

## 1 **Hnf4a mediates microbial control of intestinal gene expression and inflammation**

2 James M. Davison<sup>1,2</sup>, Colin R. Lickwar<sup>1</sup>, Lingyun Song<sup>3</sup>, Ghislain Breton<sup>4</sup>, Gregory E. Crawford<sup>3</sup>  
3 and John F. Rawls<sup>1,‡</sup>

4 <sup>1</sup>Department of Molecular Genetics and Microbiology, Center for the Genomics of Microbial  
5 Systems, Duke University, Durham, North Carolina 27710, USA; <sup>2</sup>Department of Cell Biology and  
6 Physiology, University of North Carolina, Chapel Hill, North Carolina 27599, USA; <sup>3</sup>Department of  
7 Pediatrics, Division of Medical Genetics, Center for Genomic and Computational Biology, Duke  
8 University, Durham, North Carolina 27708, USA; <sup>4</sup>Department of Integrative Biology and  
9 Pharmacology, McGovern Medical School, Houston, Texas 77030, USA; <sup>‡</sup>Corresponding Author.

### 10 **ABSTRACT**

11 Microbiota influence diverse aspects of intestinal epithelial physiology and disease in part by  
12 controlling tissue-specific transcription of host genes. However, host genomic mechanisms  
13 mediating microbial control of host gene expression are poorly understood. Using an unbiased  
14 screening strategy, we found that the zebrafish Hepatic nuclear factor 4 alpha (Hnf4a)  
15 transcription factor specifically binds and activates a microbiota-suppressed intestinal epithelial  
16 transcriptional enhancer. Genetic analysis disclosed that zebrafish *hnf4a* activates nearly half of  
17 the genes that are suppressed by microbiota, suggesting microbiota negatively regulate Hnf4a.  
18 In support, analysis of genomic architecture in mouse intestinal epithelial cells revealed that  
19 microbiota colonization leads to activation or inactivation of hundreds of enhancers along with  
20 drastic genome-wide reduction of Hnf4a and Hnf4g occupancy. Interspecies meta-analysis  
21 suggests Hnf4a may mediate microbial contributions to inflammatory bowel disease  
22 pathogenesis. These results indicate Hnf4a plays a critical conserved role in maintaining  
23 intestinal homeostasis in response to microbiota and inflammation.

24

25 **KEY WORDS:**

26 Crohn's disease, ulcerative colitis, cis-regulatory, NR2A1, NR2A2, gnotobiotic, Fiaf, H3K4me1,

27 H3K27ac, DNase hypersensitivity, microbiome

## 28 INTRODUCTION

29 All animals face the fundamental challenge of building and maintaining diverse tissues while  
30 remaining sensitive and responsive to their environment. This is most salient in the intestinal  
31 epithelium which performs important roles in nutrient absorption and barrier function while being  
32 constantly exposed to complex microbial communities (microbiota) and nutrients within the  
33 intestinal lumen. The presence and composition of microbiota in the intestinal lumen influence  
34 diverse aspects of intestinal development and physiology including dietary nutrient metabolism  
35 and absorption, intestinal epithelial renewal, and edification of the host immune system.

36 Abnormal host-microbiota interactions are strongly implicated in the pathogenesis of  
37 inflammatory bowel diseases (IBD), including Crohn's disease (CD) and ulcerative colitis (UC)  
38 (Sartor and Wu 2016). Studies in mouse and zebrafish models of IBD have established that  
39 impaired intestinal epithelial cell (IEC) responses to microbiota are a key aspect of disease  
40 progression (Bates et al. 2007; Kamada et al. 2013; Marjoram et al. 2015). Improved  
41 understanding of the molecular mechanisms by which microbiota evoke host responses in the  
42 intestinal epithelium can be expected to lead to new strategies for preventing or treating IBD  
43 and other microbiota-associated diseases.

44 The ability of IEC to maintain their physiologic functions and respond appropriately to microbial  
45 stimuli is facilitated through regulation of gene transcription. Genome-wide comparison of  
46 transcript levels in intestinal tissue or isolated IEC from mice reared in the absence of microbes  
47 (germ-free or GF) to those colonized with a microbiota (conventionalized or CV) have revealed  
48 hundreds of genes that have significantly increased or decreased mRNA levels following  
49 microbiota colonization (Camp et al. 2014). Interestingly, many mouse genes that are  
50 transcriptionally regulated by microbiota have zebrafish homologs that are similarly responsive,  
51 suggesting the existence of evolutionarily-conserved regulatory mechanisms (Rawls et al.  
52 2004). For example, the protein hormone Angiopoetin-like 4 (Angptl4, also called Fiaf) is

53 encoded by a single ortholog in the mouse and zebrafish genomes, and microbiota colonization  
54 results in significant reductions in transcript levels in the small intestinal epithelium of both host  
55 species (Bäckhed et al. 2004; Camp et al. 2012). Whereas these impacts of microbiota on host  
56 IEC transcriptomes and their downstream consequences have been extensively documented,  
57 the upstream transcriptional regulatory mechanisms remain poorly understood.

58 Specification and tuning of gene transcription proceeds in part through interactions between  
59 transcription factors (TFs) and their sequence-specific binding to *cis*-regulatory DNA. *Cis*-  
60 regulatory regions (CRRs) harbor binding sites for multiple activating or repressing TFs and are  
61 generally associated with nucleosome depletion and specific post-translational modifications of  
62 histone proteins within adjacent nucleosomes when acting as poised (H3K4me1) or active  
63 (H3K27ac) enhancers (Creyghton et al. 2010). Antibiotic administration can impact transcript  
64 levels and histone modifications in IECs (Thaiss et al. 2016), however it's unclear if these  
65 changes are indirect effects caused by alterations to microbiota composition, direct effects of  
66 the antibiotic on host cells, or by the effects of remaining antibiotic-resistant microbiota (Morgun  
67 et al. 2015). Previous studies have shown that histone deacetylase 3 is required in IECs to  
68 maintain intestinal homeostasis in the presence of microbiota (Alenghat et al. 2013), and that  
69 overall histone acetylation and methylation in the intestine is altered by microbiota colonization  
70 (Krautkramer et al. 2016). However, the direct and specific effects of the microbiota on host  
71 CRRs and subsequent transcriptional responses in IECs remain unknown.

72 Our previous work predicted key roles for one or more nuclear receptor TFs in microbial down  
73 regulation of IEC gene expression (Camp et al. 2014), but the specific TF(s) were not identified.  
74 Nuclear receptors are ideal candidate TFs for integrating microbe-derived signals, since for  
75 many their transcriptional activity can be positively or negatively regulated by the binding of  
76 metabolic or hormonal ligands (Evans and Mangelsdorf 2014). However, the roles of nuclear  
77 receptors in host responses remain poorly understood, and no previous study has defined the

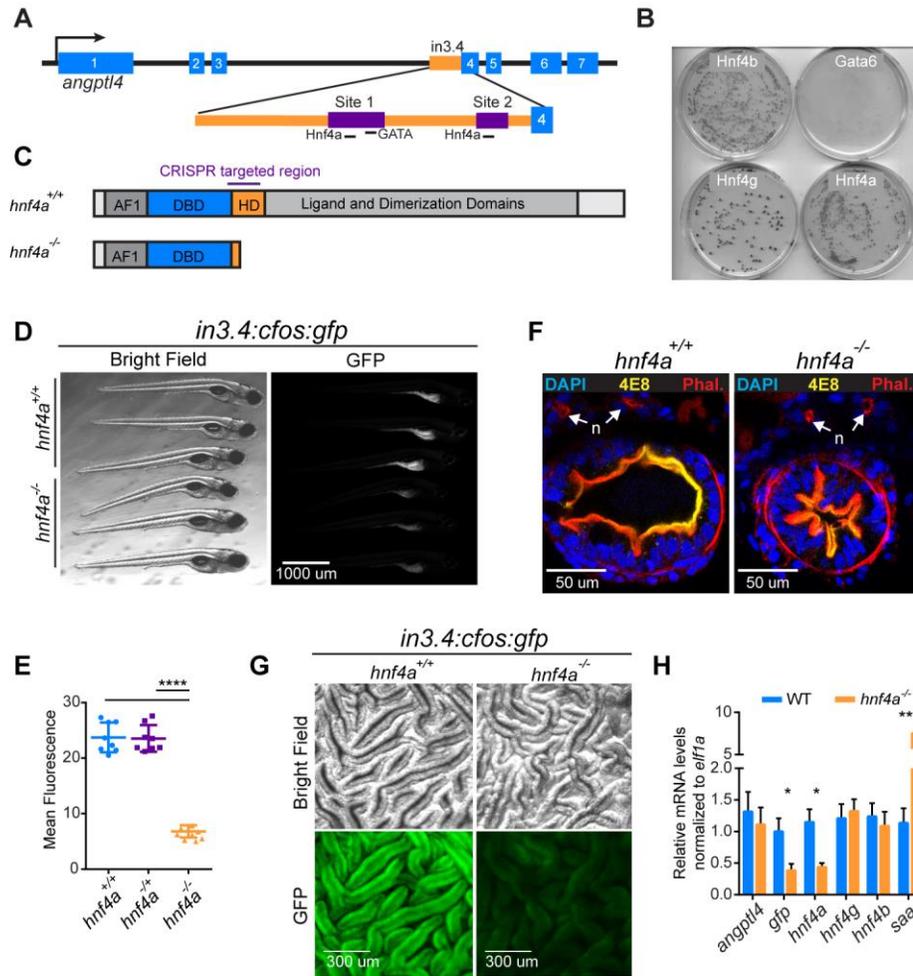
78 impact of microbiota on nuclear receptor DNA binding. Nuclear receptors are a metazoan  
79 innovation. The earliest animals encoded a single nuclear receptor orthologous to Hepatocyte  
80 nuclear factor 4 (Hnf4; nuclear receptor subfamily NR2A) (Bridgham et al. 2010). Despite  
81 subsequent duplication and diversification, distinct Hnf4 TFs remain encoded in extant animals  
82 including mammals (Hnf4a, Hnf4g) and fishes (Hnf4a, Hnf4b, Hnf4g) (Supplemental Fig. S1G).  
83 Hnf4a serves particularly important roles in IECs, where it binds CRRs and activates expression  
84 of genes involved in IEC maturation and function (Stegmann et al. 2006). IEC-specific knockout  
85 of mouse *Hnf4a* results in spontaneous intestinal inflammation similar to human IBD (Darsigny  
86 et al. 2009). In accord, genetic variants at human *HNF4A* are associated with IBD as well as  
87 colon cancer (Marcil et al. 2012; Chellappa et al. 2016). Hnf4a is predicted to bind a majority of  
88 IBD-linked CRRs and regulate IBD-linked genes (Haberman et al. 2014; Meddens et al. 2016).  
89 Similarly, genetic variants near human *HNF4G* have been associated with obesity and IBD  
90 (Franke et al. 2007; Berndt et al. 2013). Importantly, these diverse roles for Hnf4 TFs in host  
91 physiology have only been studied in host animals colonized with microbiota. Therefore, the role  
92 of Hnf4 in host-microbiota interactions and the implications for human IBD remain unknown.

