Intestinal microbial dysbiosis is associated with Crohn’s disease (CD). However, the mechanisms leading to the chronic mucosal inflammation that characterizes this disease remain unclear. In this report, we use systems-level approaches to study the interactions between the gut microbiota and host in new-onset paediatric patients to evaluate causality and mechanisms of disease. We report an altered host proteome in CD patients indicative of impaired mitochondrial functions. In particular, mitochondrial proteins implicated in H2S detoxification are downregulated, while the relative abundance of H2S microbial producers is increased. Network correlation analysis reveals that *Atopobium parvulum* controls the central hub of H2S producers. *A. parvulum* induces pancolitis in colitis-susceptible interleukin-10-deficient mice and this phenotype requires the presence of the intestinal microbiota. Administering the H2S scavenger bismuth mitigates *A. parvulum*-induced colitis *in vivo*. This study reveals that host-microbiota interactions are disturbed in CD and thus provides mechanistic insights into CD pathogenesis.
Inflammatory bowel disease (IBD) is characterized by chronic and relapsing mucosal inflammation of the gastrointestinal tract and comprises two main subtypes, Crohn’s disease (CD) and ulcerative colitis (UC)\(^1\). A prevailing hypothesis is that IBD development is a consequence of aberrant interplay between an altered intestinal microbiota with the host\(^2\). The composition of the intestinal microbiota from IBD patients is typically characterized by decreased prevalence of protective microorganisms (that is, Clostridium XIVa and IV groups) and an expansion of detrimental bacteria (that is, Enterobacteriaceae and Fusobacteria)\(^2–7\). However, there is limited information on causality and mechanism of disease. Moreover, a fundamental unanswered question is whether the observed microbial dysbiosis is a cause or a consequence of inflammation.

To address causality and microbial dysbiosis in CD pathogenesis, we conducted a systems-level study of the interaction between the intestinal microbiota and host at the mucosa-luminal interface (MLI) in newly diagnosed paediatric patients. New-onset, paediatric subjects constitute an important population to study the role of the intestinal microbiota in IBD as there are no treatment effects on the disease process, co-morbidities are rare and there are fewer confounders as compared with adults. In addition, the inflammatory phase of IBD characteristically predominates in younger subjects and thus, these patients are more likely to reveal the underlying mechanisms that promote early development of disease. To further identify key microbial drivers of inflammation, we characterized the MLI microbial community at the time of diagnostic endoscopy. In contrast to stool microbiota, the MLI microbial community is in direct contact with the site of inflammation and is therefore likely to be directly involved in the initiation and maintenance of the diseased state. To understand the interactions between the intestinal microbiome and the host, we characterized the host proteome and identified key associations between the host proteome and intestinal microbiota. We also validated the causative role of a microbe that predicts disease severity in a mouse model of colitis. The use of these approaches in the context of a phenotypically well-defined patient cohort has enabled us to elucidate a possible mechanism of disease and to identify key microbial drivers of inflammation.

Results

Characteristics of the paediatric IBD gut microbiota. Although others have previously characterized the microbiota composition at the mucosa-luminal interface (MLI) in adult cohorts\(^8,9\), the MLI of new-onset paediatric IBD patients has not been previously reported. Given the intimate location of the MLI microbiota on the host intestinal epithelium, we hypothesized that studying the microbial composition at this location would provide direct information on how microbes affect inflammation and gut physiological processes. To characterize the composition of the MLI microbiota in IBD, we collected the MLI from the ascending colon of new-onset paediatric IBD patients and controls by flushing sterile water onto the mucosa to dislodge the mucus layer from the mucosal epithelial cells followed by aspiration of these mucosal lavages through the colonoscope (Table 1 and Supplementary Table 1). We investigated the MLI microbiota using 16S rDNA-based Illumina sequencing on a total of 65 CD, 21 UC patients and 42 control subjects with the remaining samples from our cohort (29 CD, 16 UC and 21 controls) used for independent validation with alternative quantification approaches (that is, Ion Torrent/454 sequencing and quantitative PCR (qPCR)). Microbial diversity was unchanged between controls and the CD/UC patients (Fig. 1a and Supplementary Figs 1 and 2) in contrast to the current model of decreased diversity in IBD patients\(^10\), but in agreement with a previous study on new-onset paediatric IBD patients\(^5\). These results suggest that first-onset paediatric patients have a distinct microbial profile as compared with adults and post-treatment cohorts.

Phylogenotypes differentially abundant between CD and controls or UC and controls were identified using a linear discriminant analysis effect size method\(^11\) and a zero-inflated Gaussian mixture model (metagenomeSeq)\(^12\). These two statistical methods identified most of the same key microbes (Fig. 1b and Supplementary Fig. 3; and Supplementary Data 1 and 2). In addition, to ensure that our use of the V6 hypervariable region as opposed to the V4 region commonly used in microbiota studies did not unduly influence our results, we sequenced the V4 and V6 regions in a subset of our samples using Ion Torrent sequencing. Importantly, while there are obviously unique taxa identified as being differentially abundant using each approach (Supplementary Fig. 4, Supplementary Data 3), many of the same key taxa are identified using each technique. Moreover, these results are highly concordant with those obtained by Gevers et al.\(^3\), who characterized the microbiota composition of new-onset paediatric CD patients\(^5\) using rectal mucosal biopsies instead of MLI samples. These include increased Clostridiales and Bacteroidales in controls and increased Enterobacteriaceae, Veillonella, Fusobacterium, Neisseria and Haemophilus in CD patients (Supplementary Data 2). The identification of the same key players highlights the usefulness of MLI as biological samples and the high quality of our inception cohort.

Numerous reports have cautioned that 16S rDNA results can be confounded by the presence of sequencing artifacts and/or contamination from the kits used during 16S library construction\(^13\). To quantify the levels of these artifacts and contaminants in our experimental setup, we purified genomic DNA from pure bacterial cultures and subsequently processed and analysed these samples as done for our MLI aspirates. These experiments revealed that our sequencing approach yields accurate results, with each pure bacterial sample primarily composed of a single dominant operational taxonomic unit (OTU) matching the genus of the sequenced bacteria with a median relative abundance of \(\sim99.1\%\) (Supplementary Fig. 5). Using this data set, we also determined that our median PCR/sequencing error rate was 0.007, which is comparable to those obtained in other studies\(^14\) (Supplementary Fig. 6). In addition, we determined that potential contamination from our sequencing reagents/kits accounted for a median of \(\sim0.6\%\) relative abundance. Importantly, the levels of sequencing artifacts and contamination quantified in these control experiments are low enough to be confidently removed during our bioinformatic processing and statistical analyses (‘Methods’ section).

