Published in final edited form as:

J Pediatr. 2015 July ; 167(1): 138–47.e1-3. doi:10.1016/j.jpeds.2015.02.049.

Associations between gut microbial colonization in early life and respiratory outcomes in cystic fibrosis

Anne G. Hoen, PhD^{a,b,*}, Jing Li^{a,*}, Lisa A. Moulton, RN^c, George A. O'Toole, PhD^d, Molly L. Housman, MS^d, Devin C. Koestler, PhD^e, Margaret F. Guill, MD^c, Jason H. Moore, PhD^a, Patricia L. Hibberd, MD, PhD^f, Hilary G. Morrison, PhD^g, Mitchell L. Sogin, PhD^g, Margaret R. Karagas, PhD^b, and Juliette C. Madan, MD, MS^h

^aComputational Genetics Laboratory, Institute for Quantitative Biomedical Sciences, Geisel School of Medicine at Dartmouth, Dartmouth College, Hanover, NH 03755

^bDepartment of Epidemiology, Geisel School of Medicine at Dartmouth, Dartmouth College, Hanover, NH 03755

^cDivision of Allergy and Pediatric Pulmonology, Department of Pediatrics, Dartmouth Hitchcock Medical Center, Lebanon NH 03756

^dDepartment of Microbiology and Immunology, Geisel School of Medicine at Dartmouth, Hanover, NH 03755

^eDepartment of Biostatistics, University of Kansas Medical Center, Kansas City, KS 66160

^fDivision of Global Health, Department of Pediatrics, Massachusetts General Hospital, Boston, MA 02114

^gJosephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA 02543

^hDivision of Neonatology, Department of Pediatrics, Dartmouth Hitchcock Medical Center, Lebanon, NH 03756

Abstract

Objective—To examine patterns of microbial colonization of the respiratory and intestinal tracts in early life in infants with cystic fibrosis (CF) and their associations with breastfeeding and clinical outcomes.

Corresponding author: Juliette C. Madan, MD, MS., Division of Neonatology, Department of Pediatrics, Children's Hospital at Dartmouth, 1 Medical Center Drive, Lebanon, NH 03756, Phone: 603-653-6063, Fax: 603-653-6135, Juliette.C.Madan@Dartmouth.Edu.

*Contributed equally

Requests for reprints may be sent to Juliette C. Madan, MD, MS (Juliette.C.Madan@Dartmouth.Edu)

Portions of the study were presented as an abstract at the North American Cystic Fibrosis Conference, Salt Lake City, UT, October 17–19, 2013.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Study design—A comprehensive, prospective longitudinal analysis of the upper respiratory and intestinal microbiota in a cohort of infants and young children with CF followed from birth was performed. Genus-level microbial community composition was characterized using 16S-targeted pyrosequencing, and relationships with exposures and outcomes were assessed using linear mixed-effects models, time-to-event analysis and principal components analysis.

Results—Sequencing of 120 samples from 13 subjects collected from birth to 34 months revealed relationships between breastfeeding, microbial diversity in the respiratory and intestinal tracts and the timing of onset of respiratory complications, including exacerbations and colonization with *Pseudomonas aeruginosa*. Fluctuations in the abundance of specific bacterial taxa preceeded clinical outcomes, including a significant decrease in bacteria of the genus *Parabacteroides* within the intestinal tract prior to the onset of chronic *P. aeruginosa* colonization. Specific assemblages of bacteria in intestinal samples, but not respiratory samples, were associated with CF exacerbation in early life, indicating that the intestinal microbiome may play a role in lung health.

Conclusion—Our findings relating breastfeeding to respiratory outcomes, gut diversity to prolonged periods of health, and specific bacterial communities in the gut prior to respiratory complications in CF highlight a connection between the intestinal microbiome and health and point to potential opportunities for antibiotic or probiotic interventions. Further studies in larger cohorts validating these findings are needed.

Keywords

microbiome; infants; early childhood; infection; *Pseudomonas aeruginosa*

Cystic fibrosis (CF) is the most common life-limiting autosomal recessive genetic disorder among people of European descent. It is characterized by mutations in the CF transmembrane conductance regulator (*CFTR*) gene, which result in viscous epithelial secretions beginning at birth as a consequence of abnormal sodium and chloride transport. Individuals with CF experience progressive lung disease, pancreatic insufficiency and profound impacts on growth and nutrition (1, 2). Infants and young children with CF are at risk for chronic infection and inflammation, resulting in significant morbidity and early mortality (3, 4). In both animal models and in observational studies of humans, CF has been associated with atypical microbial colonization of the intestinal and the respiratory tracts, findings that are attributable to loss of *CFTR* function and the resulting altered microenvironments (4–10). The natural history of microbial acquisition in the respiratory and intestinal tracts in patients with CF beginning at birth and the impact of these communities on clinical outcomes is largely unexplored.

