**Article**

*Biﬁdobacterium breve* UCC2003 Induces a Distinct Global Transcriptomic Program in Neonatal Murine Intestinal Epithelial Cells

Raymond Kiu, Agatha Treveil, Lukas C. Harnisch, ..., Douwe van Sinderen, Tamas Korcsmaros, Lindsay J. Hall

HIGHLIGHTS

- *B. breve* administration signiﬁcantly alters the murine neonatal IEC transcriptome
- Genes/pathways involved in epithelial barrier function are particularly impacted
- *Biﬁdobacterium* may target the IEC stem cell compartment to induce regeneration

Kiu et al., iScience 23, 101336
July 24, 2020 © 2020 The Author(s).
https://doi.org/10.1016/j.isci.2020.101336

OPEN ACCESS
Article

Bifidobacterium breve UCC2003 Induces a Distinct Global Transcriptomic Program in Neonatal Murine Intestinal Epithelial Cells

Raymond Kiu,1 Agatha Treveil,1,2 Lukas C. Harnisch,1 Shabhonam Caim,1 Charlotte Leclaire,1 Douwe van Sinderen,3 Tamas Korcsmaros,1,2 and Lindsay J. Hall1,4,5,6,7,*

SUMMARY

The underlying health-driving mechanisms of Bifidobacterium during early life are not well understood, particularly how this microbiota member may modulate the intestinal barrier via programming of intestinal epithelial cells (IECs). We investigated the impact of Bifidobacterium breve UCC2003 on the transcriptome of neonatal murine IECs. Small IECs from two-week-old neonatal mice administered B. breve UCC2003 or PBS (control) were subjected to global RNA sequencing, and differentially expressed genes, pathways, and affected cell types were determined. We observed extensive regulation of the IEC transcriptome with ~4,000 genes significantly up-regulated, including key genes linked with epithelial barrier function. Enrichment of cell differentiation pathways was observed, along with an overrepresentation of stem cell marker genes, indicating an increase in the regenerative potential of the epithelial layer. In conclusion, B. breve UCC2003 plays a central role in driving intestinal epithelium homeostatic development during early life and suggests future avenues for next-stage clinical studies.

INTRODUCTION

Bifidobacterium represents a keystone member of the early life gut microbiota (Arrieta et al., 2014; O’Neill et al., 2017; Dernien et al., 2019). Certain species and strains are found at high levels in vaginally delivered breast-fed infants including Bifidobacterium longum subsp. infantis, B. longum subsp. longum, Bifidobacterium bifidum, Bifidobacterium pseudocatenulatum, and Bifidobacterium breve (Dominguez-Bello et al., 2010; Mikami et al., 2012; Nagpal et al., 2017; Stewart et al., 2018). As a dominant member of the neonatal gut microbiota, Bifidobacterium is associated with metabolism of breast milk, modulation of host immune responses, and protection against infectious diseases (Fukuda et al., 2012; Ling et al., 2016; Robertson et al., 2020; Lawson et al., 2020; Patole et al., 2016; Baucells et al., 2016; Jacobs et al., 2013; Plummer et al., 2018). However, the mechanisms driving improved health outcomes during early life are largely underexplored and are likely strain dependent.

A key interface between Bifidobacterium and the host is the intestinal epithelial cell (IEC) barrier (Thoo et al., 2019; Groschwitz and Hogan, 2009). Previous studies have indicated that certain strains of Bifidobacterium specifically modulate IEC responses during inflammatory insults, which may help protect from certain gut disorders (Hsieh et al., 2015; Srutkova et al., 2015; Grimm et al., 2015). In murine experimental models, previous work by our group has shown that infant-associated B. breve UCC2003 modulates cell death-related signaling molecules, which in turn reduces the number of apoptotic IECs (Hughes et al., 2017). This protection from pathological IEC shedding appeared to be via the B. breve exopolysaccharide (EPS) capsule and the host-immune adaptor protein MyD88. Another strain of B. breve, NumRes 204 (commercial strain) has also been shown to up-regulate the tight junction (TJ) proteins Claudin 4 and Occludin in a mouse colitis model (Zheng et al., 2014; Plantinga et al., 2011).

Many of the studies to date have focused on the role of Bifidobacterium and modulation of IECs in the context of acute or chronic gut inflammation, with expression profiling limited to specific immune or apoptosis signaling targets (Plaza-Diaz et al., 2014; Riedel et al., 2006; Liu et al., 2010; Hsieh et al., 2015). As many of these studies have involved pre-colonization of the gut with Bifidobacterium strains, followed...
by inflammatory insult, this suggests that initial priming during normal “healthy” conditions may modulate subsequent protective responses. Furthermore, these studies have often been performed in adult mice rather than exploring effects during the early life developmental window, where *Bifidobacterium* effects are expected to be most pronounced. Previous work has indicated that there is significant modulation of the neonatal IEC transcriptome in response to gut microbiota colonization, but to date no studies have probed how particular early life-associated microbiota members, like *Bifidobacterium*, may modulate neonatal IEC responses (Pan et al., 2018). Thus, to understand if and how *Bifidobacterium* may modulate IEC homeostasis during the early life developmental window, we administered *B. breve* UCC2003 to neonatal mice and profiled transcriptional responses in isolated small intestine IECs using global RNA sequencing (RNA-seq). Our analysis indicated whole-scale changes in the transcriptional program of IECs (>4,000 significantly up-regulated genes) that appear to be linked to cell differentiation/proliferation and immune development. Notably the stem cell compartment of IECs seemed to elicit the strongest gene signature. These data highlight the role of *B. breve* UCC2003 in driving early life epithelial cell differentiation and maturation, impacting intestinal integrity and immune functions, which provides a mechanistic basis for understanding associated health-promoting effects.

**RESULTS**

To examine the effects of *B. breve* UCC2003 on the transcriptional profiles of host IECs under homeostatic conditions, we extracted RNA from isolated IECs of healthy 2-week-old neonatal mice (control group) and mice gavaged with *B. breve* UCC2003 for three consecutive days (n = 5 per group). Isolated RNAs from IECs were subjected to RNA-seq to determine global mRNA expression (Figure 1). Subsequently, Differential Gene Expression analysis was performed to understand *B. breve*-associated gene regulation.

**Minimal Impact of *B. breve* UCC2003 on the Wider Neonatal Gut Microbiota**

Initially, we examined for the presence of *B. breve* UCC2003 in the gut microbiome and impact on the wider microbiota using culture and 16S rRNA microbiota profiling approaches (Figures 2A and 2B). We observed high levels of *B. breve* UCC2003 across the 4 days in fecal samples, with higher levels of viable *B. breve* UCC2003 within the colon (~10^8 CFU/g [colony-forming unit]), when compared with the small intestine (~10^5 CFU/g; Figure 2B). Based on 16S rRNA analysis, relative abundance of *Bifidobacterium* increased significantly in the UCC2003 group (p = 0.012) following bacterial administration, whereas the control group displayed very low relative *Bifidobacterium* abundance (~0.01%; Figure 2C). Principal-component analysis (PCA) on gut microbiota profiles (control versus UCC2003) showed a distinct change in microbial
community composition in the UCC2003 group primarily driven by increased relative abundance of *Bifidobacterium*, which may also correlate with increased overall microbial diversity in the UCC2003 group (Figures 2D and 2E). Linear Discriminant Analysis also indicated that *Bifidobacterium* was uniquely enriched in UCC2003 group, and that microbiota members with low relative abundance (<2%) such as *Streptococcus*, *Ruminococcus*, *Prevotella*, and *Coprococcus* were significantly lower (Figures 2F and 2G). Overall, administration of *B. breve* UCC2003 appeared to minimally impact the wider gut microbiota, without significantly...
altering relative abundance of other major resident taxa including Lactobacillus, Bacteroides, and Blautia compared with the control group.

**Impact of *B. breve* UCC2003 on the Neonatal Intestinal Epithelial Transcriptome**

To understand the distribution of samples based on IEC gene expression profiles we performed PCA analysis (Figure 3A; Table S1). All samples clustered according to group (control versus UCC2003), suggesting a significant impact of *B. breve* UCC2003 on gene expression profiles, with distance-wise clustering (Jensen-Shannon) also supporting separation of experimental groups (Figure 3B). To define differentially expressed genes (DEGs), we employed a filter of absolute log₂fold change (LFC) > 1.0 (with adjusted p < 0.05), which equates to a minimum 2-fold change in gene expression (Figures 3C–3E; Table S2). After analysis, a total of 3,996 DEGs were significantly up-regulated, whereas 465 genes were significantly down-regulated in *B. breve* UCC2003-supplemented animals when compared with controls (Figures 3A and 4A). Notably,
we also performed the same experimental protocol on healthy mice aged 10–12 weeks and did not observe any significant DEGs, suggesting that \textit{B. breve} UCC2003 modulation of IECs is strongest within the early life window under homeostatic conditions.

Figure 4. Gene Expression Analysis
(A) Heatmap comparison of gene expression profiles of 4,461 DEGs (control versus UCC2003). See also Table S2.
(B) Top 20 DEGs ranked by false discovery rate-adjusted p values (q values).
(C) Top 20 up-regulated DEGs ranked by log2FC (fold change) values.
(D) Top 20 down-regulated DEGs ranked by log2FC values.
(E) Expression of epithelial integrity associated genes in UCC2003 group (q < 0.05).
(F) Expression of integrin-associated genes in UCC2003 group. Gray dotted lines in the bar charts indicate the threshold of absolute log2FC > 1.0. Data are represented as mean ± SE.
To determine the functional role of DEGs, we examined the most significantly regulated genes ranked by false discovery rate-adjusted p-values (or, q-values). We first looked at the top 20 up-regulated DEGs in the *B. breve* UCC2003 experimental group (Figure 4B). Most genes annotated with known biological processes had cell differentiation and cell component organization functions including Ctnb1, Tjp1, Naip6, and Gm20594 (Table S3). When we ranked the top-regulated genes using LFC, we observed increased expression of Creb5, which is involved in the regulation of neuropeptide transcription (cAMP response element-binding protein; CREB) (Figure 4C). CREB is also known to regulate circadian rhythm, and we also identified additional circadian-clock-related genes that were significantly up-regulated including Per2 and Per3. We noted that several top-down-regulated DEGs were annotated as genes involved in metal binding, or metal-related genes including Mt1, Mt2, Hba-a1, Hbb-bt, and Ftl1-ps1 (Figure 4D; Table S4).