Key phylogenotypes correlating with CD inflammation severity. To identify candidate causal microbes of CD pathogenesis, we examined changes in colonic microbiota composition at different stages of the disease from mild to moderate to severe inflammation (as determined by the PCDAI (Pediatric Crohn’s Disease Activity Index) score). As taxonomy binning might mask important microbial dynamic behaviour, we identified key phylogenotypes correlating with CD inflammation severity at the OTU level (defined by sequence clustering at 97% identity as a proxy of species-level resolution). Non-supervised multidimensional scaling analysis failed to separate the microbial communities as a function of disease severity likely due to high interpersonal microbial variation (Supplementary Fig. 7) and instead segregated by microbial diversity as reported by others\(^5\). Nevertheless, we assessed the association of individual OTUs with disease severity using generalized linear models and controlling for potential...
confounders (gender and inflammation status of the sampled site). A total of 161 OTUs were found to strongly correlate, (53 OTUs positively and 108 OTUs negatively), with disease severity (Fig. 1c; Supplementary Data 4). The OTUs negatively correlating with severity include major butyrate producers such as *Blautia*, *Lachnospiraceae*, *Roseburia*, *Eubacterium rectale*, *Ruminococcus*, *Clostridium* and *Faecalibacterium*. This decrease in butyrate producers was further confirmed using a qPCR assay targeting the butyryl-CoA:acetate CoA-transferase (BCoAT) gene, which encodes a key enzyme of the major pathway for butyrate production in the gut environment. As expected, CD patients had significantly fewer copies of the BCoAT gene as compared with the control subjects indicating a decreased microbial capacity to produce butyrate (Fig. 1d). The OTUs found to positively correlate with severity include members of the order Bacteroidales, the families Enterobacteriaceae and *Veillonella*, and the genera *Atopobium*, *Fusobacterium*, *Leptotrichia*, *Sutterella*, *Vibrio*, *Parabacteroides*, *Prevotella*, *Peptostreptococcus*, *Peptococcus* and *Streptococcus*. It is noteworthy that one-fourth of these OTUs (*Atopobium*, *Fusobacterium*, *Veillonella*, *Prevotella*, *Streptococcus* and *Leptotrichia*) are known to produce H\textsubscript{2}S through the fermentation of sulfur-containing amino-acids\cite{15}. We therefore evaluated whether the relative abundance of these H\textsubscript{2}S producers could predict CD severity using receiver-operating characteristic analysis. These H\textsubscript{2}S producers predicted severe inflammation with an AUC of 0.74 (95% confidence intervals 0.61 to 0.87; Supplementary Fig. 8) indicating that this group of microbes could be used to classify CD patients with severe inflammation. Indeed, the relative abundance of these H\textsubscript{2}S producers correlated positively with the severity of inflammation (Fig. 1e). H\textsubscript{2}S can also be produced by sulfate-reducing bacteria (SRB) but we found no evidence of a link between the SRB and CD patients as shown by a qPCR assay targeting the dissimilatory sulfite reductase gene (Fig. 1f). All together, these results suggest a possible role for the H\textsubscript{2}S-producing bacteria in CD pathogenesis through the fermentation of sulfur-containing compounds and not through dissimilatory sulfate reduction.

**Atopobium parvulum is a key network hub.** To predict interactions among individual OTUs, we constructed a network of correlations between the abundance of the OTUs found to be associated with disease severity (Fig. 2a). The resulting network included 89 nodes (OTUs) and 341 edges (representing 312 co-occurrence and 29 co-exclusion interactions). Two major co-excluded modules of OTUs appeared in the network analysis separating OTUs on the basis of their relative abundance with respect to inflammation (increased or decreased in CD as compared with control subjects). One module of co-occurring OTUs consisted of *Vagococcus* (OTU-570199), *Streptococcus* (OTU-218959), OTU-535825 (*Veillonellaceae*) with *A. parvulum* (OTU-529659) as the central hub of this module. The OTUs from this module were all found to be increased in CD. The *A. parvulum*-module anti-correlated with the other major module, which is mostly composed of OTUs that are decreased in CD and belong to the class Clostridia. The relative abundance of *A. parvulum* was validated by qPCR and found to be positively correlated with CD severity (Fig. 2b). The correlation between *A. parvulum* and CD was also confirmed by 454 pyrosequencing of a subset of samples followed by linear discriminant analysis effect size method analysis (Supplementary Fig. 9).

### Functional alterations in the paediatric CD host proteome.

To gain mechanistic insights into the role of microbes in CD severity, we conducted an unbiased, shotgun, quantitative proteomic analysis of mucosal biopsies from controls (*n*=10) and CD (*n*=21) subjects with various levels of disease severity and tested for associations between these host–proteomes and the intestinal microbes. A total 3,323 proteins were quantified using the superSILAC proteomics approach described. The majority of proteins were quantifiable within a 10-fold tissue/super-SILAC ratio (85.7%; non-normalized L/H ratios, 89.3%; normalized L/H ratios) in all the samples (Supplementary Fig. 10a). Within the 475 proteins quantified by ratios outside the range of 10-fold in at least one sample (non-normalized L/H ratios), only 91 proteins had a median ratio above 10-fold (Supplementary Fig. 10b). This is comparable to a recently published superSILAC proteome study\cite{16} that showed that 83.7% of identified proteins were within a 10-fold ratio. There were 39 proteins identified in the light samples that did not have a heavy counterpart (Supplementary Fig. 10c). Of these, only one protein (Ribonuclease pancreatic; Uniprot ID: P07998) was detected in more than half of the samples without a corresponding SILAC-labelled peak from super-SILAC standard. Taken together, these results highlight the applicability of the reference proteome for use with the human biopsy sample proteomes.

Three hundred and twenty of the 3,323 quantified proteins were identified as differentially expressed by comparing the CD patients versus control (*t*-test with *q*<0.05). PCA of the proteomic profiles grouped the control subjects away from the CD patients with the first two components explaining 61.7% of the variation (Fig. 2c). KEGG pathway analysis, through the DAVID bioinformatics resources, revealed the top 10 enriched pathways (Fisher’s exact test *P*<0.01; Supplementary Fig. 11). Gene ontology analysis identified the mitochondrial proteins as the major discriminant feature of all differentially expressed proteins (Fig. 2d). Most mitochondrial proteins (93.5%; 100/107 mitochondrial proteins) were found to be significantly down-regulated in the CD patients as compared with controls (Fisher’s exact test *P*<0.05) with the extent of downregulation rising with increased severity (Supplementary Fig. 12). Interestingly, components of the mitochondrial hydrogen sulfide detoxification

### Table 1 | Subjects characteristics.

|                    | Control | Crohn’s disease | Ulcerative colitis |
|--------------------|---------|-----------------|--------------------|
| Subjects           | 63      | 94              | 37                 |
| Gender (F/M)       | 33/30   | 34/60           | 18/19              |
| Median age, y (IQ range) | 14 (10-16) | 14 (11.75-15)  | 15 (12.5-17)       |
| Disease activity (Inactive/mild/moderate/severe)* | NA  | 3/19/25/47    | 1/4/16/16          |
| Inflammation location* | NA | | |
|                    | Upper GIT | 1  | E1  | 2 |
|                    | Ileum    | 28 | E2  | 2 |
|                    | Colon    | 18 | E3  | 5 |
|                    | ileum-colon | 47 | E4  | 28 |