## 93 RESULTS

### 94 ***hnf4a* is essential for transcriptional activity from a microbiota-suppressed cis-regulatory** 95 **DNA region**

96 To identify transcriptional regulatory mechanisms underlying microbial control of host gene  
97 expression, we took advantage of a previously identified microbiota-responsive CRR termed  
98 in3.4 located within the third intron of zebrafish *angptl4* (Fig. 1A). A GFP reporter construct  
99 under control of in3.4 termed *in3.4:cfos:gfp* drives tissue specific expression of GFP in zebrafish  
100 IEC and is suppressed by microbiota colonization, recapitulating the microbial suppression of  
101 zebrafish *angptl4* (Camp et al. 2012). However, the factor(s) that mediate microbial suppression

**Figure 1**



**FIGURE 1. Zebrafish *hnf4a* is required for robust *in3.4:cfos:gfp* activity**

(A) Schematic of the microbiota-suppressed zebrafish enhancer, *in3.4*, highlighting the regions required for intestinal activity (purple) which both contain putative Hnf4 binding sites (Site 1 and Site 2) (Camp et al. 2012). (B) Image of 4 plates from the Y1H assay showing the *hnf4* family of transcription factors capable of binding *in3.4* and driving expression of the antibiotic resistance reporter gene. (C) Hnf4a<sup>+/+</sup> and Hnf4a<sup>-/-</sup> protein cartoons showing the DNA binding domain (DBD) and hinge domain (HD). We characterized the two with the largest lesions, a -41 deletion in the hinge domain and a +25 insertion in the hinge domain which both result in frame-shift early-stop codons and significantly reduced transcript. (D) Stereofluorescence GFP and bright field microscopy showing representative *hnf4a*<sup>+/+</sup> (top 3) and *hnf4a*<sup>-/-</sup> (bottom 3) 6dpf *in3.4:cfos:gfp* zebrafish. Genotype was blinded and samples were arranged by intensity of GFP fluorescence. (E) GFP fluorescence (mean ± sem) in *hnf4a*<sup>+/+</sup> (n = 8), *hnf4a*<sup>+/-</sup> (n = 8) and *hnf4a*<sup>-/-</sup> (n = 8) 6dpf *in3.4:cfos:gfp* zebrafish (Two-tailed t-test: t = 17.84, 16.51, respectively, df = 14, and \*\*\*\* p < 0.0001). (F) Confocal microscopy showing representative axial cross sections in 6dpf *hnf4a*<sup>+/+</sup> (n = 4) and *hnf4a*<sup>-41/-41</sup> (n = 4) larval zebrafish. 4e8 antibody (yellow) labels the intestinal brush border, DAPI (blue) and phalloidin (red), and nephros (n). (G) Bright field microscopy (top) and stereofluorescence GFP (bottom) for representative *hnf4a*<sup>+/+</sup> (n = 3) (left) and *hnf4a*<sup>-/-</sup> (n = 3) (right) dissected intestinal folds from adult *in3.4:cfos:gfp* zebrafish. (H) Relative mRNA levels (mean ± sem) in *hnf4a*<sup>+/+</sup> (n = 3) and *hnf4a*<sup>-/-</sup> (n = 3) adult zebrafish intestinal epithelial cell as measured by qRT-PCR. Two-tailed t-test: t = 0.93, 5.22, 6.56, 10.65, 0.75, 0.94 respectively, df = 4, and \* p < 0.05, \*\*\* p < 0.001). See also Supplemental Figures S1 and S2.

103 of in3.4 were unknown. Using a yeast one-hybrid (Y1H) assay, we tested the capacity of 150  
104 TFs expressed in the zebrafish digestive system to bind in3.4 (Supplemental Fig. S1A,B;  
105 Supplemental Table S1) and detected an interaction only with *hnf4a*, *hnf4b*, and *hnf4g* (Fig. 1B).  
106 When either of two predicted Hnf4 motifs in in3.4 are mutated, the Hnf4-in3.4 interaction in the  
107 Y1H assay and intestinal GFP expression in *in3.4:cfos:gfp* zebrafish is strongly reduced  
108 (Supplemental Fig. S1C-F). Interestingly, while *gata4*, *gata5*, and *gata6* have predicted motifs in  
109 in3.4 (Camp et al. 2012) these TFs did not interact in the Y1H assay. This suggests that HNF4  
110 TFs are capable of binding in3.4 directly and HNF4 binding sites are necessary for directing  
111 in3.4-based transcription *in vitro* and in the intestine.

112 We hypothesized that the *hnf4* transcription factor family is required to mediate microbial  
113 suppression of in3.4 activity. Although the Y1H assay demonstrated all 3 zebrafish Hnf4  
114 members are capable of binding in3.4, we concentrated our efforts on understanding the  
115 function of *hnf4a* because it is the most highly conserved Hnf4 family member (Supplemental  
116 Fig. S1G; Supplemental Table S10) and has well-documented roles in intestinal physiology  
117 (San Roman et al. 2015). To that end, we generated *hnf4a* mutant zebrafish using the  
118 CRISPR/Cas9 system (Fig. 1C; Supplemental Fig. S2A-C,E). Unlike mouse whole-animal *Hnf4a*  
119 knockout animals that fail to initiate visceral endoderm development and die during early  
120 embryogenesis (Duncan et al. 1997), zebrafish *hnf4a* mutant animals are viable and survive to  
121 sexual maturity (Supplemental Fig. S2D) providing new opportunities to study the roles of Hnf4a  
122 in host-microbiota interactions.

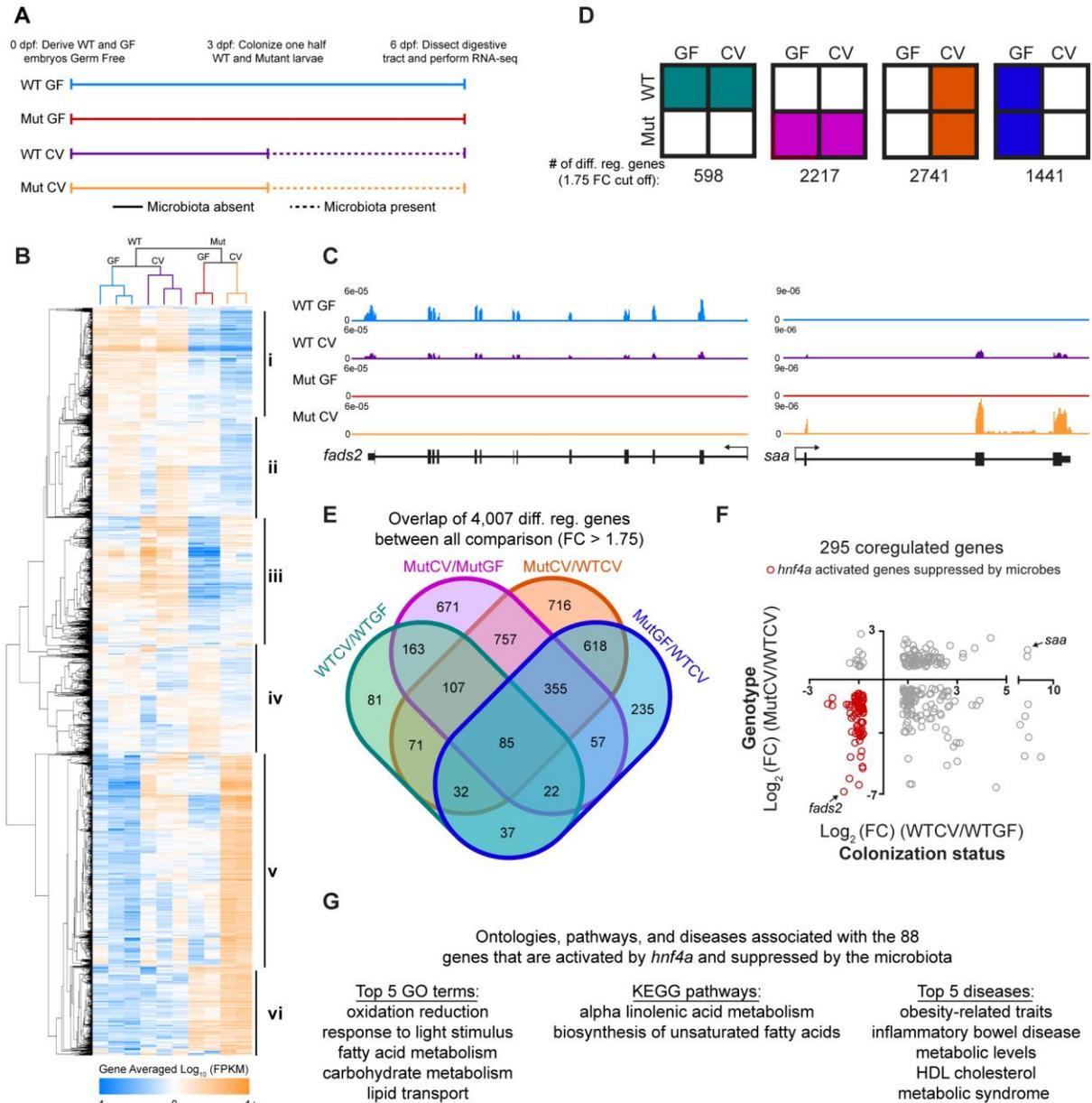
123 To determine if *hnf4a* is essential for in3.4 activity, we crossed mutant *hnf4a* alleles into the  
124 *in3.4:cfos:gfp* transgenic reporter line. GFP expression was significantly reduced in the absence  
125 of *hnf4a* suggesting that *hnf4a* activates in3.4 (Fig. 1D,E,G,H). This loss of GFP expression in  
126 *hnf4a*<sup>-/-</sup> mutants was not associated with overt defects in brush border development or epithelial  
127 polarity in larval stages (Fig. 1F), nor in the establishment of intestinal folds during adult stages

128 (Fig. 1G). However, intestinal lumen of mutant larvae was reduced in size at 6 days post  
129 fertilization (dpf) compared to WT siblings (Fig. 1F; Supplemental Fig. S2F). Together, these  
130 data indicate *hnf4a* is essential for robust activity of a microbiota-suppressed CRR. Unlike  
131 *in3.4:cfos:gfp*, *angptl4* is expressed in multiple tissues and cell types (Camp et al. 2012). To  
132 determine if intestinal *angptl4* expression is dependent on *hnf4a* function, we isolated RNA from  
133 IECs from *hnf4a*<sup>+/+</sup> and *hnf4a*<sup>-/-</sup> adult *in3.4:cfos:gfp* zebrafish and performed qRT-PCR. Adult  
134 IECs (AIECs) from *hnf4a*<sup>-/-</sup> have significant reductions in mRNA for *gfp*, *fabp2* (a known Hnf4a  
135 target in human cell lines) (Klapper et al. 2007), and *hnf4a* compared to *hnf4a*<sup>+/+</sup> controls.  
136 However, *angptl4* expression remained unchanged in *hnf4a*<sup>-/-</sup> AIECs compared to WT,  
137 suggesting *angptl4* transcript levels in the adult intestine are regulated by additional  
138 mechanisms and not solely from *in3.4* or Hnf4a (Fig. 1H). Transcript levels for *hnf4g* and *hnf4b*  
139 in *hnf4a*<sup>-/-</sup> AIEC were also unchanged. Together, these results establish that Hnf4a is required  
140 for *in3.4* activity in IECs and raises the possibility that Hnf4a may have broader roles in  
141 mediating host transcriptional and physiological responses to microbiota.

