Supplemental Information

Supplemental Materials and Methods

Mice and Animal Housing

MyD88$^{-/-}$ and *Trif*$^{-/-}$ mice were kindly provided by Dr. S Akira. Mice were bred to a C57BL/6 background and then crossed with each other to establish double heterozygote breeding lines as detailed in the main text. Animals were maintained under specific pathogen-free conditions at the animal facility of Yale University School of Medicine. Studies were approved by the Institutional Animal Care & Use Committee of Yale University.

Whole Genome Gene Expression Analysis

Tissue was homogenized using a rotor stator in TRIZOL (Life Technologies) and total RNA was isolated according to manufacturers instructions. RNA was digested with Dnase I and processed using RNAII kit (Clontech). For whole genome microarray analysis pooled samples of isolated RNA per genotype from multiple litters with equal proportions of sex-matched (both males and females included) littermates were hybridized on Illumina MouseWG-6 v2.0 Expression BeadChips for whole-genome expression profiling which covered > 19,100 unique, curated genes in the NCBI RefSeq database. This microarray whole-genome expression profiling was performed in duplicate using RNA samples pooled from different mice from different litters sacrificed 6-12 months apart from samples for each microarray. Probes included in primary data sets had to fit 3 inclusion criteria (a: 1.5 fold difference between conditions, b: considered as an "expressed" gene product based on probe *p* value in both microarray analyses, and c: obeying a+b in the duplicate microarrays). Quantitative changes are reported by unit of probe, rather than gene, as in a few situations different probes corresponded to the same gene.

Data Sets and Bioinformatics Analysis

Primary (1°) data sets: Twenty four primary (1°) data set s (A-X) were established to determine both developmental and TLR/IL-1R-regulated genes (Fig. 1). This included "Developmental" data sets used to compare gene expression profiles in the small intestine and colon (up- and down-regulated) between suckling, weaned and adult WT mice and "TLR/IL-1R-regulated per developmental stage" used to compare gene expression profiles in the small intestine and colon (WT>DKO and DKO>WT) per developmental stage (pre-weaning, post-weaning and adult) (Dataset S1; A-X).

Secondary (2°) data sets: Secondary (2°) data sets (Fig. S1, Dataset S1; Tables Y,Z, AA, AB) were established to determine coordinate developmentally- *and* TLR/IL-1R-regulated genes which were identified as transcripts requiring TLR/IL-1R signaling for developmental expression (induction or repression). Those for induction were established as developmentally up-regulated (in at least one 1° data set: weaning, adulthood or pre-

weaning/adult) and WT>DKO (in at least one developmental stage: pre-weaning, post-weaning and adult); (Fig. S1). Those for repression were established as developmentally down-regulated (in at least one 1° data set: weaning, adulthood or pre-weaning/adult) and DKO>WT (in at least one developmental stage: pre-weaning, post-weaning and adult; Fig. S1) A complete list of genes in all primary and secondary data sets in included in "Master by gene" (Dataset S1).

Tertiary (3°) data sets: Four tertiary (3°) data sets identifying developmentally and/or TLR/IL-1R-regulated genes coordinately (up-regulated by microbe for developmentally-induced and/or WT>DKO genes; down-regulated by microbe for developmentally-repressed and/or DKO>WT genes) modified by microorganisms through studies in germ-free mice were established (Fig. S1, Dataset S1; Tables AC-AF). Eight published reports of the effect of microbes (indigenous flora, intestinal pathogen or non-pathogen) on gene expression in the mouse intestine or isolated intestinal cell types (Paneth cells, epithelial progenitors) were curated. These were: comparison of adult germ-free vs conventionally reared mice gene expression in the ileum and colon (1); comparison of adult germ-free vs conventionally reared mice gene expression in the jejunum and colon(2); ileum, 10 days post-colonization by mouse cecal microbiobota or zebrafish GI microbiota (3); ileum, 10 days post-colonization of germ-free mice with mouse ileal flora (4); comparison of adult germ-free mice mono-associated with bacteria (gene expression in the ileum, 3 days post-monoassociation with *Listeria monocytogenes*, *Bacteroides thetaiotamicron*, or *Listeria innocuum* (5); ileal expression 3 days post-monoassociation with *Bacteroides thetaiotamicron* (6); comparison of adult germ-free mouse associated with a consortial microbiota (gene expression in isolated Paneth cells, 10 days post-colonization by the mouse ileal microbiota (7) and gene expression in isolated Paneth cells or intestinal epithelial cell progenitors in the middle and distal small intestine of conventionally raised mice (8).

Cross-platform analysis was performed by converting probe IDs from these various microarray platforms (Affymetrix Mouse 430 2.0, U74 Av2, Mu11K or Mu19K into Entrez Gene IDs using AILUN: array information library universal navigator (http://ailun.stanford.edu/) (9). Entrez Gene ID from Illumina probes were used from either manufactured provided or converted to Entrez Gene ID from the probe's associated GenBank Accession number using the DAVID Gene ID Conversion Tool (http://david.abcc.ncifcrf.gov/conversion.jsp) into DAVID Entrez ID (DEI).

Only Entrez Gene IDs in common between the Illumina MouseWG-6 v2.0 and Affymetrix (Mouse 430 2.0, U74 Av2, Mu11K or Mu19K) platforms for the corresponding microbial study and matched to organ sample (small intestine or colon) were considered for identifying genes *induced* with development (weaning, adulthood, pre-weaning/adult) and/or by TLR/IL-1R signaling (WT>DKO) and inducible by the indigenous microbiota or specific bacteria of the commensal microbiota (UP after colonization of adult GF or adult CONV>GF) and genes *repressed* with development (weaning, adulthood, pre-weaning/adult) and/or by TLR/IL-1R signaling (DKO>WT) and repressible by the indigenous microbiota or specific bacteria of the commensal microbiota (DOWN after colonization of adult GF or adult GF>CONV) (Fig. S1, Dataset S1, Tables AC-AF).

Inter-data set gene enrichment

To identify statistical enrichment for probes/genes present in other data sets hypergeometric distribution analysis (one-tailed Fisher's exact test) was performed between: (1) Developmental- and TLR/IL-1R-regulated 1° data sets (Fig. 1), (2) si and colon developmental data sets (Fig. S1) and (3) developmental- or TLR/IL-1R- and microbial-regulated data sets (Fig. S1). For the latter statistical analysis (3), only Entrez Gene IDs in common between the Illumina MouseWG-6 v2.0 and Affymetrix (Mouse 430 2.0, U74 Av2, Mu11K or Mu19K) platforms for the corresponding microbiota study were considered for hypergeometric distribution statistical analysis.

