Title

Age-of-diagnosis dependent ileal immune intensification and reduced alpha-defensin in older versus younger pediatric Crohn Disease patients despite already established dysbiosis

Journal

Mucosal Immunology, 12(2)

ISSN

1933-0219

Authors

Haberman, Yael
Schirmer, Melanie
Dexheimer, Phillip J
et al.

Publication Date

2019-03-01

DOI

10.1038/s41385-018-0114-4

Peer reviewed

HHS Public Access

Author manuscript

Mucosal Immunol. Author manuscript; available in PMC 2019 June 12.

Published in final edited form as:

Mucosal Immunol. 2019 March ; 12(2): 491–502. doi:10.1038/s41385-018-0114-4.

Age-of-Diagnosis Dependent Ileal Immune Intensification and Reduced Alpha-Defensin in Older versus Younger Pediatric Crohn Disease Patients despite Already Established Dysbiosis

Yael Haberman, MD, PhD[1,2], **Melanie Schirmer, PhD**[3,31], **Phillip J. Dexheimer, PhD**[1], **Rebekah Karns, PhD**[1], **Tzipi Braun, MS**[2], **Mi-Ok Kim, PhD**[1], **Thomas D. Walters, MD**[4], **Robert N. Baldassano, MD**[5], **Joshua D Noe, MD**[6], **Joel Rosh, MD**[7], **James Markowitz, MD**[8], **Wallace V. Crandall, MD**[9], **David R. Mack, MD**[10], **Anne M. Griffiths, MD**[4], **Melvin B. Heyman, MD**[11], **Susan S. Baker, MD**[12], **Richard Kellermayer, MD**[13], **Dedrick Moulton, MD**[14], **Ashish S. Patel, MD**[15], **Ajay S. Gulati, MD**[16], **Steven J Steiner, MD**[17], **Neal LeLeiko, MD**[18], **Anthony Otley, MD**[19], **Maria Oliva-Hemker, MD**[20], **David Ziring, MD**[21], **Barbara S Kirschner, MD**[22], **David J Keljo, MD**[23], **Stephen L. Guthery, MD**[24], **Stanley A Cohen, MD**[25], **Scott Snapper, MD**[26], **Jonathan Evans, MD**[27], **Marla Dubinsky, MD**[28], **Bruce Aronow, PhD**[1], **Jeffrey S. Hyams, MD**[29], **Subra Kugathasan, MD**[30], **Curtis Huttenhower, PhD**[3,31], **Ramnik J. Xavier, MD, PhD**[3,32], and **Lee A. Denson, MD**[1]

[1]Cincinnati Children's Hospital Medical Center and the University of Cincinnati College of Medicine, Cincinnati, OH, USA.

[2]Sheba Medical Center, Tel-HaShomer, affiliated with the Tel-Aviv University, Israel

[3]The Broad Institute of MIT and Harvard, Cambridge, MA, USA.

[4]Hospital for Sick Children, University of Toronto, Toronto, ON, Canada

[5]The Children's Hospital of Philadelphia, Philadelphia, PA, USA.

[6]Medical College of Wisconsin, Milwaukee, WI. USA

[7]Goryeb Children's Hospital/Atlantic Health, Morristown, NJ, USA.

[8]Cohen Children's Medical Center of New York, New Hyde Park, NY, USA.

[9]Nationwide Children's Hospital, Columbus, OH, USA.

[10]Children's Hospital of Eastern Ontario, University of Ottawa, Ottawa, ON, Canada.

[11]University of California San Francisco, San Francisco, CA, USA.

Corresponding Author: Yael Haberman, MD, PhD, Division of Pediatric Gastroenterology, Hepatology & Nutrition, Cincinnati Children's Hospital Medical Center, MLC 2010, 3333 Burnet Avenue, Cincinnati, OH 45229, Yael.haberman@cchmc.org, Tel: 1(513)8037416/ 972(544)301319.

Conflict of Interest:
The authors have no conflict of interest to declare.

Financial disclosures: The authors have no financial arrangement(s) with a company whose product figures prominently in the submitted manuscript or with a company making a competing product.

Transcript profiling: The ileal RNASeq data has been placed in the GEO repository with the following accession number: GSE101794.

[12]University at Buffalo, Buffalo, NY, USA.

[13]Texas Children's Hospital, Baylor College School of Medicine, Houston, TX,USA

[14]Vanderbilt Children's Hospital, Nashville, TN

[15]UT Southwestern Medical Center at Dallas,Dallas,TX, USA

[16]University of North Carolina, Chapel Hill, NC, USA

[17]Riley Children's Hospital, Indiannapolis, IN, USA.

[18]Hasbro Children's Hospital, Providence, RI, USA.

[19]IWK Health Centre, Halifax, NS, Canada.

[20]John Hopkins University, Baltimore, MD, USA.

[21]Cedars-Sinai Medical Center, Los Angeles, CA, USA.

[22]University of Chicago Comer Children's Hospital, Chicago, IL, USA.

[23]Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA, USA.

[24]University of Utah, Salt Lake City, UT, USA.

[25]Children's Center for Digestive HealthMedicine, Atlanta, GA, USA.

[26]Children's Hospital - Boston, Boston, MA, USA.

[27]Nemours Children's Clinic, Jacksonville, FL, USA.

[28]Mount Sinai Hospital New York, NY, USA.

[29]Connecticut Children's Medical Center, Hartford, CT, USA.

[30]Emory University, Atlanta, GA, USA.

[31]Harvard T.H. Chan School of Public Health, Boston, MA, USA

[32]Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Abstract

Age-of-diagnosis associated variation in disease location and antimicrobial sero-reactivity has suggested fundamental differences in pediatric Crohn Disease (CD) pathogenesis. This variation may be related to pubertal peak incidence of ileal involvement and Peyer's patches maturation, represented by IFNγ-expressing Th1 cells. However, direct mucosal evidence is lacking. We characterize the global pattern of ileal gene expression and microbial communities in 525 treatment-naïve pediatric CD patients and controls (Ctl), stratifying samples by their age-of-diagnosis. We show a robust ileal gene signature notable for higher expression of specific immune genes including *GM-CSF* and *INFγ*, and reduced expression of antimicrobial Paneth cell α-defensins, in older compared to younger patients. Reduced α-defensin expression in older patients was associated with higher *IFNγ* expression. By comparison, the CD-associated ileal dysbiosis, characterized by expansion of Enterobacteriaceae and contraction of *Lachnospiraceae* and *Ruminococcaceae*, was already established within the younger group and did not vary systematically with increasing age-of-diagnosis. Multivariate analysis considering individual taxa,

however did demonstrate negative associations between *Lachnospiraceae* and *IFNγ*, and positive associations between *Bacteroides* and α-defensin expression. These data provide evidence for maturation of mucosal Th1 immune responses and loss of epithelial antimicrobial α-defensins which are associated with specific taxa with increasing age-of-diagnosis in pediatric CD.

Keywords

Crohn Disease; Pediatric age-of-diagnosis; immune response; mucosal microbial profile and transcriptome

Introduction

Epidemiologic associations, clinical phenotype, and natural history differ across CD age-of-diagnosis [1–4]. Older children, adolescents, and young adults develop CD involving the ileum most commonly[5], whereas colon-only involvement is predominant in the first decade of life and in the elderly [1,3]. Variation in antimicrobial serology with increasing age-of-diagnosis within pediatric CD has suggested potential age-dependent differences in the ileal microbial community and/or immune responses [6]. These observations have informed the Paris classification system [6] that sub-divides the Montreal A1 (0–16 years) classification for CD into diagnosis prior to age 10 years (A1a) and age 10–16 years (A1b). Recently, the younger A1a group was subdivided further to children diagnosed prior to 6 years of age, the very early onset type [7,8] that was linked to monogenic forms of Inflammatory Bowel Disease (IBD) that differ from the polygenic form of CD diagnosed in older ages [4,9]. Whether there was a specific mucosal basis for the clinical-based classification of the polygenic CD to younger A1a (6 years to <10 years) and older A1b (10–16 years) was not known.

The gut-associated lymphoid Peyer's patches (PP), concentrated (50%) in the distal ileum (TI) [10,11], interact with the overlying epithelium to induce tolerance or defense against luminal antigens [12]. Importantly, PP undergoes dynamic maturation from birth, whereby the number and size of PP increase up to the second/third decade of life and then decline with age. The spatiotemporal relationship between the peak incidence of ileal CD (iCD) and PP evolution is represented primarily by INFγ-expressing Th1 cells; this has led to the hypothesis that this dynamic mucosal immune process plays a central role in iCD pathogenesis [5,10,11]. In fact, murine studies have revealed maturation of an IFN-γ associated immune signature in adult vs. preweaned mice [13], and a prior study in healthy children showed a pronounced Th1 polarization (increased IFNγ and IL-12) within PP and adjacent ileal mucosa in response to bacterial antigens [14]. However, comprehensive human mucosal-based gene expression analysis to test for maturation of a Th1 signature across the pediatric age-of-diagnosis range within CD had not been performed.