#### 142 **Hnf4a activates transcription of genes that are suppressed upon microbiota colonization**

143 To better define the roles of *hnf4a* in microbiota response and other aspects of digestive  
144 physiology, we used RNA-seq to compare mRNA levels from digestive tracts isolated from  
145 *hnf4a*<sup>+/+</sup> and *hnf4a*<sup>-/-</sup> zebrafish larvae in the presence (CV) or absence of a microbiota (GF; Fig.  
146 2A). Consistent with our previous studies (Rawls et al. 2004; Kanther et al. 2011), comparison of  
147 wildtype zebrafish reared under CV vs GF conditions revealed differential expression of 598  
148 genes that were enriched for processes such as DNA replication, oxidation reduction, and  
149 response to bacterium (Fig. 2B,D; Supplemental Fig. S3D; Supplemental Tables S2, S4).  
150 Strikingly, disruption of the *hnf4a* gene caused gross dysregulation of the transcriptional  
151 response to microbiota colonization, with the total number of microbiota responsive genes (CV

**Figure 2**



152 vs GF) increasing to 2,217. Furthermore, comparison of the *hnf4a* mutant (Mut) vs wild type  
 153 (WT) genotypes revealed differential expression of many genes in the CV condition (2,741  
 154 genes) and GF condition (1,441 genes) that inform a general role for Hnf4a in regulating genes  
 155 in the intestinal tract (Fig. 2D,E). Principal components analysis (Supplemental Fig. S3A) and  
 156 hierarchical clustering (Fig. 2B) of FPKM values indicated that *hnf4a* genotype had a complex

**FIGURE 2. *Hnf4a* activates the majority of coregulated genes that are suppressed by the microbiota**

(A) Schematic showing the experimental timeline for zebrafish digestive tract GF and CV *hnf4a*<sup>+/+</sup> and *hnf4a*<sup>-/-</sup> RNA-seq experiment. (B) Hierarchical relatedness tree and heatmap of differentially regulated genes in mutant and gnotobiotic zebrafish digestive tracts. Gene averaged log<sub>10</sub> FPKMs for the biological replicates are represented for each of the 4,007 differentially regulated genes. (C) Representative RNA-seq signal tracks at *fatty acid-desaturase 2 (fads2)*, *serum amyloid a (saa)* loci. (D) Summary of the total number of differentially expressed genes between indicated conditions (GF and CV) and genotype (WT and *hnf4a*<sup>-/-</sup> (Mut)). (E) 4-way Venn diagram showing overlaps between all 4,007 differentially regulated genes. (F) The 295 coregulated genes were plotted using the log<sub>2</sub> (FC) calculated in the WTGF/WTCV comparison (X-axis) and WTCV/MutCV (Y-axis). The 88 out of 98 genes that are activated by *hnf4a* but suppressed by the microbiota are highlighted (red) and (G) their GO term, KEGG pathway and disease associations are listed. See also Supplemental Figure S3.

157 contribution to regulating genes involved in both response to the microbiota and digestive  
158 physiology.

159 Because we found that *hnf4a* activates the microbiota-suppressed intestinal CRR, in3.4, we  
160 hypothesized that this may represent a general regulatory paradigm for other microbiota-  
161 influenced CRRs and genes across the genome. When we compared the 598 genes that were  
162 microbiota responsive in wildtype digestive tracts with the 2,741 genes that *hnf4a* regulates in  
163 CV digestive tracts we found these lists shared 295 genes that included *fads2* and *saa*, both of  
164 which have human orthologs that are either implicated (*FADS1/2*) or markers (*SAA*) of IBD  
165 (Plevy et al. 2013; Costea et al. 2014) (Fig. 2C-F). While loss of *Hnf4a* could be pleiotropic,  
166 strikingly, the overlap between these subsets reveals that a disproportionate 88 of the 98  
167 (~90%) microbiota-suppressed genes are activated by *hnf4a* (Fig. 2F; Supplemental Table S2).  
168 These 88 genes represent almost half of all 185 genes suppressed by the microbiota. These  
169 data suggest, like its role at in3.4, *hnf4a* plays a critical role in directly activating a large  
170 percentage of genes that are suppressed by microbial colonization. Interestingly, despite a clear  
171 enrichment of metabolism themed ontologies and pathways, we found IBD was among the top 2  
172 diseases associated with genes that are activated by *Hnf4a* but suppressed by the microbiota  
173 (Fig. 2G; Supplemental Table S11). Based on these results, we hypothesized that *Hnf4a* DNA

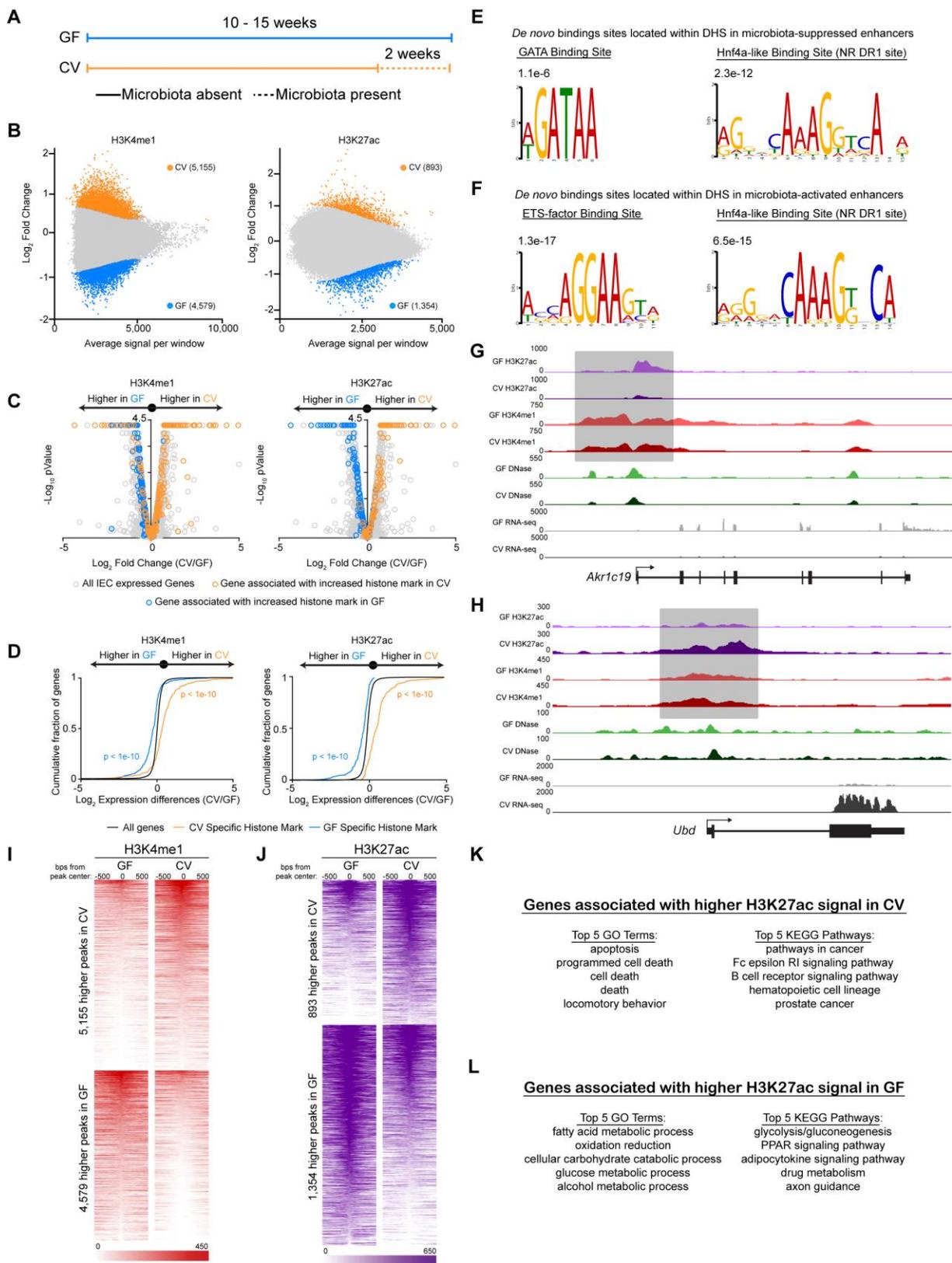
174 binding is lost upon microbial colonization within CRRs associated with microbiota-suppressed  
175 genes.