**Regulation of Intestinal Epithelial Barrier-Associated Genes**

As *B. breve* strains have been previously shown to modulate certain TJ/barrier-related proteins, we next investigated DEGs associated with intestinal epithelial barrier development/intestinal structural organization (Figure 4E). Several TJ structural-associated DEGs were observed, including Claudin-encoding gene Cldn34c1 (LFC 3.14), Junction Adhesion Molecules-encoding genes Jam2 (LFC 2.9), and TJ protein (also called Zonula Occludens protein; ZO)-encoding gene Tjp1 (LFC 1.49). Genes that encode integrins (involved in regulation of intracellular cytoskeleton) also exhibited a trend of increased expression (13/14; 92.8%). Both Piezo genes, which assist in TJ organization, Piezo1 (LFC 1.25) and Piezo2 (LFC 1.9), were significantly up-regulated in the *B. breve* UCC2003-treated group.

Over 90% cadherins, proteins associated with the assembly of adherens junctions (Figure 4E), were up-regulated, including Pcdhb14 (LFC 2.8), Pcdhg8b4 (LFC 2.7), Pcdh8 (LFC 1.3), Fat1 (LFC 1.5), and Dsg2 (LFC 1.1). Interestingly, several genes (4/7; 57.1%) involved in mucous layer generation were significantly up-regulated in the UCC2003 experimental group, including Muc2 (LFC 2.2), Muc6 (LFC 3.7), Muc5b (LFC 2.9), and Muc4 (LFC 1.24). Genes Gja1 (LFC 3.59) and Gjb8 (LFC 2.63) that encode gap junction proteins were also up-regulated. In addition, we also investigated the differential expression of genes associated with integrin assembly and downstream integrin signaling pathways (Figure 4F). Over 70% (16/21) of these genes were up-regulated, with 52.3% (11/21) significantly increased in gene expression in the UCC2003 group (LFC >1.0).

We observed increased expression of genes associated with IEC barrier development including cadherins, gap junctions, integrins, mucous layer-associated genes, and several key TJ proteins. These strongly induced gene expression profiles suggest that *B. breve* UCC2003 is involved in enhancing epithelial barrier development in neonates.

**Modulation of Cell Maturation Processes**

We next sought to understand the biological functions of up-regulated DEGs by employing PANTHER GO-Slim functional assignment and process/pathway enrichment analysis (see Figure S1; Tables S5 and S6). DEGs were predominantly involved in general biological processes including cellular process (901 genes) and metabolic process (597 genes; Table S7). At the molecular function level, DEGs were primarily assigned to binding (868 genes) and catalytic activity (671 genes; Table S8), with Olfactory Signaling Pathway and Cell Cycle (biological) pathways also found to be enriched (Table S9).

To delve further into the data, we constructed a signaling network based on up-regulated DEGs (n = 3,996) with the aim of identifying specific gene networks involved in important signaling pathways (Figure 5A). Overall, 1,491 DEGs were successfully mapped (37.3%) to a signaling network that comprised 8,180 genes. Four individual clusters of genes were detected, with functional assignment and pathway analysis implemented on these clusters (Figure 5A). All gene clusters were associated with cell differentiation and maturation, with cluster 1 (68 genes) linked specifically with DNA replication and transcription, cluster 2 (26 genes) with cell growth and immunity, cluster 3 (11 genes) with cell replication, and cluster 4 (72 genes) related to cell cycle and cell division (Table S10).

**Intestinal Cell Type Analysis on DEGs Identifies Significant Enrichment of Epithelial Stem Cells**

IECs include several absorptive and secretory cell types, namely, enterocytes, Paneth cells, goblet cells, enteroendocrine cells, tuft cells, and stem cells. As these cells perform different functions in the gut, it was
important to understand whether *B. breve* UCC2003 had a cell type-specific effect on the intestinal epithelium. Using known cell type-specific gene markers (Haber et al., 2017), we identified cell type gene signatures modulated within the UCC2003 group (Figures 5B and 5C). Importantly, all cell type markers were well represented in the expressed genes of the whole IEC transcriptomics data from both groups (control + UCC2003), thus validating the presence of all IEC types in our study data (Figure 5B). Cell type analysis of genes differentially expressed after *B. breve* UCC2003 supplementation revealed that stem cell marker genes were significantly enriched (30%; *p* < 0.05) among the six IEC types (Table S11). Signatures of other cell types were also present (linking to marker genes in the DEG list) but not significantly overrepresented.

Figure 5. Signaling Network Analysis, IEC Subtyping, and Key Regulator Analysis

(A) Cluster analysis of signaling network for significantly up-regulated genes (*n* = 3,996). Representative enriched pathways (Reactome) and GO terms (Biological Process) identified in each individual cluster were listed alongside. See also Table S10.

(B) Heat plot showing percentage of cell type signature genes in DEG and expressed genes (both control and UCC2003 groups). All expressed genes are well represented in IEC cell type signature genes.

(C) Cell type analysis on IEC DEGs using known cell-specific signature genes. Stem cells were statistically over-represented in DEGs. *p* < 0.05. See also Table S11.

(D) Key regulators of stem cell DEGs.
tuft cells (22%), enteroendocrine cells (18%), goblet cells (15%), Paneth cells (15%), and enterocytes (13%; Figure 5C). These data indicated that intestinal epithelial stem cells, cells primarily involved in cell differentiation, were the primary cell type whose numbers and transcriptomic program were regulated by B. breve UCC2003.

Further investigation of this stem cell signature revealed that of the 37 differentially expressed marker genes, 35 are up-regulated in the presence of B. breve UCC2003. This indicates an increase in the quantity of stem cells or semi-differentiated cells in the epithelium, consistent with the overrepresentation of cell cycle- and DNA replication-associated genes observed in the whole differential expression dataset. Functional analysis of the 37 stem cell signature genes revealed only one overrepresented process—Regulation of Frizzled by ubiquitination (p < 0.05), which is a subprocess of WNT signaling. WNT signaling is important in maintaining the undifferentiated state of stem cells (Nusse, 2008).

Finally, we employed a network approach to predict key transcription factor (TF) regulators of the differentially expressed stem cell marker genes, through which B. breve UCC2003 may be acting (Figure 5D). Using the TF-target gene database, DoRothEA, we identified expressed TFs known to regulate these genes (Garcia-Alonso et al., 2019; Holland et al., 2019). Five genes had no known and expressed regulator, and thus were excluded. Hypergeometric significance testing was used to identify which of these TFs are the most influential (see Methods for details). This analysis identified 32 TF regulators (Figure 5D). Of these regulators, 12 were differentially expressed in the IEC dataset (all up-regulated): Fos, Gabpa, Rcor1, Arid2, Tead1, Myb2, Met2a, Ahr, Pgr, Kmt2a, Ncoa2, and Tcf12. Functional analysis of all the TF regulators and their targeted genes together, revealed overrepresented functions relating to WNT signaling, histone methylation for self-renewal and proliferation of hematopoietic stem cells, and nuclear receptor (incl. estrogen) signaling (Table S12). These data indicate that B. breve UCC2003 directly affects key transcriptomic programs that regulate specific signaling processes, particularly within stem cells.

DISCUSSION
The early life developmental window represents a crucial time for microbe-host interactions that impacts health both in the short and longer term. Understanding how specific microbiota members modulate host responses in pre-clinical models may help the design and development of next-stage targeted microbiota therapies in humans. Here we investigated how B. breve UCC2003 induces genome-wide transcriptomic changes in small intestine IECs of neonatal mice. We observed that B. breve had a global impact on the IEC transcriptome, evidenced by the large number of significantly up-regulated genes and pathways related to cell differentiation and cell proliferation, including genes associated with epithelial barrier function. We propose that B. breve may act as a key early life microbiota member driving fundamental cellular responses in murine IECs, particularly within the stem cell compartment, and thus drives epithelial barrier development and maintenance during neonatal life stages. However, further clinical studies would be required to determine if our findings extrapolate to the human setting.

B. breve is known to confer beneficial effect on gut health; however, our knowledge related to the mechanisms underlying these responses is limited. Most studies have focused on targeted immune cells or pathways (during disease and/or inflammation), and to our knowledge no studies have probed global transcriptomic changes within IECs, the frontline physical barrier between bacteria and host (Turroni et al., 2014; Gann, 2010). Our presented findings in a pre-clinical model, namely, ~4,000 up-regulated DEGs and ~450 down-regulated DEGs within the B. breve group, indicate that this Bifidobacterium strain modulates whole-scale changes within this critical single-cell layer. Notably, we also examined how B. breve modulates adult IEC responses; however, we did not observe any significantly differently regulated genes when compared with control animals. The striking differences in DEGs between these two life points indicate that B. breve modulation of IECs is limited to the neonatal window. Dominance of Bifidobacterium in early life (including strains of B. breve) overlaps with the development and maturation of many host responses, including epithelial barrier integrity. Therefore, presence of these strains would be expected to play an over-sized role in this initial homeostatic priming, which may afford protection against inflammatory insults in later life, as has been shown previously in a mouse model of pathological epithelial cell shedding (Hughes et al., 2017). Further clinical studies would be required to probe these findings in detail to determine their importance during healthy infant development.