Paneth cells, located at the base of the intestinal crypts, produce antimicrobial peptides such as lysozymes and α-defensins to modulate the intestinal microbiota and are an important arm of the innate immune response [15]. Reduced human α-defensins (*DEFA5* and *DEFA6*) expression were previously documented in patients with iCD [16]. However, it is unclear if this deficiency is a primary event in iCD or a secondary event occurring as a consequence of inflammation [17]. Whether differences in α-defensin expression would be observed as a

function of age-of-diagnosis in pediatric iCD, and whether this in turn is associated with alterations in the microbial community, is not known. Interestingly, our group has recently shown that in a competing-risk model, older age-of-diagnosis, African American race, and ASCA IgA and CBir1 sero-positivity were associated with internal penetrating (B3) disease complications [18]. To improve our understanding of pediatric CD pathogenesis across different ages-of-diagnosis, we employed combined ileal mRNA and 16S rRNA sequencing approaches to define host transcriptome and microbial community differences in patients stratified by their age-of-diagnosis. We show that despite an age-independent microbial dysbiosis in pediatric CD, there are robust differentially expressed host ileal gene signature differences with higher expression of an inflammatory-related genes, and a lower expression of Paneth cell α-defensins, with increasing age-of-diagnosis.

Materials and Methods

The RISK cohort.

The Crohn's and Colitis Foundation sponsored RISK prospective inception cohort study [18–20] included newly diagnosed pediatric CD patients (Table 1) at 28 North American pediatric gastroenterology centers. 946 patients had inflammatory (B1) disease behavior and no disease complications (B2 or B3) at the time of diagnosis. All patients were required to undergo baseline colonoscopy and confirmation of characteristic endoscopic features and chronic active colitis/ileitis by histology prior to diagnosis and treatment, and documented initial clinical severity and disease location. Non-IBD controls were subjects suspected to have IBD, but with normal radiographic, endoscopic, and histologic findings. Ileal biopsy samples from a CD sub-cohort representative of the overall RISK cohort (age, gender, and disease phenotype and severity) and non-IBD controls (age and gender) are included in our mucosal mRNAseq analysis (254 CD and 50 control (Ctl), Table 1, 2). Ileal biopsies of CD patients (n=272) and Ctl (n=178) were also analyzed for ileal microbial profiles and were already included in our recent reports[18–20], of those 197 also had mRNAseq (160 CD and 37 Ctl). Patient-preparation for endoscopy and treatment course were according to the dictates of their physicians, not by standardized protocols. Serological determination of perinuclear anti-neutrophil cytoplasmic antibodies (pANCA), anti-Saccharomyces cerevisiae antibodies (ASCA) IgG, ASCA IgA, and anti-CBir1, was performed at Cedars-Sinai Hospital (Los Angeles, CA, USA)[21]. Granulocyte–macrophage colony-stimulating factor (GM-CSF) autoantibodies were measured at Cincinnati Children's Hospital Medical Center (Cincinnati, OH, USA). Positive serologies were defined based on specific predefined cut points. anti-GMCF is consider positive with values above 1.6mcg/mL[22], the value previously reported within pediatric-onset CD patients linked to higher risk of stricturing disease complications, and subsequently validated within adult-onset CD patients for this outcome. Other antibody levels were determined and results are expressed as ELISA units (EU/ml), which are relative to a Cedars-Sinai Laboratory standard, derived from a pool of CD patient sera with well-characterized disease found to have reactivity to this antigen; ASCA IgA is considered positive above 20 EU/ml, ASCA IgG above 40 EU/ml, anti-CBir1 above 25 EU/ml, and pANCA above 30 EU/ml.

Ileal DNA and RNA extraction and RNA-seq.

DNA and RNA were isolated from ileal biopsies as previously described [19]. Reads were quantified by kallisto [23], using Gencode v23 as the reference genome and Transcripts per Million (TPM) as an output. We included 13,206 protein-coding genes with TPM above 5 in 5 samples. Differentially expressed genes were determined in GeneSpring® software with fold change differences (FC) >=1.5 and using false discovery rate correction (FDR<0.05). Euclidean distance metric and Ward's linkage rule was used for unsupervised hierarchical clustering. ToppGene [24] and ToppCluster [25] software were used to test for functional annotation enrichment analyses. Visualization of the network was obtained using Cytoscape.v3.0.2 [26]. The RNASeq data is deposited in the GEO repository (GSE101794).

Microbial community profiling and analysis of associations testing between microbial taxa and clinical and molecular metadata.

Detailed protocols used for 16S rRNA amplification and sequencing are as previously described [18,20,27]. In brief, 16S rRNA amplicon sequencing of 450 ileal samples was performed using the Illumina MiSeq v2 platform, targeting the V4 region of the SSU rRNA gene, and generating paired-end reads of 175b. Samples with at least 3,000 reads were included. Taxonomy was assigned based on the Greengenes database [28]. Each OTU was required to occur at a relative abundance of at least 0.01% across all samples and be present in at least 5%.

Differentially abundant taxa were determined based on multivariate statistical analyses using Multivariate Analysis by Linear Models (MaAsLin) [18–20]. The following metadata were investigated in the analysis: age, gender, body mass index (BMI, as a measure of nutritional status), clinical phenotype (Ctl or CD), endoscopic severity (deep ulcers in ileum), clinical severity (Pediatric Crohn's Disease Activity Index, PCDAI), ileal gene expression of *CSF2, CXCR1, IFNG, MMP3, DEFA5, GSTA1,* and *LCT,* and *NOD2* and *ATG16L1* IBD risk allele carriage. Significant association was considered below a *q* value (Benjamini-Hochberg) threshold of 0.2.

Immunohistochemistry.

Immunohistochemistry detection of the human alpha defensin five protein encoded by the *DEFA5* gene in control and CD ileal biopsies was performed as previously described in the CCHMC Digestive Healthy Center core facility[19], using anti-*DEFA5* (mouse monoclonal anti-human alpha defensin NP5 Ab, clone nr 8C8, code ab62757, Abcam, Cambridge, UK). Staining was examined using an Olympus BX51 light microscope and digitally recorded at 20x and 40x magnification.

Ethical Considerations.

The institutional review board reviewed and approved the protocol and informed written consent was obtained. ClinicalTrials.gov identifier is NCT00790543.

Results

The RISK cohort.

RISK is a prospective inception cohort study which enrolled pediatric CD patients at diagnosis at 28 sites in North America. RISK includes 946 treatment naïve newly diagnosed CD patients that were all classified as having an inflammatory B1 phenotype without B2 (stricturing/narrowing) or B3 (internal penetrating behavior) complications at diagnosis. Change in disease behavior from B1 inflammatory to either B2 stricturing or B3 penetrating behavior was recorded during follow up. For the purpose of our study, patients were stratified based on the Paris age-of-diagnosis classification, where diagnosis prior to age 10 years is defined as A1a, and age 10–16 years is defined as A1b. The younger A1a were further subdivided to very early onset (VEO, <6 years[7]), and to early onset (EO) (6–9 years) [7,8].

Table 1 shows demographic and clinical characteristics categorized by age-of-diagnosis. Of the 946 CD patients in the RISK cohort, only 36 (4%) were classified as VEO (<6 years) CD, 196 (21%) as EO younger CD (6–9 years), and the rest (75%) as EO older CD (10–16 years). Clinical severity defined using the Pediatric Crohn's Disease Activity Index (PDCAI) was not significantly different between the groups at diagnosis. Perianal involvement was significantly higher in older CD. As previously reported[6], sero-positivity for anti-GMCSF and ASCA were significantly increased with age-of-diagnosis classifications. There were no significant differences in baseline PCDAI, early anti-TNF exposure that was previously associated with lower rate of penetrating (B3) complications[18], and PCDAI six months after diagnosis between older and younger CD cases. However, there was a significantly higher prevalence of penetrating complication (B3) during 3 years follow up in older vs. younger CD (Table 1).

Based on those differences between older (10–16 years) and younger EO CD, the relative low number of VEO cases, and the likelihood association between the VEO cases and the monogenic IBD type, we included hereafter only patients older than 6 years in the mucosal transcriptomics and microbial analyses. Ileal mucosal transcriptomics and microbial characterization were done on a representative sub-group of ileal CD (iCD, n=198), colon only CD (cCD n=56), and non-IBD Ctl (n=50) (Table 2 & Suppl. Table 1). cCD patients are subjects who met diagnostic criteria for CD but lacked ileal inflammation on endoscopy. This unique large patient population and biospecimens offered an opportunity to directly test for the association between presence of mucosal inflammation, host mucosal genes and pathways, and mucosal microbial composition by performing an age-dependent analyses in treatment naive patients.

Decreased expression of α-defensins in older iCD.