Microbial colonization patterns in infancy are influenced by environmental exposures (including delivery mode, infant diet, hospitalizations and medications), human genetics, and immune function (11–13). Microbial colonization of the respiratory tract of infants with CF has been shown to be significantly influenced by dietary exposure to breast milk, and colonization of the gut by specific assemblages of microbes appears to precede colonization of the lungs, highlighting potential interactions between nutrition, intestinal colonization and respiratory outcomes (3). Seminal studies in germ free animals highlight the importance of

gut microbial colonization for immune programming in the neonatal period, ultimately affecting lifelong systemic disease risk (14, 15). The aberrant respiratory and gastrointestinal microbial colonization patterns in young children with CF (1, 4, 16) and the success of probiotics trials that have demonstrated benefits of altering the gut microbiome in an effort to ameliorate risk of respiratory compromise in CF (17, 18) both point to a knowledge gap in our understanding of the interactions between intestinal microbial colonization and CF disease progression in early life.

The purpose of this study was to investigate the hypothesis that microbial acquisition patterns relate to risk of CF-related complications and clinical markers of CF disease progression. Associations between intestinal microflora composition and clinical outcomes in young children with CF may ultimately help to inform nutritional and probiotic treatment strategies that ameliorate colonization with pathogens, maintain a more health-promoting microbiota, and decrease morbidity and mortality.

Methods

Institutional review board approval was obtained in April 2010 (Center for the Protection of Human Subjects at Dartmouth number 21761) with yearly renewal of approval in 2011 through 2013, and parents of subjects provided written informed consent.

Eligibility criteria included: diagnosis with CF prenatally or postnatally based upon newborn screening results and subsequent confirmation with sweat chloride and genetic testing. Subjects were eligible if their care for CF would be at the Dartmouth-Hitchcock Medical Center in Lebanon, NH or Manchester, NH affiliated clinics. Enrollment occurred prior to 1 month of life with follow up prospectively at routine CF clinic visits which facilitated collection of detailed clinical information as well as samples every 3 months until the maximum age of 34 months for this data analysis. Enrollment and prospective collection is ongoing. Subjects were excluded if they had non-CF chromosomal anomalies.

Detailed clinical information was collected prospectively and included demographics, birth history, medical history at each visit both with parental interviews and medical record review, growth measurements, outcome data including information regarding outcomes of interest: CF pulmonary exacerbations (as diagnosed by an attending physician using the Akron Children's CF exacerbation score if appropriate (19), and defined by findings including wheezing, cough lasting for more than 3 days, shortness of breath, weight loss, decrease in oxygen saturations below 95%, and treatment with antibiotics), growth failure (defined as height, weight, and HC <10th percentile or a loss of >10% in growth measurements between visits), hospitalization for any CF-related complication and onset of *P. aeruginosa* colonization (as identified during routine clinic oropharyngeal culture, requiring therapy with antibiotics). Clinical culture results were collected prospectively, along with medical interventions and exposures including medications, supplements, dietary and environmental histories. Detailed antibiotic exposure data were collected and used to adjust linear mixed effects models of changes in the microbiome development over time.

Oropharyngeal and stool samples were collected at regularly scheduled CF clinic visits which occurred every 3 months beginning in the first month of life for up to 3 years. Oropharyngeal swabs were collected by doubling specimen collection swabs for routine clinical surveillance oropharyngeal culture, which allowed for the collection of a research specimen simultaneously with the routine clinical specimen with minimal additional burden for participating children. Research swabs were placed in individual 2 ml microcentrifuge tubes, stored at -20°C and transported to the lab on dry ice.

Respiratory swabs were submerged in 2 mL RNeasy lysis buffer (Qiagen) and vortexed for one minute. Swabs were removed with sterile forceps to a clean tube where they were placed cotton side up and centrifuged for 2 minutes at $14,000 \times g$ in order to recover fluid from the absorbent material of the swab, and any recovered fluid was combined with the rest of the sample. Swabs were discarded, and samples were frozen at -80°C until further processing. Bacterial DNA was extracted using the MoBio PowerSoil bacterial DNA isolation kit according to the manufacturer's instructions. Stool samples were collected by parents at home using a sterile wooden spatula and placed in a sterile collection cup in the freezer. 50 mg aliquots of stool samples were stored in sterile tubes with 1 mL RNeasy lysis buffer and frozen at -80°C . Thawed stool samples were dissolved in phosphate buffered saline, followed by DNA extraction as above.

