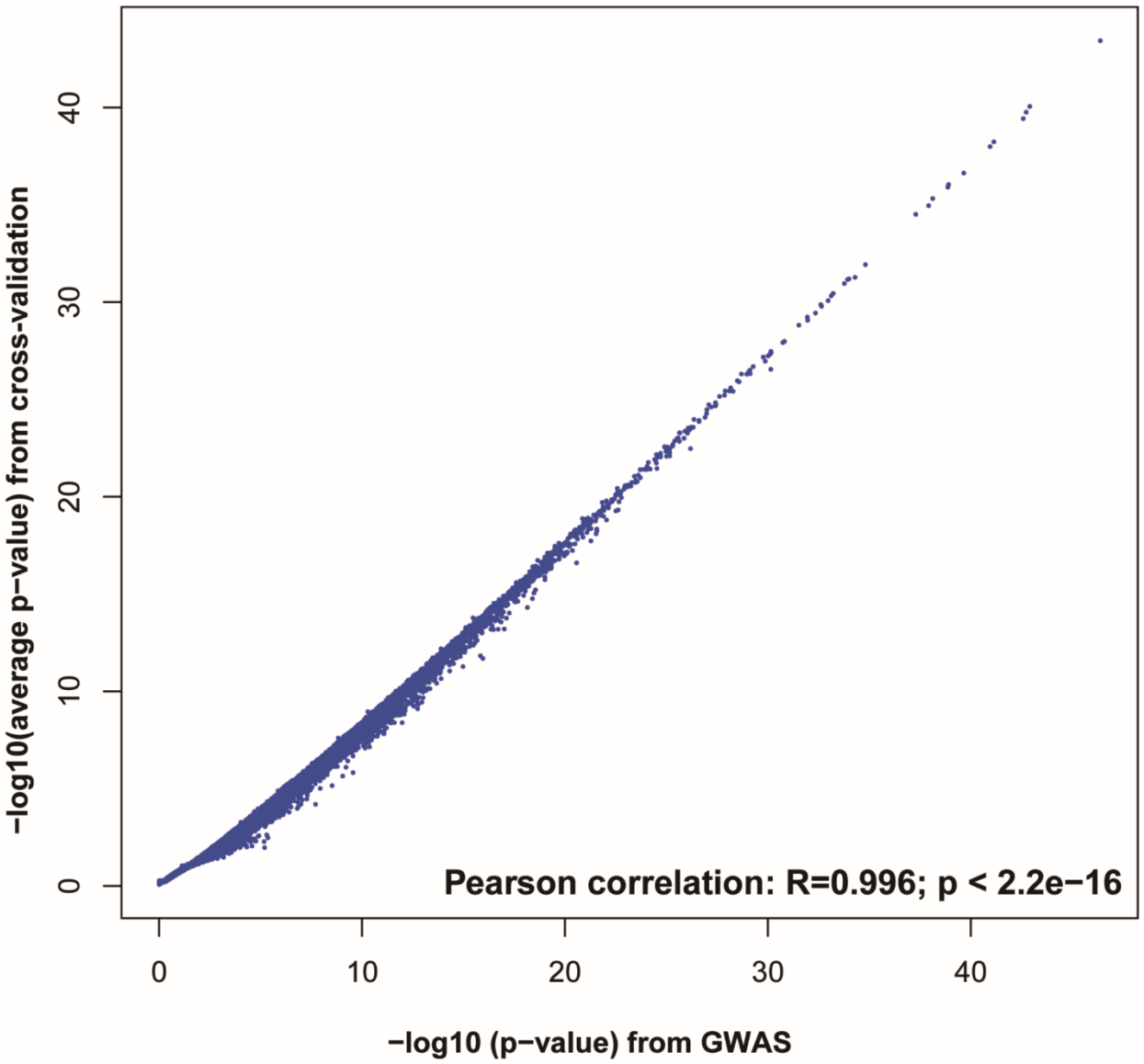
Supplementary Figures and Tables

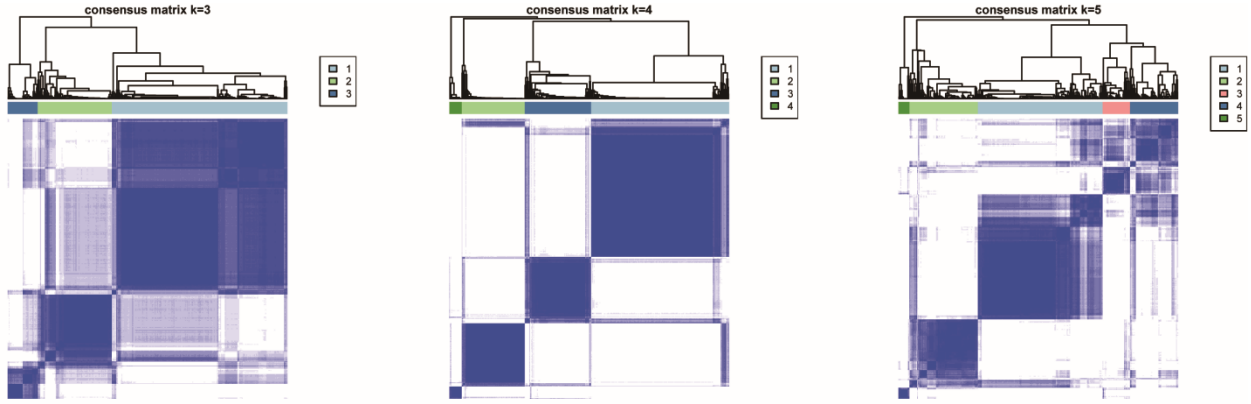
Supplementary Figures



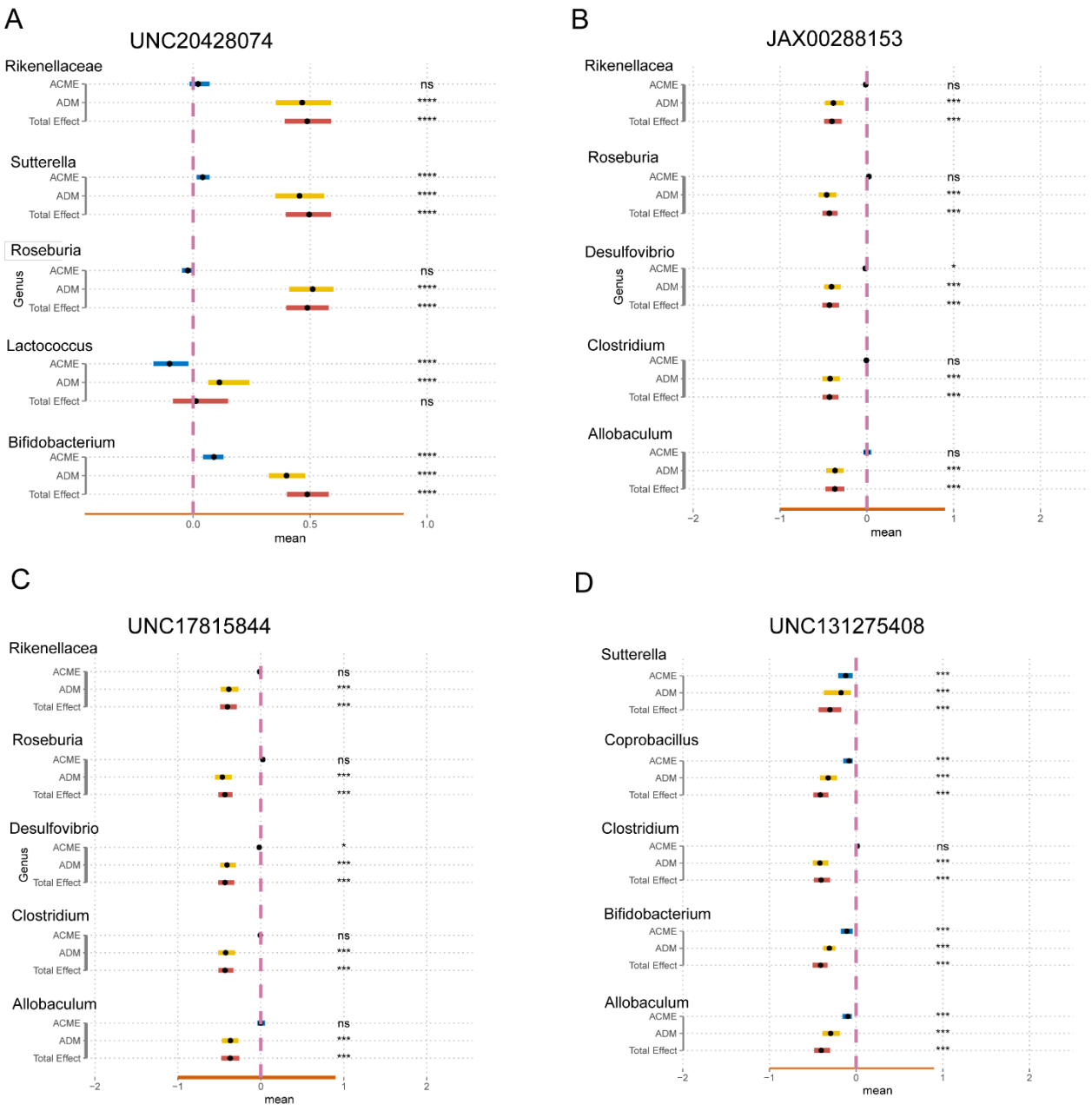
Supplementary Fig. S1 334 of 936 CTS candidate genes are transcriptionally altered in human colon adenocarcinomas. (A) Volcano plot depicting the differentially expressed genes between adenocarcinomas and normal colon tissues ($|\log_2 FC| \geq 1$ and FDR ≤ 0.001). (B) Gene expression of representative genes in normal and tumor tissues.