We identified 365 genes (Suppl. Table 2 & Fig 1) that were significantly differentially expressed between older and younger iCD. We have previously reported the identification of a core iCD gene expression signature[19]. Interestingly, the α-defensin genes (*DEFA5* & *DEFA6*) and *REG3A* were not included in the core iCD list (Fig. 1A), as these were only suppressed in older iCD compared to both younger iCD and Ctl (Fig. 1C). Results of

unsupervised clustering using the 365 differentially expressed genes are shown in Fig. 1B with the indicated location of α-defensins. 51 of 148 (34.5%) older CD and 5 of 48 (10.4%) younger CD (Chi squares p=0.0013) are in block 1, while 20 of 148 (13.5%) older, and 16 of 48 (33%) younger CD (Chi squares p=0.002) are in block 2b(ii). *DNASE1* that was negatively associated with abnormal Paneth cell morphology in iCD [29] showed a similar pattern of lower expression level in older iCD as the α-defensins and *REG3G/A* in contrast to stable increased expression of lysozyme (*LYZ*, Fig 1C & Suppl. Table 2). Consistent with this, a reduced frequency of alpha defensin 5 positive epithelial cells was detected in ileal biopsies from older CD patients, compared to younger CD patients and controls (Fig 1D). Unsupervised hierarchical clustering analysis identified groups of biopsies with similar ileal gene expression profiles. Unsupervised clustering using the top third most differentially expressed genes subset (121 of 365 genes) demonstrated that patients diagnosed at age 6, 7 8, and 9 years clustered together and separately from patients who were diagnosed at age 10, 11, 12, 13, 14, 15, and 16 years (Fig. 1E and suppl. Fig. 1).

Enhanced ileal immune responses in older iCD.

We next tested whether the ileal transcriptome of older iCD exhibited mucosal immunologic maturation represented by a Th1 gene expression profile compared with younger iCD. Functional annotation enrichment analyses using ToppGene [24] and ToppCluster [25] were used to map groups of related genes within the 365 gene signature to biologic functions and immune cell types (Fig. representative pathways A, Suppl. Table 3 & 4). Genes up-regulated in older iCD were notable for higher expression of an inflammatory-related signature including *GM-CSF (CSF2)*, *IFNγ*, matrix metalloproteinases, and collagens (Suppl. Table 2). This gene signature showed functional annotation enrichment (Fig. 2A & Suppl. Table 4) for genes induced in response to molecules of bacterial origin ($P < 4.82E-07$) including several monocyte-derived pro-inflammatory cytokines and the Th1-related cytokine *IFNγ*. Additional functional annotation enrichment was noted for genes expressed by granulocytes ($P < 2.65E-06$), B cells ($P < 2.72E-04$), and lymphoid stromal cells ($P < 8.1E-14$, Fig. 2B). The up regulated genes also showed a remarkable increase for genes encoding extracellular matrix (ECM) and extracellular matrix-associated proteins ($P < 4.95E-16$), extracellular space ($P < 1.12E-26$), and genes defining epithelial-mesenchymal transition ($P < 4.51E-15$). The functional annotation enrichment analyses included enrichment for genes associated with a specific drug or supplement, with one of the top enrichments for curcumin ($P < 1.15E-9$,Suppl. Table 4). We also noted enrichment for collagen genes ($P < 1.05E-05$) including *COL12A1, COL1A1, COL3A1, COL6A3, COL7A1*, and *COL8A1*. Of note, this age-dependent gene signature was specific, as several other mucosal inflammatory genes including the NADPH oxidase *DUOX2*, *FOLH1*, *REG1A*, and *SAA2*, that we previously identified as part of a core iCD gene signature[19], were up-regulated to the same degree in older and younger iCD and are not included in the 365 gene list. These results defined a maturation of components of both the innate and Th1 adaptive ileal gene signature in pediatric CD patients with older age-of-diagnosis.

A similar approach was applied to the down regulated genes between older and younger iCD (Fig. 2B). These analyses identified decreased genes associated with brush border localization ($P < 5.23E-24$), digestion ($P < 1.65E-13$), lipid metabolic processes ($P <$

1.58E-08), and vitamin digestion and absorption (P < 1E-07) in older iCD. Additionally, we found that 32 genes of the 365 genes had a FC ≥1.5 in older vs younger Ctl (Suppl. Table 5), including *LCT* that showed decreased expression ≥2 in the older vs. younger groups. However, only 2 passed corrected P values FDR of ≤0.05.

Decreased α-defensins expression in older CD is associated with Th1 IFNγ and inflammation but not with *NOD2* or *ATG16L1* genotype.

Our cohort also included cCD patients, who met diagnostic criteria for CD but lacked ileal inflammation on endoscopy. This unique patient population offered an opportunity to directly test for the association between mucosal inflammation and gene expression by performing age-dependent analyses. Unsupervised clustering using specifically the 365 genes that were differentially expressed between older and younger iCD demonstrated that while patients groups first clustered based on disease phenotype (CD and Ctl), older cCD clustered with younger iCD and cCD (Fig. 3A). α-defensin gene expression was not significantly decreased in the ileum of older cCD (A1b) in comparison to younger cCD or Ctl in a univariate non parametric t-test. 84 (Suppl. Table 6) of the 365 genes differentially expressed between older and younger iCD showed a FC>1.5 in the ileum of older vs. younger cCD including *DEFA5* but none passed the FDR <0.05. To capture differences between older CD with and without clinical ileal inflammation (iCD vs. cCD), we identified 1135 genes that were significantly differentially expressed between older iCD and cCD (Suppl. Table 7). Those 1135 included *CSF2* and *IFNG* that were up-regulated in all CD forms in comparison to Ctl, with further increased expression in older iCD. Of note, no genes were differentially expressed between younger iCD and cCD, showing that genes such as *CSF2* and *IFNG* were associated with overt mucosal inflammation specifically in the older age group. We next asked whether α-defensin expression would be associated with *NOD2* or *ATG16L1* genotype. We did not observe a significant difference when we compared *DEFA5* expression between patients stratified by *ATG16L1* or *NOD2* (Fig. 3B & 3C) risk allele carriage.

Th1-related IFNγ was suggested to suppress human α-defensin gene expression[30], while α-defensin expression was shown to specifically inhibit Th1-related IFNγ inflammation in a murine model [31]. We identified significant negative associations between *IFNG* and *DEFA5* gene expression (Pearson r = -0.33, P < 0.0001) in our cohort. Collectively, these results characterized an unexpected age-dependent down-regulation of α-defensin genes within pediatric CD which was in turn associated with increasing *IFNG* expression and intensified extracellular matrix and collagen expression in patient with older age-of-diagnosis.

Microbial dysbiosis in pediatric CD is already established in the younger group and lacks systematic changes observed in non-IBD Ctls.

We have previously described a pediatric CD-associated ileal dysbiosis [19,20]. Variation in microbial samples was compared by using Principal Coordinate Analysis (PCoA) with Bray-Curtis distance and samples were colored by either the Chao1 α-diversity (richness within a sample) or Paris age (Fig. 4A). The frequency of younger and older CD with PC1 values >0 was not significantly different between the two groups (41% of the younger CD and 45% of the older CD). Differences in α-diversity were the main driver of variation rather than age or

diagnosis, as was previously observed. Further, α-diversity was significantly reduced in older CD (CD A1b) vs. Ctl of any age. To capture age-dependent differences we applied a multivariate statistical framework (MaAsLin) [20] to Ctl samples (n=177, n=47 for younger and n=130 for older) and CD samples (n=272, n=64 for younger and n=208 for older), respectively. We were able to detect significant differences in microbial abundances between older vs. younger Ctl (≥10 years vs. <10 years) while controlling for body mass index (BMI), ethnicity, gender, and *NOD2*, and *ATG16L1* risk allele carriage (Fig. 4C). This included a significant decrease of Enterobacteriaceae, *Parabacteroides distasonis*, *Streptococcus*, *Veillonella*, Gemellaceae, Lachnospiraceae, and *Enterococcus* in older Ctl. Applying the same analysis to CD samples, while additionally controlling for antibiotic usage and ileal deep ulcers yielded no significant associations. Collectively these analyses demonstrated no systematic changes in CD across age-of-diagnosis suggesting that that CD-associated dysbiosis was already established at diagnosis in the younger CD patients.

Host gene expression differences associated with ileal microbial community composition.

We next applied MaAsLin [19,20] in order to test specifically for associations between age-dependent CD host gene expression and specific microbial taxonomy factor (n=229). We included in the analysis differentially expressed genes from selected representative pathways (Fig. 1 & 2) including *DEFA5*, *GSTA1*, and *LCT* that were down regulated and *CSF2*, *CXCR1, IFNG*, and *MMP1* that were up regulated in older vs. younger CD. We also included clinical phenotype (Ctl, CD), endoscopic severity (ileal deep ulcers), clinical severity (PCDAI), age-of-diagnosis (≥10 years vs. <10 years), gender, BMI, recent exposure to antibiotics, and *NOD2*, and *ATG16L1* risk allele carriage. Overall, we identified 38 significant associations for clinical phenotype, and 5 significant associations with host gene expression (Fig. 5).