DNA extracts were used to construct PCR amplicon libraries for 454 pyrosequencing targeting the V4–V6 regions of the bacterial 16S rRNA gene, performed at the Josephine Bay Paul Center and the Marine Biological Laboratory (3). Titanium amplicon sequencing initially targeted the 16S V6 hypervariable region and ultimately the V4–V6 regions. These two approaches have been assessed for their similarities and were not statistically significantly different (3), therefore the data were pooled for the analyses presented here. A mix of multiple fusion primers containing degeneracies in the 16S-specific domain designed to capture the range of diversity of rRNA sequences represented in molecular databases (20, 21) were used. Primer sequences were: 967FPP: 5'-CNACGCGAAGAACCTTANC, 967FAQ: 5' CTAACCGANGAACCTYACC, 967FU2: 5'-CAACGCGMARAACCTTACC, 967FU3: 5'-ATACGCGARGAACCTTACC, 518F: 5'-CCAGCAGCYGCGGTAAN and 1064R: 5'-CGACRRCCATGCANCACT.

Samples were amplified in triplicate and no-template controls were included for each primer set. Positive PCR controls composed of a mock community DNA preparation originally developed for the Human Microbiome Project (BEI Resources) were also included. Our quality control procedures required exact matches to the barcode and to the proximal primer and eliminated any sequence that contained an ambiguous base call at any position. Distal primers were located using the EMBOSS “fuzznuc” pattern matching program. Primer sequences were trimmed from the ends of high quality reads before analysis. UChime (22) was used to remove chimeras using both the reference mode against the gold reference set (23) and the de novo mode. We aimed for ~10,000 sequencing reads per sample; the mean number of reads that passed QC was 15,847.

Taxonomic identifiers were assigned to pyrotags by using the rRNA indexing algorithm Global Assignment of Sequence Taxonomy (GAST) (24), which compares pyrotags to

known rRNA genes that have already been placed in a phylogenetic framework of more than 1,000,000 nearly full-length rRNA reference sequences (RefSSU) on the SILVA database (25). The tag mapping methodology GAST for taxonomic assignments of environmental V6 pyrotags and the V4V6 or V6 reference databases are freely available through the VAMPS (Visualization and Analysis of Microbial Population Structure) website <http://vamps.mbl.edu/>. Operational taxonomical units defined at a 3% sequence difference cutoff were treated as the terminal taxonomical rank. A minority of sequences were assigned a terminal rank at the species level; therefore our analysis was focused toward comparisons at the genus level, for which there was considerably better sequence specificity (3). All sequence data has been deposited in the GenBank Sequence Read Archive under project number PRJNA170783.

Statistical analyses

Statistical analyses were aimed at examining the assemblages of bacterial genera and the temporal changes in intestinal and respiratory microbiome diversity associated with respiratory outcomes in developing infants with CF. The relative abundance of each genus in each sample was estimated by normalizing genus-specific sequencing read counts by the total number of reads for a given sample and rescaling by the mean total read count for all samples. Hierarchical clustering using the Euclidean distance metric and average linkage was used to cluster samples based on the $\log_{10}$ relative abundance for each bacterial taxa. In this and further analyses, the constant 0.1, which is an order of magnitude smaller than the smallest non-zero value, was added to all values prior to log transforming. Only bacterial genera present in 10% of samples were included in this cluster analysis.

Linear-mixed effects models were used to investigate changes in bacterial abundance over time. The use of linear-mixed effects models facilitates the examination of the linear relationship between two variables when one or more of the subjects being examined contain repeated measurements. We modeled, individually for each genus, $\log_{10}$ relative abundance as the response, with fixed effects terms including main effects for time (time at which the sample was collected) and whether the sample was taken before or after the first onset of each clinical outcome (CF exacerbation or *P. aeruginosa* colonization) as well as their interaction. Because the administration of antibiotics may have long term effects on microbiome composition (26), we included a binary fixed effects variable that indicated for each sample whether antibiotics had been administered prior to its collection. Additionally, we included a random effect for subject to account for within-subject repeated measures. Genera with relative abundance estimates of 0 in >90% of subjects were excluded from this analysis. We corrected for multiple comparisons by controlling the false discovery rate (27). This is a more powerful approach than conventional “family-wise” methods such as the conservative Bonferroni correction. For each hypothesis test we calculated a q-value, the false discovery rate analogue to the p-value and considered all variables associated with q-values under a threshold of 0.1 to be significant. For ease of interpretation, we report p-values, but highlight taxa that were significant only after multiple comparison correction in red.