Bioinformatic Analysis (1° and 2° data sets only)

Multiple approaches (low, intermediate and high resolution potential) to gain better insights into general trends and nuances of post-natal development and role of TLR per developmental stage (Fig. 3A). This included Gene Ontology-Biological Process (http://www.geneontology.org/): statistical enrichment in broad biological processes; Ingenuity Pathways Analysis (IPA) (http://www.ingenuity.com): statistical enrichment for curated mouse phenotypes of canonical signaling and metabolic pathways; Mouse Genome Informatics (MGI) -Mammalian Phenotype (ftp://ftp.informatics.jax.org/pub/reports/index.html#pheno) (10): statistical enrichment for curated mouse phenotypes (using one-tailed version of Fisher's exact test); DAVID-Functional Annotation Clustering and Charts (http://david.abcc.ncifcrf.gov/tools.jsp), a graph theory evidence-based method to agglomerate species-specific gene/protein identifiers from a variety of public genomic resources including NCBI, PIR and Uniprot/SwissProt (11); MetaRodent-Transcriptional Regulator (https://portal.genego.com/): statistical enrichment for curated mouse phenotypes of genes in sub-networks centered on transcription factors (regulators of levels of transcription factors and transcriptional targets (activation, inhibition, unspecified) of transcription factors based on curated interactions; and RIKEN-Transcription Factors (http://genome.gsc.riken.jp/TFdb/tf_list.html) to identify the regulation (developmental and TLR/IL-1R) of all known transcription factors (TFs) (12).

Quantitative PCR

For quantitative PCR, RNA was reverse transcribed using Superscript II (Invitrogen), and cDNAs were used for PCR with Quantitect SYBR Green reagents (Qiagen, Valencia, CA) on a Stratagene MX3000 bioanalyzer (La Jolla, CA). The abundance of each cytokine mRNA was normalized to HPRT expression and compared to levels in WT intestines to calculate the fold induction. All gene specific primers were designed using Primer3 (http://frodo.wi.mit.edu/). Sequences of primers available upon request.

Immunohistochemistry and Immunofluorescence

For measurement of the number of cells in S phase mice were administered 1 mg/ml of 5'-bromo-2'-deoxyuridine (BrDU) in PBS. Intestines were excised at 2 hrs post injection

and distal small intestine was fixed in 10% neutral formalin buffer and embedded in paraffin. Immunohistochemistry was performed using a BrDU staining kit from BD Biosciences. For immunohistochemistry for vWF Kit, c-Kit and smooth muscle actin (SMA) (all Santa Cruz Biotechnology). Tissues were counterstained with hematoxylin. For immunofluorescence, intestines were excised from mice. Luminal content was removed by flushing with ice-cold PBS using a blunted 18G needle and 10cc syringe. A small portion of the intestine was immediately embedded in OCT and snap frozen on dry ice. OCT cassettes were stored at -80°C until microtome cutting. Cuts were made at 8 μm thickness and placed on SuperFrost slides, allowed to dry overnight at room temperature, and then stored at -80°C until thawed for staining. Slides were fixed in ice-cold acetone for 3'. Sections were blocked with avidin and 5% normal donkey serum for 45' at RT, followed by biotin and 2% hydrogen peroxide (to quench endogenous peroxidase activity). For CD4, CD8, and IgA immunoflourescence, primary antibodies against either CD4, CD8, or IgA were used (60'; RT) followed by incubation with biotinylated $F(ab')_2$ Mouse-anti-IgG (1:2000) (Jackson) (60'; RT), and 30' incubation with streptavidin-HRP (1:2000) and then 8' incubation with tyramide-Cy3. Rat IgG were used as an isotype control. Tissues were counterstained with DAPI, coverslips mounted with Flouromount G, and visualized using standard fluorescence microscopy. Reagents kindly provided by Dr. Akiko Iwasaki.

16S ribosomal RNA Analysis of the Intestinal Microbiota

At time of sacrifice, colonic luminal samples were snap frozen in liquid nitrogen and stored at -80°C for later analysis. DNA was prepared from frozen colonic lumenal samples (n = 95) using a previously validated protocol (13). Amplicons spanning variable region 2 (V2) of the bacterial 16S rRNA gene, were generated by PCR using primers containing sample-specific, error-correcting barcodes. Amplicons from the different fecal DNA samples were subsequently pooled and subjected to multiplex pyrosequencing with a 454 instrument (Roche Life Sciences) using FLX Titanium chemistry. A total of 365,795 reads were generated (3621±438 reads per sample; average read length 340bp). De-multiplexed reads were subject to Operational Taxonomic Unit (OTU) picking using standard OTUPipe quality filtering as implemented in QIIME version 1.5 (14). Samples were rarefied to 400 sequences per sample, excluding one sample with insufficient sequences from subsequent analyses. Representative sequences for picked OTUs were classified taxonomically using RDP 2.4. PCoA plots were generated from a matrix of pairwise UniFrac distances. Samples were placed in three temporal bins: pre-weaning (P16-21), weaning (P25-P27) and adult (P40-P126). ANOVA was performed to identify taxa whose proportional representation in fecal microbiota varied significantly during these three phases of life, and to determine whether variation was affected by host genotype.

Supplemental References

1. Comelli EM et al. (2008) Multifaceted transcriptional regulation of the murine intestinal mucus layer by endogenous microbiota. *Genomics* 91:70–77.

2. Mutch DM et al. (2004) Impact of commensal microbiota on murine

gastrointestinal tract gene ontologies. *Physiol Genomics* 19:22–31.

3. Rawls JF, Mahowald MA, Ley RE, Gordon JI (2006) Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell* 127:423–433.

4. Stappenbeck TS, Hooper LV, Gordon JI (2002) Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *Proc Natl Acad Sci USA* 99:15451–15455.

5. Lecuit M, Sonnenburg JL, Cossart P, Gordon JI (2007) Functional genomic studies of the intestinal response to a foodborne enteropathogen in a humanized gnotobiotic mouse model. *J Biol Chem* 282:15065–15072.

6. Hooper LV et al. (2001) Molecular analysis of commensal host-microbial relationships in the intestine. *Science* 291:881–884.

7. Cash HL, Whitham CV, Behrendt CL, Hooper LV (2006) Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* 313:1126–1130.

8. Stappenbeck TS, Mills JC, Gordon JI (2003) Molecular features of adult mouse small intestinal epithelial progenitors. *Proc Natl Acad Sci USA* 100:1004–1009.

9. Chen R, Li L, Butte AJ (2007) AILUN: reannotating gene expression data automatically. *Nat Meth* 4:879.

10. Shaw DR (2009) Searching the Mouse Genome Informatics (MGI) resources for information on mouse biology from genotype to phenotype. *Curr Protoc Bioinformatics* Chapter 1:Unit1.7.

