Our Supplementary section consists of the following items:

1) Single text file containing sequences and other specific methods information, legends for 5 supplementary figures, and references for the supplementary methods.

2) Five Supplementary figures as follows:

Supplemental Figure 1. Increased volatility in microbiota of T5KO mice. Relates to Figure 3

Supplemental Figure 2. Increased susceptibility of T5KO mice to Crohn's disease associated Adherent-Invasive *E. coli* (AIEC) strain LF82 infection. Relates to figure 4.

Supplemental Figure 3. Germ-free T5KO did not exhibit any intestinal disorders Relates to figure 5.

Supplemental Figure 4. Germ-free mice are quickly conventionalized and neither WT nor T5KO mice develop colitis in response to commensal *Escherichia coli* strain F18 Relates to figure 6.

Supplemental Figure 5. Early AIEC LF82 infection induced increased expression of pro-inflammatory cytokines. Relates to figure 7.

Supplementary Methods

Fluorescent *in situ* hybridization.

Paraffin sections were dewaxed with Xylene substitute (Sigma), incubated in 99.5% Ethanol for 5 minutes and air dried. The sections were hybridized with a probe detecting Enterobacteriaceae (5'-CCCCCWCTTTGGTCTTGC-3') (Kempf et al., 2000) conjugated to Alexa 555 in hybridization buffer (20 mM Tris-HCl pH 7.4, 0.9 M NaCl, 0.1% SDS) at 50ºC over night. The sections were rinsed in wash buffer (20 mM Tris-HCl pH 7.4, 0.9 M NaCl), washed at 50ºC for 20 min and counterstaining with DAPI (Sigma). The sections were mounted using ProLong Gold Anti-fade (Invitrogen). Fluorescence images were obtained on a LSM 700 Axio Examiner.Z1 laser scanning confocal microscope, with a Plan-Apochomat 40x/1.3 Oil DIC objective, and analyzed with the ZEN 2010 software (Zeiss).

Total stool, cecal or adherent bacteria quantification

For the stool and cecal total bacteria quantification, total DNA were extracted using QIAamp DNA Stool Mini Kit (Qiagen) accordingly to the manufacturer protocol. Similarly, segments of colon were washed in HBSS and total DNA extracted using DNeasy® blood & tissue kit (Qiagen) accordingly to the manufacturer protocol. Then, DNA was subjected to qPCR using universal 16S rRNA primers: forward primer (27F) 5'-**AGAGTTTGATCCTGGCTCAG**-3', and reverse primer (338R) 5'-**TGCTGCCTCCCGTAGGAGT**-3'.

Fecal or cecal microbiota analysis by 16S rRNA gene sequencing

Six colitic T5KO mice, as previously described were selected, one with rectal prolapse and 5 exhibiting both splenomegaly and colomegaly. To avoid the confounding effects of cohousing on the diversity of fecal or cecal bacteria, we selected mice from multiple litters that

were housed separately. In parallel, 6 WT littermates and 8 non-colitic T5KO mice housed in the same cage as the selected colitic T5KO mice were picked. Bulk DNA was extracted from frozen extruded fecal or cecal contents using a PowerSoil-htp ® kit from MoBio Laboratories (Carlsbad, CA) with mechanical disruption (bead-beating). 16S rRNA genes were PCR amplified from each sample using a composite forward primer and a reverse primer containing a unique 12-base barcode, designed using the Golay error-correcting scheme, which was used to tag PCR products from respective samples (Hamady et al., 2008). We used the forward primer 5'-*GCCTTGCCAGCCCGCTCAGTC***AGAGTTTGATCCTGGCTCAG**-3': the italicized sequence is 454 Life Sciences® primer B, and the bold sequence is the broadly conserved bacterial primer 27F. The reverse primer used was 5'-*GCCTCCCTCGCGCCATCAGNNNNNNNNNNNNCA***TGCTGCCTCCCGTAGGAGT**-3': the italicized sequence is 454 Life Sciences' primer A, and the bold sequence is the broad- range bacterial primer 338R. *NNNNNNNNNNN* designates the unique twelve-base barcode used to tag each PCR product, with 'CA' inserted as a linker between the barcode and rRNA primer. PCR reactions consisted of HotMaster PCR mix (Eppendorf), 0.2 µM of each primer, 10-100 ng template, and reaction conditions were 2 min at 95°C, followed by 30 cycles of 20s at 95°C, 20s at 52°C and 60s at 65°C on an Eppendorf thermocycler. Three independent PCRs were performed for each sample, combined and purified with Ampure magnetic purification beads (Agencourt), and products visualized by gel electrophoresis. Products were quantified using Quant-iT PicoGreen dsDNA assay as described above. A master DNA pool was generated from the purified products in equimolar ratios. The pooled products were sequenced using a Roche 454 Titanium pyrosequencer at the University of South Carolina (EnGenCore).