Principal components analysis (PCA) was used to identify assemblages of bacterial genera in the stool and respiratory microbiomes that were associated with initial diagnoses of CF-related outcomes of interest. We applied a centered log-ratio transformation to the relative abundance estimates prior to downstream calculations in order to account for the dependence between variables inherent in compositional data (28). Because the length of follow up was variable between subjects, and because we were interested in early-life outcomes, we excluded from this analysis all samples collected after six months of age. We were concerned with microbiome composition prior to the onset of the outcomes of interest; therefore, we evaluated the bacterial genera in samples collected before the observation of the outcome in subjects who experienced the outcome within the first six months of life and all samples up to six months of age from subjects who did not. Taxon-specific means of transformed relative abundance estimates over the time points considered were computed and used to build principal components. Intestinal and respiratory microbiome samples were evaluated in separate models. Finally, the first two principal components from each analysis were used in logistic regression models to predict the presence or absence of each outcome. Concerns about over fitting precluded the use of more than two principal components in logistic regression models. Reported p-values were calculated using a χ^2 test of overall model significance.

Diversity was evaluated using the Simpson diversity index (SDI) (29). For the i^{th} sample,

the SDI was recorded as $1-D_i$, where $D_i = \sum_{j=1}^{N_i} \frac{n_{ij}(n_{ij}-1)}{N_i(N_i-1)}$ and where n_{ij} represents the abundance of the j^{th} genus for the i^{th} sample and N_i represents the total number of genera for the i^{th} sample. Thus, the SDI is approximately continuously distributed between 0 and 1, with values approaching 1, signifying greater microbial diversity. For examining changes in microbial diversity over time the SDI was arcsine square-root (30) transformed to satisfy the assumption of normality.

Time-to-event analyses were performed using the time to first clinical diagnosis of each outcome of interest as end points. Preliminary analyses revealed a relationship between breastfeeding and time to pulmonary exacerbation; therefore, models examining relationships between SDI and clinical end points with adjustment for breast milk exposure were explored. SDI was examined as a time-varying covariate in time-to-event analyses. Samples collected following the first clinical diagnosis of each outcome were not considered in this analysis. Significance of relationships was determined using the log rank test.

All analyses were implemented in R (<http://cran.r-project.org/>) version 3.0.

Results

A core microbiome bridges both organ systems

The gastrointestinal and upper respiratory microbiota of thirteen subjects (with the exception of one subject who declined to provide stool samples) with CF were followed prospectively from time of CF diagnosis through 6 to 34 months of age (Table; available at www.jpeds.com), and a total of 120 samples were analyzed: 66 intestinal and 54 respiratory. Hierarchical clustering of logged relative abundance of microbial genera identified a group

of bacteria exhibiting high relative abundance in both the intestinal and upper respiratory tracts and spanning the age range of the sampling. Bacteria in this core included: *Veillonella*, *Streptococcus*, *Bifidobacterium* and *Bacteroides*, genera that are identified in other studies in humans as being associated with a core microbiome (31, 32) (Figure 1). A second cluster of bacteria dominated in respiratory samples that tended to be derived from later time points; samples in this cluster were collected at a median subject age of 12 months, and subjects outside of this cluster were collected at a median age of 6 months (Figure 1).

Changes in bacterial abundance over time herald the onset of clinical outcomes

Two critical outcomes impacting long-term health in CF are the onset of CF pulmonary exacerbations and onset of *P. aeruginosa* colonization, which can lead to chronic respiratory infections. Shifts in the abundance of individual genera preceding the onset of both CF exacerbations and *P. aeruginosa* colonization were observed (Figures 2 and 3; Figure 3 available at www.jpeds.com); however, after correcting for multiple comparisons, only changes observed prior to the onset of *P. aeruginosa* colonization were significant. In intestinal samples there was a highly significant decrease in *Parabacteroides* ($p < 0.001$; Figure 2, A) before initial colonization with *P. aeruginosa*. In the respiratory tract we observed a corresponding significant increase in the relative abundance of *Salmonella* ($p = 0.002$; Figure 2, B), an enteric genus associated with pathogenicity in humans (33). Although not significant after correction for multiple testing, we also observed decreases in both *Bacteroides* and *Bifidobacterium* in intestinal samples prior to both first CF exacerbation and initial *P. aeruginosa* colonization. These two taxa have been commonly associated with the intestinal microbiota in infancy and have been linked to risk for development of asthma (34–36). *Bacteroides* has also been associated with enhanced systemic mucosal immunity (37, 38). Overall, the microbial shifts preceding initial clinical outcomes in young children with CF highlight decreases in immune-mediating bacteria in the gut, with a pathogen predominance particularly with gram negative facultative anaerobes in both organ systems.

Gut bacterial assemblages predict clinical outcomes in early life in CF

PCA of intestinal and respiratory samples revealed significant associations between CF exacerbation and gut microbiome composition ($p = 0.03$) but not respiratory microbiome composition ($p = 0.33$; Figure 4, A and B). *P. aeruginosa* colonization was not associated with microbiome composition according to this analysis (Figure 5; available at www.jpeds.com).