The bar graph (Fig. 5A) illustrates fold change differences between CD and Ctl taking into account the clinical factors and host gene expression. As was previously reported by us [19,20] and others [32], CD phenotype showed a decrease in taxa from the Firmicutes phylum including *Faecalibacterium prausnitzii, Lachnospiraceae* and *Ruminococcaceae* OTUs and an increase in taxa from the Proteobacteria and Fusobacteria phyla including *Pasteurellaceae*, *Campylobacteraceae*, *Enterobacteriaceae*, and *Fusobacteriaceae* organisms. Further, we determined a positive associations between *Bacteroides* abundance and α-defensin expression and a negative association between *Acinetobacter* and α-defensin expression. Finally, we noted a negative association between *Lachnospiraceae* abundance and *IFNγ* expression and between *Bacteroides fragilis* and *MMP1* expression. No significant associations were captured for *CSF2, CXCR1, GSTA1, and LCT* host gene expression. Altogether, these results demonstrated that while some of the age-dependent host gene expression dynamics were associated with specific microbial taxa, gene expression differences were largely not associated with systematic microbial shifts within CD age-of-diagnosis.

Discussion

Epidemiologic associations and clinical phenotype as well as natural history differ across CD age-of-diagnosis[1–4]. Using the RISK cohort, we were able to show that while there were no significant differences in baseline clinical severity (PCDAI), early exposure to anti-TNFα [18], and PCDAI six month after diagnosis between older (A1b) and younger (A1a) CD, however, the prevalence of penetrating complications (B3) during follow up was significantly higher in older CD (Table 1). By using whole transcriptome mRNA-seq analyses, we were able to identify 365 genes that were differentially expressed between older and younger iCD, with older iCD showing an increased Th1-related IFNγ profile associated with amplified inflammatory activation including an enriched innate myeloid and lymphoid stromal cell signature, and enhanced extra cellular matrix and collagen signatures. Remarkably, this signature for enhanced immune activation occurred in the older iCD group in association with a specific reduction in epithelial Paneth cell α-defensin expression, and was not associated with systematic changes in the local mucosal microbial communities across CD age-of-diagnosis (Summarizing cartoon, Figure 6). We therefore suggest that age-dependent ileal host inflammatory intensification and depression of α-defensins genes is largely intrinsic to the mucosal dynamics and is not associated with a systematic local microbial community shift.

While a reduction in α-defensins was previously described in ileal CD, there was an ongoing controversy as to whether *NOD2* and *ATG16L1* genotype and/or the presence of inflammation is the primary drive of low α-defensin levels [16,33,34]. We detected preservation of *DEFA5/DEFA6* expression in the ileum of younger iCD patients similar to Ctl levels, and a significant reduced expression in older iCD patients. We specifically show that older age-of-diagnosis is associated with decreased expression of Paneth-cell associated α-defensins and that these age–related differences cannot be explained solely by either *NOD2* or *ATG16L1* genotype. We did not however identify similar decreased expression of *TCF4* [35], *LRP6* [36], and *TCF7* [37] (also known as *TCF-1*) which were previously linked to α-defensin expression in iCD. *DNASE1* that was negatively associated with abnormal Paneth cell morphology in iCD [29] showed a similar pattern of lower expression level in older iCD as the α-defensins and *REG3G/A* in contrast to stable increased expression of lysozyme (*LYZ*) that was one of the differentially expressed core iCD genes [19] but not included in the 365 genes (Suppl. Table 2).We did identify a negative association between IFNγ and α-defensins expression, consistent with the previously reported bidirectional negative regulation of these genes[31]. Th1-related IFNγ was suggested to specifically repress α-defensin gene expression by leading to Paneth cells extrusion and subsequent cell death [30].

Early life gut microbial maturation and its implications in health and disease were previously characterized [38]. By the end of the third year of life, the microbiome composition evolves toward an adult-like configuration [39]. However, a specific understanding of the microbiome of older pediatric cohorts (>3 years) and pre-adulthood (<18 years) has been limited [40,41], whereby the largest cohort included 62 stool samples from residents of the United States [39]. Our cohort [19,20] includes 178 Ctl with ileal mucosal microbial composition. We were able to show that dynamic microbial changes occur in the ileum of older Ctl cases in comparison to younger Ctl cases, with no similar significant association

detected in in older vs. younger CD patients. Collectively these data show that the CD dysbiosis characterized by increased abundance of pro-inflammatory taxa including taxa from Enterobacteriaceae, Pasteurellaceae, Veillonellaceae, and Fusobacteriaceae families, and decreased abundance of anti-inflammatory taxa from the Erysipelotrichaceae, Ruminococcaceae and Lachnospiraceae families was largely independent of CD age-of-diagnosis.

To test for associations between the microbiota, and age-dependent ileal gene expression, we specifically performed multivariate analysis that included Ctl versus CD clinical phenotype, age-of-diagnosis, and host genes that were differentially expressed in older vs. younger CD. Indeed, we were able to identify a significant association with expression of specific Th1 and tissue remodeling genes (*IFNγ* and *MMP1*), which were up-regulated in older CD, and *Lachnospiraceae* and *Bacteroides fragilis* OTU abundance respectively. We also identified a significant positive association between *Bacteroides* abundance and negative association between *Acinetobacter,* and the antimicrobial gene *DEFA5.* Importantly, *DEFA5*-transgenic mice also showed significant increase in taxa from the Bacteroidetes phylum [15]. These data demonstrate that while some of the mucosal gene expression differences were associated with the abundance of specific microbial taxa, much of the robust age-dependent mucosal immune gene signature maturation observed is likely driven by host mucosal factors. However, with further defining of the changes in intestinal microbiota, specific strains may be found to play important in mucosal gene expression. For instance, a *Faecalibacterium prausnitzii* out was found to be decreased but there are also other *F. prausnitzii* OTUs that are increased in CD suggesting potential strain-specificity [42].

Our study has several strengths, but also some limitations. Although it is reasonable to assume that puberty influences the observed gut maturation, we lacked consistent Tanner Stage information to specifically address this hypothesis. However, the clinical practice approach, namely the Paris classification, is specifically based on age-of-diagnosis and not on Tanner Stage. Future studies will need to test whether Tanner Stage and puberty-related hormones can further define primary pathways driving mucosal gene expression maturation in association with disease onset at specific ages. Because we require OTUs to occur at a relative abundance of at least 0.01% across all samples, it remains possible that a rare but infrequent immunomodulatory organism might explain the differences in the younger and older group. Moreover, we lacked power to test for differences in mucosal gene expression between sub-sets of younger and older patients stratified by measures of microbial diversity. Additionally, we used whole biopsies, composed of a mixture of cellular components rather than single cell transcriptomics. However, in order to capture the overall pathogenic process, and as a potential future diagnostic/prognostic tool, there are also substantial advantages in using whole mucosal biopsies, which are the basis for diagnosis and follow up in the clinical setting. Another limitation is the inability to replicate this results in an independent large treatment naïve human cohort not affected by treatment regimens. RISK is the largest treatment naïve inception cohort, involving 28 sites in the US, and other such large human cohorts are not yet available. However, there are several murine and mechanistic studies supporting our findings. Age-specific response to enteric *Salmonella* infection, with developmentally regulated intestinal expression of IFN-γ and its target genes was already noted [13]. Paneth cells degranulation does not directly occur upon stimulation with microbial

antigens or bacteria, but IFN-γ induces rapid and complete loss of Paneth cells granules, coupled with induction of apoptosis, luminal extrusion, and death of Paneth cells [30]. Finally, similar to our findings, Firmicutes was significantly lower and Bacteroidetes was significantly higher in the DEFA5 transgenic (+/+) mice than in controls [15].

In summary, our data identify important clinical and biologic differences between older and younger pediatric CD. We would suggest that pediatric A1b CD phenotype and mucosal gene expression signature, is likely very similar to adult (A2) CD [16,43,44], and different from the younger pediatric age group. Interestingly, in mice and humans the initial formation of Peyer's patches occurs before birth in a relative sterile environment and hence is likely independent of microbiota[45]. Consistent with this observation, our data suggest that much of the age-dependent differentially expressed gene expression signature in pediatric CD (≥10 years) is intrinsic to the host ileal mucosa, and not primarily associated with systematic shifts in the local microbial community. These specific host and microbial factors may offer the potential to tailor future age-based therapy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

This work was supported by the Crohn's and Colitis Foundation, the Gene and Protein Expression and Bioinformatics cores of the National Institutes of Health (NIH)-supported Cincinnati Children's Hospital Research Foundation Digestive Health Center (1P30DK07839–201), NIH grant U54 DE023798 and HMP2 (R.J.X., and C.H), the Leona M. and Harry B. Helmsley Charitable Trust (R.J.X., and C.H), the European Crohn's and Colitis Organization (ECCO, Y.H), the Israel Science Foundation (grant No 908/15), the I-CORE program (Y.H), and European Research Council starting grant (grant No 758313, Y.H). We thank the Crohn's and Colitis Foundation RISK study publication committee for critical review of this manuscript.