176 **Hnf4a binding sites are enriched in promoters near genes associated with microbiota-**  
177 **regulated H3K27ac marks**

178 Previous attempts to identify microbial responsive enhancers genome-wide were complicated  
179 by the lack of significant changes in chromatin DNase accessibility between GF and CV IECs  
180 from mouse colon and ileum (Camp et al. 2014). These previous findings suggested other  
181 chromatin dynamics may be involved in regulating the IEC response to microbiota. We therefore  
182 sought to provide a genomic context for understanding how the microbiota alter Hnf4a activity  
183 and chromatin modifications in IECs by performing RNA-seq, DNase-seq, and ChIP-seq for the  
184 enhancer histone modifications H3K4me1 and H3K27ac, and the Hnf4 TF family members  
185 Hnf4g and Hnf4a in CV and GF conditions totaling 35 datasets. We conducted these  
186 experiments in jejunal IEC from gnotobiotic mice because: (1) ChIP-grade antibodies for mouse  
187 Hnf4a and Hnf4g are available, (2) the larger organ size provided sufficient numbers of IECs for  
188 ChIP-seq experiments, and (3) we speculated that the roles of Hnf4a in host response to  
189 microbiota may be conserved to mammals. We first performed DNase-seq in jejunal IEC from  
190 mice reared GF or colonized for two weeks with a conventional mouse microbiota (CV) to  
191 determine the impact of microbiota colonization on chromatin accessibility (Fig. 3A). In accord  
192 with previous studies that tested for chromatin accessibility in ileal or colonic IECs from GF  
193 or CV mice (Camp et al. 2014), we similarly found no differential DNase hypersensitivity sites  
194 (DHSs) in GF or CV jejunum (data not shown, but see Supplemental Fig. S4A; Supplemental  
195 Tables S6, S8). These data indicate that gross accessibility changes in chromatin do not  
196 underlie the transcription of microbiota-responsive genes in IECs.

197 To test if other metrics of chromatin utilization were dynamically regulated by microbiota, we  
198 performed ChIP-seq from GF and CV mouse jejunal IECs for histone marks H3K4me1 and

**Figure 3**



**FIGURE 3. Microbiota selectively induce enhancer activity near genes that are upregulated upon microbiota colonization**

(A) Schematic showing the gnotobiotic experimental timeline for testing mRNA levels and chromatin architecture in GF and CV. (B) MA plots from DEseq2 analysis (FDR < 0.01) of H3K4me1 (n = 3 per condition) (left) and H3K27ac (n = 2 per condition) (right) ChIP-seq from GF and CV mouse jejunal IECs. Colored dots signify regions significantly enriched for a histone mark in GF (blue) or CV (orange). We found 4,579 unique H3K4me1 and 1,354 unique H3K27ac peaks in GF and 5,155 unique H3K4me1 and 893 unique H3K27ac peaks in CV. (C) Volcano plots showing pairwise comparison of RNA expression between GF and CV jejunal IECs. Blue and orange dots represent genes associated with a region enriched for H3K4me1 (left) or H3K27ac (right) signal in GF or CV. (D) Two-sided Kolmogorov-Smirnov goodness-of-fit test shows a positive relationship on average between the presence of a region enriched for H3K4me1/H3K27ac signal in a specific colonization state and increased transcript abundance of a neighboring gene in that same colonization state. (E) Top *de novo* binding site motifs found in DHSs that are flanked by regions enriched with H3K27ac signal in GF (E) or CV (F). Representative ChIP-seq tracks highlighting a microbiota-regulated gene associated with differential histone marks in GF (G) (*Akr1c19*, *Aldo-keto reductase 1c19*) or CV (H) (*Ubd*, *Ubiquitin D*). Heatmaps showing the average GF and CV H3K4me1 (I) or H3K27ac (J) signal at the 1000 bps flanking differential sites. (K-L) GO terms and KEGG pathways enriched in genes associated with differential H3K27ac sites shown in J. See also Supplemental Figure S4.

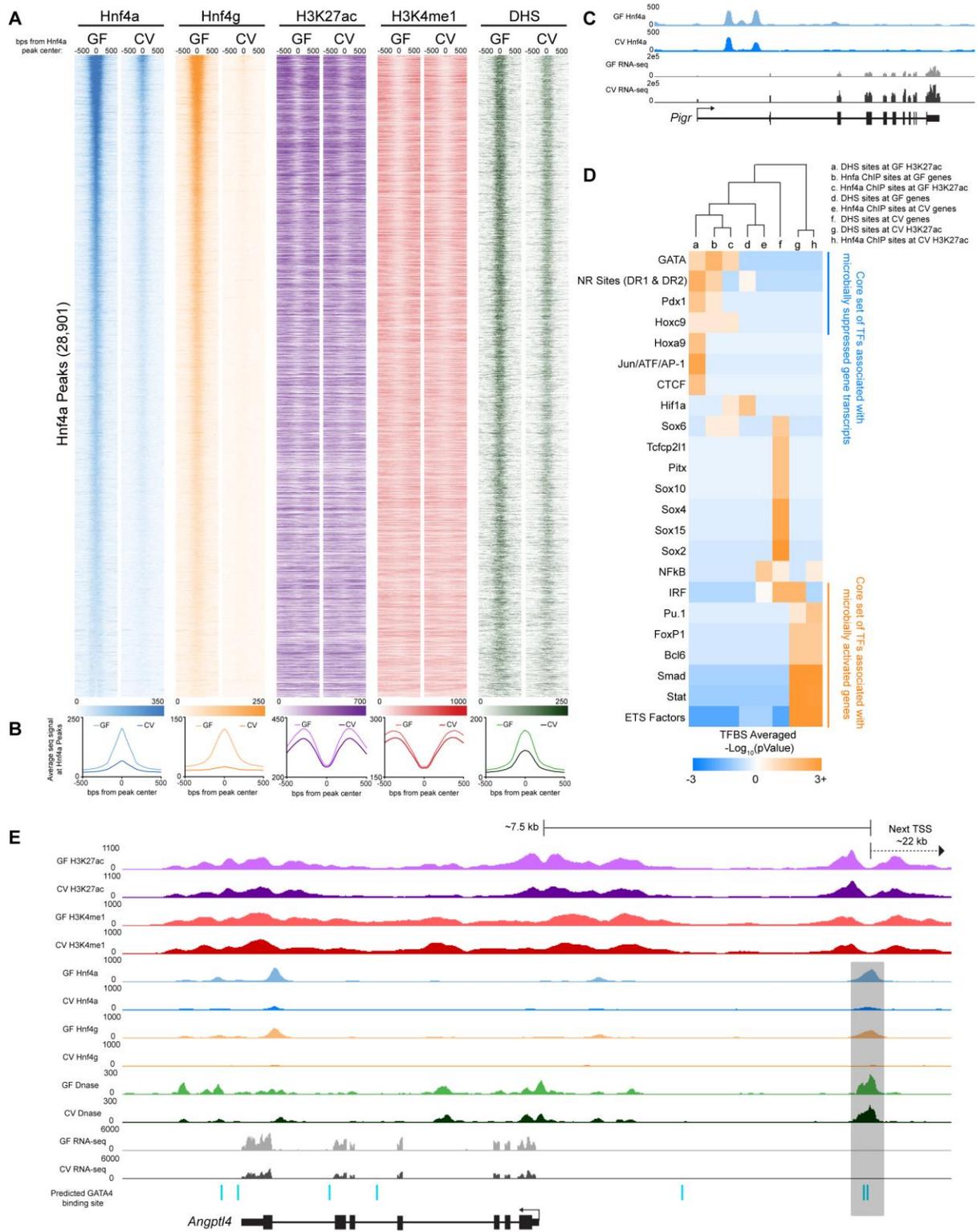
201 H3K27ac that are enriched at poised enhancers and active enhancers, respectively (Fig. 3B).  
202 By determining the single-nearest gene TSS within 10kb of the differential histone marks and  
203 overlaying these data with our new RNA-seq datasets, we found that regions that gain poised  
204 (H3K4me1) and activated (H3K27ac) enhancers upon colonization are associated with genes  
205 that have increased transcript levels upon colonization (Fig. 3C,H-K; Supplemental Fig. S4I;  
206 Supplemental Tables S3, S6, S8). Similarly, regions that lose poised and active enhancers upon  
207 colonization are associated with microbiota-suppressed genes (Fig. 3C,G,I,J,L; Supplemental  
208 Fig. S4J; Supplemental Tables S3, S6, S8). A two-sided Kolmogorov-Smirnov goodness-of-fit  
209 test shows a positive relationship between differential H3K4me1/H3K27ac region and increased  
210 transcript abundance of nearby genes in the same colonization state (Fig. 3D). Collectively, we  
211 identified for the first time a genome-wide map of hundreds of newly identified microbial  
212 regulated CRRs, suggesting that microbiota regulation of host genes is mechanistically linked to  
213 histone modifications changes more than gross chromatin accessibility changes (Camp et al.  
214 2014).

215 We leveraged this novel atlas of microbiota-regulated enhancers and accessible chromatin to  
216 determine which TFs are predicted to bind to these regions. An unbiased analysis found that  
217 Hnf4a binding site motifs were enriched in promoters of genes associated with microbiota-  
218 suppressed enhancers (Supplemental Fig. S4E), and STAT1 binding site motifs were enriched  
219 in promoters of genes associated with microbiota-activated enhancers (Supplemental Fig. S4F).  
220 Interestingly, DHS sites associated with differentially active enhancers were enriched for two  
221 different sets of TF binding sites. DHSs flanked by microbiota-inactivated enhancers were  
222 enriched for nuclear receptor DR1 sites and GATA binding sites (Fig. 3E). DHS sites associated  
223 with microbiota-activated enhancers were similarly enriched for the nuclear receptor DR1  
224 binding sites but also for STAT/IRF-like and ETS binding sites (Fig. 3F). These data suggest  
225 that nuclear receptors like Hnf4a may play a central role in IEC responses to microbial  
226 colonization.

### 227 **Microbiota colonization is associated with a reduction in Hnf4a and Hnf4g cistrome** 228 **occupancy**

229 To directly evaluate the impact of microbiota on Hnf4a activity, we tested the plasticity of the  
230 genome wide distribution of Hnf4s in response to microbial colonization. Hnf4a bound 28,901  
231 and Hnf4g bound 21,875 across the genome in GF conditions in jejunal IECs with ~80% of  
232 these sites being bound by both TFs. In striking contrast, the number of sites bound by Hnf4a  
233 and Hnf4g in CV conditions was ~10 fold less (Fig. 4A,B; Supplemental Tables S5, S8). Of the  
234 3,964 Hnf4a binding sites detected in CV there were only 267 Hnf4a sites that were specific to  
235 the CV condition (Supplemental Fig. S5A,C; Supplemental Table S8). Yet, the genes associated  
236 with these Hnf4a sites that are retained in CV are enriched for ontologies and pathways  
237 fundamental to intestinal epithelial biology (Supplemental Fig. S5B). However, we did find that  
238 the average CV Hnf4a signal was significantly increased at Hnf4a sites associated with  
239 microbiota-induced genes relative to those Hnf4a sites associated with microbiota-suppressed

**Figure 4**



**FIGURE 4. Microbiota colonization results in extensive loss of Hnf4a and Hnf4g DNA binding in IEC**

(A) Heatmaps showing the average GF and CV ChIP-seq or DNase-seq signal at the 1000 bps flanking Hnf4a sites found in GF. (B) Line plots showing the average GF (light-colored line) and CV (dark-colored line) ChIP-seq and DNase-seq RPKM-normalized signal for the indicated TF, histone mark or DHS at the 1000 bps flanking Hnf4a sites found in GF (Hnf4a: n = 3 per condition; Hnf4g: n = 4 per condition; H3K27ac n = 2 per condition; H3K4me1: n = 3 per condition; DNase: N = 3 for CV, n = 2 for GF). (C) Representative signal tracks highlighting a microbiota-induced gene (*Pigr*, *Polymeric immunoglobulin receptor*) that is associated with an Hnf4a peak with similar signal in both GF and CV jejunal IECs. (D) Heatmap showing the enrichment of TFBS motifs within 50 bps of the DHS or Hnf4a peak maxima. (E) Representative signal track at *Angptl4* highlighting two GATA4 sites within an Hnf4a bound region. See also Supplemental Figure S5.