Exploring the murine transcriptional responses in more detail revealed that expressions of key genes associated with formation of epithelial barrier components were up-regulated, including major cell junction
protein-encoding genes (75%; 42/56 genes). More specifically, several integrin-associated genes were up-regulated in the presence of UCC2003. Integrins facilitate cell-cell and cell-extracellular matrix adhesion and binding and assembly of the fibronectin matrix that is pivotal for cell migration and cell differentiation (Harburger and Calderwood, 2009; Qin et al., 2004; Mosher et al., 1991). Integrins also play an important role in downstream intracellular signaling that controls cell differentiation, proliferation, and cell survival, including the Raf-MEK-ERK signaling pathway (we also observed enrichment of genes involved in this pathway) (Chernyavsky et al., 2005; Li et al., 2016). Another key intestinal barrier component is represented by TJs, linking complexes between intercellular spaces, and comprise transmembrane proteins including occludins, claudins, zona occludens, and junctional adhesion molecules (Edelblum and Turner, 2009; Groschwitz and Hogan, 2009). Dysfunctional TJ may lead to a “leaky” gut, which is characteristic of numerous intestinal disorders including inflammatory bowel diseases (Krug et al., 2014). Notably, previous work has suggested early life microbiota disruptions (via antibiotic usage) and reductions in numerous intestinal disorders including inflammatory bowel diseases (Krug et al., 2014). Notably, previous work has suggested early life microbiota disruptions (via antibiotic usage) and reductions in bifidobacterial-associated metabolites on these responses would require further studies to confirm metabolic pathway(s) connecting with putative signaling pathways.

Further network and functional analysis indicated that clusters of up-regulated DEGs were associated with cell maturation and cell differentiation (as confirmed by cell type-specific analysis), suggesting that neonatal B. breve exposure positively modulates IEC cell differentiation, growth, and maturation. Somewhat surprisingly, we did not observe the same type of striking responses in immune pathways, which may be masked by the sheer number of DEGs involved in cellular differentiation and processes. However, pathways such as TLR1 and TLR2 pathways do appear to be enriched (cluster 2 of signaling network analysis). This may link to previous work indicating that the UCC2003 EPS signals via TLR2 to induce MyD88 signaling cascades to protect IECs during intestinal inflammation (Hughes et al., 2017). B. breve M-16V was also shown to interact with TLR2 to up-regulate ubiquitin-editing enzyme A20 expression that correlated with increased tolerance to a TLR4 cascade in porcine IECs, further supporting the involvement of B. breve in programming key host immunoregulation receptors (Tomesada et al., 2013).

Cell type-specific analysis of DEGs revealed stem cells as the IEC type most affected by B. breve, with absorptive enterocytes least affected despite being most accessible to bacteria in the gut. It could be hypothesized that B. breve or their secreted metabolites may reach the crypts of the small intestinal epithelium. This has been previously suggested by in situ hybridization histology in vivo and by Bifidobacterium-conditioned media altering the expression of hundreds of host epithelial genes linked to immune response, cell adhesion, cell cycle, and development in IECs in vitro (Hughes et al., 2017; Guo et al., 2015). However, the direct impact of bifidobacterial-associated metabolites on these responses would require further studies to confirm metabolic activity of B. breve within the small intestine (via transcriptomics and metabolomics), although daily
supplementation with live bacteria may also provide a source of these metabolites in our model. Interestingly, certain Bifidobacterium and Lactobacillus strains that have been heat killed have also been shown to induce host responses, indicating that surface structures alone may play a role in downstream effects (Pique et al., 2019). All but two of the 37 differentially expressed stem cell marker genes were up-regulated in the presence of B. breve UCC2003, indicating an activating effect resulting in increased pluripotency of stem cells, increased quantity of stem cells, and/or an increased quantity of semi-differentiated cells. Single-cell sequencing of IECs could be used to further investigate this finding. Thirty-two TFs were predicted to regulate these stem cell signature genes, providing possible targets for future investigation of the mechanisms underlying these responses. Functional analysis of the stem cell signature genes and their regulators suggests that B. breve increases pluripotency of stem cells and/or semi-differentiated epithelial cells through WNT signaling and nuclear hormone signaling (Jeong and Mangelsdorf, 2009). Furthermore, the overrepresentation of the process “RUNX1 regulates transcription of genes involved in differentiation of HSCs” indicates a possible role for histone methylation in response to B. breve UCC2003 (Imperato et al., 2015). Further determination of host and bacterial metabolome and proteome after B. breve exposure may allow identification of the specific underlying molecular mechanisms (Guo et al., 2015).

In conclusion, B. breve UCC2003 plays a central role in orchestrating global neonatal IEC gene responses in a distinct manner as shown in our murine model, modulating genes involved in epithelial barrier development, and driving universal transcriptomic alteration that facilitates cell replication, differentiation, and growth, particularly within the stem cell compartment. This study enhances our overall understanding of the benefits of specific early life microbiota members in intestinal epithelium development, with prospective avenues to probe further health-promoting mechanisms of Bifidobacterium in humans. Further work exploring time-dependent transcriptional responses and impact of other Bifidobacterium species and strains (and use of mutants and transcriptionally active strains as positive controls), in tandem with metabolomic and proteomic approaches, is required to advance our understanding on the key host pathways and bifidobacterial molecules governing development and maturation of the intestinal barrier during the early life window. Nevertheless, further clinical studies would be essential to explore if these responses and findings are similar to those observed in humans.

Limitations of the Study
As we only observed low relative abundance of Bifidobacterium in our control neonatal animals this may suggest that induction of responses may be linked to the introduction of a new microbiota member (i.e., B. breve UCC2003), therefore results should be carefully interpreted. However, we did not observe associated global transcriptional inflammatory immune changes that would be expected if this was the case, but rather global changes in barrier function transcripts and pathways. Furthermore, Bifidobacterium has previously been isolated from C57BL/6 mice (including from our mouse colony), and therefore appears to be a resident rodent gut microbiota member, although it is found at varying abundances in different animal units and suppliers (Grimm et al., 2015; Hughes et al., 2020). Indeed, one particular study has shown that high levels of resident Bifidobacterium in mice directly correlated with improved immune responses to cancer immunotherapies (Sivan et al., 2015). In addition, we did not explore if B. breve UCC2003 is potentially driving more nuanced microbe-microbe interactions, and that, indirectly, these may also be stimulating IEC responses. Therefore, further studies probing these aspects in more detail, and comparing other Bifidobacterium strains, to compare and contrast responses, would be of interest.

B. breve UCC2003 is a model strain that was previously isolated from the stool of a breast-fed infant (National Collection of Industrial Food and Marine Bacteria, 2020, Sheehan et al., 2007). Although a human-associated strain, it has not been used in clinical studies, so directly extrapolating to human-specific settings should be cautiously considered. Further large-scale clinical studies would be required to confirm any positive strain-level impacts; however, in-depth analysis of, e.g., small IECs would be unethical in a healthy infant cohort, which emphasizes the importance of preclinical models.

Previous studies have shown that this strain can efficiently colonize (long term) the mouse gastrointestinal tract; however, we could not confirm this in our short-term, daily supplementation study (Cronin et al., 2008, O’Connell Motherway et al., 2011). Therefore the IEC responses observed may occur as a result of transient interactions with B. breve UCC2003 as it passes through the small intestine. Nevertheless, although at lower levels (~10^5 CFU/g), we did observe viable B. breve UCC2003 in the small intestine, linking to our subsequent observations of significant impacts on the IEC transcriptome from this intestinal region.
Very-low-abundance microbiota members (<2% relative abundance), including Streptococcus, Rumino-
coccus, Prevotella, and Coprococcus, were significantly reduced in relative abundance compared with con-
trols, raising the question whether supplementation of Bifidobacterium could have reduced these taxa. Regrettably, we could not determine if this is a bifidobacterial effect due to the lack of longitudinal sam-
pies, and we did not quantify bacterial titer, which is an important consideration for future work. We also did not profile microbial community composition within the small intestines, which is known to differ from fecal samples.

Resource Availability
Lead Contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by
the Lead Contact, Lindsay J. Hall (Lindsay.Hall@quadram.ac.uk).

Materials Availability
This study did not generate new unique reagents.

Data and Code Availability
The code generated for RNA-seq analysis during this study is available at GitHub https://github.com/
raymondkiu/Bifidobacterium-IEC-transcriptomics. The accession number for the raw sequencing reads
(both RNA-seq and 16S rRNA amplicon sequencing) reported in this paper is European Nucleotide Archive
(ENA): PRJEB36661.

METHODS
A l lm e th o ds c a nb ef o u n di nt h ea c c o m p a n y Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION
Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101336.

ACKNOWLEDGMENTS
This research was supported in part by the Norwich Bioscience Institutes (NBI) Computing infrastructure for
Science (CiS) group through the provision of a High-Performance Computing (HPC) Cluster. L.J.H. is
funded by a Wellcome Trust Investigator award (100974/C/13/Z) and together with T.K. by a BBSRC ISP
grant for Gut Microbes and Health BB/R012490/1 and its constituent project(s), BBS/E/F/000PR10353
and BBS/E/F/000PR10355. T.K. is also funded by the Genomics for Food security CSP grant from the BBSRC
(BB/CSP17270/1). A.T. is supported by the BBSRC Norwich Research Park Biosciences Doctoral Training
Partnership (grant BB/M011216/1). D.v.S. is supported by Science Foundation Ireland (SFI/12/RC/2273-
P1 and SFI/12/RC/2273-P2).