Abbreviations:

A1a	age < 10 at diagnosis
A1b	age ≥ 10 at diagnosis
BMI	body mass index
CD	Crohn disease
cCD	colon-only CD (L2 Montreal classification)
Ctl	non-IBD control
CCFA	The Crohn's & Colitis Foundation of America
DC	dendritic cells
DU	deep ulcers
IBD	inflammatory bowel diseases
iCD	ileal CD (L1 or L3 Montreal classification)

MaAsLin	Multivariate Analysis by Linear Models
OUT	operational taxonomic units
PCDAI	Pediatric Crohn's Disease Activity Index
UC	ulcerative colitis

References

1. Ruel J, Ruane D, Mehandru S, Gower-Rousseau C & Colombel JF IBD across the age spectrum-is it the same disease? Nature reviews. Gastroenterology & hepatology 11, 88–98, doi:10.1038/nrgastro.2013.240 (2014). [PubMed: 24345891]
2. Siegel CA et al. Real-time tool to display the predicted disease course and treatment response for children with Crohn's disease. Inflammatory bowel diseases 17, 30–38, doi:10.1002/ibd.21386 (2011). [PubMed: 20812335]
3. Gower-Rousseau C et al. Epidemiology of inflammatory bowel diseases: new insights from a French population-based registry (EPIMAD). Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver 45, 89–94, doi:10.1016/j.dld.2012.09.005 (2013).
4. Heyman MB et al. Children with early-onset inflammatory bowel disease (IBD): analysis of a pediatric IBD consortium registry. The Journal of pediatrics 146, 35–40, doi:10.1016/j.jpeds.2004.08.043 (2005). [PubMed: 15644819]
5. Meinzer U et al. Ileal involvement is age dependent in pediatric Crohn's disease. Inflammatory bowel diseases 11, 639–644 (2005). [PubMed: 15973117]
6. Levine A et al. Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification. Inflammatory bowel diseases 17, 1314–1321, doi:10.1002/ibd.21493 (2011). [PubMed: 21560194]
7. Uhlig HH et al. The diagnostic approach to monogenic very early onset inflammatory bowel disease. Gastroenterology 147, 990–1007 e1003, doi:10.1053/j.gastro.2014.07.023 (2014). [PubMed: 25058236]
8. Muise AM, Snapper SB & Kugathasan S The age of gene discovery in very early onset inflammatory bowel disease. Gastroenterology 143, 285–288, doi:10.1053/j.gastro.2012.06.025 (2012). [PubMed: 22727850]
9. de Bie CI et al. Diagnostic workup of paediatric patients with inflammatory bowel disease in Europe: results of a 5-year audit of the EUROKIDS registry. Journal of pediatric gastroenterology and nutrition 54, 374–380, doi:10.1097/MPG.0b013e318231d984 (2012). [PubMed: 21857248]
10. Cornes JS Number, size, and distribution of Peyer's patches in the human small intestine: Part I The development of Peyer's patches. Gut 6, 225–229 (1965). [PubMed: 18668776]
11. Van Kruiningen HJ, Ganley LM & Freda BJ The role of Peyer's patches in the age-related incidence of Crohn's disease. Journal of clinical gastroenterology 25, 470–475 (1997). [PubMed: 9412954]
12. Jung C, Hugot JP & Barreau F Peyer's Patches: The Immune Sensors of the Intestine. International journal of inflammation 2010, 823710, doi:10.4061/2010/823710 (2010). [PubMed: 21188221]
13. Rhee SJ, Walker WA & Cherayil BJ Developmentally regulated intestinal expression of IFN-gamma and its target genes and the age-specific response to enteric Salmonella infection. J Immunol 175, 1127–1136 (2005). [PubMed: 16002714]
14. Salvati VM et al. Enhanced expression of interferon regulatory factor-1 in the mucosa of children with celiac disease. Pediatric research 54, 312–318, doi:10.1203/01.PDR.0000079184.70237.9C (2003). [PubMed: 12788988]
15. Salzman NH et al. Enteric defensins are essential regulators of intestinal microbial ecology. Nature immunology 11, 76–83, doi:10.1038/ni.1825 (2010). [PubMed: 19855381]

16. Wehkamp J et al. Reduced Paneth cell alpha-defensins in ileal Crohn's disease. Proceedings of the National Academy of Sciences of the United States of America 102, 18129–18134, doi:10.1073/pnas.0505256102 (2005). [PubMed: 16330776]
17. Ramasundara M, Leach ST, Lemberg DA & Day AS Defensins and inflammation: the role of defensins in inflammatory bowel disease. Journal of gastroenterology and hepatology 24, 202–208, doi:10.1111/j.1440-1746.2008.05772.x (2009). [PubMed: 19215333]
18. Kugathasan S et al. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. Lancet, doi:10.1016/S0140-6736(17)30317-3 (2017).
19. Haberman Y et al. Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. The Journal of clinical investigation 124, 3617–3633, doi:10.1172/JCI75436 (2014). [PubMed: 25003194]
20. Gevers D et al. The treatment-naive microbiome in new-onset Crohn's disease. Cell host & microbe 15, 382–392, doi:10.1016/j.chom.2014.02.005 (2014). [PubMed: 24629344]
21. Dubinsky MC et al. Increased immune reactivity predicts aggressive complicating Crohn's disease in children. Clin Gastroenterol Hepatol 6, 1105–1111, doi:10.1016/j.cgh.2008.04.032 (2008). [PubMed: 18619921]
22. Han X et al. Granulocyte-macrophage colony-stimulating factor autoantibodies in murine ileitis and progressive ileal Crohn's disease. Gastroenterology 136, 1261–1271, e1261–1263, doi:10.1053/j.gastro.2008.12.046 (2009).23 [PubMed: 19230854]
23. Bray NL, Pimentel H, Melsted P & Pachter L Near-optimal probabilistic RNA-seq quantification. Nature biotechnology 34, 525–527, doi:10.1038/nbt.3519 (2016).
24. Chen J, Bardes EE, Aronow BJ & Jegga AG ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic acids research 37, W305–311, doi:10.1093/nar/gkp427 (2009). [PubMed: 19465376]
25. Kaimal V, Bardes EE, Tabar SC, Jegga AG & Aronow BJ ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. Nucleic acids research 38, W96–102, doi:10.1093/nar/gkq418 (2010). [PubMed: 20484371]
26. Saito R et al. A travel guide to Cytoscape plugins. Nature methods 9, 1069–1076, doi:10.1038/nmeth.2212 (2012). [PubMed: 23132118]
27. Caporaso JG et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The* ISME journal 6, 1621–1624, doi:10.1038/ismej.2012.8 (2012). [PubMed: 22402401]
28. McDonald D et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME* journal 6, 610–618, doi:10.1038/ismej.2011.139 (2012). [PubMed: 22134646]
29. Vandussen KL et al. Genetic variants synthesize to produce paneth cell phenotypes that define subtypes of Crohn's disease. Gastroenterology 146, 200–209, doi:10.1053/j.gastro.2013.09.048 (2014). [PubMed: 24076061]
30. Farin HF et al. Paneth cell extrusion and release of antimicrobial products is directly controlled by immune cell-derived IFN-gamma. The Journal of experimental medicine 211, 1393–1405, doi:10.1084/jem.20130753 (2014). [PubMed: 24980747]
31. Biswas A et al. Induction and rescue of Nod2-dependent Th1-driven granulomatous inflammation of the ileum. Proceedings of the National Academy of Sciences of the United States of America 107, 14739–14744, doi:10.1073/pnas.1003363107 (2010). [PubMed: 20679225]
32. Frank DN et al. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proceedings of the National Academy of Sciences of the United States of America 104, 13780–13785, doi:10.1073/pnas.0706625104 (2007). [PubMed: 17699621]
33. Simms LA et al. Reduced alpha-defensin expression is associated with inflammation and not NOD2 mutation status in ileal Crohn's disease. Gut 57, 903–910, doi:10.1136/gut.2007.142588 (2008). [PubMed: 18305068]

34. Shanahan MT et al. Mouse Paneth cell antimicrobial function is independent of Nod2. Gut, doi: 10.1136/gutjnl-2012-304190 (2013).
35. Perminow G et al. Defective paneth cell-mediated host defense in pediatric ileal Crohn's disease. The American journal of gastroenterology 105, 452–459, doi:10.1038/ajg.2009.643 (2010). [PubMed: 19904243]
36. Koslowski MJ et al. Association of a functional variant in the Wnt co-receptor LRP6 with early onset ileal Crohn's disease. PLoS genetics 8, e1002523, doi:10.1371/journal.pgen.1002523 (2012). [PubMed: 22393312]
37. Beisner J et al. TCF-1 mediated Wnt Signaling regulates Paneth cell innate immune defense effectors HD-5 and -6: implications for Crohn's disease. *American journal of* physiology. Gastrointestinal and liver physiology, doi:10.1152/ajpgi.00347.2013 (2014).
38. Yatsunenko T et al. Human gut microbiome viewed across age and geography. Nature 486, 222–227, doi:10.1038/nature11053 (2012). [PubMed: 22699611]
39. Gordon JI, Dewey KG, Mills DA & Medzhitov RM The human gut microbiota and undernutrition. Science translational medicine 4, 137ps112, doi:10.1126/scitranslmed.3004347 (2012).
40. Harris RA et al. Colonic Mucosal Epigenome and Microbiome Development in Children and Adolescents. J Immunol Res 2016, 9170162, doi:10.1155/2016/9170162 (2016). [PubMed: 27006956]
41. Hollister EB et al. Structure and function of the healthy pre-adolescent pediatric gut microbiome. Microbiome 3, 36, doi:10.1186/s40168-015-0101-x (2015). [PubMed: 26306392]
42. Assa A et al. Mucosa-Associated Ileal Microbiota in New-Onset Pediatric Crohn's Disease. Inflammatory bowel diseases 22, 1533–1539, doi:10.1097/MIB.0000000000000776 (2016). [PubMed: 27271491]
43. Noble CL et al. Characterization of intestinal gene expression profiles in Crohn's disease by genome-wide microarray analysis. Inflammatory bowel diseases 16, 1717–1728, doi:10.1002/ibd.21263 (2010). [PubMed: 20848455]
44. Arijs I et al. Predictive value of epithelial gene expression profiles for response to infliximab in Crohn's disease. Inflammatory bowel diseases 16, 2090–2098, doi:10.1002/ibd.21301 (2010). [PubMed: 20848504]
45. Renz H, Brandtzaeg P & Hornef M The impact of perinatal immune development on mucosal homeostasis and chronic inflammation. Nature reviews. Immunology 12, 9–23, doi:10.1038/nri3112 (2012).