16S rRNA gene sequence analysis

Sequences were analyzed using the open source software package Quantitative Insights Into Microbial Ecology (QIIME (Caporaso et al., 2010)). Sequences that passed quality filtering using the default parameters in QIIME were checked for chimeras and assigned to operational taxonomic units (OTUs) using OTUpipe (Edgar et al., 2011) using 97% pair-wise identity, then classified taxonomically using the RDP classifier (Wang et al., 2007) retrained with Greengenes (McDonald et al., 2012). A single representative sequence for each OTU was aligned using PyNAST (Caporaso et al., 2010)**,** then a phylogenetic tree was built using FastTree (Price et al., 2009). The phylogenetic tree was used for computing the UniFrac distances between samples (Lozupone and Knight, 2005). Semivariogram plots were used to plot days dissimilarity (Euclidean, x axis) against community dissimilarity (UniFrac, y axis) divided by treatment, which show the volatility of the samples through time (Caporaso et al., 2010; Curran et al., 2000; McBratney et al., 1997). To fully assess this variability we applied a linear regression on the points that showed how the colitic mice are the most similar at any given point but that the community dissimilarity considerably grows when compared against distant (time) samples; contrastingly Lab WT are almost equally dissimilar at any giving time point and through time.

We used "nearest shrunken centroid" (Knights et al., 2011; Tibshirani et al., 2002) analysis (Predictive Analysis of Microarrays package under R software) to search for discriminating OTUs between colitic and non-colitic mice at all taxonomic levels. At each of the taxonomic levels we picked a threshold which allowed minimum cross validated misclassification error rate with a minimum number of OTUs. We used these OTUs to predict the health state of each mouse at each taxonomic level at the different weeks and calculated the overall error rates. Jackknifing PCoA plots were used to assess the variation between genotypic replicates and the effect of rarefaction level to measure the robustness of the clusters (Lozupone et al., 2007). For the comparisons of mean phylotype abundances in WT and T5KO mice, significance levels were adjusted for multiple comparisons using Bonferroni's correction.

Supplemental Figure Legends

Supplemental Figure 1. Increased volatility in microbiota of T5KO mice (relates to Figure 3). Stool from wild-type, non-colitic and colitic T5KO mice (n=5-8 mice per group) were collected weekly for 9 weeks after weaning (from 3-week to 11-week old). Stool microbiota composition was analyzed via 16S rRNA analysis. (A) After removing the Enterobacteria OTUs from the QIIME analysis and clustering of mouse cecal bacterial communities using principal coordinates analysis (PCoA) of the UniFrac unweighted distance matrix, the average of the UniFrac unweighted distance for each category (WT, non-colitic and colitic T5KO) between consecutive time points has been calculated. (B) OTUs summarized at family level and found to discriminate between colitic and non-colitic mice. (C) Overall misclassification error rates using OTUs summarized at the class level and (D) overall misclassification error rates using OTUs summarized at the order level at each week. Analysis was done by ANOVA and statistical significance (P<0.01) is denoted by asterisk (*).