Breast milk feeding, microbial diversity and time to clinical complications

Previous analysis of the microbiome in a subset of this cohort revealed any breast milk exposure (defined as exclusive breast milk feeding in infancy, or breast milk in addition to formula supplementation) to be an important determinant of microbial diversity in the respiratory tract (3). Analyses of the full cohort with longer follow up revealed a striking trend toward prolonged time to first CF exacerbation among subjects exposed to breast milk ($p = 0.06$; Figure 6).

Therefore, we adjusted for breast milk exposure in further analyses using a time-to-event approach with microbial diversity as a time-varying covariate to analyze breastfeeding-adjusted relationships between diversity and both CF exacerbation and *P. aeruginosa* colonization. Though not statistically significant, we observed relationships between microbial diversity and time to first clinical complication. Namely, subjects with increased microbial diversity in their intestinal tracts tended to experience longer time to both CF exacerbation ($p=0.12$) and *P. aeruginosa* colonization ($p=0.62$; Figure 7, A and B available at www.jpeds.com). Interestingly, in the respiratory microbiota, the reverse pattern was found: higher microbial diversity in the respiratory tract was non-significantly associated with the shortest time to clinical outcomes ($p=0.20$ and 0.85 ; Figure 7, C and D).

Discussion

We have described in detail the developing microbiome in CF, and identified important relationships between exposure to breast milk, gut microbiome composition and diversity, and respiratory health. We found that a core microbiota bridges both organ systems that persists from birth to as long as 34 months. The core is populated by common human microbial taxa that have been identified in other culture-dependent studies of patients with CF (39), as well as in a culture-independent, earlier evaluation of our cohort (3). Many of the organisms identified throughout the repeat sampling in both organ systems are anaerobes, which are expected in infants' intestines after the initial colonization with facultative anaerobes (40, 41), and in the CF lung, where anaerobic bacteria predominate (39, 42, 43). A cross-sectional analysis of individuals from infancy throughout adulthood on 3 continents revealed the greatest inter-individual variability in the gut microbiota occurs in the first 3 years, after which time subjects had similarity to others in terms of microbial patterns (12). This variability is attributable to the fact that infants have a nearly sterile gut at birth, and are acquiring bacteria in the first weeks and months of life from environmental exposures including delivery mode, diet, medications, and their caregivers (11, 13, 44, 45). The variability of the microbial patterns in early life and its relationship to exposures, many of which are alterable, highlights opportunities for intervention, both nutritional and with medications.

Our study is limited by the small size of the cohort, however we were able to lay the groundwork for larger studies to better characterize the developing microbiome in infants and young children with CF and the interrelationship between the gut and respiratory tract in systemic disease progression. The use of oropharyngeal sampling to investigate microbial patterns in the upper respiratory tract in CF affords safe, non-invasive means of sampling microbes in non-sputum producing infants and young children, and is utilized as a standard practice in clinical care in CF to identify lung pathogens commonly found in disease progression in CF (16, 46). Studies evaluating the relationship between upper and lower respiratory sampling show a range of correlation but reveal the high sensitivity, specificity and negative predictive value of oropharyngeal sampling (46–48). The field of respiratory microbiome investigation is new, and links between diversity and health outcomes, or core microbial communities and disease, akin to the gut microbiome, are preliminary. One study identified a possible core airway microbiome in healthy adults which included *Pseudomonas*, *Streptococcus*, *Prevotella*, *Fusobacterium*, *Veillonella*, *Porphyromonas* and

Haemophilus (49) with slight variation in another cohort (50). We identified a respiratory core that persisted as long as 34 months, dominated by *Streptococcus* and *Veillonella*.

Of particular interest in our findings is the relationship between dietary exposures, particularly breast milk, and respiratory outcomes. We discovered that even after adjusting for exposure to breast milk, increased diversity of the gut microbiota is associated with prolonged periods of health, with delays in the time to initial *P. aeruginosa* colonization and first CF exacerbation. In many studies, increased diversity in the gut microbiome is associated with health, and a multitude of diseases have been associated with decreased diversity in disorders ranging from inflammatory bowel disease to obesity (51), and more recently with systemic diseases including respiratory disease such as asthma, COPD and CF (52, 53), linking the hygiene hypothesis and aberrations in the gut microbiota to immune-mediated systemic disease (54, 55). In respiratory samples, we observed a pattern of decreased diversity associated with prolonged time to CF exacerbation and *P. aeruginosa* colonization. This finding is despite our previous observation in the same cohort that both the intestinal and respiratory tracts exhibit comparable rates of increase in microbial diversity over time (3). In older populations of subjects with CF, decreased microbial diversity in the respiratory tract is associated with increased illness severity, however these findings are confounded by the relationship between increasing age and worsening lung function and infection in CF, as well as chronic antibiotic exposure (4, 56, 57).