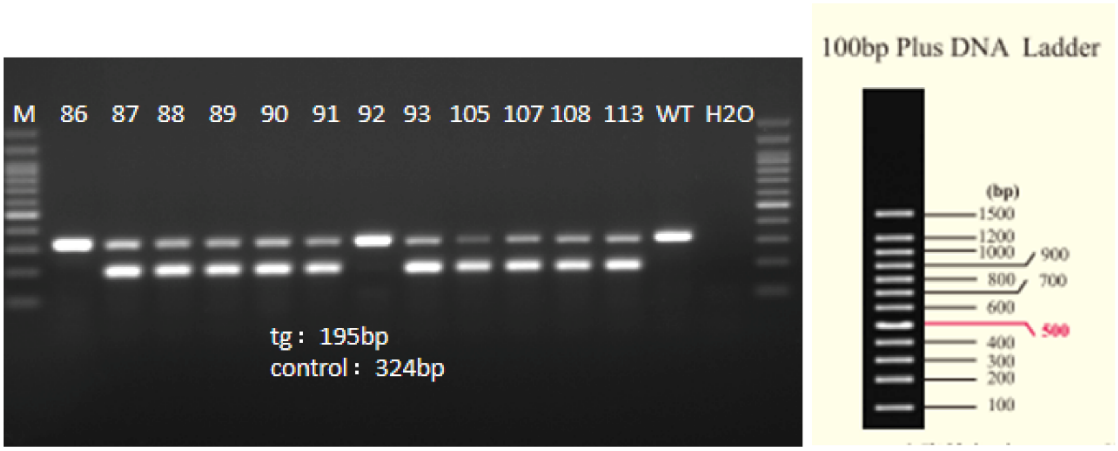
Supplementary Fig. S2 Correlations between the p-values obtained from GWAS analysis and cross-validation analysis confirmed the findings from GWAS.



Supplementary Fig. S3 Classification of patients with colon cancer into different groups based on expression of 334 CTS candidate genes in TCGA-COAD cohort. Consensus classification heatmaps based on three (left panel) and four (right) clusters.