*Estimated based on PCDAI/PUCAI; Pediatric CD/UC Activity Index.
*1*, ulcerative proctitis; E2, left-sided UC; E3, extensive UC; and E4, pancolitis.
Figure 1 | Assessment of the microbiota composition at the mucosa-luminal interface of new-onset paediatric IBD. (a) Diversity of the intestinal microbiota in control and CD paediatric patients as a function of disease severity. Shannon index was calculated using 500,000 reads per sample. Crosses indicate the mean while horizontal lines indicate the median. (b) Histogram of linear discriminant analysis (LDA) effect size score for CD differentially abundant taxa compared with controls (n = 65 and 42 for CD and controls, respectively); only OTUs meeting an LDA significant threshold ≥ 2 with a P < 0.05 (pairwise Wilcoxon test) are shown and are denoted with their lowest defined taxonomy. (c) Phylogenetic tree of the differentially abundant OTUs identified by metagenomeSeq analysis (fold change ≥ 2 and P < 0.05); an increasing red intensity indicates OTUs whose relative abundance increased, whereas an increasing blue intensity indicates OTUs whose relative abundance decreased in CD patients with severe inflammation as compared with mild (outer circle) or severe as compared with moderate (inner circle); the maximum colour output for this figure was set at a Log2 value of ± 3. (d) The qPCR quantification of butyrate producing bacteria in control and CD microbiota. Butyryl-CoA: acetyl-CoA transferase gene (BCoAT); n = 20 and 34 for control and CD, respectively; statistical comparison by Mann-Whitney two-tailed test; **P < 0.01. (e) Log2 sum of rarefied reads assigned to taxa known to produce H2S through amino acid fermentation were plotted for the control and as a function of CD severity (n = 39, 11, 11, 43 for control subjects and mild, moderate and severe CD, respectively). (f) The qPCR quantification of sulfate reducing bacteria in control and CD patient microbiota as a function of disease severity. Dissimilatory sulfide reductase gene (DSR); n = 16, 9, 9 and 8 for control, mild, moderate and severe CD, respectively. (a,e,f) Statistical comparison by Kruskal-Wallis test using Dunn's post hoc test and followed by a Bonferroni correction for the significance level; *P < 0.05; **P < 0.01. Crosses indicate the mean while horizontal lines indicate the median.
Figure 2 | Enrichment of A. parvulum and altered mitochondrial proteome define the severity of Crohn’s disease. (a) Interaction network of differentially abundant OTUs; each node represents an OTU and are sized according to their number of interactions; each edge denotes a significant co-exclusion (red) or co-occurrence (grey) relationship between OTUs. Nodes are coloured by their number of significant co-exclusions. (b) The qPCR quantification of A. parvulum in control and CD microbiota as a function of disease severity. ΔCt was calculated by subtracting average Ct values of universal 16S rDNA from average Ct values of A. parvulum specific primers (n = 18 for controls, nine for mild CD and seven for moderate and severe CD). **P < 0.01 estimated using Kruskal–Wallis followed by Dunn’s post hoc test. Crosses indicate the mean while horizontal lines indicate the median. (c) Principal component analysis of the differentially expressed proteins among controls and CD patients categorized as a function of disease severity. (d) Functional annotation (cellular component) analysis of the differentially expressed proteins; the 10 most significantly enriched functional groups (GO terms) are shown (Fisher’s exact test P < 10–13). The illustrated P values are for classifications that were significantly enriched compared to the whole proteomic data set.

complex, namely the sulfur dioxygenase (ETHE1), the thiosulfate sulfurtransferase (TST) and the components of complexes III and IV of the mitochondrial respiratory chain, were downregulated in the CD patients compared with controls (two-tailed Mann–Whitney test P < 0.05). Secondary validation by qRT–PCR confirmed the repression of tst, cytochrome c oxidase subunit IV (hcox41) and the sulfide dehydrogenase genes (SQRDL) transcripts, all of which detoxify H2S, in CD and/or UC patients (Supplementary Fig. 13). These findings indicate that the decreased abundance of these H2S-detoxification proteins is a hallmark of CD disease activity and possibly UC as well.

We next performed ‘transkingdom’ correlation analysis between the levels of differentially expressed mitochondrial proteins and the OTUs that correlated with CD inflammation severity (Fig. 3). In total, 46 OTUs displayed significant associations with 96 host mitochondrial proteins with a P < 0.05 in using both Kendall and Spearman correlations. Interestingly, OTUs that were found to be decreased in CD displayed strong positive correlations with mitochondrial proteins while OTUs found to be increased mainly displayed negative correlations. Of particular note, OTUs displaying positive correlations are known butyrate producers. OTU-580521 (Ruminococcaceae), OTU-182190 (Lachnospiraceae), OTU-158660 (Bacteroides), OTU-53985 (Parabacteroides distasonis) and OTU-64396 (Fusobacterium) were identified as the top candidates exhibiting reverse correlations with up to 31 mitochondrial proteins including proteins from the respiratory chain and the H2S detoxification complex.

A. parvulum causes pancolitis in Il10−/− mice. Altogether these observations suggest that A. parvulum might play a key role
in the development or progression of inflammation severity in the CD patients. To evaluate the colitogenic potential of *A. parvulum*, we utilized colitis-susceptible Il10^{−/−} mice.18,19 Germ-free Il10^{−/−}--NF-kBEGFP mice were transferred to specific pathogen free (SPF) housing and after 2 weeks acclimatization, gavaged with *A. parvulum* (10^{6} colony-forming units (CFUs) per mouse). Thereafter the mice were gavaged with *A. parvulum* weekly for 6 weeks and killed 6 weeks later. Compared with control uninfected Il10^{−/−}--NF-kBEGFP mice, *A. parvulum*-colonized Il10^{−/−} mice displayed colitis (crypt hyperplasia, ulcers, goblet cell depletion and immune cell infiltration) with histologic inflammation scores significantly higher compared with controls (Mann–Whitney U-test *P* < 0.05; Fig. 4a,b). Colonoscopy imaging revealed mucosal erythema, friability and mucosal ulceration in *A. parvulum*-colonized Il10^{−/−} mice compared with the healthy mucosa observed in control mice (Fig. 4c). At the molecular level, the colon of *A. parvulum*-colonized Il10^{−/−} mice showed increased expression of Cxcl1 and Il17 compared with controls (Mann–Whitney U-test *P* < 0.01; Supplementary Fig. 14). To determine whether the colitogenic effect of *A. parvulum* was dependent on the presence of commensal bacteria, we mono-associated germ-free Il10^{−/−} mice with *A. parvulum* for 8 weeks. Interestingly, *A. parvulum* mono-associated GF Il10^{−/−} mice did not cause significant colitis compared with controls suggesting that this bacterium cannot induce disease on its own (Fig. 4d; Histological score: 1.6 versus 0.67 respectively; not significant).

Bismuth (III)-subsalicylate (hereafter denoted as ‘bismuth’) is a known H_{2}S scavenger.20 To test the potential role of H_{2}S in *A. parvulum*-mediated colitis in Il10^{−/−} mice, gnotobiotic Il10^{−/−} mice were transferred to SPF conditions and randomized into four groups; SPF only, SPF plus *A. parvulum*, SPF plus bismuth and finally, SPF plus *A. parvulum* and bismuth. As expected *A. parvulum* worsened the development of colitis (Fig. 5a–c). Interestingly, treatment with bismuth significantly improved the colitis scores in *A. parvulum*-colonized mice (Fig. 5a–c) and prevented expression of Il1β, Il12p40 and Cxcl1 messenger RNA (mRNA; Supplementary Fig. 15). Although these findings suggest a role for H_{2}S production in *A. parvulum*-induced colitis, the reduced colitis observed with bismuth administration could also be due to bismuth’s potential antimicrobial activity on *A. parvulum* and/or the intestinal microbiota.

**Discussion**

In this study, we found key alterations in the microbiota–host interplay in newly diagnosed children with IBD providing insight into the mechanism of IBD pathogenesis and potentially opening new avenues for treatment interventions. Our study found that mitochondrial proteins were the primary proteins downregulated in CD indicating a central role for the mitochondria in CD pathogenesis. The significant downregulation of these proteins was associated with a depletion of butyrate producers together with a bloom of pathobionts many of which are known to be potent H_{2}S producers. These results suggest a regulatory relationship between the microbiota and mitochondrial and a disturbance of this relationship in the CD patients.