AUTHOR CONTRIBUTIONS
Conceptualization, R.K., L.C.H., and L.J.H.; Methodology, R.K., A.T., L.C.H., and L.J.H.; Software, R.K.,
A.T., and S.C.; Validation, R.K., A.T., T.K., and L.J.H.; Formal analysis, R.K. and A.T.; Investigation, R.K.,
A.T., L.C.H., and C.L.; Resources, S.C.; Data Curation, R.K.; Writing – Original Draft Preparation,
R.K., A.T., and L.J.H.; Writing – Review and Editing, R.K., A.T., D.v.S, T.K., and L.J.H.; Visualization, R.K.
and A.T.; Supervision, T.K. and L.J.H.; Project Administration, R.K.; Funding Acquisition, T.K., D.v.S., and
L.J.H.

DECLARATION OF INTERESTS
The authors declare no competing interests.

Received: March 30, 2020
Revised: June 5, 2020
Accepted: June 30, 2020
Published: July 24, 2020
REFERENCES

Amietz, M.C., Stensma, L.T., Amenyogbe, N., Brown, E.M., and Finlay, B. (2014). The intestinal microbiome in early life: health and disease. Front. Immunol. 5, 427.

Baucells, B.J., Mercadal, Hally, M., Alvarez Sanchez, A.T., and Figueras, A. (2016). Probiotic associations in the prevention of necrotising enterocolitis and the reduction of late-onset sepsis and neonatal mortality in preterm infants under 1,500g: a systematic review. An. Pediatr. (Barc.) 85, 247–255.

Chernysvsky, A.I., Arredondo, J., Karlsson, E., Wessler, I., and Grando, S.A. (2005). The Ras/Raf-1/MEK/ERK signaling pathway coupled to integrin expression mediates cholinergic regulation of keratinocyte directional migration. J. Biol. Chem. 280, 39220–39228.

Cronin, M., Sleator, R.D., Hill, C., Fitzgerald, G.F., and van Sinderen, D. (2008). Development of a luminal-based reporter system to monitor Bifidobacterium breve UCC2003 persistence in mice. BMC Microbiol. 8, 161.

Dersjen, M., Alvarez, A.S., and de Vos, W.M. (2019). The gut microbiota in the first decade of life. Trends Microbiol. 27, 997–1010.

Ding, A.U., Hassan, A., and Zhu, Y. (2020). Inhibitory effect of Bifidobacterium Animalis Bifidum ATCC25921 on colitis and its mechanism. J. Nutr. Biochem. 79, 108353.

Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc. Natl. Acad. Sci. U S A 107, 11971–11975.

Edelblum, K.L., and Turner, J.R. (2009). The tight junction in inflammatory disease communication breakdown. Curr. Opin. Pharmacol. 9, 715–720.

Ewaschuk, J.B., Diaz, H., Meddings, L., Diedrichs, B., Dmytrash, A., Backer, J., Looijer-van Langen, M., and Marwani, K.L. (2008). Secreted bioactive factors from Bifidobacterium infantis enhance epithelial Cell barrier function. Secreted bioactive factors from Bifidobacterium infantis enhance epithelial Cell barrier function. Biochim. Biophys. Acta Gene Regul. Mech. 1863, 194431.

Hsieh, C.Y., Osaka, T., Moriyama, E., Date, Y., Kikuchi, J., and Tsuneda, S. (2015). Strengthening of the intestinal epithelial tight junction by Bifidobacterium bifidum. PLoS Rep. 3, e12327.

Hughes, K.R., Harmsch, L.C., Alcon-Giner, C., Mitra, S., Wright, C.J., Ketskemety, J., van Sinderen, D., Watson, A.J., and Hall, J.J. (2017). Bifidobacterium breve reduces apoptotic epithelial cell shedding in an exopoxysaccharide and MyD88-dependent manner. Open Biol. 7, 160155.

Hughes, K.R., Schofield, Z., DaIby, M.J., Cairn, S., Chalikin, L., Bernuzzi, F., Alcon-Giner, C., Le Gall, G., Watson, A.J.M., and Hall, J.L. (2020). The early life microbiota protects neonatal mice from pathological small intestinal epithelial cell shedding. FASEB J. 34, 7075–7088.

Imperato, M.R., Cauchy, P., Obier, N., and Boniver, C. (2015). The RUNX1-PU1 axis in the control of hematopoiesis. Int. J. Hematol. 101, 319–329.

Jacobs, S.E., Tobin, J.M., Opie, G.F., Donath, S., Tabrizi, H.N., Pirotta, M., Morley, C.J., Garland, S.M., and Proprems, G. (2013). Probiotic effects on late-onset sepsis in very preterm infants: a randomized controlled trial. Pediatrics 132, 1055–1062.

Jeong, Y., and Mangelsdorf, D.J. (2009). Nuclear receptor regulation of stemness and stem cell differentiation. Exp. Mol. Med. 41, 525–537.

Kronman, M.P., Zaatouis, T.E., Haynes, K., Feng, R., and Coffin, S.E. (2012). Antibiotic exposure and IBD development among children: a population-based cohort study. Pediatrics 130, e794–803.

Krug, S.M., Schulze, J.D., and Fromm, M. (2014). Tight junction, selective permeability, and inflammatory diseases. Semin. Cell Dev. Biol. 36, 166–176.

Lawson, M.A.E., O’neill, I.J., Kujawa, M., Gowinadah Javadi, S., Wijeyesekera, A., Fleaga, Z., Chalklen, L., and Hall, J.I. (2020). Breast milk-derived human milk oligosaccharides promote Bifidobacterium interactions within a single ecosystem. ISME J. 14, 635–648.

Lewis, M.C., Merrifield, C.A., Berger, B., Cloarec, O., Dunker, S., Mercenier, A., Nicholson, J.K., Holmes, J., and Bailey, M. (2017). Early intervention with Bifidobacterium lactis nCC2818 modulates the host-microbe interface independent of the sustained changes induced by the neonatal environment. Sci. Rep. 7, 5310.

Li, L., Zhao, G.D., Shi, Z., Qi, L.L., Zhou, L.Y., and Fu, Z.X. (2016). The Ras/Raf/MEK/ERK signaling pathway and its role in the occurrence and development of gut bifidobacteria in a population-based cohort study. Pediatrics 137, 3–20, quiz 21–2.

Liu, C., Zhang, Z.Y., Dong, K., and Guo, X.K. (2010). Adhesion and immunomodulatory effects of Bifidobacterium lactis HNO19 on intestinal epithelial cells INT-407. World J. Gastroenterol. 16, 2283–2290.

Mikami, K., Kimura, M., and Takahashi, H. (2012). Influence of maternal bifidobacteria on the development of gut bifidobacteria in preterm infants. Pharmaceuticals (Basel) 5, 629–642.

Mosher, D.F., Fogerty, F.J., Chernousov, M.A., and Barry, E.L. (1991). Assembly of fibronectin into extracellular matrix. Ann. N. Y Acad. Sci. 614, 167–180.

Nagpal, R., Kurakawa, T., Tsui, H., Takahashi, T., Kawashima, K., Nagata, S., Nómoto, K., and Yamashiro, Y. (2017). Evolution of gut Bifidobacterium population in healthy Japanese infants over the first three years of life: a quantitative assessment. Sci. Rep. 7, 10099.

National Collection of Industrial Food and Marine Bacteria (NCIMB) (2020). NCIMB 8807 General Info (Aberdeen).

Ng, S.C., Benjamin, J.L., Mccarthy, N.E., Hedin, C.R., Koutsoumpas, A., Plamondon, S., Price, C.L., Hart, A.L., Kammi, M.A., Forbes, A., Knight, S.C., Lindsay, J.O., Whelan, K., and Stagg, A.J.
(2011). Relationship between human intestinal dendritic cells, gut microbiota, and disease activity in Crohn’s disease. Inflamm. Bowel Dis. 17, 2027–2037.

Nusse, R. (2008). Wnt signaling and stem cell activity in Crohn’s disease. Inflamm. Bowel Dis. 20, 15632–15649.

Plummer, E.L., Bulach, D.M., Murray, G.L., Jacobs, S.E., Tabrizi, S.N., Garland, S.M., and Proprems Study, G. (2018). Gut microbiota of preterm infants supplemented with probiotics: sub-study of the ProPrems trial. BMC Microbiol. 18, 184.

Qin, J., Virogradaova, O., and Plov, E.F. (2004). Integrin bidirectional signaling: a molecular view. PLoS Biol. 2, e169.

Riedel, C.U., Foata, F., Philippo, D., Adolfo, E., Ekman, B.J., and Blum, S. (2006). Anti-inflammatory effects of bifidobacteria by inhibition of LPS-induced NF-kappaB activation. World J. Gastroenterol. 12, 3729–3735.

Robertson, C., Savva, G.M., Clapuci, R., Jones, J., Marmion, H., Brown, E., Minohca, A., Hall, L.J., and Clarke, P. (2020). Incidence of necrotising enterocolitis before and after introducing routine prophylactic Lactobacillus and Bifidobacterium probiotics. Arch. Dis. Child. Fetal Neonatal Ed. 105, 380–386.

Shaw, S.Y., Blanchard, J.F., and Bernstein, C.N. (2010). Association between the use of antibiotics in the first year of life and pediatric inflammatory bowel disease. Am. J. Gastroenterol. 105, 2667–2692.

Sheehan, V.M., Sleator, R.D., Hill, C., and Fitzgerald, G.F. (2007). Improving gastric transit, gastrointestinal persistence and therapeutic efficacy of the probiotic strain Bifidobacterium breve UCC2003. Microbiology 153, 3563–3571.

Shirazi, T., Longman, R.J., Corfield, A.P., and Probert, C.S. (2000). Mucins and inflammatory bowel disease. Postgrad. Med. J. 76, 473–478.