Although we observed intriguing relationships between microbial diversity and the timing of respiratory complications, in our cohort these were not significant; however, we were able to identify specific groups of bacteria in the respiratory tract and the gut that predict the onset of respiratory complications, implicating, perhaps, a role for specific microbial taxa in CF disease progression. Even though it should be noted that there exists the potential for critical shifts in the abundance of certain rare bacterial genera that this analysis was underpowered to detect, our findings highlight a potential role for fluctuations in gut microbiota in immune development and respiratory health and point to opportunities for further larger studies and interventions in this population with dietary, altered antibiotic or probiotic regimens. We found an increase in the respiratory tract of *Salmonella*, a common human enteric pathogen (33), prior to onset of *P. aeruginosa* colonization. In the intestinal samples, we observed a highly significant decline in the abundance of *Parabacteroides* prior to initial *P. aeruginosa* colonization. Although not significant, we also observed decreases of two important gut colonizers, *Bacteroides* and *Bifidobacterium*, in stool samples prior to first CF exacerbation and initial *P. aeruginosa* colonization. *Parabacteroides* was in the past grouped with *Bacteroides*, and they are related organisms (59). *Bacteroides* is an early colonizer in neonatal life, and is associated with risk for development of asthma (34–36) as well as increased systemic mucosal immunity (37, 38). *Bifidobacterium* has been shown previously to be depleted in the intestinal tracts of CF subjects relative to their healthy siblings (53). *Bifidobacterium* is a dominant taxon in healthy, breast fed infants, and is reliant upon breast milk oligosaccharides and glycosylated proteins for the production of short chain fatty acids that protect against colonization with pathogens (60) and infection with pathogenic enteric bacteria including *Escherichia coli* (61). Thus, it may be that the decrease in these bacteria prior to first CF exacerbation in our cohort is related to systemic

immune-mediated changes in gut flora composition or metabolic products promoting the health of commensals, or inhibiting growth of pathogens.

Analyses of associations between bacterial communities and clinical complications in our cohort revealed statistically significant patterns in the gut microbiome relating to the onset of a pulmonary exacerbation in the first six months of life, whereas respiratory outcomes appear to be driven less by respiratory microbiome. The pulmonary immune response is thought to relate to commensal gut bacteria through mechanisms such as direct recognition of commensals in the airways by immune cells, microbial metabolites that are distributed systemically, and cell-mediated immune programming (52).

Our findings relating breast milk diet to respiratory outcomes, intestinal tract microbial diversity to prolonged periods of health, and bacterial communities colonizing the gut prior to respiratory complications in CF highlight a connection between the gut microbiome and systemic health. The neonatal period is a time of rapid acquisition of microbes in the gut, and in breast fed infants these microbes are initially facultative anaerobes, followed by anaerobes, including *Bifidobacteria*, whose growth is promoted by oligosaccharides in breast milk (11, 40, 44, 62). Competent evolution of the innate and adaptive immune system, beginning in infancy, has evolved to require interactions with the gut microbiome (63–69), and alterations in the gut microbiome are associated with increased risk of allergy/atopy, asthma, and obesity (70–73). The gut microbiota in CF has been demonstrated to be altered and less diverse with a paucity of health promoting bacteria, in murine models (74) and in humans (75, 76), even in the absence of antibiotic exposure. Several animal models have identified benefits of enteral probiotics and improvements in respiratory disease (77–79), and small randomized controlled trials of enteral probiotics, including *Lactobacillus* GG, have shown respiratory benefit in older children and young adults with CF (17, 18). Our findings suggest that the gut microbiome in CF in early life is a determinant of respiratory and systemic disease progression, affording opportunities to investigate nutritional and probiotic interventions to prolong a healthy microbiota and prevent morbidity in this important disease.

Acknowledgments

Supported by the CF Foundation and the Harry Shwachman Career Development Award, The Joshua Burnett Career Development Award through the Hitchcock Foundation, the Geisel School of Medicine at Dartmouth Department of Pediatrics (to J.M.), the Neukom Institute (to G.O. and J.M.), the CF Foundation Research Development Program (STANTO07R0 [to G.O.]) the National Institutes of Health (K01LM011985 [to A.H.], R01AI59694 [J.M.], R37AI83256-06 [G.O.], 4UH3DK083993 [M.S. and H.M.], K24AT003683 [to P.H.], P20GM104410 [to A.H., M.K., J.M.], P01ES022832 [to M.K.]), and EPA (RD-83459901 [to M.K.]).