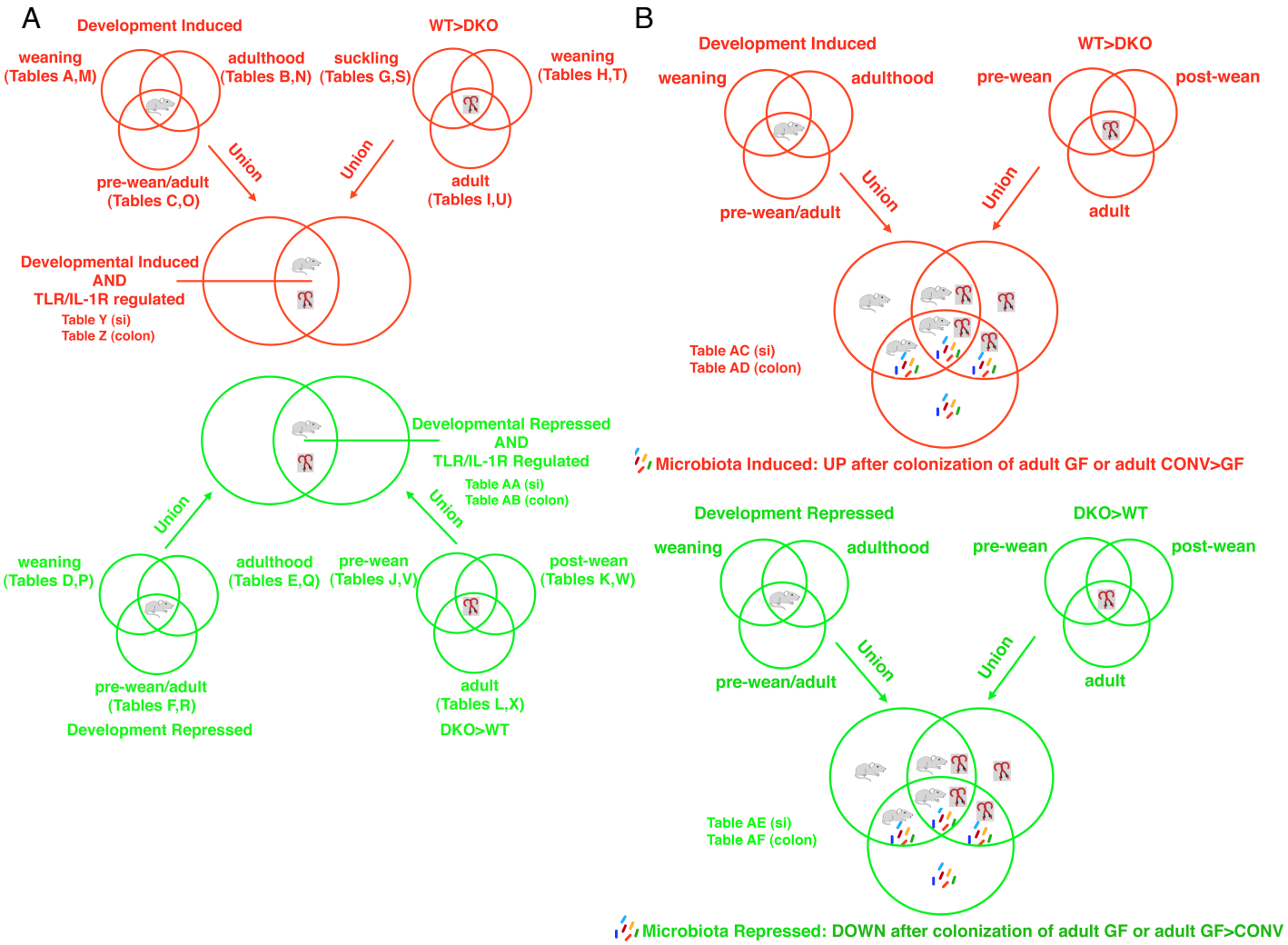
11. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.

12. Roach JC et al. (2007) Transcription factor expression in lipopolysaccharide-activated peripheral-blood-derived mononuclear cells. *Proc Natl Acad Sci USA* 104:16245–16250.

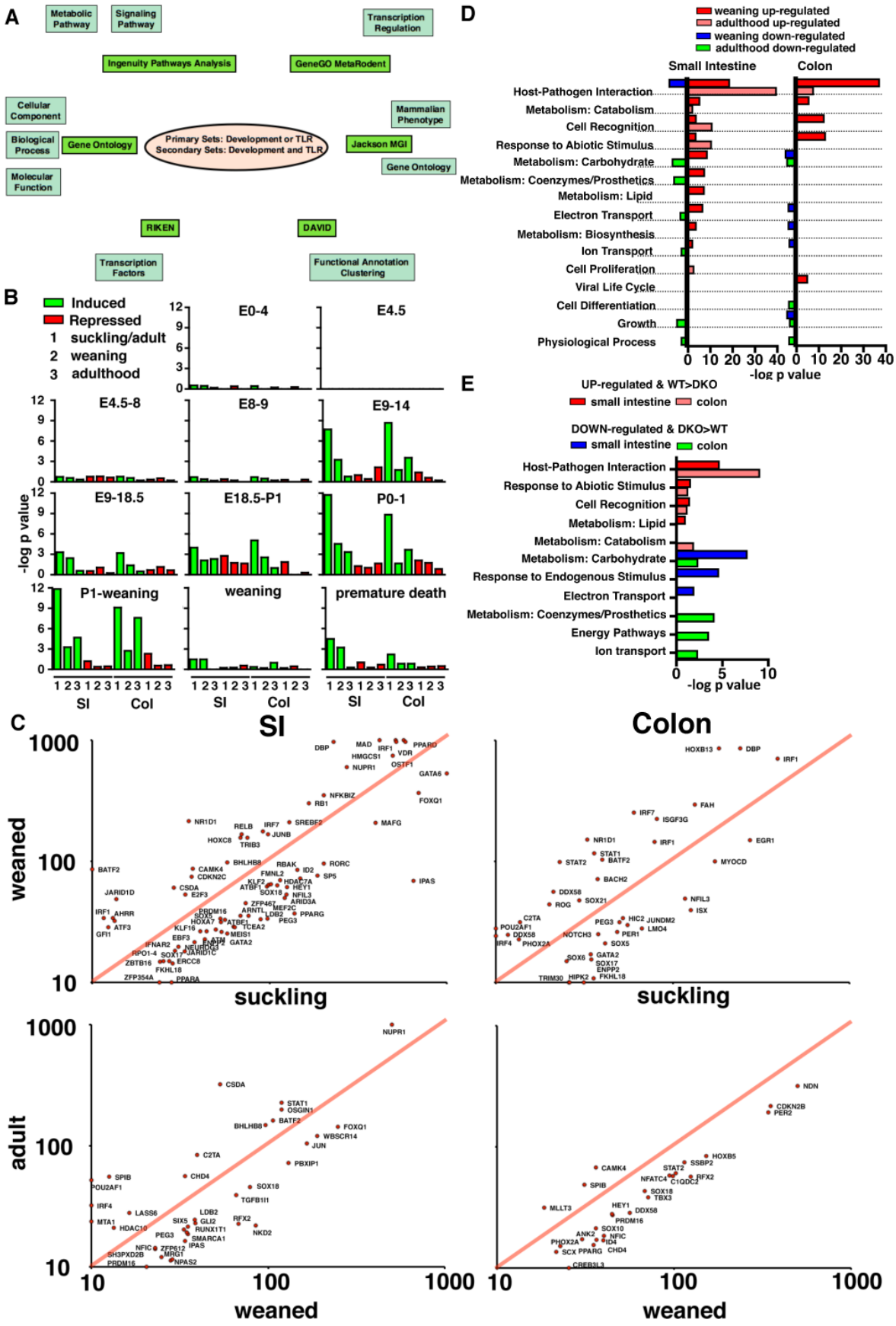
13. Ridaura VK et al. (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 341:1241214.