Sivan, A., Corrales, L., Hubert, N., Williams, J.B., Aquino-Michales, K., Earley, Z.M., Benyamin, F.W., Lee, Y.M., Jabri, B., Alegre, M.L., Chang, E.B., and Gajewski, T.F. (2015). Commensal Bifidobacterium promotes antitumor immunity and facilitates anti-PD-L1 efficacy. Science 350, 1084–1089.

Srivatva, D., Schwarzer, M., Hudcovic, T., Zakostelska, Z., Drab, V., Spanova, A., Rittich, B., Kozakova, H., and Schabasova, I. (2015).

Bifidobacterium longum CCM 7952 promotes epithelial barrier function and prevents acute DSS-induced colitis in strictly strain-specific manner. PLoS One 10, e0134050.

Steed, H., Macfarlane, G.T., Blackett, K.L., Bahrami, B., Reynolds, N., Walsh, S.V., Cummings, J.H., and Macfarlane, S. (2010). Clinical trial: the microbiological and immunological effects of symbiotic consumption—a randomized double-blind placebo-controlled study in active Crohn’s disease. Aliment. Pharmacol. Ther. 32, 872–883.

Stewart, C.J., Ajami, N.J., O’Brien, J.L., Hutchinson, D.S., Smith, D.P., Wong, M.C., Ross, M.C., Lloyd, R.E., Doddsapani, H., Metcalf, G.A., Muzny, D., Gibbs, R.A., Vatanen, T., Huttonower, C., Xavier, R.J., and Rovers, M. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. Nature 562, 583–588.

Thoo, L., Noti, M., and Krebs, P. (2019). Keep calm: the intestinal barrier at the interface of peace and war. Cell Death Dis. 10, 849.

Tomasu, A., Villena, J., Murata, K., Chiba, E., Shimazu, T., Aso, H., Iwabuchi, N., Xiao, J.Z., Saito, T., and Kitazawa, H. (2013). Immunoregulatory effect of bifidobacteria strains in porcine intestinal epithelial cells through modulation of ubiquitin-editing enzyme A20 expression. PLoS One 8, e52925.

Turrini, F., Taverniti, V., Ruas-Madiedo, P., Duranti, S., Guglielmetti, S., Lucili, A., Gioiosa, P., Palanza, P., Margolles, A., van Sinderen, D., and Ventura, M. (2014). Bifidobacterium bifidum PRL2010 modulates the host innate immune response. Appl Environ Microbiol. 80, 730–740.

Van der Sluis, M., de Konin, B.A., de Bruin, A.C., Velich, A., Meijerink, J.P., van Goudoever, J.B., Buller, H.A., Dekker, J., van den Nuding, I., Benes, I.B., and Einerhand, A.W. (2006). Muc2 deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection. Gastroenterology 131, 117–129.

Zheng, B., van Bergenhenegouwen, J., Overbeek, S., van de Kort, H.J., Garssen, J., Folkerts, G., Vos, P., Morgan, M.E., and Kranekeld, A.D. (2014). Bifidobacterium breve attenuates murine dextran sodium sulfate-induced colitis and increases regulatory T cell subset responses. PLoS One 9, e95441.
Supplemental Information

*Bifidobacterium breve* UCC2003 Induces a Distinct Global Transcriptomic Program in Neonatal Murine Intestinal Epithelial Cells

Raymond Kiu, Agatha Treveil, Lukas C. Harnisch, Shabhonam Caim, Charlotte Leclaire, Douwe van Sinderen, Tamas Korcsmaros, and Lindsay J. Hall
Figure S1. Functional analysis on differentially expressed genes. Related to Figure 4. (A) Panther Slim GO-term major categories of significantly up-regulated genes (n=3,996). Related to Table S7 and Table S8. (B) Functional and pathway enrichment analysis on significantly up-regulated genes (Panther Slim GO-term). Only top 20 FDR-ranked enriched pathways (Reactome pathways) are shown. Statistical significance cut-offs: FDR<0.05. Statistical significance: Fisher’s Exact Test. Fold Enrichment was calculated against all expressed genes in IECs as the background (n=21,537). Related to Table S5, Table S6 and Table S9.
| Gene          | Ensembl ID             | Chromosome | Description/ Putative function                                      | Biological process (GO)                                                                 |
|--------------|------------------------|------------|---------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Gm27149      | ENSMUSG0000098426      | 7          | Unknown                                                             | Unknown                                                                               |
| Ccnb1p1      | ENSMUSG0000071470      | 14         | Cyclin B1 interacting protein 1                                      | Cell differentiation, Cellular component organisation, Protein metabolic process         |
| Gm10359      | ENSMUSG0000094708      | 5          | glyceraldehyde-3-phosphate dehydrogenase pseudogene                  | Unknown                                                                               |
| Gm12671      | ENSMUSG0000095937      | 4          | glyceraldehyde-3-phosphate dehydrogenase pseudogene                  | Unknown                                                                               |
| Gm48216      | ENSMUSG0000114367      | 13         | Unknown                                                             | Unknown                                                                               |
| Gm17131      | ENSMUSG0000085328      | 5          | Unknown                                                             | Unknown                                                                               |
| Tmem72       | ENSMUSG0000048108      | 6          | transmembrane protein 72                                           | Unknown                                                                               |
| Ccdc107      | ENSMUSG0000028461      | 4          | coiled-coil domain containing 107                                   | Unknown                                                                               |
| CR936839.3   | ENSMUSG0000111855      | Unknown    | Unknown                                                             | Cell differentiation, Cellular component organisation, Immune system process, system development |
| Hist1h4b     | ENSMUSG0000069266      | 13         | H4 clustered histone 2                                              | Cell differentiation, Cellular component organisation, Immune system process, system development |
| Gm48836      | ENSMUSG0000113523      | 12         | Unknown                                                             | Unknown                                                                               |
| Gm42669      | ENSMUSG0000106631      | 5          | Unknown                                                             | Unknown                                                                               |
| Gprin3       | ENSMUSG0000045441      | 6          | GPRIN family member 3                                               | Unknown                                                                               |
| Fgd4         | ENSMUSG0000022788      | 16         | FYVE, RhoGEF and PH domain containing 4                             | Cellular component of organisation, Protein metabolic process, Response to stimulus, Signaling |
| D10Wsu102e   | ENSMUSG0000020255      | 10         | DNA segment, Chr 10, Wayne State University 102, expressed          | Unknown                                                                               |
| Gm20594      | ENSMUSG0000096887      | 6          | Unknown                                                             | Cell death, Response to stimulus, Signaling                                           |
| Vps13b       | ENSMUSG0000037646      | 15         | vacuolar protein sorting 13B                                        | Establishment of localisation                                                          |
| Gm48366      | ENSMUSG0000113523      | 12         | Unknown                                                             | Unknown                                                                               |
| Gm48054      | ENSMUSG0000113921      | 13         | Unknown                                                             | Unknown                                                                               |
| Naip6        | ENSMUSG0000078942      | 13         | NLR family, apoptosis inhibitory protein 6                           | Cell death, Immune system process, Response to stimulus                                |

Table S3. Annotation of top 20 significantly up-regulated genes. Related to Figure 4.
| Gene    | Ensembl ID       | Chromosome | Description/ Putative function              | Biological process (GO)                                                                 |
|---------|------------------|------------|---------------------------------------------|----------------------------------------------------------------------------------------|
| Gm7849  | ENSMUSG00000079114 | 8          | Defensin                                    | Immune system process, response to stimulus                                            |
| Rps27rt | ENSMUSG00000050621 | 9          | ribosomal protein S27, retrogene            | Unknown                                                                                  |
| Gm6158  | ENSMUSG00000090381 | 14         | Unknown                                     | Unknown                                                                                  |
| mt-Co2  | ENSMUSG00000064354 | MT         | mitochondrially encoded cytochrome c oxidase II | Carbohydrate derivative metabolism, cell death                                             |
| Gm3650  | ENSMUSG00000097891 | 18         | Unknown                                     | Unknown                                                                                  |
| Gm5831  | ENSMUSG00000111133 | 9          | Unknown                                     | Unknown                                                                                  |
| Gm7331  | ENSMUSG00000059461 | x          | Unknown                                     | Unknown                                                                                  |
| Polr3h  | ENSMUSG00000022476 | 15         | polymerase (RNA) III (DNA directed) polypeptide H | Immune system process, response to stimulus                                             |
| Gm14698 | ENSMUSG00000071748 | X          | Unknown                                     | Unknown                                                                                  |
| Ftl1-ps1| ENSMUSG00000062382 | 13         | erritin light polypeptide 1                 | Unknown                                                                                  |
| Rps16   | ENSMUSG00000037563 | 7          | ribosomal protein S16                       | Protein metabolic process, response to stimulus                                          |
| Mt2     | ENSMUSG00000031762 | 8          | metallothionein 2                           | Homeostatic process, response to stimulus, signaling                                    |
| Mt1     | ENSMUSG00000031765 | 8          | metallothionein 1                           | Cell death, homeostatic process, response to stimulus, signaling                         |
| Hba-a1  | ENSMUSG00000069919 | 11         | hemoglobin alpha                            | Cell differentiation, homeostatic process, immune system process, response to stimulus, signaling, system development |
| Hbb-bt  | ENSMUSG00000073940 | 7          | hemoglobin, beta adult t chain              | Unknown                                                                                  |
| Rps2-ps10| ENSMUSG00000091957 | 18         | ribosomal protein S2                        | Unknown                                                                                  |
| Gpx3    | ENSMUSG00000018339 | 11         | glutathione peroxidase 3                    | Response to stimulus                                                                    |
| Smcp    | ENSMUSG00000074435 | 3          | sperm mitochondria-associated cysteine-rich protei | Unknown                                                                                  |
| Gm49384 | ENSMUSG00000113786 | 12         | Predicted gene                              | Unknown                                                                                  |
| Gm7336  | ENSMUSG00000078636 | 7          | Predicted gene                              | Unknown                                                                                  |