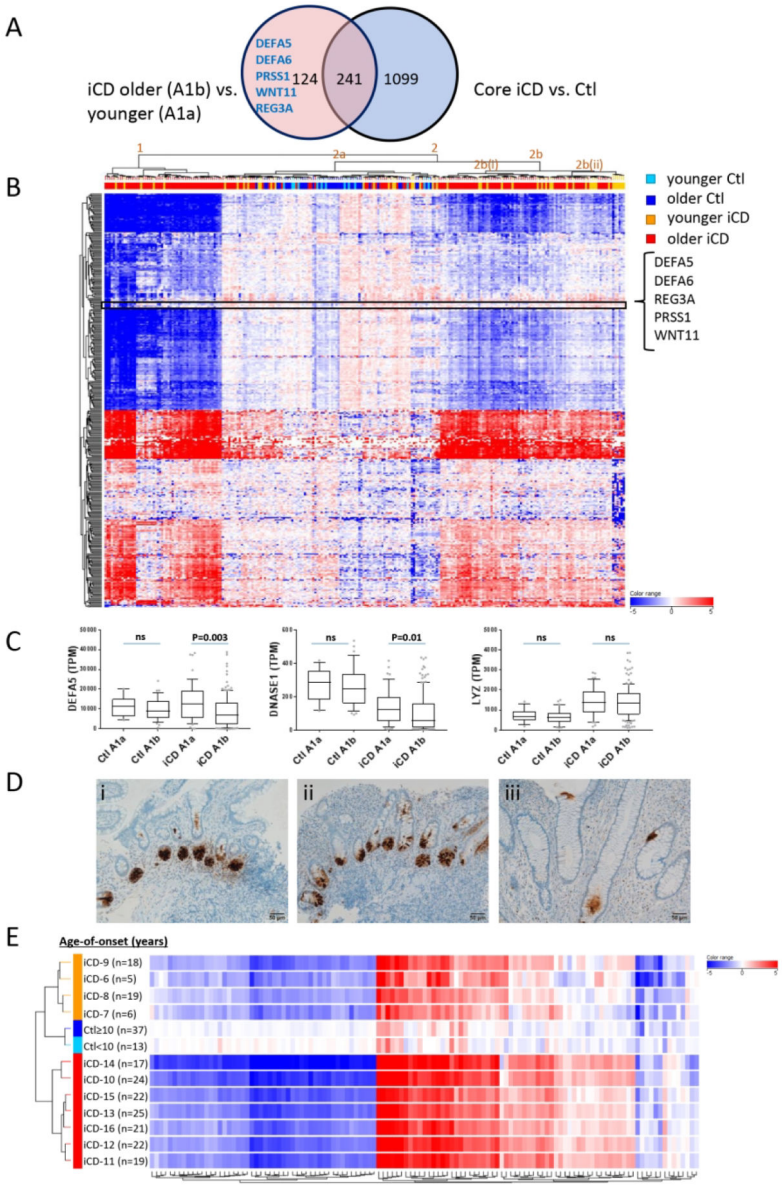

Figure 1. Decreased epithelial Paneth cell α-defensins signature in pediatric iCD ≥ 10 years at diagnosis.

(A) Venn diagram shows an overlap between the previously reported 1340 core iCD gene signature [19] and the 365 genes that were differentially expressed in the ileum between older (A1b) and younger (A1a) iCD (FDR correction [0.05], fold-change≥1.5). 124 of these 365 genes including α-defensins and *REG3A* were not included in the core iCD gene list. (B) Unsupervised hierarchical clustered heat map of the 365 genes differentially expressed genes between older and younger iCD with up-regulated genes in red and down-regulated genes in blue. Above the heat map, younger Ctl (light blue), older Ctl (dark blue), younger CD (orange), and older CD (red) samples are indicated. (C) TPM ileal gene expression is shown for the indicated genes for the indicated groups stratified by age-of-diagnosis with Kruskal-Wallis test with Dunn's multiple comparisons. (D) Immunohistochemistry in representative CD patients and non-IBD controls. Relatively high *DEFA5* staining is shown for older non-

IBD control (i) and younger CD (ii) that correlated with transcripts per million (TPM) values of >10,000 and 6143 for gene expression by RNASeq, respectively, and relatively low *DEFA5* staining in older CD that correlated with transcripts per million (TPM) values of 1239 for gene expression by RNASeq. Images were captured using an Olympus BX51 light microscope and digitally recorded at 20x magnification. (E) Averaged and unsupervised hierarchical clustering heat map of 121 of 365 (top third differentially expressed genes between older and younger iCD) stratified by yearly age-of-diagnosis as indicated. Number of patients for each age group is indicated adjacent to the heat map.

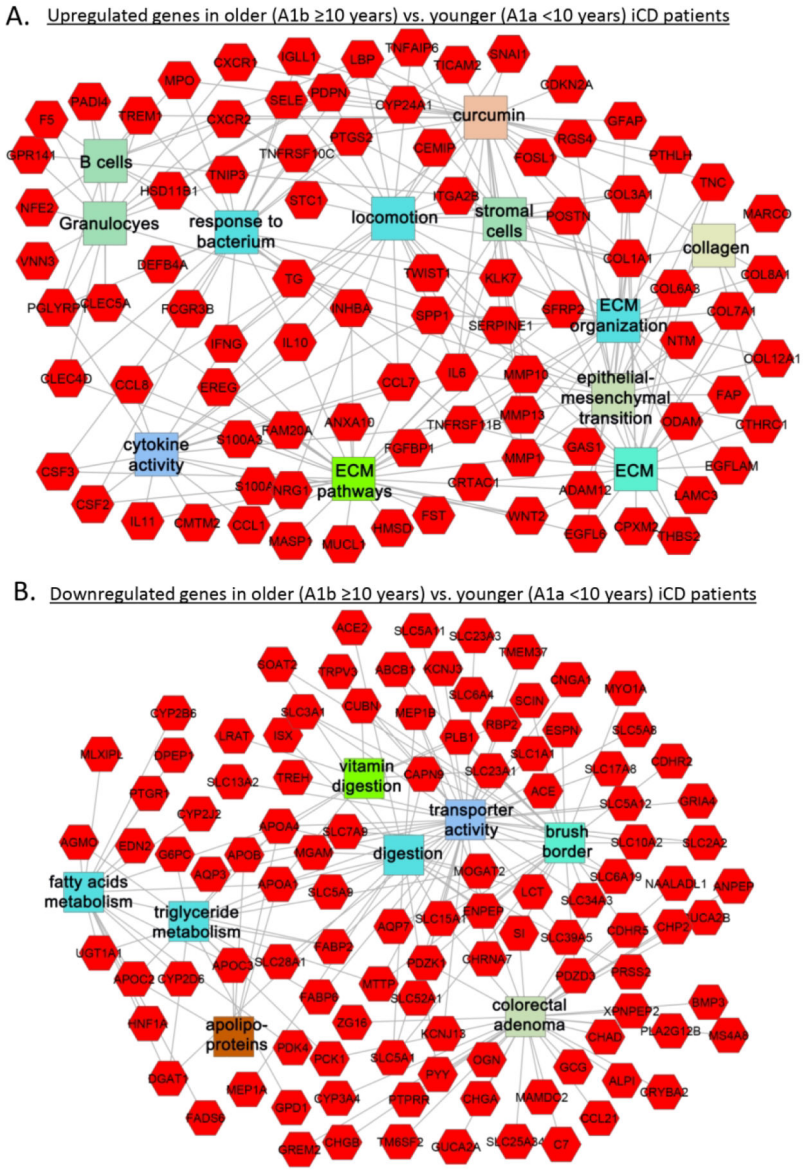

Figure 2. Enhanced ileal innate and adaptive Th1 immune responses in pediatric iCD ≥10 years at diagnosis.