14. Caporaso JG et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Meth* 7:335–336.

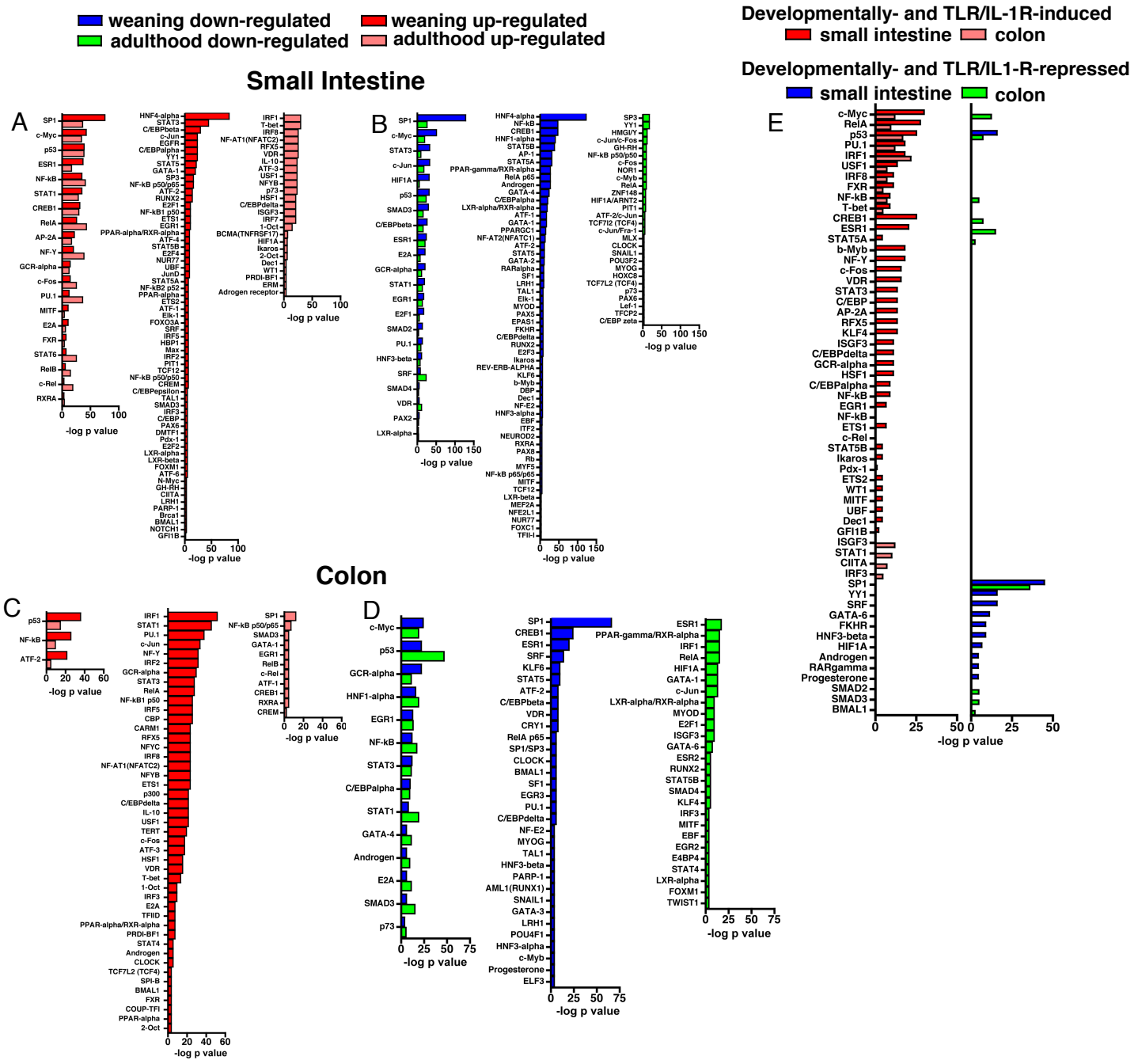
Supplemental Figure 1



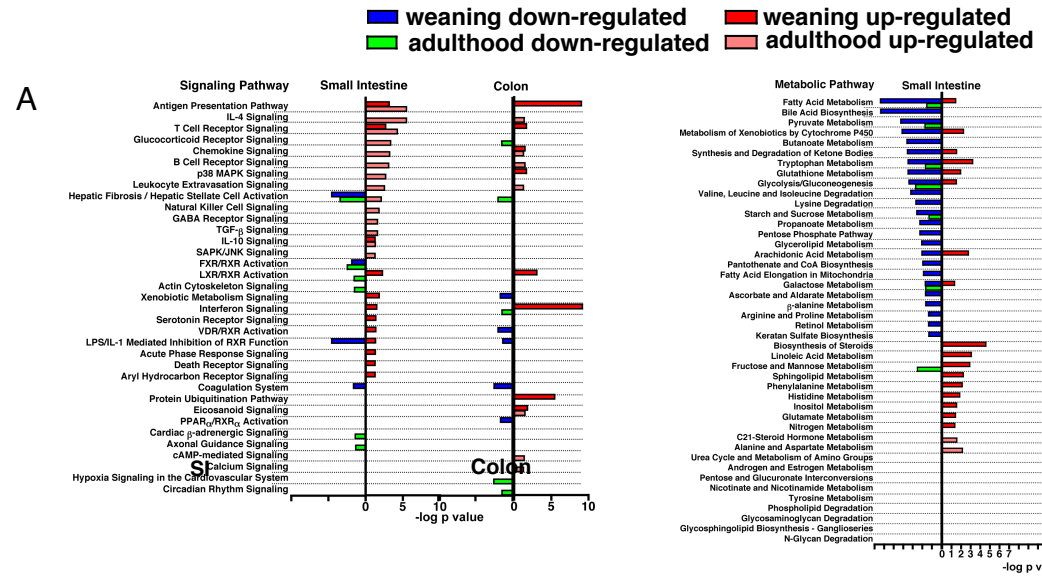
Supplemental Figure 2



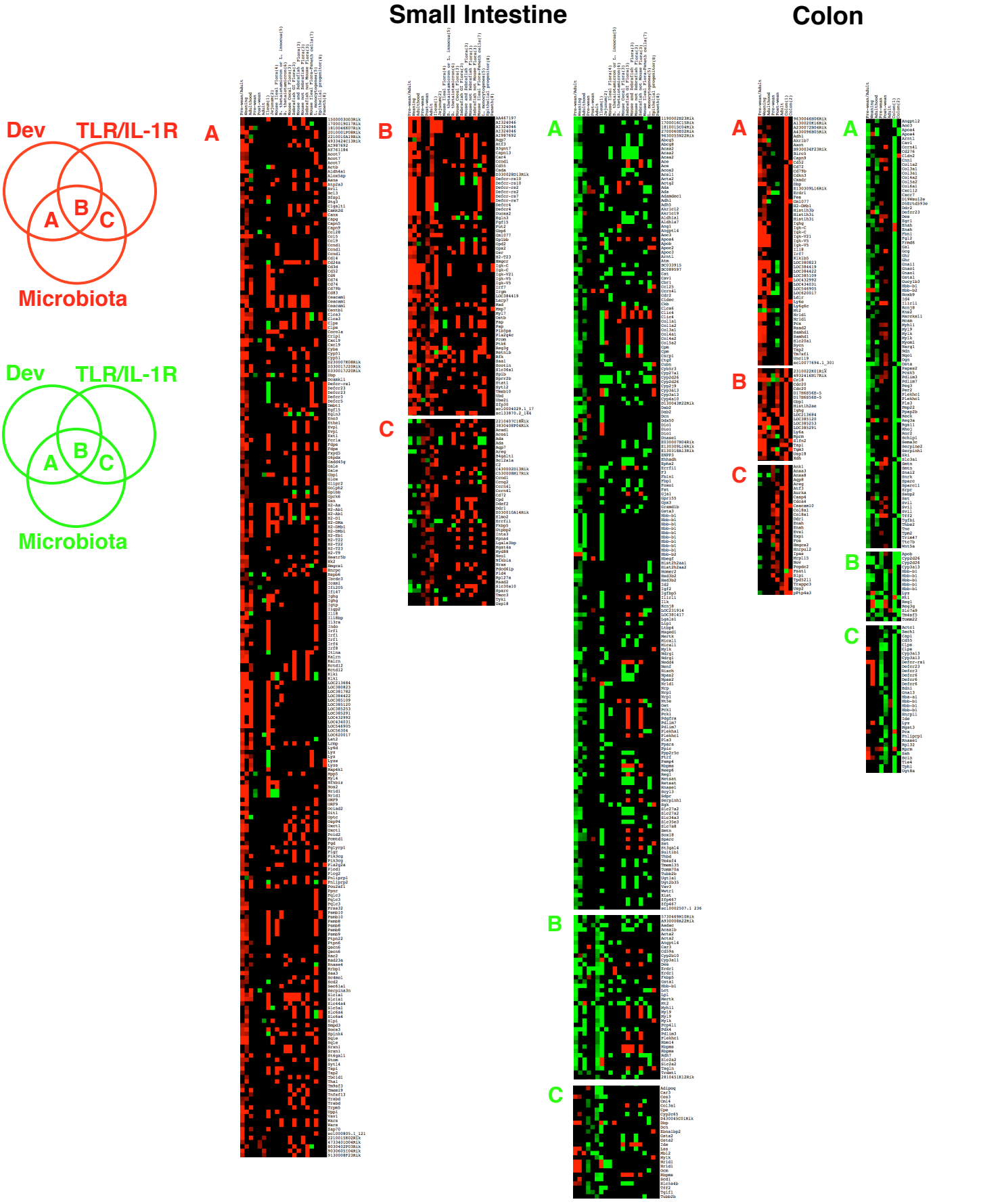
Supplemental Figure 3



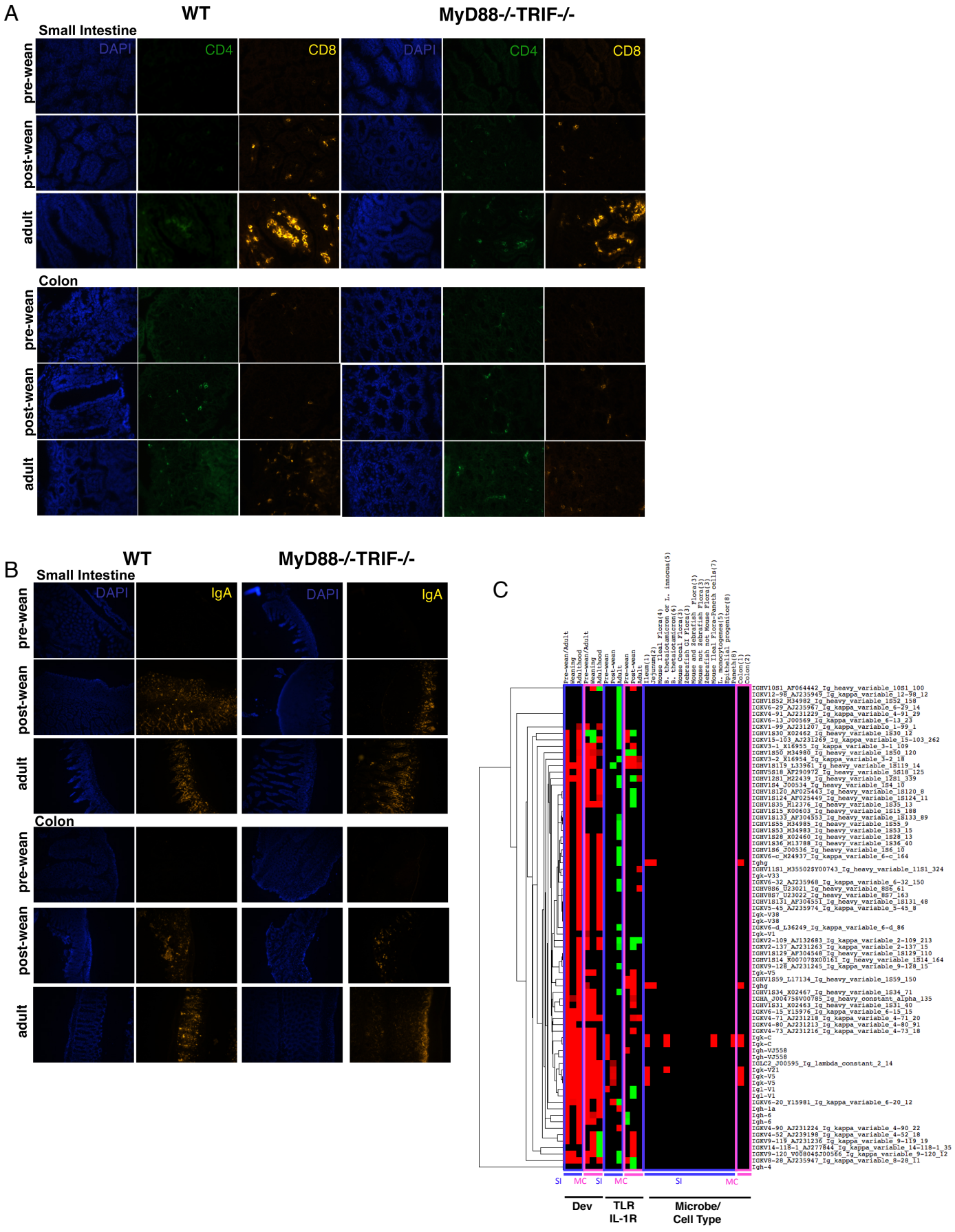
Supplemental Figure 4



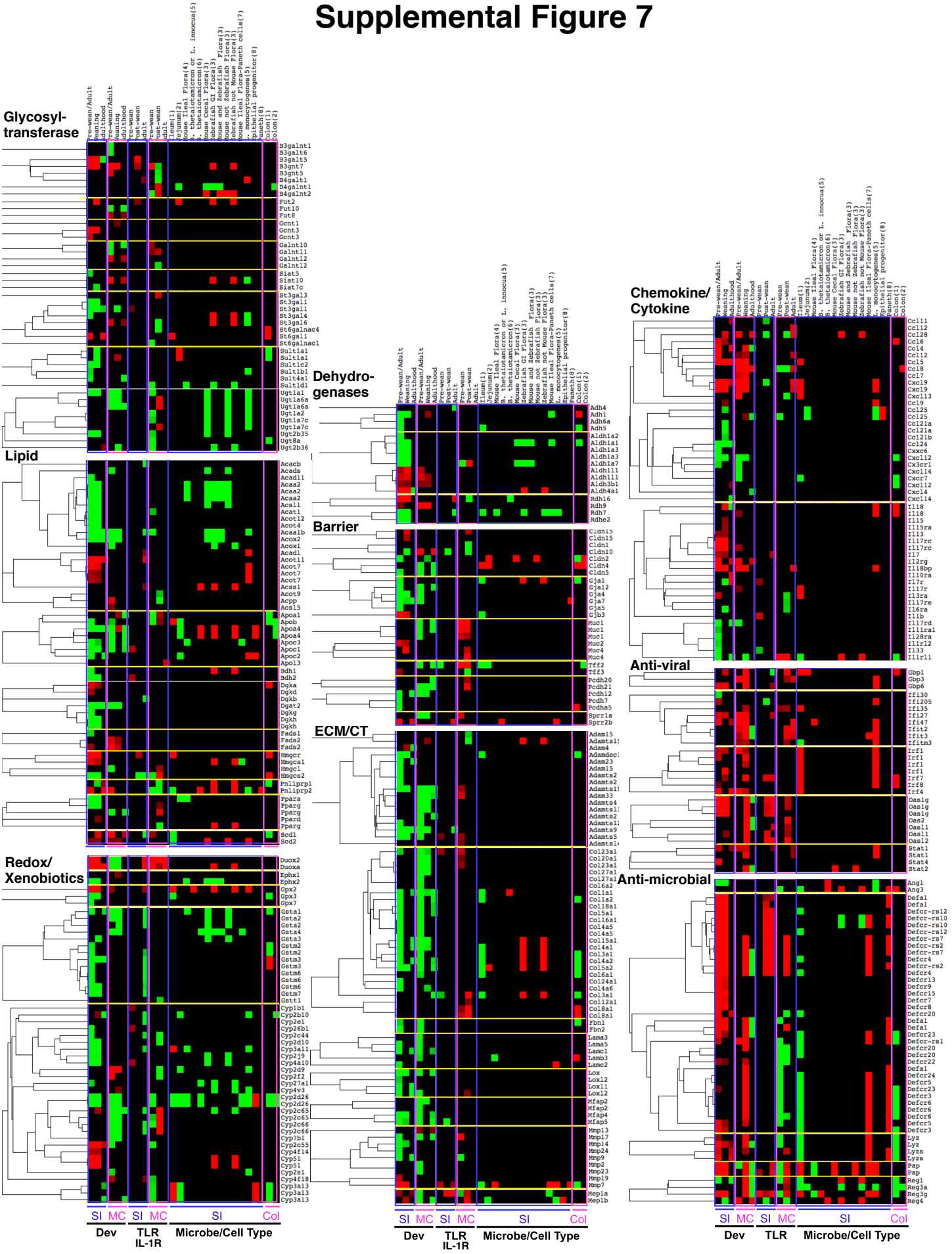
Supplemental Figure 5



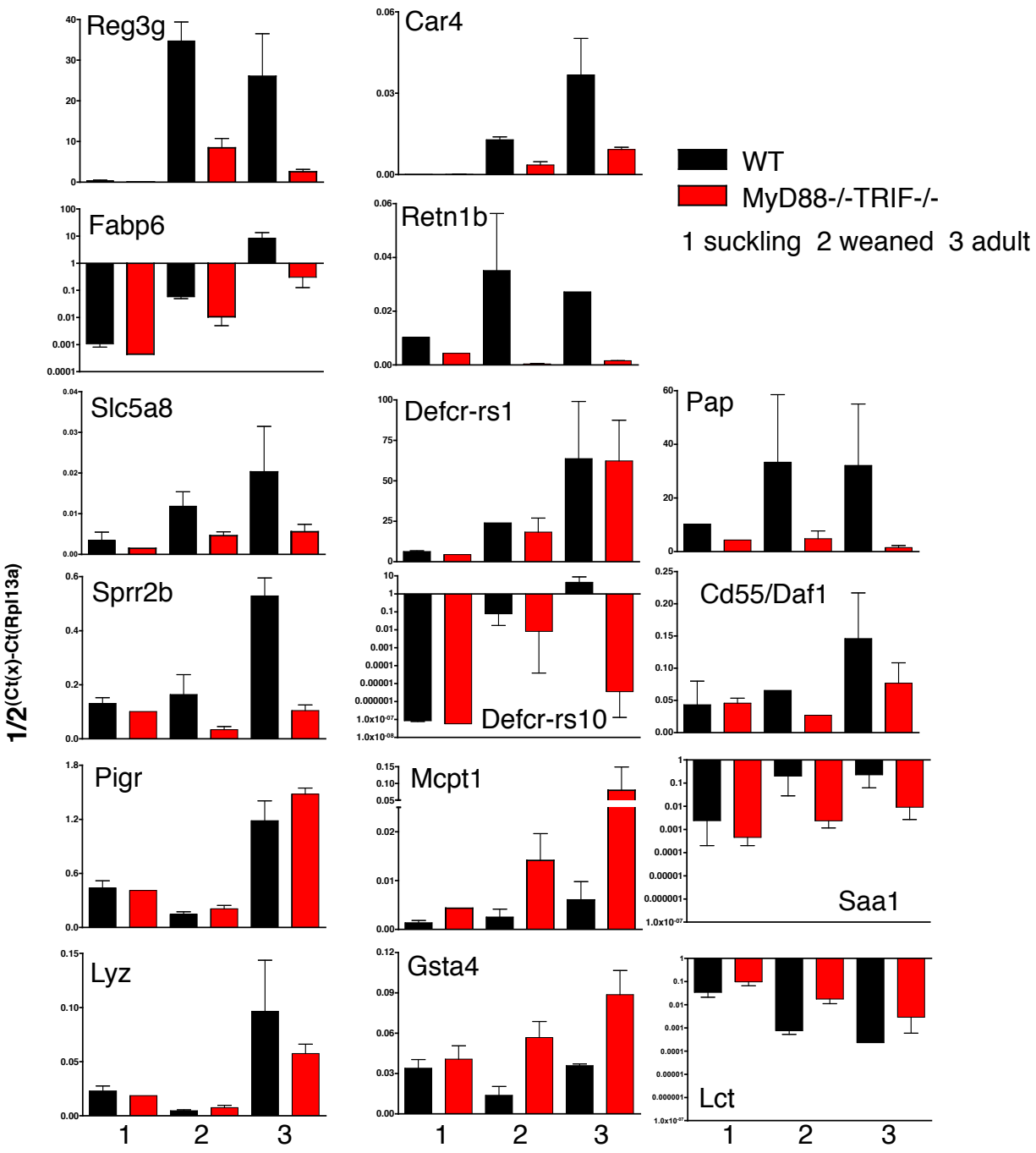
Supplemental Figure 6



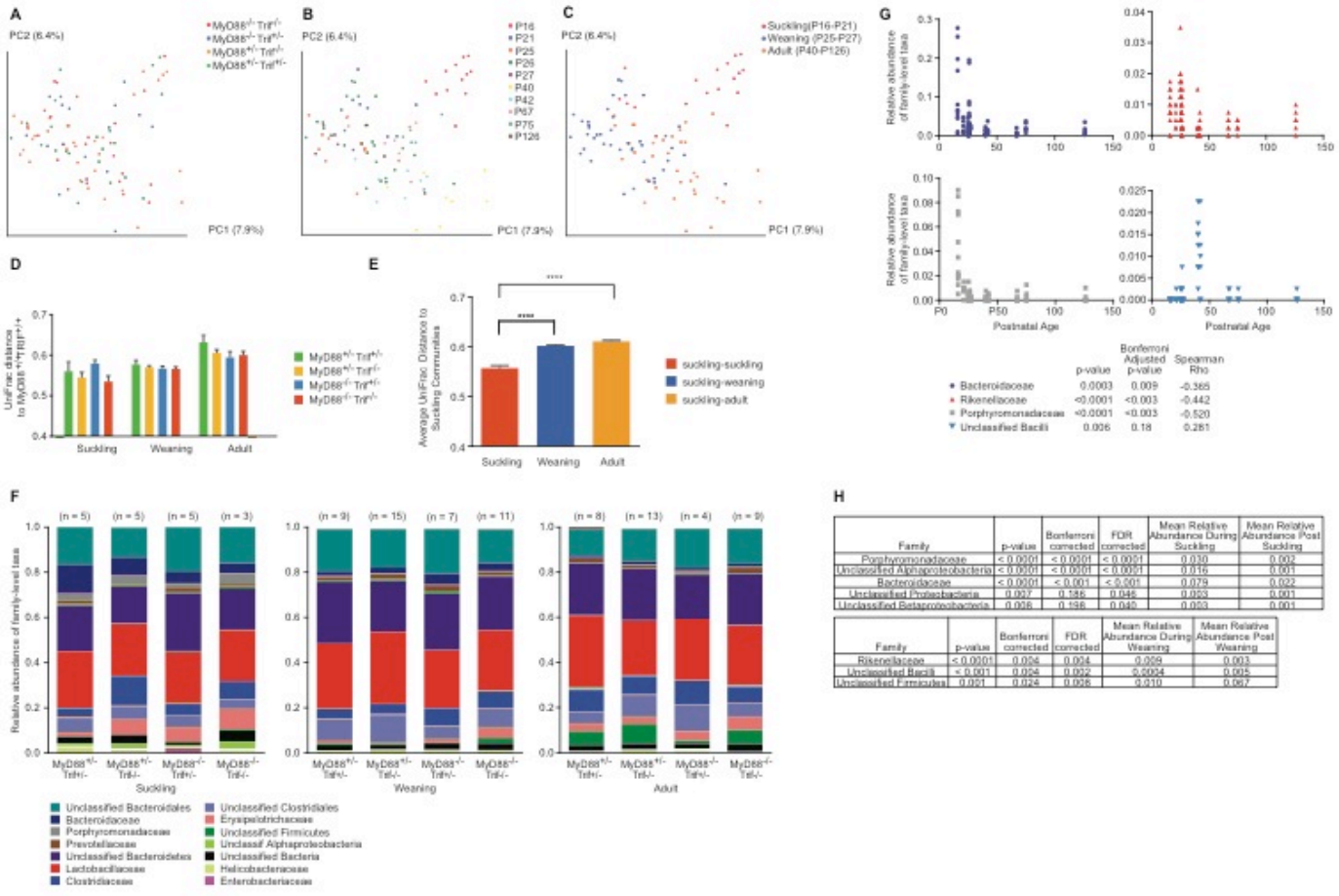
Supplemental Figure 7



Supplemental Figure 8



Supplemental Figure 9



H

Family	p-value	Bonferroni corrected	FDR corrected	Mean Relative Abundance During Suckling	Mean Relative Abundance Post Suckling
Porphyromonadaceae	< 0.0001	< 0.0001	< 0.0001	0.030	0.002
Unclassified Alphaproteobacteria	< 0.0001	< 0.0001	< 0.0001	0.016	0.001