Table S4. Annotation of top 20 significantly down-regulated genes. Related to Figure 4.
Table S5. GO Biological Process enrichment analysis in up-regulated DEGs. Related to Figure 4 and Figure S1.

| GO Biological Process                                         | Gene Count | Fold Enrichment | FDR      |
|---------------------------------------------------------------|------------|-----------------|----------|
| Transcription, DNA-templated                                  | 252        | 1.41            | 1.15E-03 |
| Detection of chemical stimulus involved in sensory perception | 97         | 1.82            | 1.92E-03 |
| Sensory perception of chemical stimulus                       | 98         | 1.76            | 1.97E-03 |
| Cellular response to DNA damage stimulus                      | 66         | 1.84            | 1.54E-02 |
| Sensory perception                                            | 111        | 1.59            | 1.57E-02 |
| Gene expression                                                | 350        | 1.25            | 2.66E-02 |
| Chromosome organization                                       | 57         | 1.86            | 2.66E-02 |
| Transcription by RNA polymerase II                            | 193        | 1.35            | 3.09E-02 |
| Regulation of metabolic process                                | 262        | 1.28            | 4.48E-02 |

Table S6. GO Molecular Functions enrichment analysis in up-regulated DEGs. Related to Figure 4 and Figure S1.

| GO Molecular Functions                                        | Gene Count | Fold Enrichment | FDR      |
|---------------------------------------------------------------|------------|-----------------|----------|
| Transmembrane signaling receptor activity                     | 279        | 1.56            | 5.07E-08 |
| G-protein coupled receptor activity                           | 224        | 1.63            | 8.61E-08 |
| Molecular transducer activity                                 | 321        | 1.43            | 1.76E-06 |
| Signaling receptor activity                                   | 302        | 1.43            | 4.06E-06 |
| DNA binding                                                   | 225        | 1.49            | 1.20E-05 |
| Nucleic acid binding                                          | 344        | 1.33            | 1.14E-04 |
| Heterocyclic compound binding                                 | 352        | 1.29            | 7.08E-04 |
| Microtubule binding                                           | 43         | 2.2             | 3.01E-03 |
| Protein kinase activity                                       | 133        | 1.46            | 7.97E-03 |
| Tubulin binding                                               | 45         | 1.94            | 1.24E-02 |
| Protein serine/threonine kinase activity                      | 94         | 1.47            | 4.74E-02 |
| Phosphotransferase activity, alcohol group as acceptor       | 144        | 1.36            | 5.04E-02 |
| Biological Process | Genes | Percentage |
|--------------------|-------|------------|
| cellular process (GO:0009987) | 901 | 31.70% |
| metabolic process (GO:0008152) | 597 | 21.00% |
| biological regulation (GO:0065007) | 492 | 17.30% |
| localization (GO:0051179) | 239 | 8.40% |
| multicellular organismal process (GO:0032501) | 233 | 8.20% |
| response to stimulus (GO:0050896) | 117 | 4.10% |
| developmental process (GO:0032502) | 55 | 1.90% |
| biological adhesion (GO:0022610) | 55 | 1.90% |
| immune system process (GO:0002376) | 52 | 1.80% |
| cellular component organization or biogenesis (GO:0071840) | 37 | 1.30% |
| reproduction (GO:0000003) | 37 | 1.30% |
| cell proliferation (GO:0008283) | 15 | 0.50% |
| biological phase (GO:0044848) | 10 | 0.40% |
| rhythmic process (GO:0048511) | 4 | 0.10% |
| multi-organism process (GO:0051704) | 2 | 0.10% |

Table S7. GO Biological Process functional assignment to 3,996 upregulated genes. Related to Figure 4 and Figure S1.

| Molecular Function | Genes | Percentage |
|--------------------|-------|------------|
| binding (GO:0005488) | 868 | 37.60% |
| catalytic activity (GO:0003824) | 671 | 29.00% |
| molecular transducer activity (GO:0060089) | 321 | 13.90% |
| transcription regulator activity (GO:0140110) | 156 | 6.80% |
| molecular function regulator (GO:0098772) | 127 | 5.50% |
| transporter activity (GO:0005215) | 108 | 4.70% |
| structural molecule activity (GO:0005198) | 53 | 2.30% |
| translation regulator activity (GO:0045182) | 4 | 0.20% |
| cargo receptor activity (GO:0038024) | 2 | 0.10% |

Table S8. GO Molecular Function functional assignment to 3,996 upregulated genes. Related to Figure 4 and Figure S1.
| Reactome Pathway                                               | Gene Count | Fold Enrichment | FDR       |
|---------------------------------------------------------------|------------|-----------------|-----------|
| Olfactory Signaling Pathway                                   | 108        | 1.88            | 3.20E-05  |
| Cell Cycle                                                    | 154        | 1.63            | 1.13E-04  |
| Gene expression (Transcription)                              | 238        | 1.45            | 1.18E-04  |
| Chromatin organization                                        | 67         | 2.08            | 2.54E-04  |
| Chromatin modifying enzymeS                                   | 67         | 2.08            | 2.82E-04  |
| DNA Double-Strand Break Repair                                | 51         | 2.3             | 4.19E-04  |
| G alpha (s) signalling events                                 | 122        | 1.62            | 6.50E-04  |
| RNA Polymerase II Transcription                              | 202        | 1.39            | 3.06E-03  |
| Generic Transcription Pathway                                | 173        | 1.39            | 8.50E-03  |
| Signal Transduction                                           | 461        | 1.21            | 8.98E-03  |
| G2/M DNA damage checkpoint                                    | 29         | 2.48            | 9.07E-03  |
| DNA Repair                                                    | 77         | 1.66            | 1.06E-02  |
| Cell Cycle, Mitotic                                           | 122        | 1.48            | 1.06E-02  |
| HATs acetylate histones                                       | 23         | 2.64            | 1.85E-02  |
| Epigenetic regulation of gene expression                     | 28         | 2.28            | 2.38E-02  |
| Homology Directed Repair                                      | 30         | 2.23            | 2.39E-02  |
| Nonhomologous End-Joining                                     | 19         | 2.82            | 2.46E-02  |
| Mitotic Prometaphase                                          | 58         | 1.73            | 2.48E-02  |
| HDR through Homologous Recombination                          | 29         | 2.26            | 2.63E-02  |
| Recruitment and ATM-mediated phosphorylation of repair        | 22         | 2.53            | 2.78E-02  |

Table S9. Reactome pathway enrichment analysis in up-regulated DEGs. Related to Figure 4 and Figure S1.

| Reactome Pathway                                               | Gene Count | Fold Enrichment | FDR       |
|---------------------------------------------------------------|------------|-----------------|-----------|
| Regulation of FZD by ubiquitination (R-MMU-4641263)            | 3          | 35.6            | 2.39E-02  |
| Endogenous sterols (R-MMU-211976)                             | 3          | 27.06           | 3.63E-02  |
| Estrogen-dependent gene expression (R-MMU-9018519)            | 6          | 19.61           | 4.95E-04  |
| ESR-mediated signaling (R-MMU-8939211)                        | 6          | 18.04           | 4.69E-04  |
| RUNX1 regulates transcription of genes involved in differentiation of HSCs (R-MMU-8939236) | 4          | 14.78           | 3.41E-02  |
| Signaling by Nuclear Receptors (R-MMU-9006931)                | 6          | 11.66           | 4.23E-03  |
| Transcriptional regulation by RUNX1 (R-MMU-8878171)           | 6          | 9.73            | 9.64E-03  |
| Generic Transcription Pathway (R-MMU-212436)                  | 15         | 5.88            | 4.71E-05  |
| RNA Polymerase II Transcription (R-MMU-73857)                 | 15         | 5               | 1.89E-04  |
| Gene expression (Transcription) (R-MMU-74160)                 | 15         | 4.38            | 4.93E-04  |
| Signal Transduction (R-MMU-162582)                            | 19         | 2.37            | 3.65E-02  |

Table S12. Reactome pathway enrichment analysis of differentially expressed stem cell signature genes and their expressed regulators. Related to Figure 5.
TRANSPARENT METHODS

Animals
All animal experiments and related protocols were performed in accordance with the Animals (Scientific Procedures) Act 1986 (ASPA) under project licence (PPL: 80/2545) and personal licence (PIL: I68D4DCCF), approved by UK Home Office and University of East Anglia (UEA) FMH Research Ethics Committee. C57BL/6J two-week-old neonatal female mice (n=10) were housed in two separate cages with their mothers within UEA Disease Modelling Unit. Mice were euthanised via ASPA Schedule 1 protocol (CO₂ and cervical dislocation).

Bacterial culturing, inoculum preparation and CFU enumeration
*B. breve* UCC2003 (also known as NCIMB 8807) was streaked from frozen glycerol stocks onto autoclaved Reinforced Clostridial Agar (RCA) plates (Oxoid, UK) and incubated in an anaerobic chamber (miniMACS, Don Whitley Scientific) at 37°C for 48 h prior to picking single colonies for inoculation in prewarmed sterilised Reinforced Clostridial Medium (Oxoid, UK).

For preparation of gavage inoculums, 5 ml of inoculated broth was incubated overnight followed by subculturing into 5 ml De Man, Rogosa and Sharpe (MRS) medium (Oxoid). After an additional overnight incubation, another sub-culturing into 40 ml RCM was performed. Inoculums were prepared from cultures by 3 rounds of centrifugation at 3220 g for 10 min followed by three PBS washes before dilution in 4 ml (adult mice) or 2 ml (neonatal mice) sterile PBS. Bacterial concentration of inoculum was enumerated by plating serial dilutions in sterile PBS on RCA plates and enumerating colonies following two-day incubation to calculate CFU/ml.