241 genes, suggesting Hnf4a may play a limited role in genes upregulated by colonization  
242 (Supplemental Fig. S5F). In contrast, GF Hnf4a ChIP signal was equivalent at Hnf4a sites  
243 associated with microbiota-suppressed and induced genes (Supplemental Fig. S5F). We do not  
244 believe that the reduction of Hnf4a binding is the result of chromatin quality in a particular  
245 condition since there are genomic locations where GF and CV Hnf4a sites appeared to have  
246 equivalent signal (Fig. 4C). Furthermore, ChIP enrichment in these IEC preparations for another  
247 zinc finger TF, CTCF, was unaffected by microbiota colonization (Supplemental Fig. S5D). This  
248 indicates that the observed reduction of Hnf4 ChIP signal in CV IECs is a result of microbiota on  
249 Hnf4 binding, and is not the result of altered ChIP efficiency or sample quality in the different  
250 conditions. The dramatic loss of Hnf4a and Hnf4g DNA binding upon colonization is consistent  
251 with Hnf4a acting as a potent activator of microbiota-suppressed genes.

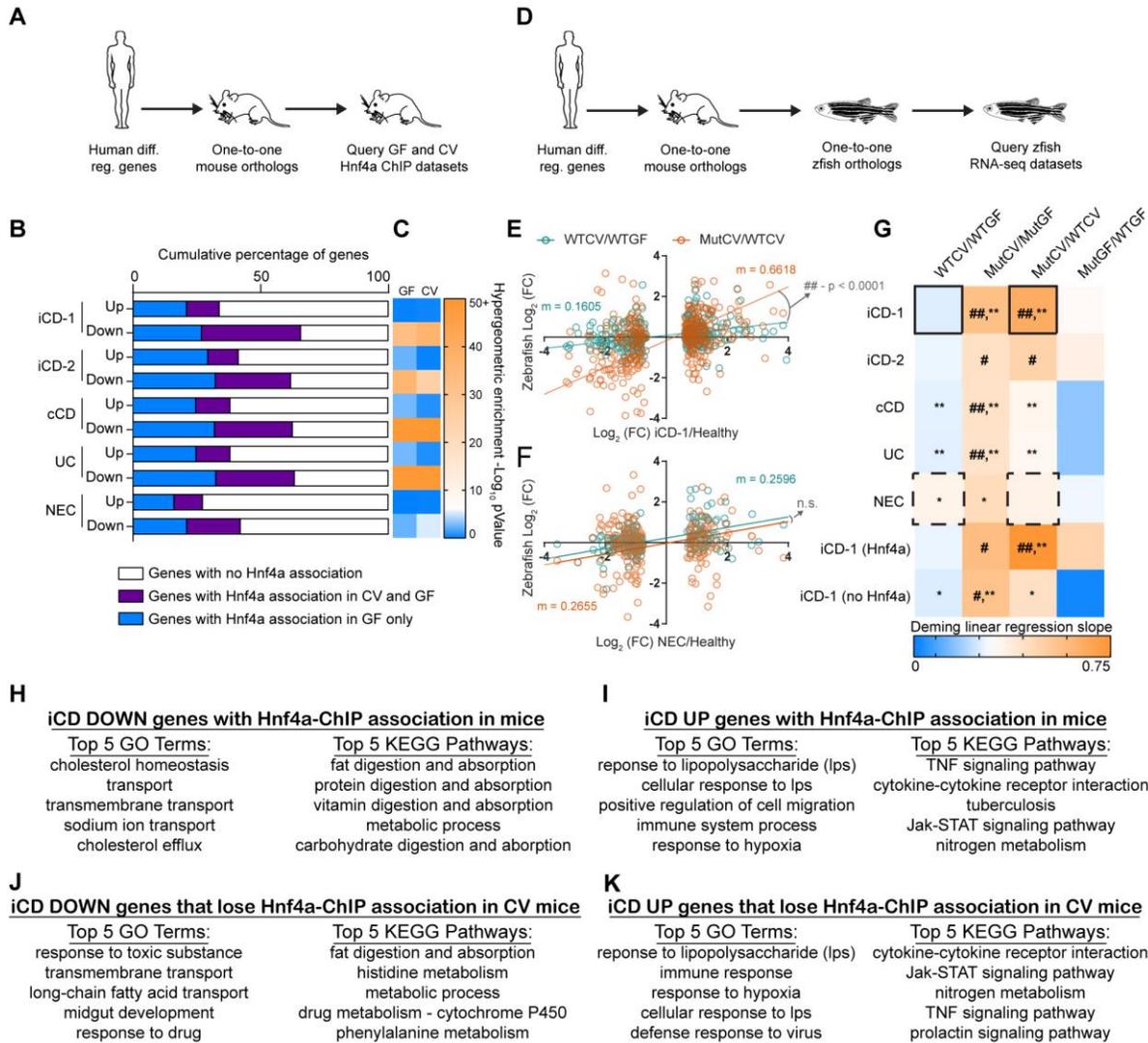
252 We further speculated that certain coregulatory sequence-specific transcription factors may also  
253 contribute to regulating transcription with Hnf4 at these sites. To explore this possibility, we  
254 searched for TF motifs associated with Hnf4a ChIP sites and found an enrichment of putative  
255 binding sites for TFs known to be involved in small intestinal physiology (GATA and Hoxc9) as  
256 well as nutrient metabolism (Pdx1) at both Hnf4a bound regions associated with genes and  
257 enhancers suppressed by microbes (Fig. 4D). We similarly found GATA sites located within an  
258 Hnf4a-bound CRR near murine *Angptl4* (Fig. 4E), similar to the coincident Hnf4 and GATA  
259 motifs in in3.4 (Camp et al. 2012). Furthermore, binding sites for TFs known to be involved in

260 cell proliferation and cell death (ETS transcription factor family) are enriched near Hnf4a bound  
261 regions that intersect microbiota-induced enhancers (Fig. 4D). Collectively our integrative  
262 analyses of these novel ChIP-seq, DNase-seq, and RNA-seq datasets identifies a core set of  
263 putative microbiota-responsive TFs that may interact with Hnf4a to mediate microbial control of  
264 IEC gene expression. These results suggest Hnf4a plays a major role in integrating microbial  
265 signals to regulate gene expression, and raise the possibility that this novel microbiota-Hnf4a  
266 axis might contribute to human disease.

### 267 **Microbiota-mediated suppression of Hnf4a may contribute to IBD pathogenesis**

268 Both Hnf4a and the intestinal microbiota have been separately implicated in the pathogenesis of  
269 inflammatory bowel diseases such as Crohn's disease (CD) and ulcerative colitis (UC) (Ahn et  
270 al. 2008; Sartor and Wu 2016). However, a mechanistic link between microbiota and Hnf4a in  
271 the context of IBD pathogenesis has not been established. Previous transcriptomic studies have  
272 identified genes differentially expressed in ileal (iCD) and colonic CD (cCD) and UC (Arijs et al.  
273 2009; Haberman et al. 2014) biopsies. We queried these human gene lists to identify one-to-  
274 one orthologs in mice, and referenced them against our new gnotobiotic mouse jejunal Hnf4a  
275 ChIP-seq data (Fig. 5A). Strikingly, the majority of human genes downregulated in all of these  
276 IBD datasets have mouse orthologs that are associated with an Hnf4a-bound region (Fig. 5B,C;  
277 Supplemental Table S7). Focusing on the iCD dataset from the largest of these previous studies  
278 (Haberman et al. 2014), we found differential iCD genes associated with Hnf4a sites are  
279 enriched for distinct ontologies and pathways that are dysregulated in IBD (Fig. 5H-K). In  
280 contrast to IBD, analysis of a transcriptomic dataset from human necrotizing enterocolitis (NEC)  
281 (Tremblay et al. 2016) revealed significantly less Hnf4a-bound regions near genes both induced  
282 and suppressed in NEC (Fig. 5B,C). These data indicate microbiota-dependent and microbiota-  
283 independent suppression of Hnf4a activity may play an important role in IBD pathologies, but  
284 not NEC.

**Figure 5**



286

287 To assess if microbiota suppression of Hnf4a activity regulates genes differentially expressed in

288 IBD, we queried the published human IBD and NEC gene expression datasets to identify

289 human-mouse-zebrafish one-to-one-to-one orthologs that were differentially expressed in our

290 RNA-seq analysis of gnotobiotic zebrafish *hnf4a* mutants (Fig. 5D). We found ortholog

291 expression fold changes in human IBD/healthy comparisons most closely resembled the

292 expression fold changes of MutCV/MutGF and MutCV/WTCV (Fig. 5E-G). Neither the

293 WTCV/WTGF nor the MutGF/WTGF comparisons faithfully recapitulate the expression profiles

**FIGURE 5. Microbiota suppression of Hnf4a activity is highly correlated with genes and intestinal processes suppressed in human IBD and conserved in zebrafish**

(A) Flow chart showing the experimental design and filters used to identify IBD or NEC gene orthologs associated with mouse Hnf4a ChIP sites. (B) Bar chart showing the proportion of Hnf4a associations in GF and CV mouse jejunal IECs near human-to-mouse one-to-one gene orthologs differentially regulated in human pediatric ileal Crohn's Disease (iCD-1), adult iCD (iCD-2), adult colonic Crohn's Disease (cCD), adult ulcerative colitis (UC), or neonatal necrotizing enterocolitis (NEC). (C) Heatmap representing the  $-\text{Log}_{10}$  (pValue) of the enrichment of GF or CV Hnf4a-associated genes that are differentially regulated genes in the indicated IBD datasets. pValues were calculated using a hypergeometric enrichment analysis and converting all Hnf4a ChIP-associated mouse genes to human orthologs (GF = 5863 genes and CV = 2119 genes). (D) Flow chart showing the experimental design and filters used to identify correlations between gnotobiotic WT or mutant zebrafish gene expression and gene orthologs differentially expressed in human IBD or NEC. Because loss of *hnf4a* function in zebrafish appeared to more closely resemble iCD signature than cCD or UC, we performed pairwise comparisons of gene orthologs that are (1) differentially regulated in human iCD and (2) have a mouse Hnf4a ChIP association. Example of Deming linear regression analysis showing the correlation of  $\text{Log}_2$  (FC) between WTCV/WTGF (E) or MutCV/WTGF (F) zebrafish and pediatric iCD or NEC.  $m$  = slope of the line. (G) Heatmap representing slopes of Deming linear regression lines showing positive correlative relationships between the  $\text{log}_2$  gene expression fold changes of one-to-one orthologs from human diseases compared to  $\text{log}_2$  fold changes in zebrafish WTCV/WTGF, MutCV/MutGF, MutCV/WTCV, and MutGF/WTGF. Hash signs indicate slope of Deming linear regression lines is significantly greater than WTCV/WTGF comparison (#,  $p < 0.05$ ; ##,  $p < 0.0001$ ). Asterisks indicate slope of Deming linear regression line is significantly greater than MutGF/WTGF (\*,  $p < 0.05$ ; \*\*,  $p < 0.001$ ). Solid boxes correspond to slope of lines in panel 5D, and dashed boxes correspond to slope of lines in panel 5E. (H-K) The top 5 GO terms and the top 5 KEGG pathways for indicated gene lists.