We are deeply grateful to the children and families that made this study possible. We would also like to acknowledge the efforts of the nurses and staff in the Dartmouth-Hitchcock pulmonology clinics, as well as the pulmonology research nurses, including Lynn Feenan and Dana B. Dorman. We thank Tom Caldwell, MD, for database creation and management and Thomas Hampton for helpful discussions. We gratefully acknowledge the Dartmouth Translational Research Core, supported by the National Center for Research Resources (P20RR018787) and the National Institute of General Medical Sciences (P20GM103413), for collecting and processing clinical samples.

References

1. Emerson J, Rosenfeld M, McNamara S, Ramsey B, Gibson RL. *Pseudomonas aeruginosa* and other predictors of mortality and morbidity in young children with cystic fibrosis. *Pediatr Pulmonol*. 2002; 34:91–100. [PubMed: 12112774]
2. Bronstein MN, Sokol RJ, Abman SH, Chatfield BA, Hammond KB, Hambidge KM, et al. Pancreatic insufficiency, growth, and nutrition in infants identified by newborn screening as having cystic fibrosis. *J Pediatr*. 1992; 120:533–40. [PubMed: 1552390]
3. Madan JC, Koestler DC, Stanton BA, Davidson L, Moulton LA, Housman ML, et al. Serial analysis of the gut and respiratory microbiome in cystic fibrosis in infancy: interaction between intestinal and respiratory tracts and impact of nutritional exposures. *MBio*. 2012; 3:e000251–12.
4. Cox MJ, Allgaier M, Taylor B, Baek MS, Huang YJ, Daly RA, et al. Airway microbiota and pathogen abundance in age-stratified cystic fibrosis patients. *PLoS One*. 5:e11044. [PubMed: 20585638]
5. Rosenecker J. Relations between the frequency of the DeltaF 508 mutation and the course of pulmonary disease in cystic fibrosis patients infected with *Pseudomonas aeruginosa*. *Eur J Med Res*. 2000; 5:356–9. [PubMed: 10958769]
6. McKone EF, Emerson SS, Edwards KL, Aitken ML. Effect of genotype on phenotype and mortality in cystic fibrosis: a retrospective cohort study. *Lancet*. 2003; 361:1671–6. [PubMed: 12767731]
7. Polizzi A, Francavilla R, Castaldo G, Santostasi T, Tomaiuolo R, Manca A, et al. Phenotypic expression of genotype-phenotype correlation in cystic fibrosis patients carrying the 852del22 mutation. *Am J Med Genet A*. 2005; 132:434–40. [PubMed: 15633186]
8. Sagel SD, Gibson RL, Emerson J, McNamara S, Burns JL, Wagener JS, et al. Impact of *Pseudomonas* and *Staphylococcus* infection on inflammation and clinical status in young children with cystic fibrosis. *J Pediatr*. 2009; 154:183–8. [PubMed: 18822427]
9. Green DM, McDougal KE, Blackman SM, Sosnay PR, Henderson LB, Naughton KM, et al. Mutations that permit residual CFTR function delay acquisition of multiple respiratory pathogens in CF patients. *Respir Res*. 11:140. [PubMed: 20932301]
10. Collaco JM, Blackman SM, McGready J, Naughton KM, Cutting GR. Quantification of the relative contribution of environmental and genetic factors to variation in cystic fibrosis lung function. *J Pediatr*. 2010; 157:802–7. [PubMed: 20580019]
11. Eggesbo M, Moen B, Peddada S, Baird D, Rugtveit J, Midtvedt T, et al. Development of gut microbiota in infants not exposed to medical interventions. *APMIS*. 2011; 119:17–35. [PubMed: 21143523]
12. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012; 486:222–7. [PubMed: 22699611]
13. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A*. 2010; 107:11971–5. [PubMed: 20566857]
14. Meurens F, Berri M, Siggers RH, Willing BP, Salmon H, Van Kessel AG, et al. Commensal bacteria and expression of two major intestinal chemokines, TECK/CCL25 and MEC/CCL28, and their receptors. *PLoS One*. 2007; 2:e677. [PubMed: 17653288]
15. Backhed F. 99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: the normal gut microbiota in health and disease. *Clin Exp Immunol*. 2010; 160:80–4. [PubMed: 20415855]
16. Klepac-Ceraj V, Lemon KP, Martin TR, Allgaier M, Kembel SW, Knapp AA, et al. Relationship between cystic fibrosis respiratory tract bacterial communities and age, genotype, antibiotics and *Pseudomonas aeruginosa*. *Environ Microbiol*. 2010; 12:1293–303. [PubMed: 20192960]
17. Weiss B, Bujanover Y, Yahav Y, Vilozni D, Fireman E, Efrati O. Probiotic supplementation affects pulmonary exacerbations in patients with cystic fibrosis: a pilot study. *Pediatr Pulmonol*. 2010; 45:536–40. [PubMed: 20503277]
18. Bruzzese E, Raia V, Spagnuolo MI, Volpicelli M, De Marco G, Maiuri L, et al. Effect of *Lactobacillus* GG supplementation on pulmonary exacerbations in patients with cystic fibrosis: a pilot study. *Clin Nutr*. 2007; 26:322–8. [PubMed: 17360077]