(A) Top functional annotation enrichment analyses using Toppgene/ToppCluster [25] platforms of the 171 up-regulated and differentially expressed genes between older and younger iCD and Cytoscape [26] was used for visualization. (B) Top functional annotation enrichment analyses using Toppgene/ToppCluster [25] platforms using the 194 down-regulated and differentially expressed genes between older and younger iCD and Cytoscape [26] was used for visualization.

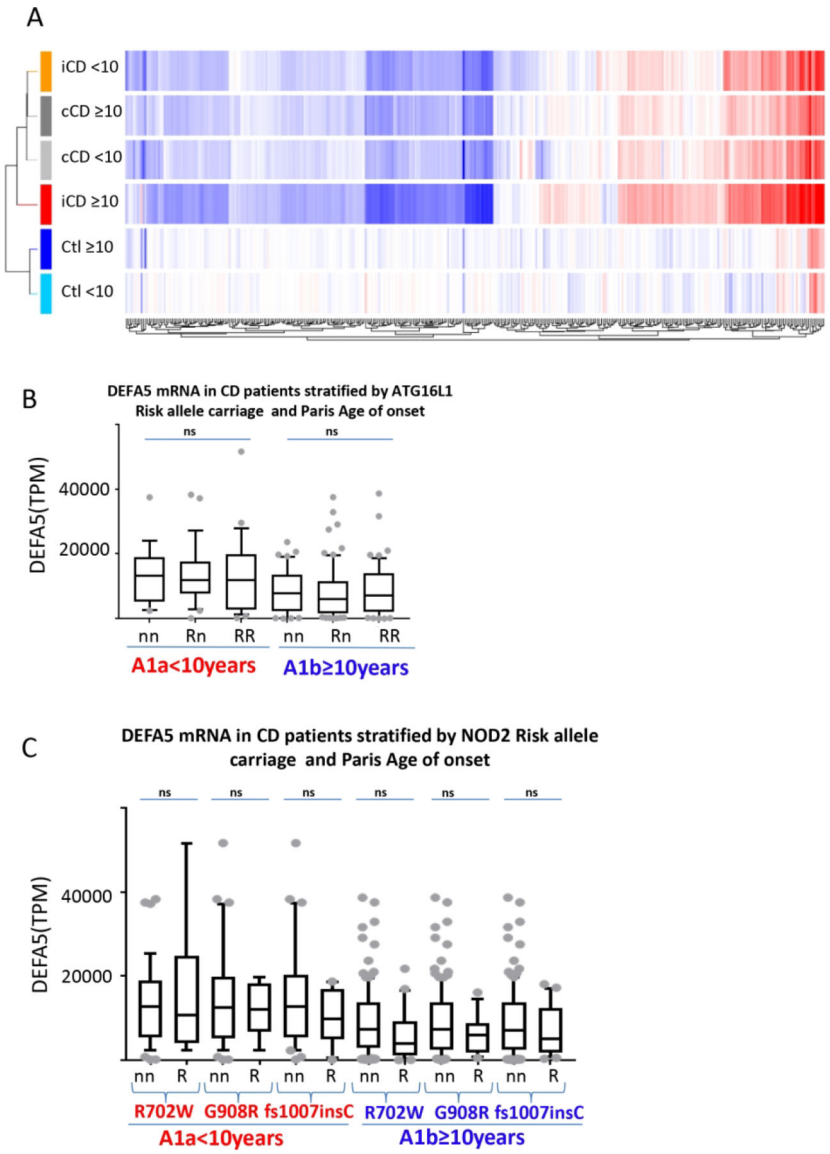

Figure 3. Decreased epithelial Paneth cell signature in pediatric CD ≥10 years at diagnosis is associated with clinical ileal inflammation.

(A) Unsupervised averaged hierarchical clustered genes heat map of 365 genes differentially expressed between older and younger iCD is shown for each clinical sub-group [Ctl, cCD and iCD divided to <10 and ≥10 years]. (B & C) DEFA5 TPM ileal gene expression is shown for the indicated younger and older CD age-of-diagnosis groups as indicated, stratified by their ATG16L1 (B) or NOD2 (C) genotype. R; risk allele (homozygote or heterozygotes), nn; no risk allele, nR; heterozygote for the risk allele. RR; homozygote for the risk allele.

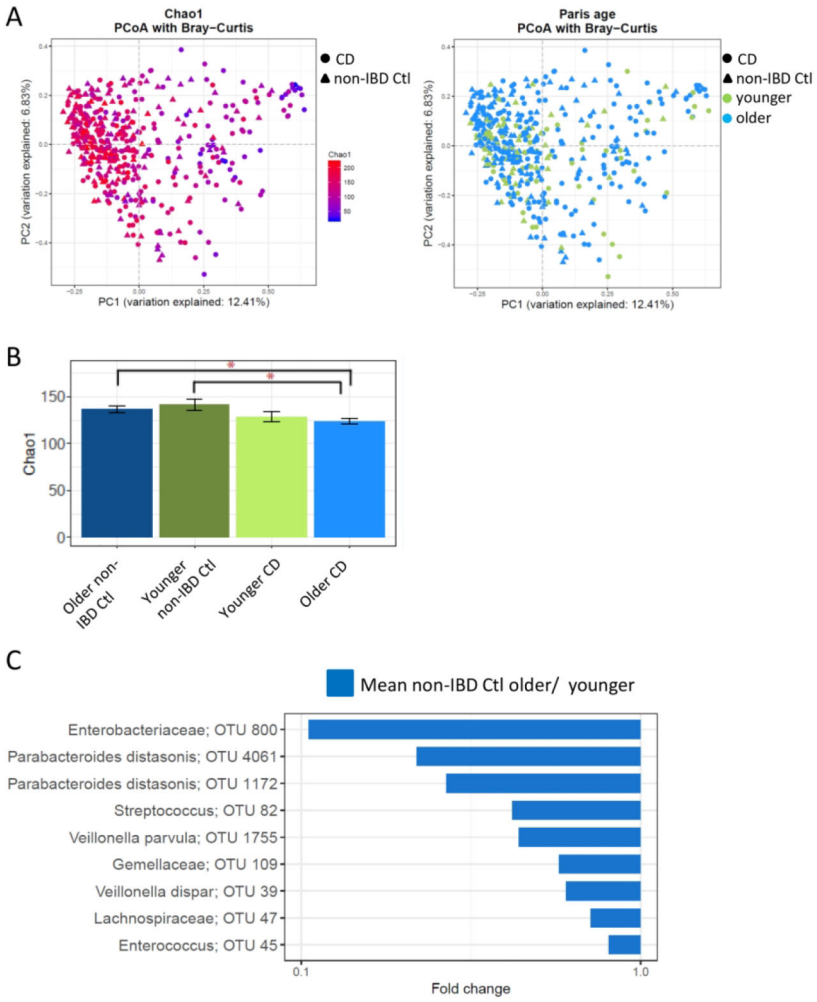

Figure 4. Age-associated shifts in the ileal microbial community composition detected in non-IBD controls are not present within CD.

(A) PCoA with Bray−Curtis distance comparing microbial community diversity in samples from CD patients (n=272) and Ctl (n=178). Left panel, samples are colored by the Chao1 diversity index. Right panel, samples are colored by <10 and ≥10 years. Triangular shape indicates Ctl, filled circles indicate CD samples. (B) Mean and standard deviation of Chao1 α-diversity is shown for Ctl and CD divided in younger <10 and older ≥10 years. *p < 0.01, a two-sided t-test was used. (C) The bar graph shows fold change (mean older A1b Ctl/mean younger Ctl) for significant associations between the indicated taxa as determined by MaAsLin while taking Paris age, gender, and body mass index (BMI) into account.