294 of IBD/healthy comparisons. This indicates that both the microbiota and loss of *hnf4a* function  
295 are necessary to induce an IBD-like gene expression profile in zebrafish. Strikingly, the positive  
296 correlation and significant resemblance to the iCD-like gene signatures in the colonized *hnf4a*<sup>-/-</sup>  
297 compared to colonized *hnf4a*<sup>+/+</sup> zebrafish digestive tracts become even stronger when we  
298 limited our analysis to one-to-one orthologs that have an association with an Hnf4a bound  
299 region in mouse IECs (Fig. 5G). Together, these results suggest that Hnf4a protects the host  
300 and maintains transcriptional homeostasis in the presence of a microbiota and protects against  
301 an IBD-like gene expression signature.

## 302 DISCUSSION

303 Over the course of animal evolution, the intestinal epithelium has served as the primary barrier  
304 between animal hosts and the complex microbial communities they harbor. IECs maintain this

305 barrier and perform their physiological roles in nutrient transport and metabolism through  
306 dynamic transcriptional programs. The regulatory mechanisms that orchestrate these  
307 transcriptional programs represent potential therapeutic targets for a variety of human intestinal  
308 diseases including IBD. Here we discovered that Hnf4a activity and its transcriptional network  
309 are suppressed by microbiota. Hnf4a is the oldest member of the nuclear receptor TF family  
310 (Bridgham et al. 2010), and our findings in fish and mammals suggest that microbial  
311 suppression of Hnf4a may be a conserved feature of IEC transcriptional programs present in the  
312 common ancestor.

313 Our results suggest important new links between Hnf4a and microbiota in the context of human  
314 IBD. IBD patients, particularly those suffering from Crohn's disease, often present with  
315 decreased serum low-density lipoprotein levels and reduced total cholesterol levels compared to  
316 healthy individuals (Hrabovsky et al. 2009; Agouridis et al. 2011). These serum levels are  
317 consistent with reduced transcript levels for genes involved in intestinal absorption and transport  
318 of lipid and cholesterol in ileal and colonic biopsies from IBD patients (Arijs et al. 2009;  
319 Haberman et al. 2014). Transcription factors, including nuclear receptors like Hnf4a and Fxr, are  
320 known to regulate bile acid production, lipid mobilization and cholesterol absorption have  
321 already been implicated in IBD (Ahn et al. 2008; Nijmeijer et al. 2011). However, our work is the  
322 first to demonstrate a role for microbiota in suppressing Hnf4a that may ultimately affect IBD  
323 pathogenesis. Hnf4a has been shown to play key roles in anti-oxidative and anti-inflammatory  
324 defense mechanisms (Marcil et al. 2010) so aberrant microbial suppression could promote a  
325 chronic inflammatory state. Hnf4a target genes are downregulated in human IBD (Arijs et al.  
326 2009; Haberman et al. 2014) and mouse experimental colitis, and the Hnf4a target ApoA1 has  
327 been shown to be protective against intestinal inflammation in mice (Gkouskou et al. 2016). We  
328 speculate that the genes governed by this novel microbiota-Hnf4a axis may include additional  
329 anti- and pro-inflammatory factors that could provide new targets for IBD therapy.

330 Our results reveal similar effects of microbiota colonization and experimental colitis on Hnf4a  
331 cistrome occupancy, but the underlying molecular mechanisms are unresolved. DSS induced  
332 colitis results in reduced Hnf4a protein levels and altered cellular localization (Chahar et al.  
333 2014), however our results indicate the microbiota neither reduce Hnf4a protein levels nor  
334 impact its nuclear localization in jejunal IECs two weeks after colonization (Supplemental Fig.  
335 S5H,I). Colonization of GF mice with microbiota initiates a transcriptional adaptation in the  
336 intestine that progresses for several weeks before reaching homeostasis (El Aidy et al. 2012),  
337 so further studies are needed to determine if Hnf4a activity is similarly altered at other stages of  
338 this adaptive process. However, these data collectively suggest that microbiota suppress Hnf4a  
339 activity in the jejunum through mechanisms distinct from those utilized in experimental colitis.

340 Despite the importance of Hnf4a in digestive and metabolic health (Palanker et al. 2009;  
341 Frochet et al. 2012), relatively little insight into its regulation in a biological context has been  
342 reported. *In vitro* and cell culture studies have identified possible suppressors and activators of  
343 Hnf4a including phosphorylation by AMPK and acetylation by CBP, both of which have been  
344 shown to induce Hnf4a activity (Soutoglou et al. 2000; Hong et al. 2003). Other studies have  
345 shown that Hnf4a regulation is controlled by alternative promoters which generate different  
346 isoforms (Huang et al. 2009). However, we did not detect differential *Hnf4a* exon usage by  
347 DEXseq (Li et al. 2015) in GF and CV mouse jejunal IECs (data not shown). Another facet of  
348 Hnf4a biology that has not been adequately explored is the identity of its endogenous ligand(s).  
349 Although historically considered an orphan nuclear receptor, several fatty acids including linoleic  
350 acid have been identified as ligands for Hnf4a (Hertz et al. 1998; Palanker et al. 2009; Yuan et  
351 al. 2009). Fatty acids are an attractive class of putative regulators of Hnf4a since the microbiota  
352 are known to augment FA absorption in zebrafish enterocytes (Semova et al. 2012).  
353 Identification of Hnf4a ligands or other co-regulatory factors that mediate microbial control of  
354 Hnf4a could provide novel therapeutic targets for IBD (Mbodji et al. 2013).

355 In our attempt to understand how the microbiota regulate Hnf4a activity and host gene  
356 transcription, we were motivated to investigate if microbiota impact histone modification and  
357 chromatin accessibility in the mouse jejunum. Our findings support the model that microbiota  
358 alter IEC gene expression by affecting TF binding and histone modification at tissue-defined  
359 open chromatin sites (Camp et al. 2014). We provide the genomic addresses of hundreds of  
360 microbiota-regulated enhancers as well as the genes associated with these enhancers and  
361 Hnf4a binding sites. Similar to other findings in intraepithelial lymphocytes (Semenkovich et al.  
362 2016), our work demonstrates a clear microbial contribution to the modification of the histone  
363 landscape in IECs and provides another important layer of regulation that orchestrate microbiota  
364 regulation of host genes involved in intestinal physiology and human disease. We were also  
365 able to establish a link between microbiota-regulated genes and enhancers and NR binding  
366 sites. These NR binding sites are coincident with a core set of TFs that are enriched near  
367 microbiota-suppressed enhancers/genes (GATA) or induced enhancers/genes (ETS-factors and  
368 IRF) (Supplemental Fig. S6). Coregulation by other TFs represents one possible mode of Hnf4a  
369 regulation by which the microbiota could suppress Hnf4a activity without impacting the gene  
370 transcription of all Hnf4a-associated genes.

## 371 **METHODS**

### 372 **Yeast 1-Hybrid ORFeome Screen:**

373 The yeast 1-hybrid ORFeome screen was performed using the Clontech Matchmaker™ Gold  
374 Yeast One-hybrid Library Screening System (cat. 630491) protocol with the following  
375 exceptions: The Y1HGold yeast strain was transformed using standard yeast transformation  
376 procedures with BstBI digested pBait-AbAi containing either the WT or a SDM in3.4 or the p53  
377 binding site sequence, and positive transformants were selected on SD/-URA media. In  
378 addition, a ORFeome library consisting of 148 zebrafish transcription factors cloned from adult

379 zebrafish liver (Supplemental Table S1) into pDEST22 prey vectors containing an N-terminal  
380 GAL4-activation domain was utilized (Boyle et al. 2016). For additional information, see  
381 Supplemental Methods.

### 382 **Zebrafish Transgenesis and Imaging:**

383 Co-injections of Tol2 SDM or WT in3.4:cfos:gfp plasmid and transposase mRNA were  
384 performed as described (Camp et al., 2012) with the following exceptions: 50 – 100 zebrafish  
385 embryos were injected at the 1–2 cell stage with approximately 69 pg of plasmid DNA at a  
386 DNA:transposase ratio of 1:2. Three mosaic patches within a given tissue of an imaged fish  
387 were quantified for mean fluorescence intensity and averaged. Statistical significance was  
388 analyzed using Kruskal-Wallis one-way analysis of variance and Dunn's multiple comparison  
389 test using GraphPad Prism software. For additional information, see Supplemental Methods.

### 390 **Zebrafish Mutagenesis:**

391 Targeted gene deletion of the *hnf4a* gene was performed using CRISPR/Cas9 nuclease RNA-  
392 guided genome editing targeting the fourth exon of *hnf4a*. The guide RNA sequences were  
393 designed using "CRISPR Design Tool" (<http://crispr.mit.edu/>). Guide RNAs (Supplemental  
394 Table S9) were generated from BamHI (New England Biolabs R0136L) digested pT7-gRNA  
395 plasmid (a gift from Wenbiao Chen and available from Addgene:  
396 <http://www.addgene.org/46759/>) and by performing an *in vitro* transcription reaction using  
397 MEGAscript T7 kit (Ambion/Invitrogen AM1354). For additional information, see  
398 Supplemental Methods.