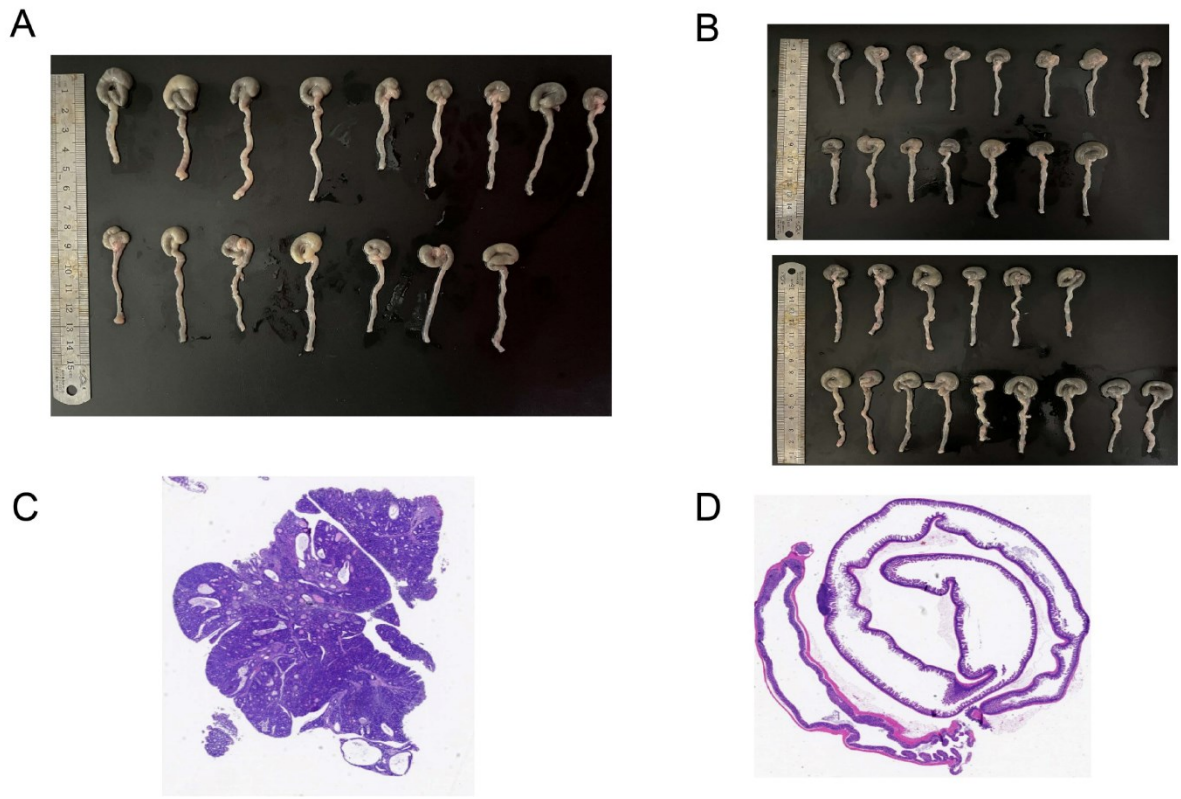


Supplementary Fig. S4 Microbes mediate the genetic effects on CTS. (A) SNP UNC20428074. (B) SNP JAX00288153. (C) SNP UNC17815844. (D) SNP UNC131275408.