Bacteroidaceae	< 0.0001	< 0.001	< 0.001	0.079	0.022
Unclassified Proteobacteria	0.007	0.186	0.046	0.003	0.001
Unclassified Betaproteobacteria	0.008	0.196	0.040	0.003	0.001

Family	p-value	Bonferroni corrected	FDR corrected	Mean Relative Abundance During Weaning	Mean Relative Abundance Post Weaning
Rikenellaceae	< 0.0001	0.004	0.004	0.009	0.003
Unclassified Bacilli	< 0.001	0.024	0.022	0.0004	0.005
Unclassified Firmicutes	0.001	0.024	0.026	0.010	0.067

Supplemental Figure Legends

Fig. S1. Secondary and tertiary datasets. (*A*) Schematic of creation of secondary data sets from primary developmental and TLR/IL-1R regulated data sets (Tables A-X in Dataset S1). (*B*) Schematic of creation of tertiary data sets from primary developmental and TLR/IL-1R regulated data sets with the inclusion of meta-analysis of microbiota regulated genes. Datasets can be found in Dataset S1.

Fig. S2. Bioinformatic analysis of postnatal developmentally regulated genes. (*A*) Schematic of bioinformatics tools used to analyze primary and secondary data sets (see *SI* Materials and Methods for details. (*B*) Enrichment of developmentally regulated genes within Mammalian Phenotype lethality phenotypes at pre- and post- natal stages. (*C*) Expression of genes encoding for RIKEN TF factors per postnatal developmental stage. Axes are expression in array hybridization units. Enrichment of developmentally induced and repressed genes (*D*) and those coordinately regulated by TLR/IL-R (TIR) signaling (*E*) by Gene Ontology analysis.

Fig. S3. Analysis of developmental and TLR/IL-1R regulated data sets for transcriptional subnetworks and RIKEN transcription factors. Primary data sets of developmentally induced (*A*) and repressed (*B*) genes in the small intestine and induced (*C*) and repressed (*D*) in the colon, in addition to secondary data sets of coordinate regulation of developmental and TLR/IL-1R regulated genes (*E*) were analyzed by GeneGo MetaRodent bioinformatics analysis to identify statistically enriched transcriptional subnetworks.

Fig. S4. Analysis of developmental and TLR/IL-1R regulated data sets by Gene Ontology and Ingenuity Pathways Analysis (IPA). Primary data sets of developmentally induced and repressed (*A*), TLR/IL-1R regulated genes at each postnatal developmental stage (*C*) and secondary data sets of coordinate regulation of developmental and TLR/IL-1R regulated genes (*D*) were analyzed by Gene Ontology (panel *B*) and IPA Signaling and Metabolic Pathways (panels *A*, *C* and *D*).