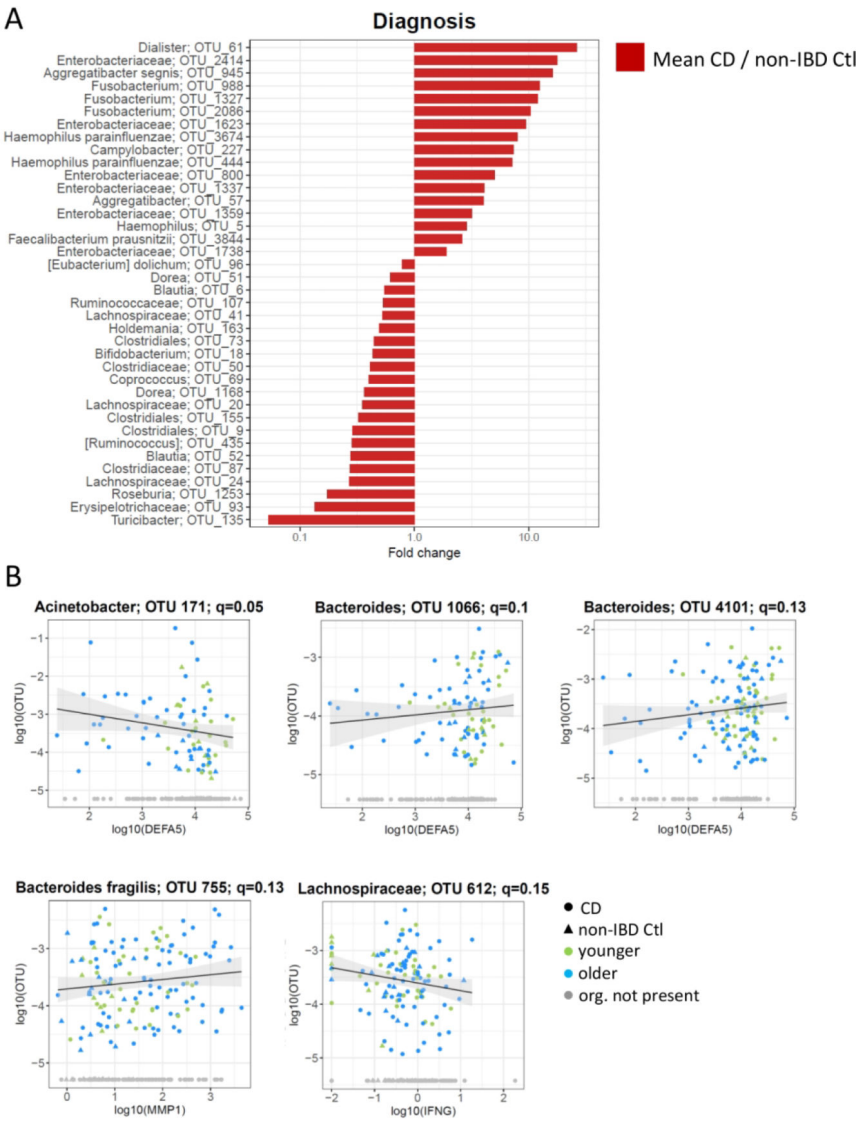

Figure 5. Co-variation of the ileal microbial community structure with host gene expression.

(A) The bar graph shows fold change (mean CD/mean Ctl) for significant associations between the indicated taxa and clinical phenotype (Ctl, CD) as determined by MaAsLin while taking Paris age, gender, body mass index (BMI, as a measure of nutritional status), endoscopic severity (deep ulcers in ileum), clinical severity (Pediatric Crohn's Disease Activity Index, PCDAI), antibiotics, ileal gene expression of *CSF2, CXCR1, IFNG, MMP3, DEFA5, GSTA1,* and *LCT,* and *NOD2*, and *ATG16L1* IBD risk allele carriage into account. (B) Scatter plots are shown for significant associations between the indicated taxa (*y-axis*) and host gene expression (*x-axis*) based on the multivariate statistical analysis described in A.

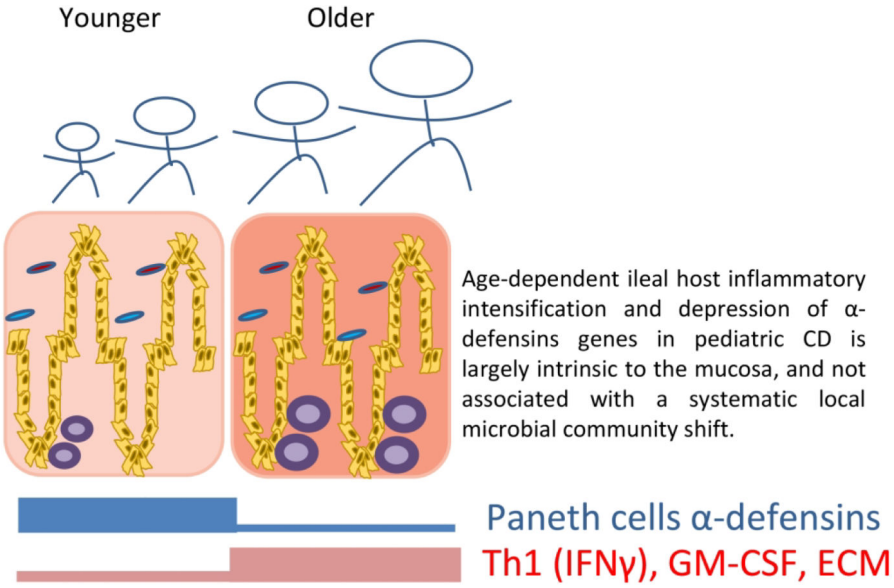

Figure 6. Summarizing cartoon.

Using high throughput mRNA and 16S rRNA amplicon sequencing of treatment naïve newly diagnosed ileal biopsies from CD cases we identify 365 genes that were robustly differentially expressed between older iCD (10–16 years) and younger iCD (<10 years). Those differentially expressed genes showed increased Th1-related IFNγ profile associated with amplified innate myeloid inflammatory activation, and enhanced extra cellular matrix and collagen signatures in older iCD cases. Remarkably, these signatures for enhanced immune activation was associated with a specific reduction in epithelial Paneth cell α-defensin expression in older iCD, and was not associated with systematic changes in the local mucosal microbial communities.

Table 1:

Clinical and Demographic Characteristics stratified by age-of-diagnosis of the RISK cohort

	VEO CD A1a (0-<6years) N=36	EO younger CD (A1a, 6-<10years) N=196	EO older CD(A1b,≥10years) N=714	P valueVEO A1a vs. EO A1b	P valueEO A1a vs. EOA1b
Mean (SD) Age (years)	4.7(1)	8.6(1)	13.5(2)		
Male gender %(n)	67%(24)	59%(116)	62%(446)	0.51	0.45
Location at diagnosis					
Location L1%(n)	19%(7)	16%(32)	22%(155)	0.83	0.12
Location L2 %(n)	31%(11)	24%(46)	23%(168)	0.49	0.98
Location L3 %(n)	50%(18)	60%(118)	55%(391)	0.34	0.2
Perianal involvement (%)	11%(4)	9%(18)	16%(112)	0.96	**0.02**
PCDAI at diagnosis					
<=10 [inactive, %(n)]	19%(7)	12%(23)	12%(88)	0.32	0.92
11 to 30 [mild, %(n)]	39%(14)	40%(78)	42%(301)	0.92	0.61
>30 [moderate-severe, %,(n)]	42%(15)	48%(95)	46%(325	0.57	0.51
Baseline diagnostic serology					
Positive anti-GMCSF (>1. 6mcg/mL)	22%(8)	38%(75)	50%(342)	**0.003**	**0.016**
Positive ASCA IgA (>20 EU/ml)	8%(3)	19%(37)	26%(188)	**0.016**	**0.03**
Positive ASCA IgG (>40 EU/ml)	6%(2)	16%(32)	26%(188)	**0.005**	**0.004**
Positive anti CBir1 (>25 EU/ml)	56%(20)	39%(76)	36%(257)	**0.002**	**0.53**
Positive pANCA (>30 EU/ml)	8%(3)	13%(25)	16%(115)	0.59	0.27
Therapy within 180 days					
Neither IM nor anti-TNFα	56%(20)	28%(54)	28%(199)	**0.004**	0.929
IM	36%(13)	52%(102)	43%(304)	0.444	**0.018**
Anti-TNFα	3%(1)	13%(26)	19%(139)	**0.012**	0.46
Both IM & anti-TNFα	6%(2)	7%(14)	10%(72)	0.374	0.213
PCDAI 6 month after Diagnosis	N=33	N=187	N=678		
<=10 (clinical remission)	45%(15)	60%(113)	69%(465)	**0.004**	0.054
Complications During Follow up					
B1 (inflammatory) to B2 (stricturing) within 3 years of F/U	8.3%(3)	4.1%(8)	7.3%(52)	0.33	0.11
B1 (inflammatory) to B3 (penetrating) within 3 years of F/U	0%(0)	0.5%(1)	4.6%(33)	0.67	**0.007**

VEO: Very early onset (<6 years); EO: Early onset (6 to <17years); PCDAI: Pediatric Crohn Disease Activity Index at diagnosis prior to treatment; L1: ileal location; L2: colon-only location; L3: ileo-colonic location; IM: Immunomodulators. B1: inflammatory behavior; B2: stricturing behavior: B3: penetrating behavior. pANCA: perinuclear anti-neutrophil cytoplasmic antibodies. ASCA: anti-Saccharomyces cerevisiae antibodies.

Table 2:

Clinical and Demographic Characteristics

	Ctl n=50	cCD n=56	iCD n=198	iCD A1a n=48	iCD A1b n=150
Mean (SD) Age (years)	11.8(3)	12.4(3)	12.2(3)	8.6(1)	13.4(2)
Male gender (%)	62	59	63	67	59
Location L1 (%)	-	-	28	15*	33
Perianal involvement (%)	-	21	16	10	17
Ileal deep ulcers (%)	-	-	45	45	46
PCDAI at diagnosis (n)		n=52	n=197	n=47	n= 150
<=10 (inactive, %)	-	6	9	11	8
11 to 30 (mild, %)	-	48	52	46	53
>30 (moderate-severe, %)	-	46	39	42	39

PCDAI: Pediatric Crohn Disease Activity Index at diagnosis prior to treatment; L1: ileal location. Younger vs older

*p=0.015