Bacterial treatment and administration
Neonatal mice were orally gavaged with *B. breve* UCC2003 inoculations of 10⁸ CFU/ml in 50 μl every 24 h for 3 consecutive days. Control mice received oral gavages of sterile PBS. *B. breve* UCC2003 viable presence/transition through the gut was confirmed by collection of fresh faeces or intestinal content homogenised with 1 ml sterile PBS followed by serial-dilution plating in sterile PBS on RCA supplemented with 50 mg/L mupirocin and counting of colonies following two-day incubation to calculate CFU/mg.

Gut microbiota profiling by 16S rRNA amplicon sequencing and analysis
Genomic DNA extraction of mouse faecal samples on day 4 was performed with FastDNA Spin Kit for Soil (MP Biomedicals) following manufacturer’s instructions and extending the bead-beating step to 3 min as described previously (Alcon-Giner et al., 2019). Extracted DNA was quantified, normalised and sequenced on Illumina MiSeq platform using a read length of 2 × 300 bp. After quality pre-filtering and removals of chimeras, sequencing reads were analysed using open-reference OTU clustering strategy (QIIME v1.9.1) to assign bacterial taxonomy based on SILVA_132 database (Quast et al., 2013). OTU tables in BIOM format was converted to genus counts in MEGAN6 and visualised using R library ggplot2 as described previously (Kiu et al., 2019, Caporaso et al., 2010, Huson et al., 2016).

Tissue collection and isolation of small intestinal epithelial cells (IECs)
Upon tissue harvesting, 0.5 cm² sections of small intestines were collected in 200 μl RNAlater™ (Thermo Fisher Scientific) at the animal unit prior to IECs isolation (from fresh samples) via an adapted Weisser method as described below (Hughes et al., 2017). Sections of small intestines were placed in ice-cold PBS in 200 ml Duran bottles. Faecal matter was washed off by inverting 10 times in 0.154M NaCl and 1mM DTT. Liquid was drained and mucus layer removed through incubation of samples in 1.5mM KCl, 96mM NaCl, 27 mM Tri-sodium citrate, 8mM NaH₂PO₄ and 5.6mM Na₂HPO₄ for 15 min at 220 rpm and 37°C. IECs were separated from basal membrane by incubation in 1.5 mM EDTA and 0.5 mM DTT for 15 min at 200 rpm and 37 °C followed by shaking vigorously 20 times. IECs were collected from solution by centrifugation at 500 g for 10 min at 4 °C. Supernatant was then discarded and cell pellet resuspended in 3 ml of ice-cold PBS. Cell concentrations of
isolated IEC samples calculated by labelling dead cell with trypan blue at a 1:1 v/v ratio and enumeration of viable cells using a Neubauer haemocytometer on an inverted microscope (ID03, Zeiss).

**RNA extraction and sequencing**

RNA was extracted from IECs by adding a volume containing 2 x 10^6 cells in PBS to QIAshredder spin columns (QIAGEN) followed by centrifugation at 9,300 g for 1 min. Follow-through was mixed with 600 μl RLT lysis buffer and used for subsequent RNA isolation. Homogenised samples in RLT buffer from IECs were processed by adding 700 μl of 70% ethanol and mixing by pipetting. Samples were then added into RNasey spin column and spun at 8,000 g for 15 s. Flow-through was discarded and process repeated until all of sample was filtered through column. Then 700 μl of buffer RW1 was added to column and centrifuge at 8,000 g for 30 s. Again, flow through was discarded and filter placed in a new collection tube. To the filter, 500 μl RPE was added and spun at 8,000 g for 2 min. Spin column was then placed in a new collection tube and centrifuged at 8,000 g for 2 min. Columns were transferred to a RNA low-bind Eppendorf tube and 30 μl of RNase free water added to directly to the filter. After an incubation of 1 min at RT, sample was centrifuged at 8,000 g for 1 min and flow through containing RNA stored at -80°C.

Purified RNA was quantified, and quality controlled using RNA 6000 Nano kit on a 2100 Bioanalyser (Agilent). Only samples with RIN values above 8 were sequenced. RNA sequencing was performed at the Wellcome Trust Sanger Institute (Hinxton, UK) on paired-end 75 bp inserts on an Illumina HiSeq 2000 platform. Isolated RNA was processed by poly-A selection and/or Ribo-depletion.

**Sequence pre-processing and Differential Gene Expression (DGE) analysis**

Sequencing quality of raw FASTQ reads were assessed by FastQC software (v0.11.8). FASTQ reads were subsequently quality-filtered using fastp v0.20.0 with options -q 10 (phred quality <10 was discarded) followed by merging reads into single read file for each sample (merge-paired-reads.sh) and rRNA sequence filtering via SortMeRNA v2.1 based on SILVA rRNA database optimised for SortMeRNA software (Chen et al., 2018, Kopylova et al., 2012). Filtered reads were then unmerged (unmerge-paired-reads.sh) and ready for transcript quantification.

Transcript mapping and quantification were performed using Kallisto v0.44.0 (Bray et al., 2016). Briefly, *Mus musculus* (C57BL/6 mouse) cDNA sequences (GRCm38.release-98_k31) were retrieved from Ensembl database and built into an index database with Kallisto utility index at default parameter that was used for following transcript mapping and abundance quantification via Kallisto utility quant at 100 bootstrap replicates (-b 100) (Zerbino et al., 2018).

DGE analysis was performed using R library Sleuth (v0.30.0) (Pimentel et al., 2017). Gene transcripts were mapped to individual genes using Ensembl BioMart database with Sleuth function sleuth_prep with option gene_mode = TRUE. Genes with an absolute log_2(fold change) >1.0 (based on Wald test statistics) and q value <0.05 (or, FDR-adjusted p value; based on likelihood ratio test) were considered to be significantly regulated (Kinsella et al., 2011).

**Functional annotation and enrichment analysis**

Functional assignment and enrichment analysis was performed using PANTHER Classification System (Mi et al., 2019a). Briefly, for functional assignment analysis, a list of genes of interest in Ensembl IDs were supplied to the webserver to be mapped to the Mouse Genome Database (MGD) to generate functional classification on those genes of interest (Bult et al., 2019). For functional enrichment analysis, a gene list was supplied together with a background gene list in Ensembl IDs to Panther web server, then selected ‘functional overrepresentation test’ and chose a particular function class in the drop-down menu. Recommended by the database developers, Fisher’s exact test and False Discovery Rate (FDR) correction were used to perform enrichment analysis (Mi et
al., 2019b). FDR <0.05 was used as the default cut-off for significant enrichment. Functional annotation of top 20 up/down-regulated genes was assigned manually via Ensembl and/or MGI (Mouse Genome Informatics) databases (Bult et al., 2019, Cunningham et al., 2019).

Network, cluster and signalling pathway analysis
A signalling network of all up-regulated DEGs and their first neighbours was built using all available biological signalling databases in the Cytoscape (v3.7.2) OmniPath app (v1, *Mus musculus*) (Turei et al., 2016, Shannon et al., 2003). Modules of highly connected genes within the signalling network were identified using the MCODE plug-in within Cytoscape (Bader and Hogue, 2003). Settings below were applied to obtain clusters in the network: degree cutoff = 3, haircut = true, fluff = false, node score cutoff = 0.5, k-core = 3 and max depth = 100.

The nodes of each individual module were tested for functional enrichment based on both Reactome and PANTHER annotations using PANTHER Classification System as described in previous sub-section ‘Functional annotation and enrichment analysis’ (Mi and Thomas, 2009, Croft et al., 2011, Mi et al., 2019a).

Enrichment of cell type specific marker genes
Cell type signature gene sets for murine intestinal epithelial cells were obtained from Haber et al. (Haber et al., 2017). Both droplet and plate-based results were used. Gene symbols were converted to Ensembl IDs using db2db (Mudunuri et al., 2009). Hypergeometric significance calculations were applied to test the presence of cell type specific signatures in the list of differentially expressed genes using all expressed genes as the statistical background (normalised counts > 1 in ≥ 1 sample). Bonferroni multiple correction was applied and any corrected p < 0.05 was deemed significant. Genes with normalised counts > 1 in ≥ 1 sample per condition (*B. breve* UCC2003 treated or control) were used to identify cell type signature genes expressed per condition.

Key regulator analysis
All mouse transcription factor - target gene interactions with quality scores A-D were obtained from DoRothEA v2 via the OmniPath Cytoscape app (Garcia-Alonso et al., 2019, Shannon et al., 2003, Turei et al., 2016). A subnetwork was generated consisting of differentially expressed stem cell signature genes and all their upstream TFs which were expressed in the transcriptomics dataset (normalised counts > 1 in ≥ 1 sample). These TFs were further filtered for their relevance in the network. Here all expressed genes and their upstream expressed TFs were extracted from the DoRothEA network. A hypergeometric significance test was carried out on any node with degree ≥ 5 to determine if the proportion of connected nodes which are differentially expressed is higher than in the whole network. Any TF with p < 0.05 following Benjamini-Hochberg correction were deemed significant and used to filter the stem cell signature gene subnetwork. Network visualisation was carried out in Cytoscape (Shannon et al., 2003). Functional enrichment carried out against Reactome pathways as described in previous sub-sections.