Fig. S5. Coordinate regulation of developmentally-, TLR/IL-1R- and microbiota-mediated transcription. Venn diagram schematically representing the classification of genes in primary data sets (Fig. S1) that also demonstrate coordinate regulation by the microbiota based on meta-analysis (see *SI* Material and Methods for details) (left panel). Heat maps of genes/probes that fall into one of three categories of coordinate developmental and microbiota (A), TLR/IL-1R and microbiota (B) or developmental, TLR/IL-1R and microbiota (C) regulation. “Pre-wean/Adult”, “Weaning” and “Adulthood” columns are developmentally primary data sets where red boxes indicate probes/genes induced with that developmental transition and green boxes indicate probes/genes repressed with that developmental transition. “Pre-wean”, “post-wean” and “adult” are TLR/IL-1R primary data sets where red boxes indicate probes/genes greater

expressed in WT vs. DKO intestines at that developmental stage and green boxes indicate probes/genes greater expressed in DKO vs. WT intestines at that developmental stage. Columns ending with numbers in parenthesis (numbers refer to data sets from the study of corresponding Supplemental Reference) are microbe-regulated data sets. Red boxes indicate genes up-regulated by the microbiota or greater expressed in conventionally raised vs. germ-free mice in the particular study and where green boxes indicate genes down-regulated by the microbiota or lesser expressed in conventionally raised vs. germ-free mice (see *SI* Material and Methods for details).

Fig. S6. Developmental and TLR/IL-1R regulation of adaptive immune responses. Immunofluorescence staining against CD4 (green) and CD8 (yellow) (*A*) or IgA (yellow) (*B*) per developmental stage and genotype in the small intestine and colon. (C) Heatmap of probes in primary data sets encoding immunoglubulin light and heavy chain genes. "Pre-wean/Adult", "Weaning" and "Adulthood" columns are developmentally primary data sets where red boxes indicate probes/genes induced with that developmental transition and green boxes indicate probes/genes repressed with that developmental transition. "Pre-wean", "post-wean" and "adult" are TLR/IL-1R primary data sets where red boxes indicate probes/genes greater expressed in WT vs. DKO intestines at that developmental stage and green boxes indicate probes/genes greater expressed in DKO vs. WT intestines at that developmental stage. Columns ending with numbers in parenthesis (numbers refer to data set from study of corresponding Supplemental Reference) are microbe-regulated data sets. Red boxes indicate genes up-regulated by the microbiota or greater expressed in conventionally raised vs. germ-free mice in the particular study and where green boxes indicate genes down-regulated by the microbiota or lesser expressed in conventionally raised vs. germ-free mice (see *SI* Material and Methods for details).

Fig. S7. Developmental, TLR/IL-1R and microbiota regulation of transcription by biological process. Heatmap of probes in primary data sets manually curated into biological processes/gene family as indicated. "Pre-wean/Adult", "Weaning" and "Adulthood" columns are developmentally primary data sets where red boxes indicate probes/genes induced with that developmental transition and green boxes indicate probes/genes repressed with that developmental transition. "Pre-wean", "post-wean" and "adult" are TLR/IL-1R primary data sets where red boxes indicate probes/genes greater expressed in WT vs. DKO intestines at that developmental stage and green boxes indicate probes/genes greater expressed in DKO vs. WT intestines at that developmental stage. Columns ending with numbers in parenthesis (numbers refer to data set from study of corresponding Supplemental Reference) are microbe-regulated data sets. Red boxes indicate genes up-regulated by the microbiota or greater expressed in conventionally raised vs. germ-free mice in the particular study and where green boxes indicate genes down-regulated by the microbiota or lesser expressed in conventionally raised vs. germ-free mice (see *SI* Material and Methods for details). ECM/CT = Extracellular Matrix/Connective Tissue.

Fig. S8. RT-QPCR confirmation of selected genes. Relative mRNA expression of indicated genes (compared to Rpl13a) at the small intestine over developmental stage in WT and MyD88-/-TRIF-/- littermates (N=4-6 mice). Error bars = SEM.

Fig. S9. 16S rRNA-based analyses of the impact of postnatal development and TLR/IL-1R deficiency on colonic microbiota configuration. PCR amplicons from the V2 region of bacterial 16S rRNA genes present in the fecal microbiota of WT ($MyD88^{+/-}TRIF^{+/-}$), $MyD88^{-/-}TRIF^{+/-}$, $MyD88^{+/-}TRIF^{-/-}$ and $MyD88^{-/-}TRIF^{-/-}$ mice of various ages were subjected to multiplex pyrosequencing. Communities were compared using unweighted UniFrac. (*A-C*) Principal Coordinates analysis (PCoA) across all ages (panel A), at the indicated postnatal days (panel B) and at three phases of life (panel C). (*D*) Average UniFrac distance (±SEM) between the fecal microbiota of WT mice (reference control) versus the fecal microbiota of mice with the indicated genotype, at the indicated life phase. None of the distances are significantly different from the reference control (*E*) Average (±SEM) unweighted UniFrac distance of weaning and adult mice relative to suckling animals (data from mice from the four genotypes combined in each age bin) ****, $p<0.0001$ (Unpaired two-tailed *t-test* with 1,000 Monte Carlo simulations) (*F*) Family-level taxonomic profiles for mice of the indicated genotype at the indicated life phase, classified using RDP 2.4. (*G*) Spearman correlations showing family-level taxa whose relative abundance significantly correlates with postnatal age (data from mice belonging to all four genotypes combined for the analysis). (H) Statistically significant differences in family-level taxa abundances between samples collected during suckling and those collected at weaning and adult, (top table) and between weaning and adult time-points (bottom table). Analyses were performed using ANOVA followed by multiple hypothesis correction as indicated.

Dataset S1. Primary, secondary and tertiary data sets. Primary developmental and TLR/IL-1R data sets (Tables A-X), secondary coordinate developmental and TLR/IL-1R regulated data sets (Tables Y, Z, AA, AB) and tertiary coordinate developmental, TLR/IL-1R and microbiota data sets (Tables AC-AF). DEI = DAVID Entrez ID. Raw hybridization expression across all samples for microarray 1 and 2, and average ratio (both microarrays) per probe for developmental primary data sets are included. For example, column "WT.si.d26.16" refers to the small intestinal weaning transition with the ratio of weaned/suckling in WT mice. Table "Master by gene" is a compilation of all probes that were significantly present in one of the primary developmental or TLR/IL-1R data sets.

Dataset S2. GeneGo MetaRodent and RIKEN TF data sets. Tables starting with "MR" are exports from GeneGo MetaRodent analysis of the indicated data set. "Riken table" is a table of RIKEN TF genes manually curated into those showing coordinate regulation in developmental, TLR/IL-1R and microbe-regulated datasets. Red lettering indicates developmental induction, WT>DKO expression or microbe up-regulated or increased expression in conventionally raised compared to germ-free mice. Green lettering indicates developmental repression, DKO>WT expression or microbe down-regulated or increased expression in germ-free mice compared to conventionally raised mice. "RIKEN statistics" is a table of the percentage of probes encoding TF compared to all probes in a particular data set.