### 399 **Mouse IEC Isolation for DNase, ChIP and RNA-seq:**

400 Mice were euthanized under CO<sub>2</sub> and cervical dislocation and placed on a chilled wax  
401 dissection pad. The small intestine was removed from the mouse and the jejunum was excised

402 from the duodenum and ileum. Duodenum was defined as the anterior 5 cm of the midgut and  
403 ileum was defined as posterior 6 cm of midgut as described (Camp, et al 2014). Adipose and  
404 vasculature were removed from the tissue. The jejunum was opened longitudinally along the  
405 length of the tissue, exposing the lumen and epithelial cell layer. Luminal debris was washed  
406 away from the epithelia with ice cold sterile PBS. The tissue was temporarily stored in 10 ml of  
407 ice cold sterile PBS with 1x Protease Inhibitor (Complete EDTA-Free, Roche 1187350001) and  
408 10  $\mu$ M Y-27632 (ROCK I inhibitor, Selleck Chemicals S1049) to inhibit spontaneous apoptosis.  
409 The jejunum was moved into a 15 ml conical tube containing 3 mM EDTA in PBS with 1x  
410 protease inhibitor and 10  $\mu$ M Y-27632. The tissue was placed on a nutator in a cold room for 15  
411 minutes. The jejunum was removed from the 3 mM EDTA and placed on an ice cold glass petri  
412 dish with PBS containing 1mM MgCl<sub>2</sub> and 2 mM CaCl<sub>2</sub> with protease inhibitors and 10  $\mu$ M Y-  
413 27632. Villi were scraped off of the tissue using a sterile plastic micropipette and placed into a  
414 new 15 ml conical tube. The isolated IECs were then crosslinked for ChIP or used for DNase-  
415 seq or RNA-seq. For additional information, see Supplemental Methods.

#### 416 **Mouse Intestine Immunofluorescence and Western Blot:**

417 Mid-jejunal tissue was dissected and cleaned as in the IEC villi isolation above. The whole,  
418 splayed open tissue was pinned to 3% agarose and fixed in 4% PFA overnight with gentle  
419 agitation at 4°C. The fixed tissue was washed 4 times with PBS for 15 minutes. The tissue was  
420 then permeabilized in PBS with 0.5% Tween 20 for 1.5 hours at room temperature. Following  
421 permeabilization, the tissue was blocked in 5% donkey serum in PBS with 0.1% Tween 20 for 2  
422 hours at room temperature, incubated in primary and secondary antibodies and imaged on a  
423 Leica SP8 confocal microscope. Western blots were performed on non-crosslinked IEC lysates  
424 (see below) using standard chemoluminescence Western blot protocols. The western blot  
425 shown in Supplemental Fig. S5H is a representative of two experiments. For additional  
426 information, see Supplemental Methods.

427 **Bioinformatic and Statistical analysis:**

428 Sample sizes for zebrafish experiments (noted in figure legends) were selected based on  
429 genotype availability and transgenesis efficiency. All sample collection was performed two or  
430 more times on independent days. For sequencing experiments, statistical calls for differential  
431 gene expression were made by Cuffdiff using parameters stated above. For the zebrafish RNA-  
432 seq experiment Next-Gen sequencing was performed once and at the same time to avoid batch  
433 effects: WTGF and WTCV (n = 3); MutGF and MutCV (n = 2). We originally collected n = 3  
434 MutGF and MutCV biological replicates, however, using pre-established criteria and to avoid  
435 RNA contamination, we excluded one biological replicate from all analysis from these groups  
436 because of sequencing reads that mapped within the deleted *hnf4a* exon in the *hnf4a*<sup>-/-</sup>  
437 genotype.

438 GF mice were randomly chosen by gnotobiotic staff for microbiota colonization (CV) based on  
439 their availability and litter sizes. All sample collection was performed two or more times per  
440 condition on independent days. GF and CV mouse samples were collected on different days.  
441 For sequencing experiments, statistical calls for differential gene expression and differential  
442 peak calls were made by Cuffdiff, MACS2, and DEseq2 using parameters stated above. For the  
443 mouse RNA-seq experiment Next-Gen sequencing was performed once and at the same time  
444 to avoid batch effects: GF (n = 2) and CV (n = 2). Paired GF and CV ChIP and library  
445 amplification was performed simultaneously. Typically, biological ChIP replicates were  
446 sequenced on different days and were always paired with the other condition (i.e. CV and GF  
447 were always sequenced together). The number of biological ChIP replicates (noted in figure  
448 legends) was dependent on reproducibility between ChIP samples and/or our ability to  
449 determine statistical differential sites using DEseq2 (for H3K4me1 and H3K27ac).

450 All statistical metrics (except where otherwise noted) were performed in Graphpad Prism 7.01.

451 Deming linear regression was used for Fig. 5 because it is a stronger and more accurate  
452 assessment of correlation when both the x and y variables have experimental error. Details  
453 regarding the other statistical tests used in this study can be found in the figure legends or  
454 above.

455 For additional information, see Supplemental Methods.

456 **Dataset Accessions:**

457 Sequence datasets are available at GEO accession number GSE90462.

458 **ACKNOWLEDGEMENTS**

459 We thank Balfour Sartor, Scott Magness, and Maureen Bower for assistance with gnotobiotic  
460 mice, and Wenbiao Chen and Stacy Horner for sharing reagents. We also thank the Genomic  
461 Sequencing Laboratory at HudsonAlpha Institute for Biotechnology and the Duke Sequencing  
462 and Genomic Technologies Facility. This work was supported by grants from the National  
463 Institutes of Health (R01-DK081426, U24-DK097748, P01-DK094779, R24-OD016761, and  
464 P30-DK34987).

465

## 466 REFERENCES

- 467 Agouridis AP, Elisaf M, Milionis HJ. 2011. An overview of lipid abnormalities in patients with  
468 inflammatory bowel disease. *Ann Gastroenterol* **24**: 181-187.
- 469 Ahn SH, Shah YM, Inoue J, Morimura K, Kim I, Yim S, Lambert G, Kurotani R, Nagashima K,  
470 Gonzalez FJ et al. 2008. Hepatocyte nuclear factor 4alpha in the intestinal epithelial cells  
471 protects against inflammatory bowel disease. *Inflamm Bowel Dis* **14**: 908-920.
- 472 Alenghat T, Osborne LC, Saenz SA, Kobuley D, Ziegler CG, Mullican SE, Choi I, Grunberg S,  
473 Sinha R, Wynosky-Dolfi M et al. 2013. Histone deacetylase 3 coordinates commensal-bacteria-  
474 dependent intestinal homeostasis. *Nature* **504**: 153-157.
- 475 Arijis I, De Hertogh G, Lemaire K, Quintens R, Van Lommel L, Van Steen K, Leemans P,  
476 Cleynen I, Van Assche G, Vermeire S et al. 2009. Mucosal gene expression of antimicrobial  
477 peptides in inflammatory bowel disease before and after first infliximab treatment. *PLoS One* **4**:  
478 e7984.
- 479 Bäckhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A, Semenkovich CF, Gordon JI. 2004.  
480 The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S*  
481 *A* **101**: 15718-15723.
- 482 Bates JM, Akerlund J, Mittge E, Guillemin K. 2007. Intestinal alkaline phosphatase detoxifies  
483 lipopolysaccharide and prevents inflammation in zebrafish in response to the gut microbiota.  
484 *Cell Host Microbe* **2**: 371-382.
- 485 Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, Feitosa MF, Justice AE, Monda KL, Croteau-  
486 Chonka DC, Day FR et al. 2013. Genome-wide meta-analysis identifies 11 new loci for  
487 anthropometric traits and provides insights into genetic architecture. *Nat Genet* **45**: 501-512.
- 488 Boyle G, Richter K, Priest HD, Traver D, Mockler TC, Chang JT, Kay SA, Breton G. 2016.  
489 Comparative analysis of vertebrate diurnal/circadian transcriptomes. *PLoS One* **In Press**.
- 490 Bridgham JT, Eick GN, Larroux C, Deshpande K, Harms MJ, Gauthier ME, Ortlund EA, Degnan  
491 BM, Thornton JW. 2010. Protein evolution by molecular tinkering: diversification of the nuclear  
492 receptor superfamily from a ligand-dependent ancestor. *PLoS Biol* **8**.
- 493 Camp JG, Frank CL, Lickwar CR, Guturu H, Rube T, Wenger AM, Chen J, Bejerano G,  
494 Crawford GE, Rawls JF. 2014. Microbiota modulate transcription in the intestinal epithelium  
495 without remodeling the accessible chromatin landscape. *Genome Res* **24**: 1504-1516.
- 496 Camp JG, Jazwa AL, Trent CM, Rawls JF. 2012. Intronic cis-regulatory modules mediate tissue-  
497 specific and microbial control of *angptl4/fiaf* transcription. *PLoS Genet* **8**: e1002585.
- 498 Chahar S, Gandhi V, Yu S, Desai K, Cowper-Sal-lari R, Kim Y, Perekatt AO, Kumar N,  
499 Thackray JK, Musolf A et al. 2014. Chromatin profiling reveals regulatory network shifts and a  
500 protective role for hepatocyte nuclear factor 4alpha during colitis. *Mol Cell Biol* **34**: 3291-3304.
- 501 Chellappa K, Deol P, Evans JR, Vuong LM, Chen G, Briancon N, Bolotin E, Lytle C, Nair MG,  
502 Sladek FM. 2016. Opposing roles of nuclear receptor HNF4alpha isoforms in colitis and colitis-  
503 associated colon cancer. *Elife* **5**.