Supplemental Figure 2. Increased susceptibility of T5KO mice to Crohn's disease associated Adherent-Invasive *E. coli* (AIEC) strain LF82 infection (relates to figure 4). Eight week old wild-type and T5KO (n=6-8 mice per group) were pretreated with 10 mg of streptomycin and 24 h later infected orally infected with 10^9 flagellate AIEC strain LF82 bacteria or its isogenic flagellin deficient mutant (LF82-Δ*fliC*). (A) Numeration of AIEC LF82 or LF82-Δ*fliC* present in the WT or T5KO mouse stool from day 1 to day 10 post infection. (B) WT or T5KO mice clearance of AIEC LF82 or LF82-Δ*fliC* by numbering the bacteria in the stool. (C) Gross picture of cecum 48 h post infection. (D) Following euthanasia (48h post infection), cecum was isolated and cecum MPO activity was measured. (E-F) Inflammation

100 severity has been monitored in the cecum by calculating a histological score as described in
101 Methods and representative H&E stained histological observations of cecum following treatment
102 (magnification, 100×). The data is representative of 3 independent experiments.* $p<0.05$.
103
104 **Supplemental Figure 3. Germ-free T5KO did not exhibit any intestinal disorders (relates**
105 **to figure 5).** Twelve-week old wild-type and T5KO mice (n=7-8 mice per group) were received
106 from Taconic as germ-free. (A) Body mass. (B) Colon mass. (C) Spleen mass. (D) Colon
107 MPO. (E) Mice were bled retroorbitally. Serum was separated and used for Lcn-2 ELISA. (F)
108 Quantitation of total fecal bacteria by qPCR analysis using universal *16s rRNA* primers. (G)
109 Stool was collected and diluted in 500 μL of PBS. Then, supernatant was assayed for Lcn-2
110 expression by ELISA.
111
112 **Supplemental Figure 4. Germ-free mice are quickly conventionalized and neither WT nor**
113 **T5KO mice develop colitis in response to commensal *Escherichia coli* strain F18 (relates to**
114 **figure 6).** (A) Germ-free wild-type (n=6 mice) were transferred to a sterile cage upon delivery
115 and kept in a conventional animal house. As control, germ-free mice were also gavaged with
116 cecal content from a conventional mouse. Stools were collected daily and total fecal bacteria
117 were quantified by qPCR analysis using universal *16s rRNA* primers. (B-J) Germ-free wild-type
118 and T5KO mice (n=4 mice per group) were orally infected with 10^7 commensal *E. coli* F18
119 bacteria. (B) Body mass was monitored daily during the treatment. (C) Numeration of *E. coli*
120 F18 present in the WT or T5KO mouse stool from day 1 to day 7 post infection. (D) Following
121 euthanasia, spleen was isolated and mass measured. (E) Colon mass. (F) Colon length. (G)
122 Gross picture of colon. (H) Colon MPO activity. (I-J) Histological score and representative
123 H&E stained colon (magnification, 100×). * $p<0.05$.
124

Supplemental Figure 5. Early AIEC LF82 infection induced increased expression of pro-inflammatory cytokines (relates to figure 7). Germ-free wild-type and T5KO mice (n=4 mice per group) were orally infected with 10^7 flagellate AIEC LF82 bacteria. After 7 days post infection, mice were bled retroorbitally. Serum was separated and used to assay for several pro-inflammatory cytokines, namely IL-1β (A), TNF-α (B), and Lcn-2 (C) by ELISA. * $p<0.05$.

References

Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., and Knight, R. (2010). PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics *26*, 266-267.

Curran, P.J., Atkinson, P.M., Foody, G.M., and Milton, E.J. (2000). Linking remote sensing, land cover and disease. Adv Parasitol *47*, 37-80.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics *27*, 2194-2200.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J., and Knight, R. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nat Methods *5*, 235-237.

Kempf, V.A., Trebesius, K., and Autenrieth, I.B. (2000). Fluorescent In situ hybridization allows rapid identification of microorganisms in blood cultures. Journal of clinical microbiology *38*, 830-838.

Knights, D., Kuczynski, J., Koren, O., Ley, R.E., Field, D., Knight, R., DeSantis, T.Z., and Kelley, S.T. (2011). Supervised classification of microbiota mitigates mislabeling errors. ISME J *5*, 570-573.