Accumulating data support a role for mitochondrial dysfunction in IBD pathogenesis. Reduced mitochondrial functions have been reported in UC patients with an up to 60% decreased activity of the respiratory chain complexes II, III and IV (ref. 21). Morphological changes in the mitochondria have been observed in the IBD patients22–24 and functional defects at complexes III and IV have been described in a paediatric CD patient.25 Variants in mitochondrial DNA, which result in increased metabolic activities, protect mice from colitis.26 The decreased expression

---

**Figure 3 | Microbiota–mitochondria correlation analysis.** Clustered heatmap of Spearman–Kendall correlation analysis between differentially abundant OTUs and mitochondrial proteins; red and blue colours indicate a positive or negative Spearman correlation respectively (*P* value < 0.05 for both the Kendall Tau and the Spearman’s rank statistics was used to define significance denoted by ‘+’). Note: only OTUs and mitochondrial proteins with at least one significant correlation are shown. Whether the OTU was increased or decreased in mild versus severe CD patients is denoted on the bar to the left of the OTU name (blue decreased, red increased).
of the mitochondrial proteins observed here in paediatric CD patients would perturb mitochondrial functions. Such mitochondrial perturbations have been shown to result in increased epithelial permeability through ROS production and promote transcytosis of bacteria across the epithelial layer\textsuperscript{27}. Increased intestinal permeability has been reported in active CD further suggesting a role for mitochondrial dysfunction in the pathogenesis of IBD\textsuperscript{28,29}.

Butyrate is the main energy source of colonocytes and is produced by the bacterial fermentation of unabsorbed carbohydrate. Here, we demonstrate that the intestinal microbiota of paediatric CD patients is characterized by a significant decrease in the relative abundance of butyrate producers as has been reported previously in the context of paediatric IBD\textsuperscript{3}. Our data also revealed a positive correlation between the relative abundance of the butyrate producers and mitochondrial proteins, suggesting a signalling role for butyrate in host mitochondrial gene expression. In support of this observation, butyrate is known to activate the signalling role for butyrate in host mitochondrial gene expression. Butyrate exhibits anti-inflammatory properties, whereas H\textsubscript{2}S is an important mediator of many physiological and pathological processes and has been previously associated with UC and colorectal cancer\textsuperscript{17,31}. Interestingly, \textit{A. parvulum} was identified by network analysis as the most prominent microbe associated with mitochondrial dysfunction. Importantly, the increased abundance of \textit{A. parvulum} as a function of inflammation severity in paediatric CD patients was not observed in UC and therefore is not simply a consequence of inflammation, suggesting a causal effect in CD.

Although causal relationships would be better identified by monitoring changes in microbiota composition before the onset of the disease while following disease establishment, logistical constraints preclude this experiment. Instead we tested the hypothesis that changes in microbial composition between distinct and progressive stages of the disease would unravel causal interactions among the host and specific members of the gut microbes. Indeed, the observation that the relative abundance correlates positively with inflammation severity in paediatric CD patients is characterized by a significant decrease in the relative abundance of butyrate producers as has been reported previously in the context of paediatric IBD\textsuperscript{3}. Our data also revealed a positive correlation between the relative abundance of the butyrate producers and mitochondrial proteins, suggesting a signalling role for butyrate in host mitochondrial gene expression.
association conditions. These results indicate a key role for both *A. parvulum* and the gut microbiota in colitis development. In support to our findings, a recent report on CD patient response to exclusive enteral nutrition (EEN) therapy has reported that a putative *A. parvulum* OTU was enriched in CD patients before treatment and was negatively associated with days on EEN. Although a role for H₂S production by *A. parvulum* in colitis development remains to be proven directly, excess H₂S has been recently shown to act as an autocrine T-cell activator, potentially contributing to unwanted T-cell responses against commensal bacteria, consistent with our observation that the intestinal microbiota is required for *A. parvulum*-induced experimental colitis. Accordingly, potential H₂S trapping by bismuth may have contributed to the mitigation of *Atopobium*-induced colitis development in our mouse model. This finding is in agreement with a prospective mechanistic role for H₂S in inflammation. A role for H₂S in colitis development is supported by the recent finding that microbial H₂S formation contributes to mucus degradation opening the intestinal barrier to toxic compounds and pathobionts.

H₂S is an important mediator of many biological processes (angiogenesis, cytoprotection, metabolism, inflammation) and can be pro- or anti-inflammatory depending on its concentration and the particular circumstance. In light of these opposite effects, a recent study reported that pediatric CD patients undergoing EEN treatment and presenting with reduced colonic inflammation had increased levels of fecal H₂S. The findings from the latter study challenge the concept of a pro-inflammatory role for H₂S in CD. However, as EEN feed is enriched in sulfur-containing amino-acids, the observed increase in H₂S levels might simply be a consequence of a change in diet. Clearly, complex modes of action of H₂S need to be further explored to reconcile these apparent controversial observations. Furthermore, the impact of bacterial-derived exogenous H₂S on the intestine may be quite different than endogenous cell-based H₂S.

Altogether, our results emphasize the importance of the microbial community and its interaction with the host in the pathogenesis of IBD. We propose a central role for mitochondrial dysfunction in colitis development. Although it is difficult to identify the cause underpinning observed downregulation of
mitochondrial proteins, the loss of butyrate producers in paediatric CD patients will lead to decreased butyrate production which in turn will impair mitochondrial functions likely resulting in ROS production, increased epithelial permeability and translocation of commensal microbes across the epithelial barrier. Intriguingly, antibiotic exposure has been shown to reduce the abundance of butyrate producers and also markedly decreases

Intriguingly, antibiotic exposure has been shown to reduce the translocation of commensal microbes across the epithelial barrier. In ROS production, increased epithelial permeability and mitochondrial proteins, the loss of butyrate producers in paediatric CD patients with severe inflammation suggests a key role for this microbe in driving mucosal inflammation and disease exacerbation. Our results point toward a mechanism of CD pathogenesis involving a disruption of the mitochondria–microbiota relationship leading to dysfunctional mitochondria in the presence of microbes themselves involved in a complex interactive microenvironment at the MLI required for mucosal inflammation.

**Methods**

**Ethics statement.** The collection of samples from paediatric patients was approved by the Research Ethics Board of the Children’s Hospital of Eastern Ontario (CHEO). The protocol for the use of Il10−/− mice was approved by the Institutional Animal Care and Use Committee of University of North Carolina at Chapel Hill and the University of Florida. Written informed consent was obtained from the participants or parents.

**Study design and sampling.** We conducted a cross-sectional study of all eligible patients under 18 years of age scheduled to undergo colonoscopy for their initial diagnostic work-up for suspected IBD. Those who were not diagnosed with IBD acted as controls. Exclusion criteria relating to known conditions affecting the intestinal microbiota composition included: (1) a body mass index greater than the 95th percentile for age; (2) diabetes mellitus (insulin and non-insulin dependent); (3) infectious gastroenteritis within the preceding 2 months; and (4) use of any antibiotics within 2 weeks before colonoscopy, or (5) irritable bowel syndrome. These same exclusion criteria were applied to the non-IBD control group. All IBD cases met the standard diagnostic criteria following thorough clinical, endoscopic, histologic and radiological evaluation. Disease location was based on endoscopic and radiologic appearance and characterized using the Paris modification of the Montreal Classification for IBD. Clinical disease activity of CD was determined using the PCDAI and of UC using the PUCAI (Pediatric Ulcerative Colitis Activity Index). All the controls had macroscopically and histologically normal mucosa, and did not carry a diagnosis for any known chronic intestinal disorder (for example, celiac disease, eosinophilic enterocolitis, irritable bowel syndrome or inflammatory bowel disease) within the preceding 24 months. The collection of samples from paediatric patients was considered to be potential sequencing errors. The reads that were matched against the Illumina paired-end sequencing adapters and the barcode trimmed using NovoBarCode (Novocraft.com). Afterwards, the DNA concentration in each reaction was quantified using a Montage PCR96 Cleanup Kit (Millipore, Billerica, MA, USA). The DNA was then purified using the FastDNA Spin Kit as described above (hereafter referred to as ‘pure’ samples) and Sanger sequenced to confirm the identity of each bacterial species. In addition, aliquots of the extracted DNA were repurified a second time using the FastDNA kits to assess the impact of potential kit contamination (hereafter referred to as ‘re-extract’ samples). 16S rDNA-V6 libraries for all the samples were constructed, with triplicate libraries constructed for the re-extract samples and sent for Illumina sequencing as described above.