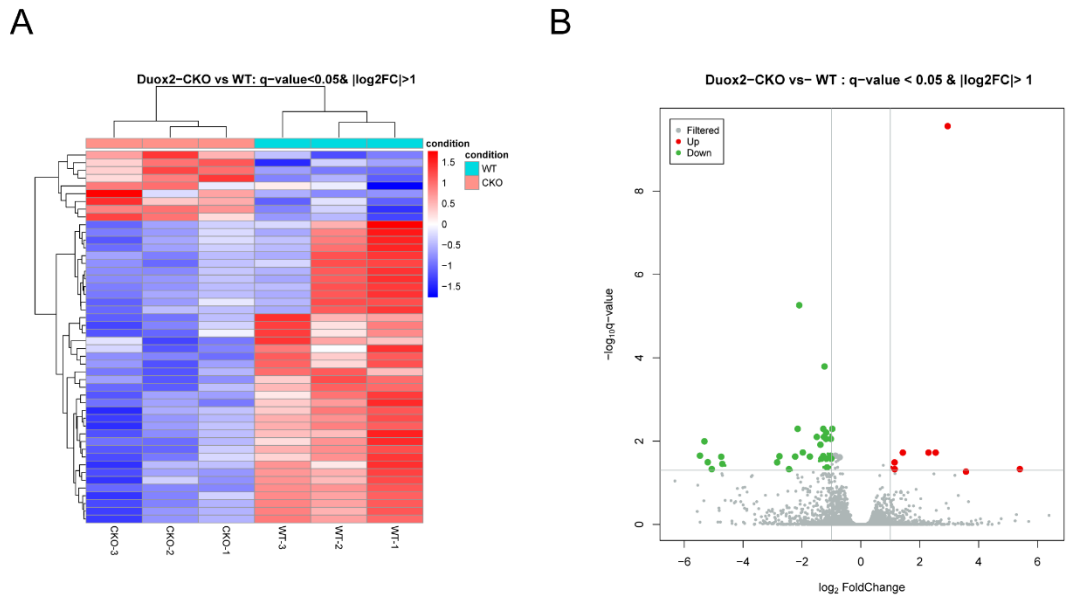


ID	86	87	88	89	90	91	92	93	B105	B107	B108	B113
Villin-Cre	-	+	+	+	+	+	-	+	+	+	+	+

Supplementary Fig. S5 Establishment and Characterization of intestine-specific Duox2 deficient mice by polymerase chain reaction (PCR). (+) are Duox2 $^{fl/fl}$ Villin-Cre mice, (-) are Duox2$^{fl/fl}$ mice.



Supplementary Fig. S6 Representative photographs of colorectal tumors in WT and CKO mice. (A) Representative images of the colon in mice from the WT mice. (B) Representative images of the colon in mice from the and CKO mice. (C) Representative H&E image of colorectal tumor showing hyperplastic glands with low-grade dysplasia. (D) Representative H&E image of colon without tumors.



Supplementary Fig. S7 Differentially expressed genes were identified between WT samples and CKO samples. (A) Heatmap of the differentially expressed genes between WT samples and CKO samples, with red indicating higher expression and blue indicating lower expression. (B) volcano map of the differentially expressed genes between WT samples and CKO samples, with red indicating higher expression and green indicating lower expression.

Supplementary Tables

Supplementary Table S1 List of collaborative cross mice used in this study and tumor status in each mouse (Note: mice highlighted in red were excluded for analysis).

Supplementary Table S2 List of SNPs and p-values obtained from the genome-wild association study of tumor susceptible phenotype.

Supplementary Table S3 Human homolog and their alteration in colon adenocarcinoma and overlap with human GWAS. Significantly downregulated genes (FDR $\leq$ 0.001 and fold change $\leq$ 0.5) highlighted in Green and significantly upregulated genes (FDR $\leq$ 0.001 and fold change $\geq$ 2.0) highlighted in red.

Supplementary Table S4 List of overrepresented KEGG pathways among 334 candidate susceptibility genes that significantly differentially expressed in human colon adenocarcinoma.

Supplementary Table S5 Number of sequencing reads at genus level in each sample.

Supplementary Table S6 List of the p-values obtained by Mann-Whitney test between low and high colon tumor susceptibility (CTS) groups and the adjusted p-value by Bonferroni for each genus.

Supplementary Table S7 The p-values for each SNP associated with the abundance of genus (Significant p-value <1.0E-06 is highlighted in red).

Supplementary Table S8 List of SNPs significantly associated with the abundance of multiple genera.

Supplementary Table S9 List of the p-values obtained by Mann-Whitney test between *Duox2* wildtype (WT) and conditional knockout (CKO) mice the adjusted p-value by Bonferroni for each genus.

Supplementary Table S10 List of differentially expressed genes between Duox2 wildtype (WT) and conditional knockout (CKO) mice.

Supplementary Table S11 The use of DUOX2 expression and cancer microbiome data from TCGA-COAD cohort to generate Figure 8B, C and D.