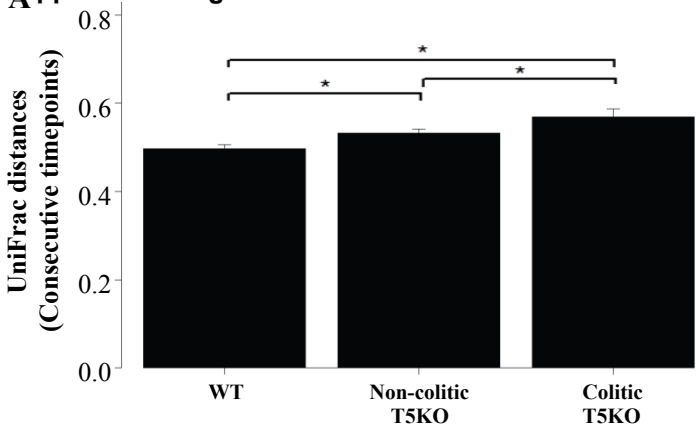
149 Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing
150 microbial communities. Appl Environ Microbiol *71*, 8228-8235.
151 Lozupone, C.A., Hamady, M., Kelley, S.T., and Knight, R. (2007). Quantitative and qualitative
152 beta diversity measures lead to different insights into factors that structure microbial
153 communities. Appl Environ Microbiol *73*, 1576-1585.
154 McBratney, A.B., Whelan, B.M., and Shatar, T.M. (1997). Variability and uncertainty in spatial,
155 temporal and spatiotemporal crop-yield and related data. Ciba Found Symp *210*, 141-160.
156 McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., Desantis, T.Z., Probst, A., Andersen,
157 G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit
158 ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J *6*, 610-618.
159 Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution
160 trees with profiles instead of a distance matrix. Mol Biol Evol *26*, 1641-1650.
161 Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. (2002). Diagnosis of multiple cancer
162 types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A *99*, 6567-6572.
163 Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007). Naive Bayesian classifier for rapid
164 assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol *73*,
165 5261-5267.

166

167

Supplemental Figures

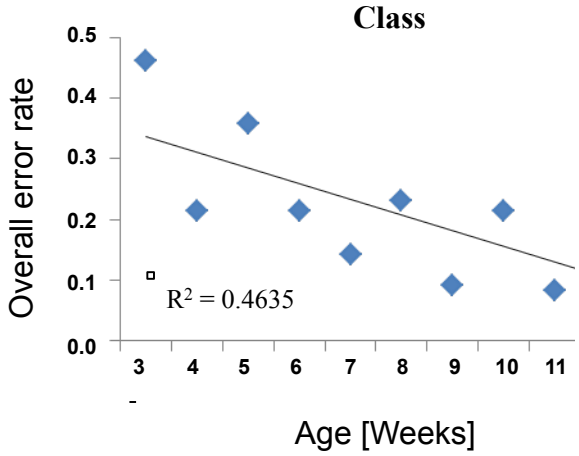
A



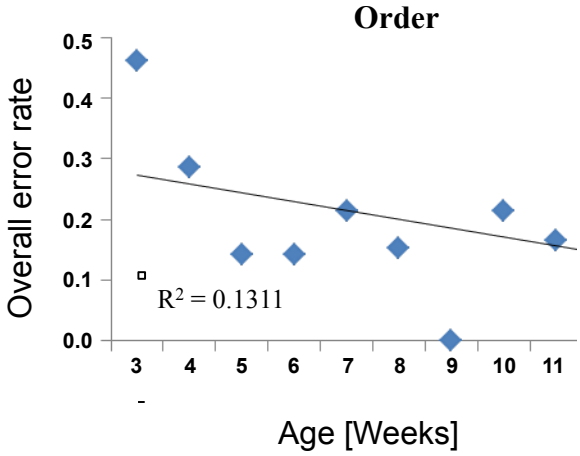
B

Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae
Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae
Tenericutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae
Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae
Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae
Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae
Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae
TM7;c__TM7−3;o__CW040;f__F16
Tenericutes;c__Mollicutes;o__Mycoplasmatales;f__Mycoplasmataceae