**Illumina microbiota sequencing data analysis.** Raw data were first analysed for decreasing quality along the read. As expected, we noted decreasing quality near the end of the reads. Given that our paired end reads are expected to overlap by ~21–25 nt (Supplementary Fig. 16), we trimmed four nucleotides from the 3′ end of the sequences using the fastx_trimmer script (http://hannonlab.cshl.edu/) to remove error-prone bases that could result in incorrect read merging. To note, this trimming step reduced the total number of singletons/doubletons in the final rarefied OTU table by 5% (data not shown). The trimmed paired-end sequences were merged using Fast Length Adjustment of Short reads. More than 95% of the reads were successfully merged, while the sequences that failed to merge were discarded. The merged reads were then quality filtered with a minimum quality score of 20 over 95% of the sequences. The remaining reads were demultiplexed by OTU using a closed-reference OTU picking workflow against the Greengenes reference set (release 4 February 2011) based on an average percentage of identity of 97%. Note that the same Greengenes database was used for the Illumina, Ion Torrent and 454 pyrosequencing data sets discussed below. Singletons and doubletons were removed and a table of OTU counts per sample was generated. Next, the OTU table was randomly subsampled to a total number of 200,000 reads per sample of 200,000. The resulting rarefied OTU table was used to analyse the microbiota structure and diversity using the microbial ecology tools available in the QIME package and as input for all other downstream analyses. Any further analysis of specific filtering of the Illumina OTU table was done with an individual ‘Methods’ section. The phylogenetic tree was constructed using the FastTree method of QIME 1.8.0 and the generated file was exported to the Interactive Tree of Life (iTOL) software to view and format the constructed tree.

**Illumina pure culture sequencing data analysis.** The raw reads obtained for the pure culture sequencing experiments were processed as described above for the Illumina metagenome analysis to generate an OTU table rarefied to 200,000 reads per sample (Supplementary Table 1). Each pure culture consisted of single ‘dominate’ OTU representing with a median of ~99% of the total reads and with a taxonomic classification matching the expected bacteria at the genus level (Supplementary Fig. 5, Supplementary Data 5). Any OTUs present with a taxonomic classification matching the expected bacteria at the family level were considered as potential sequencing errors. The reads that were matched against these erroneous OTUs were extracted and re-aligned using BLASTN against the dominant OTU to determine the PCR/sequencing error rate (Supplementary Fig. 6).
Supplementary Data 5). OTUs present in >90% of the samples (in either the pure or re-extract sample sets) were considered to be background contamination originating from either the DNA extraction kits or reagents used during PCR amplification (Supplementary Fig. 5).

**Library construction for Ion Torrent sequencing.** The V4 or V6 region of 16S rRNA was amplified in a single PCR reaction that incorporated both the Ion Torrent adaptors and an 11 bp barcode sequence using modified universal 16S rDNA-V4 (ref. 31) and V6 (ref. 45) primers (Supplementary Table 3). Note that the conserved V6 sequence used for the Ion Torrent sequencing is identical to those used in the Illumina library construction method. For the V6 reactions, the PCR cycling conditions were identical to the first PCR reaction described in the 16S rDNA-V6 Illumina library construction method. For the V4 reactions, the PCR cycling conditions were prepared in a total volume of 50 μl containing 5 μg of extracted metagenomic DNA, 0.5 μM of each primer, and 1× Phusion Flash High-Fidelity PCR Master Mix. For the V4 reactions, the PCR cycling conditions were identical to the first PCR reaction described in the 16S rDNA-V6 Illumina library construction method. The PCR cycling conditions were then subjected to 34 cycles of 98°C for 30 s, 55°C for 30 s and 72°C for 90 s, followed by a final extension at 72°C for 5 min as previously described31. The samples were purified using Purelink PCR purification columns and quantified using a Qubit fluorometer (Life Technologies). PCR products were visualized on a 1.5% agarose gel to ensure successful amplification and that the products were the expected size. Two hundred nanograms of each sample was subsequently pooled together for both the V4 and V6 libraries. The pooled libraries were size-selected using Agencourt AMPure XP DNA purification beads to remove primer dimers. The pool concentrations of each library pool were determined using an Agilent BioAnalyzer (Agilent Technologies). The libraries were combined in an equimolar ratio. This equimolar V4/V6 pool was subjected to emulsion PCR using the Ion OneTouch HiQ Template Kit and sequencing of the templated beads was completed on a single Proton Chip on an Ion Torrent Proton using an Ion Proton HiQ sequencing kit according to the manufacturer’s instructions (Life Technologies).

**Ion Torrent sequencing data analysis.** Ion Torrent sequencing was done for both the V4/V6 regions from 36 right colon samples (n = 13, 12, 11 for control, CD and UC). The unaligned BAM file generated by the Ion Torrent Proton (note that adaptor sequences/primer dimers are automatically removed by the Proton analysis pipeline) was downloaded, converted to fastq format and reads split according to the expected sizes for the V4/V6 regions plus the barcodes (234–266 and 162–194, respectively, representing a ± 1 bp window), using samtools32 in conjunction with bioawk33. Reads were quality trimmed using an in house modified Mott trimming algorithm implemented in seqtk (https://github.com/lh3/seqtk) and reads that were shorter than the expected size for either the V4 or V6 regions (as described above) after trimming were discarded. The V4/V6 reads were then demultiplexed using FastqMultiX34, converted from fastq to fasta and fed to QiIME, in order to determine the taxonomic profiles of the samples. The reads were clustered into OTUs using USEARCH with an alignment similarity level of 97% and a minimum length of 100 bp, (2) exact matching to the sequencing primers, (3) no ambiguous nucleotides and (4) a minimum average quality score of 20. Next, sequences were de novo clustered using UCLUST35 based on average percentage of identity of 97%. The most abundant read from each OTU was picked as a representative sequence for that cluster. The representative sequences were then aligned using PyNAST with a minimum alignment length of 100 and a minimum percentage identity of 75%, followed by checking the chimeric OTUs with the blast fragments approach. Only six representative sequences were identified as chimeric and therefore were removed. Taxonomy assignments were made with the Ribosomal Database Project (RDP)56 Blast algorithm using the RDP56 database (release 4 February 2011) with an e value < 1e–8 and a confidence score of ≥ 0.5. Next, singletons (OTUs that had only one matching sequence) were filtered out from the resulting OTU table. The OTU table was then used to determine the alpha and beta diversity within and between the samples using the QIIME’s default criteria. To identify the community bacterial biodiversity, the relative abundance of different phylogenetic levels computed by QIME was analysed by linear discriminant analysis effect size algorithm36 using its default parameters. The 454-pyrosequencing reads assigned as Atoxopobium by QIIME analysis were retrieved and found to match to A. parvulum following alignment of the reads against the RDP56 and NCBI databases (the aligned region covered the entire 454 sequence length with >99% sequence identity to A. parvulum).