Statistical analyses and graphing
Student t-tests were performed using Rv.3.6.0, details of which were provided in the results and figure legends (R Development Core Team, 2010). LDA statistical tests for microbiome analysis was performed via LEfSe on Galaxy platform using default parameters (Segata et al., 2011, Jalili et al., 2020). PCA was performed via R library ggfortify function autoplot and prcomp, while Shannon diversity index was computed via R library vegan (Dixon, 2003, Tang et al., 2016, R Development Core Team, 2010). All other relevant statistical analyses (including enrichment analysis) were performed within specific software and details were provided in figure legends or as described in the previous sections.

All statistical graphs were either plotted using R library ggplot2 or Sleuth (Wickham, 2016, Pimentel et al., 2017). Heatmaps were graphed using R library gplots function heatmap.2 (Warnes et al., 2016).
Ethics approval
Animal experiments were performed under the UK Regulation of Animals (Scientific Procedures) Act of 1986. The project licence (PPL 80/2545) under which these studies were carried out was approved by the UK Home Office and the UEA Ethical Review Committee. Mice were sacrificed by CO₂ and cervical dislocation.

SUPPLEMENTAL REFERENCES
ALCON-GINER, C., DALBY, M. J., CAIM, S., KETSKEMETY, J., SHAW, A., SIM, K., LAWSON, M., KIU, R., LECLAIRE, C., CHALKLEN, L., KUJAWSKA, M., MITRA, S., FARDUS-REID, F., BELTEKI, G., MCCOLL, K., SWANN, J. R., KROLL, J. S., CLARKE, P. & HALL, L. J. 2019. Microbiota supplementation with Bifidobacterium and Lactobacillus modifies the preterm infant gut microbiota and metabolome. bioRxiv, 698092.

BADER, G. D. & HOGUE, C. W. 2003. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics, 4, 2.

BRAY, N. L., PIMENTEL, H., MELSTED, P. & PACHTER, L. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol, 34, 525-7.

BULT, C. J., BLAKE, J. A., SMITH, C. L., KADIN, J. A., RICHARDSON, J. E. & MOUSE GENOME DATABASE, G. 2019. Mouse Genome Database (MGD) 2019. Nucleic Acids Res, 47, D801-D806.

CAPORASO, J. G., KUCZYNSKI, J., STOMBAUGH, J., BITTINGER, K., BUSHMAN, F. D., COSTELLO, E. K., FIERER, N., PENA, A. G., GOODRICH, J. K., GORDON, J. I., HUTTLEY, G. A., KELLEY, S. T., KNIGHTS, D., KOENIG, J. E., LEY, R. E., LOZUPONE, C. A., MCDONALD, D., MUEEGGE, B. D., PIRRUNG, M., REEDER, J., SEVINSKY, J. R., TURNBAUGH, P. J., WALTERS, W. A., WIDMANN, J., YATSUNENKO, T., ZANEVELD, J. & KNIGHT, R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods, 7, 335-6.

CHEN, S., ZHOU, Y., CHEN, Y. & GU, J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34, i884-i890.

CROFT, D., O’KELLY, G., WU, G., HAW, R., GILLESPIE, M., MATTHEWS, L., CAUDY, M., GARAPATI, P., GOPINATH, G., JASSAL, B., SU, S., KALATSKAYA, I., MAHAJAN, S., MAY, B., NDEGWA, N., SCHMIDT, E., SHAMOVSKY, V., YUNG, C., BUCNEY, E., HERMJAKOB, H., D’EUSTACHIO, P. & STEIN, L. 2011. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res, 39, D691-7.

CUNNINGHAM, F., ACHUTHAN, P., AKANNI, W., ALLEN, J., AMODE, M. R., ARMEAN, I. M., BENNETT, R., BHAL, J., BILLIS, K., BODDU, S., CUMMINS, C., DAVIDSON, C., DODIYA, K. J., GALL, A., GIRON, C. G., GIL, L., GREGIO, T., HAGGERTY, L., HASKELL, E., HOUERLIE, T., IZOUOGU, O. G., JANACEK, S. H., JUETTEMANN, T., KAY, M., LAIRD, M. R., LAVIDAS, I., LIU, Z., LOVELAND, J. E., MARUGAN, J. C., MAUREL, T., MCMAHON, A. C., MOORE, B., MORALES, J., MUDGE, J. M., NUHN, M., OGEH, D., PARKER, A., PARTON, A., PATRICIO, M., ABDUL SALAM, A. I., SCHMITT, B. M., SCHULENBURG, H., SHEPPARD, D., SPARRROW, H., STAPLETON, E., SZUBA, M., TAYLOR, K., THREADGOLD, G., THORMANN, A., VULLO, A., WALTS, B., WINTERBOTTOM, A., ZADIASA, A., CHAKIACHVILI, M., FRANKISH, A., HUNT, S. E., KOSTADIMA, M., LANGRIDGE, N., MARTIN, F. J., MUFFATO, M., PERRY, E., RUFFIER, M., STAINES, D. M., TREVANION, S. J., AKEN, B. L., YATES, A. D., ZERBINO, D. R. & FLICEK, P. 2019. Ensembl 2019. Nucleic Acids Res, 47, D745-D751.

DIXON, P. 2003. VEGAN, a package of R functions for community ecology. J Veg Sci, 14, 927-930.

HUSON, D. H., BEIER, S., FLADE, I., GORSKA, A., EL-HADIDI, M., MITRA, S., RUSCHEWEYH, H. J. & TAPPU, R. 2016. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLoS Comput Biol, 12, e1004957.

JALILI, V., AFGAN, E., GU, Q., CLEMENTS, D., BLANKENBERG, D., GOECKS, J., TAYLOR, J. & NEKRUTENKO, A. 2020. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. Nucleic Acids Res.
KINSELLA, R. J., KAHARI, A., HAIDER, S., ZAMORA, J., PROCTOR, G., SPUDICH, G., ALMEIDA-KING, J., STAINES, D., DERWENT, P., KERHORNOU, A., KERSEY, P. & FLICEK, P. 2011. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database (Oxford), 2011, bar030.

KIU, R., BROWN, J., BEDWELL, H., LECLAIRE, C., CAIM, S., PICKARD, D., DOUGAN, G., DIXON, R. A. & HALL, L. J. 2019. Genomic analysis on broiler-associated Clostridium perfringens strains and exploratory caecal microbiome investigation reveals key factors linked to poultry necrotic enteritis. Animal Microbiome, 1, 12.

KOPYLOVA, E., NOE, L. & TOUZET, H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics, 28, 3211-7.

MI, H., MURUGANUJAN, A., EBERT, D., HUANG, X. & THOMAS, P. D. 2019a. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res, 47, D419-D426.

MI, H., MURUGANUJAN, A., HUANG, X., EBERT, D., MILLS, C., GUO, X. & THOMAS, P. D. 2019b. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). Nat Protoc, 14, 703-721.

MI, H. & THOMAS, P. 2009. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. Methods Mol Biol, 563, 123-40.

MUDUNURI, U., CHE, A., YI, M. & STEPHENS, R. M. 2009. bioDBnet: the biological database network. Bioinformatics, 25, 555-6.

PIMENTEL, H., BRAY, N. L., PUENTE, S., MELSTED, P. & PACHTER, L. 2017. Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods, 14, 687-690.

QUAST, C., PRUESSE, E., YILMAZ, P., GERKEN, J., SCHWEER, T., YARZA, P., PEPLIES, J. & GLOCKNER, F. O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res, 41, D590-6.

R DEVELOPMENT CORE TEAM 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

SEGATA, N., IZARD, J., WALDRON, L., GEVERS, D., MIROPOLSKY, L., GARRETT, W. S. & HUTTENHOWER, C. 2011. Metagenomic biomarker discovery and explanation. Genome Biol, 12, R60.

SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N. S., WANG, J. T., RAMAGE, D., AMIN, N., SCHWIKOWSKI, B. & IDEKER, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res, 13, 2498-504.

TANG, Y., HORIKOSHI, M. & LI, W. X. 2016. ggfortify: Unified Interface to Visualize Statistical Result of Popular R Packages. The R Journal, 8.

TUREI, D., KORCSMAROS, T. & SAEZ-RODRIGUEZ, J. 2016. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. Nat Methods, 13, 966-967.

WARNES, G. R., BOLKER, B., BONEBAKKER, L., GENTLEMAN, R., HUBER, W., LIAM, A., LUMLEY, T., MAECHLER, M., MAGNUSSON, A., MOELLER, S., SCHWARTZ, M. & VENABLES, B. 2016. gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1 ed.

WICKHAM, H. 2016. ggplot2: Elegant Graphics for Data Analysis. New York, Springer-Verlag.

ZERBINO, D. R., ACHUTHAN, P., AKANNI, W., AMODE, M. R., BARRELL, D., Bhai, J., BILLS, K., CUMMINS, C., GALL, A., GIRON, C. G., GIL, L., GORDON, L., HAGGERTY, L., HASKELL, E., HOURLIER, T., IZUOGU, O. G., JANACEK, S. H., JUETTEMANN, T., TO, J. K., LAIRD, M. R., LAVIDAS, I., LIU, Z., LOVELAND, E. J., MAUREL, T., MCLAIREN, W., MOORE, B., MUDGE, J., MURPHY, D. N., NEWMAN, V., NUHN, M., OGEH, D., ONG, C. K., PARKER, A., PATRICIO, M., RIAT, H. S., SCHULENBURG, H., SHEPPARD, D., SPARROW, H., TAYLOR, K., THORMANN, A., VULLO, A., WELTS, B., ZADISSA, A., FRANKISH, A., HUNT, S. E., KOSTADIMA, M., LANGRIDGE, N., MARTIN, F. J., MUFFATO, M., PERRY, E., RUFFIER, M., STAINES, D. M., TREVANION, S. J., AKEN, B. L., CUNNINGHAM, F., YATES, A. & FLICEK, P. 2018. Ensembl 2018. Nucleic Acids Res, 46, D754-D761.