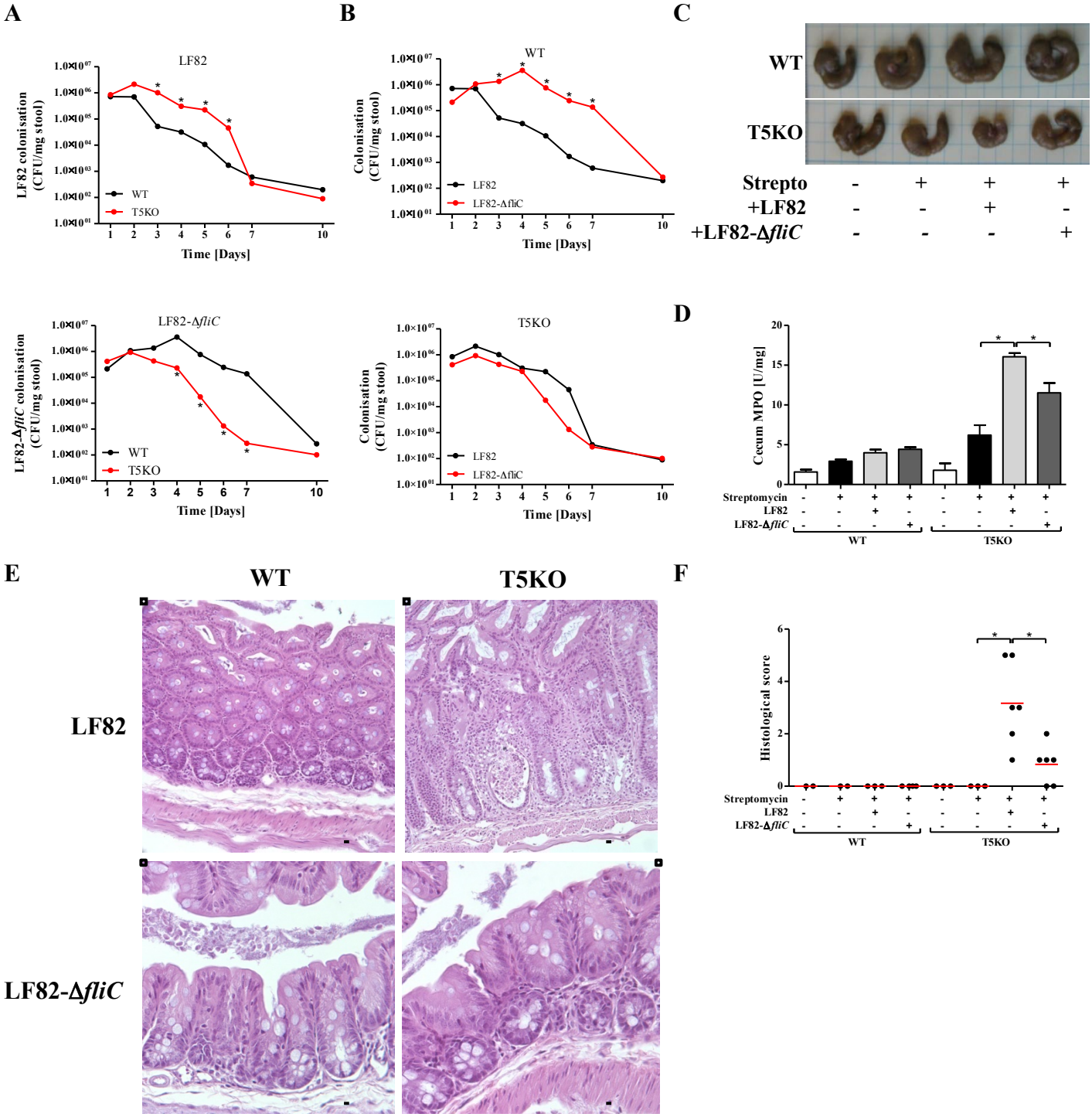
C



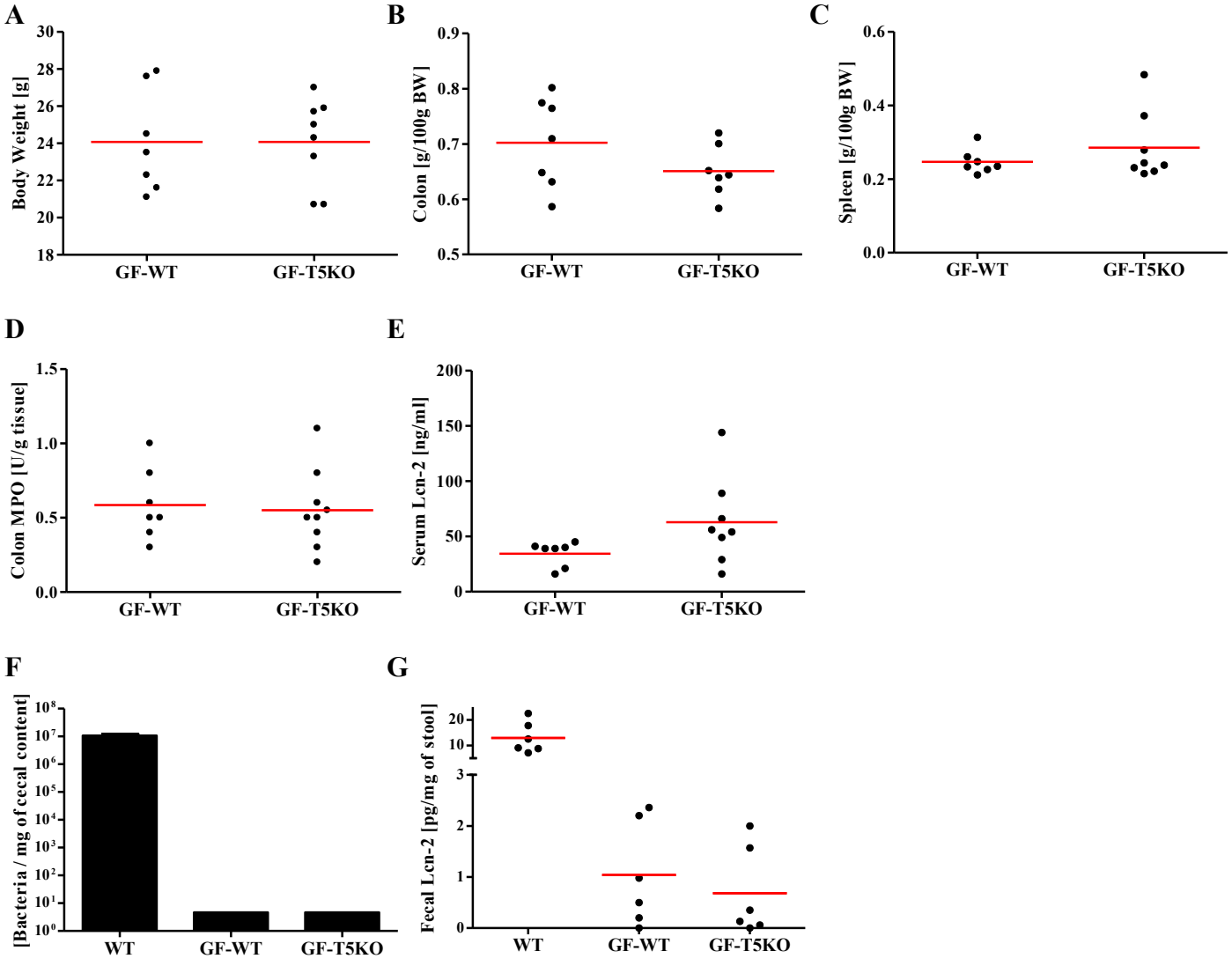
D