**Statistical analysis of the microbiota data.** Several statistical approaches were used to identify taxa significantly associated with disease status and severity. Identification of OTUs exhibiting differential abundance between patients with different disease severity (mild versus moderate versus severe) was performed using the metagenomeSeq R package—a statistical approach that accounts for confounding factors and accounts for undersampling12. Briefly, the OTU counts were normalized using a cumulative sum scaling approach and OTUs with a normalized count < 5 in at least one of the groups were discarded. OTUs exhibiting differential abundance as a function of disease severity were identified using a zero-inflated Gaussian mixture model. Nonparametric testing was conducted to compare the abundance of the taxa identified across the three disease statuses using a Kruskal-Wallis test with a post hoc Dunn’s test was performed to compare the relative abundance of specific taxa or groups of bacteria as a function of disease status (CD versus UC versus control) and disease severity (mild versus moderate versus severe). The relative abundance of the taxa identified was also analysed by statistical testing at all the phyla levels. Six different clones were extracted from the amplified products using the QIAprep Spin Miniprep kit (Qiagen, Hilden, Germany) following its standard protocol and eluted in 25 μl molecular biology-grade water. Two microliters of the pooled purified amplicons were then used as a template for the second PCR following the same PCR conditions except using 30 amplification cycles instead of 15 cycles. The second PCR added the titanium tails to the 16S-V6 amplicons, which are essential for the 454 sequencing procedure. The amplicons from the second PCR were pooled and purified using Qiagen’s MiniElute PCR purification columns (Qiagen, Hilden, Germany) following its standard protocol and eluted in 25 μl molecular biology-grade water. Two microliters of the pooled purified amplicons were then used as a template for the second PCR following the same PCR conditions except using 30 amplification cycles instead of 15 cycles. The second PCR added the titanium tails to the 16S-V6 amplicons, which are essential for the 454 sequencing procedure. The amplicons from the second PCR were pooled and purified again using Qiagen MiniElute PCR purification columns. Negative control reactions (no DNA template) were included in all the experiments. PCR products were visualized on a 1.5% agarose gel to check the amplification. Amplification fragments from different samples were normalized to the same concentration (100 ng μl–1), captured to streptavidin coated sepharose beads and exposed to emulsion PCR.

Afterwards, the immobilized double stranded amplicons were denatured into single-stranded DNA, which was then annealed to the sequencing primer by heating a mix 55°C for 5 min. The resulting single-stranded DNA was sequenced on a 454 Genome Sequencer FLX System (Roche Diagnostics GmbH) using GS Titanium chemistry according to the standard amplicon sequencing protocol. Samples covering the three tested phenotypes (two to four samples each) were sequenced on the same run with each being sequenced in a 1/16 sector of 70 × 75 πlitter plate. Note that as the samples were physically separated from each other during the sequencing reaction, there was no need to use barcoded primers.

**454-pyrosequencing data analysis.** A total of 346,160 reads were generated from the 454 pyrosequencing of 16S rDNA-V6 region from 26 right colon samples. The raw pyrosequencing reads were de-noised and processed to remove low quality and short reads using Quantitative Insights Into Microbial Ecology pipeline release 1.8.0. Demultiplexing was performed using an in house code, short reads using Quantitative Insights Into Microbial Ecology pipeline release 1.8.0. After demultiplexing, reads that did not match any of the following parameters: (1) minimum read length of 100 bp, (2) exact matching to the sequencing primers, (3) no ambiguous nucleotides and (4) a minimum average quality score of 20. Next, sequences were de novo clustered using UCLUST35 based on average percentage of identity of 97%. The most abundant read from each OTU was picked as a representative sequence for that cluster. The representative sequences were then aligned using PyNAST with a minimum alignment length of 100 and a minimum percentage identity of 75%, followed by checking the chimeric OTUs with the blast fragments approach. Only six representative sequences were identified as chimeric and therefore were removed. Taxonomy assignments were made with the Ribosomal Database Project (RDP)56 Blast algorithm using the RDP56 database (release 4 February 2011) with an e value < 1e–8 and a confidence score of ≥ 0.5. Next, singletons (OTUs that had only one matching sequence) were filtered out from the resulting OTU table. The OTU table was then used to determine the alpha and beta diversity within and between the samples using the QIIME’s default criteria. To identify the community bacterial biodiversity, the relative abundance of different phylogenetic levels computed by QIME was analysed by linear discriminant analysis effect size algorithm36 using its default parameters. The 454-pyrosequencing reads assigned as Atoxopobium by QIIME analysis were retrieved and found to match to A. parvulum following alignment of the reads against the RDP56 and NCBI databases (the aligned region covered the entire 454 sequence length with >99% sequence identity to A. parvulum).

**Statistical analysis of the microbiota data.** Several statistical approaches were used to identify taxa significantly associated with disease status and severity. Identification of OTUs exhibiting differential abundance between patients with different disease severity (mild versus moderate versus severe) was performed using the metagenomeSeq R package—a statistical approach that accounts for confounding factors and accounts for undersampling12. Briefly, the OTU counts were normalized using a cumulative sum scaling approach and OTUs with a normalized count < 5 in at least one of the groups were discarded. OTUs exhibiting differential abundance as a function of disease severity were identified using a zero-inflated Gaussian mixture model. Nonparametric testing was conducted to compare the abundance of the taxa identified across the three disease statuses using a Kruskal-Wallis test with a post hoc Dunn’s test was performed to compare the relative abundance of specific taxa or groups of bacteria as a function of disease status (CD versus UC versus control) and disease severity (mild versus moderate versus severe). The relative abundance of the taxa identified was also analysed by statistical testing at all the phyla levels. Six different clones were extracted from the amplified products using the QIAprep Spin Miniprep kit (Qiagen, Hilden, Germany) following its standard protocol and eluted in 25 μl molecular biology-grade water. Two microliters of the pooled purified amplicons were then used as a template for the second PCR following the same PCR conditions except using 30 amplification cycles instead of 15 cycles. The second PCR added the titanium tails to the 16S-V6 amplicons, which are essential for the 454 sequencing procedure. The amplicons from the second PCR were pooled and purified again using Qiagen MiniElute PCR purification columns. Negative control reactions (no DNA template) were included in all the experiments. PCR products were visualized on a 1.5% agarose gel to check the amplification. Amplification fragments from different samples were normalized to the same concentration (100 ng μl–1), captured to streptavidin coated sepharose beads and exposed to emulsion PCR.
Quantification of butyrate-producers by qPCR. The overall abundance of butyrate-producing bacteria was determined by quantifying the amount of the butyryl CoA:acetate CoA-transferase (BcOAT) gene using the primers BcOATcsfR/BcOATcsfL (Supplementary Table 4) as described elsewhere. The BcOAT gene was amplified from 30 ng of metagenomic DNA in a 25 μl qPCR reaction containing 1× QuantiTect SYBR Green PCR master mix (Qiagen) and 2.5 μM of BcOATcsfL/BcOATcsfR primers. The amplification conditions were as follows: one cycle of 95 °C for 15 min; 40 cycles of 94 °C for 15 s, 53 °C and 72 °C each for 30 s with data acquisition at 72 °C. The 16S rDNA sequence was used to normalize total bacterial content in each sample and was amplified using adapted universal primers UniF/UniR (Supplementary Table 4). The assay was done in duplicate for each sample. The Ct values of each target was calculated by subtracting the CT of the 16S rDNA from the BcOAT Ct. Then, the ΔCt values were compared between the CD and control groups using a Mann–Whitney two-tailed test with a Dunn’s post hoc test and P values below 0.05 were considered significant.

Microbial correlation network analysis. A correlation network was constructed to identify co-occurrence and mutual exclusion among OTUs from the control subjects and CD patients. To focus on bacteria potentially associated with disease, only the OTUs exhibiting differential abundance as a function of disease severity were chosen for this network analysis. In addition, OTUs that were found to be absent in more than two-thirds of the samples or with a minimum occurrence <15 across all the samples were removed from the data set, leaving a matrix of 97 OTUs. The network was built as previously described using the CoNet Cytoscape plug-in. The correlation scores were calculated for each OTU pair using a combination of two correlation (Pearson and Spearman) and two dissimilarity measures (Bray–Curtis and Kullback–Leibler). The 400 top- and bottom-ranking edges from each method were retrieved. Edge- and method-specific permutations (100) were used to control potential false-positive correlations. The resulting distribution was used to calculate statistical significance for each edge using 100 bootstrap iterations. The P value was adjusted for multiple comparisons using a Benjamini–Hochberg correction and a value of <0.05 was deemed significant. The P value was obtained for each correlation or dissimilarity measure and only edges supporting at least two significant measures were included. The network consisted of 89 nodes (OTUs) and 341 edges (correlations). The resulting network was visualized using Cytoscape 3.2.1 (ref. 61).