- 504 Costea I, Mack DR, Lemaitre RN, Israel D, Marcil V, Ahmad A, Amre DK. 2014. Interactions  
505 between the dietary polyunsaturated fatty acid ratio and genetic factors determine susceptibility  
506 to pediatric Crohn's disease. *Gastroenterology* **146**: 929-931.
- 507 Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato  
508 MA, Frampton GM, Sharp PA et al. 2010. Histone H3K27ac separates active from poised  
509 enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* **107**: 21931-21936.
- 510 Darsigny M, Babeu JP, Dupuis AA, Furth EE, Seidman EG, Levy E, Verdu EF, Gendron FP,  
511 Boudreau F. 2009. Loss of hepatocyte-nuclear-factor-4alpha affects colonic ion transport and  
512 causes chronic inflammation resembling inflammatory bowel disease in mice. *PLoS One* **4**:  
513 e7609.
- 514 Duncan SA, Nagy A, Chan W. 1997. Murine gastrulation requires HNF-4 regulated gene  
515 expression in the visceral endoderm: tetraploid rescue of Hnf-4(-/-) embryos. *Development* **124**:  
516 279-287.
- 517 El Aidy S, van Baarlen P, Derrien M, Lindenbergh-Kortleve DJ, Hooiveld G, Levenez F, Dore J,  
518 Dekker J, Samsom JN, Nieuwenhuis EE et al. 2012. Temporal and spatial interplay of  
519 microbiota and intestinal mucosa drive establishment of immune homeostasis in  
520 conventionalized mice. *Mucosal Immunol* **5**: 567-579.
- 521 Evans RM, Mangelsdorf DJ. 2014. Nuclear Receptors, RXR, and the Big Bang. *Cell* **157**: 255-  
522 266.
- 523 Franke A, Hampe J, Rosenstiel P, Becker C, Wagner F, Hasler R, Little RD, Huse K, Ruether A,  
524 Balschun T et al. 2007. Systematic association mapping identifies NELL1 as a novel IBD  
525 disease gene. *PLoS One* **2**: e691.
- 526 Frochot V, Alqub M, Cattin AL, Carriere V, Houllier A, Baraille F, Barbot L, Saint-Just S, Ribeiro  
527 A, Lacasa M et al. 2012. The transcription factor HNF-4alpha: a key factor of the intestinal  
528 uptake of fatty acids in mouse. *Am J Physiol Gastrointest Liver Physiol* **302**: G1253-1263.
- 529 Gkouskou KK, Ioannou M, Pavlopoulos GA, Georgila K, Siganou A, Nikolaidis G, Kanellis DC,  
530 Moore S, Papadakis KA, Kardassis D et al. 2016. Apolipoprotein A-I inhibits experimental colitis  
531 and colitis-propelled carcinogenesis. *Oncogene* **35**: 2496-2505.
- 532 Haberman Y, Tickle TL, Dexheimer PJ, Kim MO, Tang D, Karns R, Baldassano RN, Noe JD,  
533 Rosh J, Markowitz J et al. 2014. Pediatric Crohn disease patients exhibit specific ileal  
534 transcriptome and microbiome signature. *J Clin Invest* **124**: 3617-3633.
- 535 Hertz R, Magenheim J, Berman I, Bar-Tana J. 1998. Fatty acyl-CoA thioesters are ligands of  
536 hepatic nuclear factor-4alpha. *Nature* **392**: 512-516.
- 537 Hong YH, Varanasi US, Yang W, Leff T. 2003. AMP-activated protein kinase regulates  
538 HNF4alpha transcriptional activity by inhibiting dimer formation and decreasing protein stability.  
539 *J Biol Chem* **278**: 27495-27501.
- 540 Hrabovsky V, Zadak Z, Blaha V, Hyspler R, Karlik T, Martinek A, Mendlova A. 2009. Cholesterol  
541 metabolism in active Crohn's disease. *Wien Klin Wochenschr* **121**: 270-275.

- 542 Huang J, Levitsky LL, Rhoads DB. 2009. Novel P2 promoter-derived HNF4alpha isoforms with  
543 different N-terminus generated by alternate exon insertion. *Exp Cell Res* **315**: 1200-1211.
- 544 Kamada N, Seo SU, Chen GY, Nunez G. 2013. Role of the gut microbiota in immunity and  
545 inflammatory disease. *Nat Rev Immunol* **13**: 321-335.
- 546 Kanther M, Sun X, Muhlbauer M, Mackey LC, Flynn EJ, 3rd, Bagnat M, Jobin C, Rawls JF.  
547 2011. Microbial colonization induces dynamic temporal and spatial patterns of NF-kappaB  
548 activation in the zebrafish digestive tract. *Gastroenterology* **141**: 197-207.
- 549 Klapper M, Bohme M, Nitz I, Doring F. 2007. The human intestinal fatty acid binding protein  
550 (hFABP2) gene is regulated by HNF-4alpha. *Biochem Biophys Res Commun* **356**: 147-152.
- 551 Krautkramer KA, Kreznar JH, Romano KA, Vivas EI, Barrett-Wilt GA, Rabaglia ME, Keller MP,  
552 Attie AD, Rey FE, Denu JM. 2016. Diet-Microbiota Interactions Mediate Global Epigenetic  
553 Programming in Multiple Host Tissues. *Mol Cell* **64**: 982-992.
- 554 Li Y, Rao X, Mattox WW, Amos CI, Liu B. 2015. RNA-Seq Analysis of Differential Splice  
555 Junction Usage and Intron Retentions by DEXSeq. *PLoS One* **10**: e0136653.
- 556 Marcil V, Seidman E, Sinnett D, Boudreau F, Gendron FP, Beaulieu JF, Menard D, Precourt LP,  
557 Amre D, Levy E. 2010. Modification in oxidative stress, inflammation, and lipoprotein assembly  
558 in response to hepatocyte nuclear factor 4alpha knockdown in intestinal epithelial cells. *J Biol*  
559 *Chem* **285**: 40448-40460.
- 560 Marcil V, Sinnett D, Seidman E, Boudreau F, Gendron FP, Beaulieu JF, Menard D, Lambert M,  
561 Bitton A, Sanchez R et al. 2012. Association between genetic variants in the HNF4A gene and  
562 childhood-onset Crohn's disease. *Genes Immun* **13**: 556-565.
- 563 Marjoram L, Alvers A, Deerhake ME, Bagwell J, Mankiewicz J, Cocchiari JL, Beerman RW,  
564 Willer J, Sumigray KD, Katsanis N et al. 2015. Epigenetic control of intestinal barrier function  
565 and inflammation in zebrafish. *Proc Natl Acad Sci U S A* **112**: 2770-2775.
- 566 Mbodji K, Charpentier C, Guerin C, Querec C, Bole-Feysot C, Aziz M, Savoye G, Dechelotte P,  
567 Marion-Letellier R. 2013. Adjunct therapy of n-3 fatty acids to 5-ASA ameliorates inflammatory  
568 score and decreases NF-kappaB in rats with TNBS-induced colitis. *J Nutr Biochem* **24**: 700-705.
- 569 Meddens CA, Harakalova M, van den Dungen NA, Foroughi Asl H, Hijma HJ, Cuppen EP,  
570 Bjorkegren JL, Asselbergs FW, Nieuwenhuis EE, Mokry M. 2016. Systematic analysis of  
571 chromatin interactions at disease associated loci links novel candidate genes to inflammatory  
572 bowel disease. *Genome Biol* **17**: 247.
- 573 Morgun A, Dzutsev A, Dong X, Greer RL, Sexton DJ, Ravel J, Schuster M, Hsiao W, Matzinger  
574 P, Shulzhenko N. 2015. Uncovering effects of antibiotics on the host and microbiota using  
575 transkingdom gene networks. *Gut* **64**: 1732-1743.
- 576 Nijmeijer RM, Gadaleta RM, van Mil SW, van Bodegraven AA, Crusius JB, Dijkstra G, Hommes  
577 DW, de Jong DJ, Stokkers PC, Verspaget HW et al. 2011. Farnesoid X receptor (FXR)  
578 activation and FXR genetic variation in inflammatory bowel disease. *PLoS One* **6**: e23745.

- 579 Palanker L, Tennessen JM, Lam G, Thummel CS. 2009. Drosophila HNF4 regulates lipid  
580 mobilization and beta-oxidation. *Cell Metab* **9**: 228-239.
- 581 Plevy S, Silverberg MS, Lockton S, Stockfisch T, Croner L, Stachelski J, Brown M, Triggs C,  
582 Chuang E, Princen F et al. 2013. Combined serological, genetic, and inflammatory markers  
583 differentiate non-IBD, Crohn's disease, and ulcerative colitis patients. *Inflamm Bowel Dis* **19**:  
584 1139-1148.
- 585 Rawls JF, Samuel BS, Gordon JI. 2004. Gnotobiotic zebrafish reveal evolutionarily conserved  
586 responses to the gut microbiota. *Proc Natl Acad Sci U S A* **101**: 4596-4601.
- 587 San Roman AK, Aronson BE, Krasinski SD, Shivdasani RA, Verzi MP. 2015. Transcription  
588 factors GATA4 and HNF4A control distinct aspects of intestinal homeostasis in conjunction with  
589 transcription factor CDX2. *J Biol Chem* **290**: 1850-1860.
- 590 Sartor RB, Wu GD. 2016. Roles for Intestinal Bacteria, Viruses, and Fungi in Pathogenesis of  
591 Inflammatory Bowel Diseases and Therapeutic Approaches. *Gastroenterology*  
592 doi:10.1053/j.gastro.2016.10.012.
- 593 Semenkovich NP, Planer JD, Ahern PP, Griffin NW, Lin CY, Gordon JI. 2016. Impact of the gut  
594 microbiota on enhancer accessibility in gut intraepithelial lymphocytes. *Proc Natl Acad Sci U S*  
595 *A* doi:10.1073/pnas.1617793113.
- 596 Semova I, Carten JD, Stombaugh J, Mackey LC, Knight R, Farber SA, Rawls JF. 2012.  
597 Microbiota regulate intestinal absorption and metabolism of fatty acids in the zebrafish. *Cell*  
598 *Host Microbe* **12**: 277-288.
- 599 Soutoglou E, Katrakili N, Talianidis I. 2000. Acetylation regulates transcription factor activity at  
600 multiple levels. *Mol Cell* **5**: 745-751.
- 601 Stegmann A, Hansen M, Wang Y, Larsen JB, Lund LR, Ritte L, Nicholson JK, Quistorff B,  
602 Simon-Assmann P, Troelsen JT et al. 2006. Metabolome, transcriptome, and bioinformatic cis-  
603 element analyses point to HNF-4 as a central regulator of gene expression during enterocyte  
604 differentiation. *Physiol Genomics* **27**: 141-155.
- 605 Thaiss Christoph A, Levy M, Korem T, Dohnalová L, Shapiro H, Jaitin Diego A, David E, Winter  
606 Deborah R, Gury-BenAri M, Tatrovsky E et al. 2016. Microbiota Diurnal Rhythmicity Programs  
607 Host Transcriptome Oscillations. *Cell* **167**: 1495-1510.e1412.
- 608 Tremblay E, Thibault MP, Ferretti E, Babakissa C, Bertelle V, Bettolli M, Burghardt KM,  
609 Colombani JF, Grynspan D, Levy E et al. 2016. Gene expression profiling in necrotizing  
610 enterocolitis reveals pathways common to those reported in Crohn's disease. *BMC Med*  
611 *Genomics* **9**: 6.
- 612 Yuan X, Ta TC, Lin M, Evans JR, Dong Y, Bolotin E, Sherman MA, Forman BM, Sladek FM.  
613 2009. Identification of an endogenous ligand bound to a native orphan nuclear receptor. *PLoS*  
614 *One* **4**: e5609.
- 615