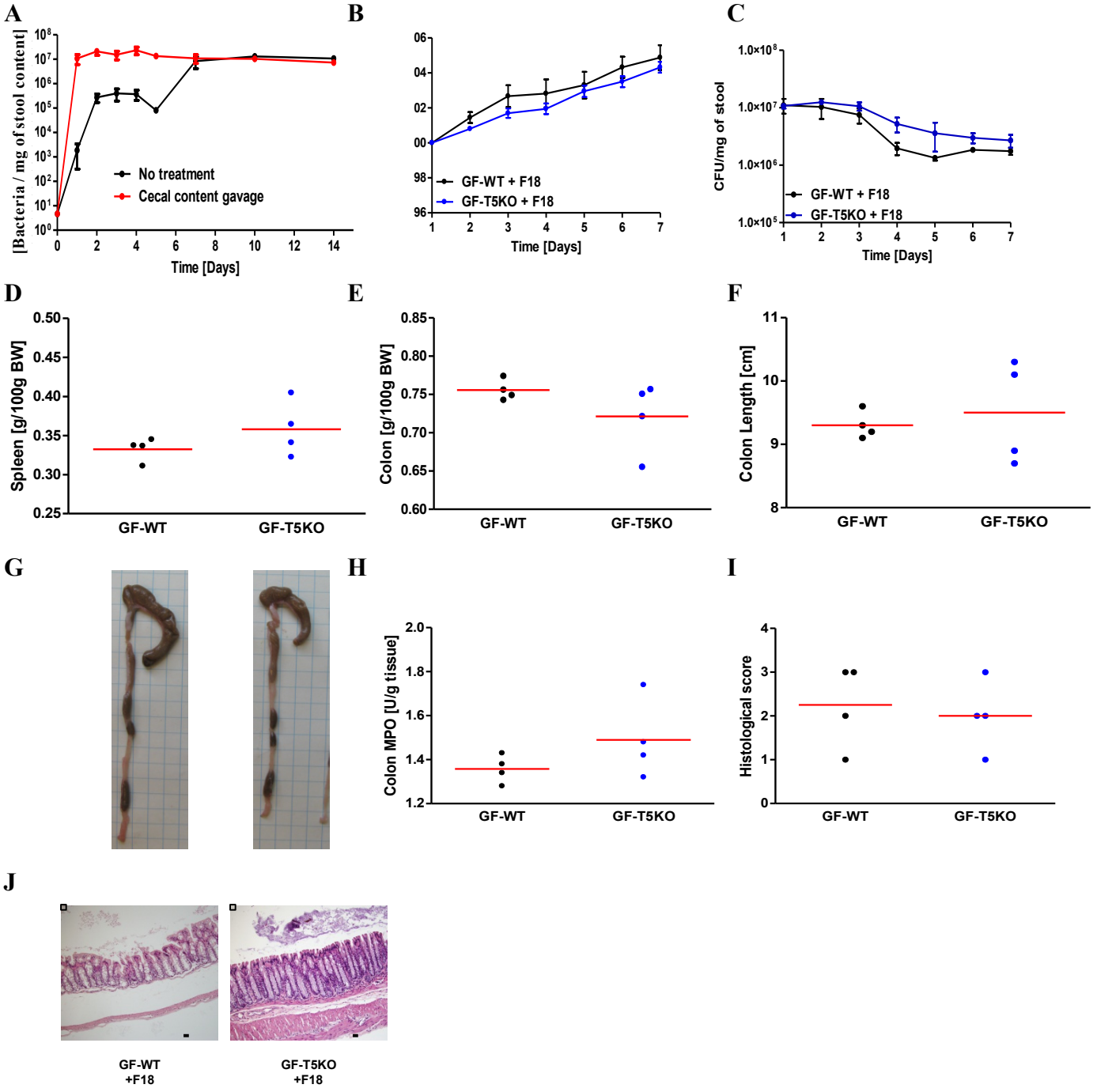
Supplemental Figure 1



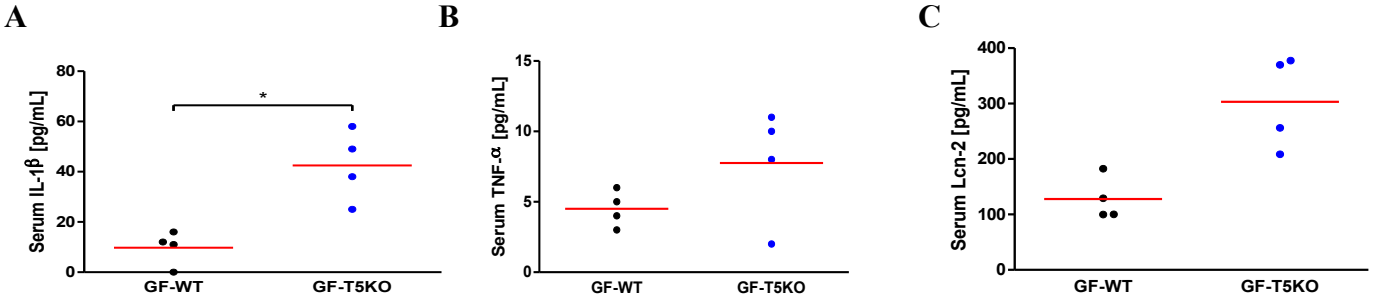
Supplemental Figure 2



Supplemental Figure 3



Supplemental Figure 4



Supplemental Figure 5