Stable isotope labelling by amino acids in cell culture. The human hepatic HuH7 cells (HuH-7) and human embryonic kidney 293 cells (HEK-293) were originally acquired from ATCC and the human colorectal cancer 116 cells (HCT-116) were obtained from the Japanese Collection of Research Bioresources Cell Bank. HuH-7, HEK-293 and HCT-116 were individually grown at 37 °C in a 5% CO2 humidified incubator (to note, all three cell lines used in this study were found to be mycoplasma positive). Stable isotope labelling by amino acids in cell culture (SILAC) medium was prepared as follows: DMEM lacking lysine, arginine and methionine was custom prepared with AthenaES (Baltimore, MD, USA) and supplemented with 16 μM unlabeled lysine (BD Bioscience, Franklin Lakes, NJ, USA) and 8 μM unlabeled arginine (BD Bioscience, Oakville, ON, Canada). 10% (v/v) dialysed FBS (Gibco-Invitrogen; Burlington, ON, Canada), 1 mM sodium pyruvate (Gibco-Invitrogen), 28 μg ml−1 gentamicin (Gibco-Invitrogen), and [15N2,15N2]-L-lysine, [15N4,15N4]-L-arginine (heavy isotope form of amino acids; MedchemExpress, Aickburk, ON, Canada) at least two independent measurements were included. The cells were grown at least 10 doublings in SILAC media to allow for complete incorporation of the isotopically labelled amino acids into the cells.

Determination of the rate of SILAC amino acids incorporation. The cells were grown to 80% confluency in SILAC medium (5 × 106 cells were plated in 10-cm dishes). Next, the cells were washed twice with ice-cold phosphate-buffered saline and lysed by addition of 1 ml of 10X RIPA buffer (50 mM Tris (pH 7.6), 150 mM NaCl, 1% (v/v) NP-40, 0.5% (w/v) deoxycholate, 0.1% (w/v) SDS with protease inhibitor cocktail (Complete Mini Roche; Mississauga, ON, Canada) and phosphatase inhibitor (PhosStop Roche tablet)). The lysates were then transferred to 15 ml conical tubes and the proteins were precipitated by addition of 5 ml ice-cold acetone followed by incubation at −20 °C overnight. The proteins were collected by centrifugation (3,000 g, 10 min, 4 °C), washed with ice-cold acetone two times and the protein pellets were re-solubilized in 300 μl of a 50 mM NH4HCO3 solution containing 8 M urea. The protein concentrations were determined by the Bradford dye-binding method using Bio-Rad’s Protein Assay Kit (Mississauga, ON, Canada). For the general in-solution digestion, 200 μg of protein lysates were reconstituted in 50 mM NH4HCO3, (200 μl) and the proteins were reduced by mixing with 5 μl of 4 mM DTT at 56 °C for 15 min. The proteins were then subjected to alkylation by mixing with 20 μl of 400 mM iodoacetamide in darkness (15 min at room temperature) followed by addition of 800 μl of 50 mM Tris-HCl (pH 8.0), containing 1% (v/v) of Trypsin (TPCK Treated, Worthington) on a shaker (250 rpm) at 37 °C overnight. Finally, 200 μl of 50 mM Tris-HCl pH 8.0 was added to elute the peptides by centrifugation (twice). The peptides were fractionated, using an in-house constructed SCX column with five pH fractions (pH 4.0, 6.0, 10.0 and 12.0). The buffer composition was 20 mM boric acid, 20 mM phosphoric acid and 20 mM acetic acid, with the pH adjusted by using 1 M NaOH. Finally, the fractionated samples were desalted using in-house C18 desalting cartridges and dried in a speed-vac before LC-MS analysis.

Mass-spectrometry analyses. All resulting peptide mixtures were analysed by high-performance liquid chromatography/electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS). The HPLC-ESI-MS/MS consisted of an automatic nano-TripleQ-Trap nanocLC 400 system (Waters, Dublin, CA, USA) coupled with an LTQ Velos Pro Orbitrap Elite mass spectrometer (ThermoFisher Scientific, San Jose, CA, USA) equipped with a nanoelectrospray interface operated in positive ion mode. Briefly, each peptide mixture was reconstituted in 20 μl of 0.5% (v/v) formic acid and 12 μl was loaded on a 200 μm × 50 mm fritted silica capillary column packed in-house with reverse phase Magic C18 AQ resins (5 μm; 200 Å pore size; Dr Maisch GmbH, Ammerbuch, Germany). The separation of peptides was performed on an analytical column (75 μm × 10 cm) packed with reverse phase beads (3 μm; 120 Å pore size; Dr Maisch GmbH, Ammerbuch, Germany) using a 120 μm iden Dionex Prominence liquid chromatography system (ThermoFisher Scientific, San Jose, CA, USA) equipped with an LTQ Velos Pro Orbitrap Elite mass spectrometer (ThermoFisher Scientific, San Jose, CA, USA) equipped with a nanoelectrospray interface operated in positive ion mode. Briefly, each peptide mixture was reconstituted in 20 μl of 0.5% (v/v) formic acid and 12 μl was loaded on a 200 μm × 50 mm fritted silica capillary column packed in-house with reverse phase Magic C18 AQ resins (5 μm; 200 Å pore size; Dr Maisch GmbH, Ammerbuch, Germany). The separation of peptides was performed on an analytical column (75 μm × 10 cm) packed with reverse phase beads (3 μm; 120 Å pore size; Dr Maisch GmbH, Ammerbuch, Germany) using a 120 μm iden Dionex Prominence liquid chromatography system (ThermoFisher Scientific, San Jose, CA, USA) equipped with a nanoelectrospray interface operated in positive ion mode. Briefly, each peptide mixture was reconstituted in 20 μl of 0.5% (v/v) formic acid and 12 μl was loaded on a 200 μm × 50 mm fritted silica capillary column packed in-house with reverse phase beads (3 μm; 120 Å pore size; Dr Maisch GmbH, Ammerbuch, Germany) using a 120 μm iden Dionex Prominence liquid chromatography system (ThermoFisher Scientific, San Jose, CA, USA) equipped with a nanoelectrospray interface operated in positive ion mode. Briefly, each peptide mixture was reconstituted in 20 μl of 0.5% (v/v) formic acid and 12 μl was loaded on a 200 μm × 50 mm fritted silica capillary column packed in-house with reverse phase beads (3 μm; 120 Å pore size; Dr Maisch GmbH, Ammerbuch, Germany) using a 120 μm iden Dionex Prominence liquid chromatography system (ThermoFisher Scientific, San Jose, CA, USA) equipped with a nanoelectrospray interface operated in positive ion mode. Briefly, each peptide mixture was reconstituted in 20 μl of 0.5% (v/v) formic acid and 12 μl was loaded on a 200 μm × 50 mm fritted silica capillary column packed in-house with reverse phase beads (3 μm; 120 Å pore size; Dr Maisch GmbH, Ammerbuch, Germany) using a 120 μm iden Dionex Prominence liquid chromatography system (ThermoFisher Scientific, San Jose, CA, USA) equipped with a nanoelectrospray interface operated in positive ion mode.
Quantification of H₂S detoxification genes expression level. The quantification of the expression level of TST, SQORDL (Sulfide Quinone Reductase Like) and COX4-I (Cytochrome C oxidase subunit IV isoform 1) relative to hGAPDH (Glyceraldehyde-3-Phosphate Dehydrogenase) was determined using an Applied Biosystems 7300 and Quantitect SYBR Green RT-PCR kit (Qiagen, Hilden, Germany). The primers used were either designed using the NCBI Primer-BLAST or incorporated into the chow (Teklan Global 18% Protein Rodent Diet) at a concentration of 7 kg⁻¹ (Harlan Laboratories, Madison, WI, USA) and then irradiated for gnotobiotic experiments. The mice were fed with this diet 1 week before the colonization with A. parvulum. The tissue samples from the colon were collected for RNA and histology as previously described62. All the animal procedures were approved by the University of North Carolina at Chapel Hill Animal Care and Use Committee. The histological images were acquired using a DFC310 FX (LEICA) coupled to LEICA Application Suite AFv4.6 software. Intestinal inflammation was blindly scored as previously described63. The tissue was divided into four quarters to allow individual scoring of the mid-section, which was assessed by a blinded investigator. A score was given to each quarter separately and then added to generate a final colitis score on a scale of 0–16.

Statistical analyses of I10⁻/⁻ mice results. Unless specifically noted, the statistical analyses were performed using GraphPad Prism version 6 (GraphPad, La Jolla, CA, USA). The comparisons of mouse studies were made with a nonparametric analysis of variance, and then a Mann–Whitney U-test. The experiments were considered statistically significant if P < 0.05.

Mouse endoscopy. Colonoscopy was performed using a ‘Cololview System’ (Karl Storz Veterinary Endoscopy) as described previously64. The mice were anaesthetized using 1.5 to 2% isoflurane and ~4 cm of the colon from the anal verge to the splenic flexure was visualized. The procedures were digitally recorded on an AIDA Comp PC.

qRT-PCR on mouse intestinal samples. Total RNA from intestinal tissues was extracted from the distal part of the colon using TRIzol (Invitrogen, Carlsbad, CA, USA) following the manufacturer’s protocol. The complementary DNA was reverse-transcribed using M-MLV (Invitrogen, Carlsbad, CA, USA) and mRNA expression values were measured using SYBR Green PCR Mastermix (Applied Biosystems) on an ABI 7900HT Fast Real-Time PCR System and normalized to β-actin. The primers used are listed in Supplementary Table 5. The PCR reactions were performed for 40 cycles according to the manufacturer’s recommendations, and RNA fold changes were calculated using the ΔΔCt method65.
16. Tyanova, S.
36. Wang, R. Physiological implications of hydrogen sulfide: a whiff exploration
35. Guo, F. F., Yu, T. C., Hong, J. & Fang, J. Y. Emerging roles of hydrogen sulfide

NATURE COMMUNICATIONS | 7:13419 | DOI: 10.1038/ncomms13419 | www.nature.com/naturecommunications

14. Schloss, P. D., Gevers, D. & Westcott, S. L. Reducing the effects of PCR

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms13419 ARTICLE

11. Segata, N.
12. Paulson, J. N., Stine, O. C., Bravo, H. C. & Pop, M. Differential abundance

23. Soderholm, J. D.
21. Sifroni, K. G.
20. Suarez, F. L., Furne, J. K., Springfield, J. & Levitt, M. D. Bismuth subsalicylate

29. Zeissig, S.
32. Quince, C.
21. Sifroni, K. G.
20. Suarez, F. L., Furne, J. K., Springfield, J. & Levitt, M. D. Bismuth subsalicylate

23. Soderholm, J. D.
21. Sifroni, K. G.
20. Suarez, F. L., Furne, J. K., Springfield, J. & Levitt, M. D. Bismuth subsalicylate

29. Zeissig, S.
32. Quince, C.
21. Sifroni, K. G.
20. Suarez, F. L., Furne, J. K., Springfield, J. & Levitt, M. D. Bismuth subsalicylate

23. Soderholm, J. D.
21. Sifroni, K. G.
20. Suarez, F. L., Furne, J. K., Springfield, J. & Levitt, M. D. Bismuth subsalicylate

29. Zeissig, S.
32. Quince, C.
21. Sifroni, K. G.
20. Suarez, F. L., Furne, J. K., Springfield, J. & Levitt, M. D. Bismuth subsalicylate

23. Soderholm, J. D.
21. Sifroni, K. G.
20. Suarez, F. L., Furne, J. K., Springfield, J. & Levitt, M. D. Bismuth subsalicylate

29. Zeissig, S.
32. Quince, C.
Acknowledgements
This work was funded by the Government of Canada through Genome Canada and the Ontario Genomics Institute (OGI-067), CIHR grant number GPH-129340, CIHR grant number MOP-114872 and the Ontario Ministry of Economic Development and Innovation (REG1-4450). We also acknowledge funding from the IBD foundation, Crohn’s and Colitis Canada (CCC), the CHEO foundation, the Children’s Hospital of Eastern Ontario Research Institute and the Faculty of Medicine of the University of Ottawa. D.F. acknowledges a Canada Research Chair in Proteomics and Systems Biology and funding from CIHR grant number 117960-1. C.J. acknowledges funding from NIH (R01 DK73338, R01 DK47700 and R21 CA195226). We thank Maureen Bower of the National Gnotobiotic Rodent Resource Center at the University of North Carolina for assistance with gnotobiotic mice (NIH 5P40OD010995, CGIBD grant P30 DK034987). We thank A. Mack, M. Boland, C. Jimenez-Rivera and J. Barkey for their assistance in clinical data collection and making this research project possible. W.M. acknowledges funding from Mansoura University, Egypt. T.A. was supported by a scholarship from King Abdulaziz University, through the Saudi Arabian Cultural Bureau in Canada. M.M. acknowledges funding from CCFA. E.B. was supported by a Career Development Award from the Canadian Child Health Clinician Scientist Program, and a New Investigator Award from the Canadian Institutes of Health Research, Canadian Association of Gastroenterology, and Crohn’s and Colitis Canada. J.B. is the recipient of a CIHR/CAG/CCC Postdoctoral Award.

Author contributions
W.M., C.-K.C. and M.M. were involved in the experimental plan and data analysis. W.M. processed the biological samples and generated the microbiota data, C.-K.C. and A.E.S. generated and analysed the proteomic data, M.M. performed the mice experiments. S.A.D. and H.Z. helped perform the proteomic analysis. T.A., J.B., S.S. and M.H. participated in characterizing the microbiota. A.B. performed the pure culture experiments. E.B. and R.S. participated in patient diagnosis and enrolment, sample isolation and collection of clinical data. A.S., D.R.M., D.F. and C.J. contributed equally to the work. A.S. was involved in all the aspects of the study from experimental design to data analysis. D.M. recruited the patients, performed samples collection and clinical study. D.F. was involved in the experimental design and analysis of the proteomes. C.J. was involved in the experimental design and analysis of the mice experiments. W.M., J.B. and A.S. wrote the paper. All the authors have revised and approved the manuscript submission.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: A.S., D.M., D.F., W.M., T.A. and C.-K.C. have a corresponding patent application ‘Methods for the diagnosis and treatment of inflammatory bowel disease’ filed in 2014 (WO/2014/138999). A.S., D.M. and D.F. have co-founded Biogentics, a clinical microbiomics company. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Mottawea, W. et al. Altered intestinal microbiota–host mitochondria crosstalk in new onset Crohn’s disease. Nat. Commun. 7, 13419 doi: 10.1038/ncomms13419 (2016).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/