19. Kraynack NC, Gothard MD, Falletta LM, McBride JT. Approach to treating cystic fibrosis pulmonary exacerbations varies widely across US CF care centers. *Pediatr Pulmonol.* 2011; 46:870–81. [PubMed: 21465675]
20. MS. Characterizing microbial population structures through massively parallel sequencing. *Uncultivated Microorganisms.* 2009; 2009:19–34.
21. Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* 2007; 8:R143. [PubMed: 17659080]
22. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics.* 2011; 27:2194–200. [PubMed: 21700674]
23. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome research.* 2011; 21:494–504. [PubMed: 21212162]
24. Huse SM, Dethlefsen L, Huber JA, Mark Welch D, Relman DA, Sogin ML. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet.* 2008; 4:e1000255. [PubMed: 19023400]
25. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 2007; 35:7188–96. [PubMed: 17947321]
26. Jakobsson HE, Jernberg C, Andersson AF, Sjolund-Karlsson M, Jansson JK, Engstrand L. Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *Plos One.* 2010; 24:5.
27. Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful Approach to multiple yesting. *J Roy Stat Soc B Met.* 1995; 57:289–300.
28. Aitchison, J. The statistical analysis of compositional data. London, New York: Chapman and Hall; 1986.
29. Hill M. Diversity and evenness: a unifying notation and its consequences. *Ecology.* 1973; 54:427–32.
30. Rocke D. On the beta transformation family. *Technometrics.* 1993; 35:72–81.
31. Huse SM, Ye Y, Zhou Y, Fodor AA. A core human microbiome as viewed through 16S rRNA sequence clusters. *PLoS One.* 2012; 7:e34242. [PubMed: 22719824]
32. Structure function and diversity of the healthy human microbiome. *Nature.* 2012; 486:207–14. [PubMed: 22699609]
33. Chen HM, Wang Y, Su LH, Chiu CH. Nontyphoid salmonella infection: microbiology, clinical features, and antimicrobial therapy. *Pediatr Neonatol.* 2013; 54:147–52. [PubMed: 23597525]
34. Bjorksten B, Naaber P, Sepp E, Mikelsaar M. The intestinal microflora in allergic Estonian and Swedish 2-year-old children. *Clin Exp Allergy.* 1999; 29:342–6. [PubMed: 10202341]
35. Bjorksten B, Sepp E, Julge K, Voor T, Mikelsaar M. Allergy development and the intestinal microflora during the first year of life. *J Allergy Clin Immunol.* 2001; 108:516–20. [PubMed: 11590374]
36. Kalliomaki M, Kirjavainen P, Eerola E, Kero P, Salminen S, Isolauri E. Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol.* 2001; 107:129–34. [PubMed: 11150002]
37. Mazmanian SK, Round JL, Kasper DL. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature.* 2008; 453:620–5. [PubMed: 18509436]
38. Marques TM, Wall R, Ross RP, Fitzgerald GF, Ryan CA, Stanton C. Programming infant gut microbiota: influence of dietary and environmental factors. *Curr Opin Biotechnol.* 2010; 21:149–56. [PubMed: 20434324]
39. Rogers GB, Carroll MP, Hoffman LR, Walker AW, Fine DA, Bruce KD. Comparing the microbiota of the cystic fibrosis lung and human gut. *Gut Microbes.* 2010; 1:85–93. [PubMed: 21326915]
40. Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, et al. Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics.* 2006; 118:511–21. [PubMed: 16882802]

41. Adlerberth I, Wold AE. Establishment of the gut microbiota in Western infants. *Acta Paediatr.* 2009; 98:229–38. [PubMed: 19143664]
42. Rogers GB, Daniels TW, Tuck A, Carroll MP, Connett GJ, David GJ, et al. Studying bacteria in respiratory specimens by using conventional and molecular microbiological approaches. *BMC Pulm Med.* 2009; 9:14. [PubMed: 19368727]
43. Tunney MM, Field TR, Moriarty TF, Patrick S, Doering G, Muhlebach MS, et al. Detection of anaerobic bacteria in high numbers in sputum from patients with cystic fibrosis. *Am J Respir Crit Care Med.* 2008; 177:995–1001. [PubMed: 18263800]
44. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. Development of the human infant intestinal microbiota. *PLoS Biol.* 2007; 5:e177. [PubMed: 17594176]
45. Favier CF, Vaughan EE, De Vos WM, Akkermans AD. Molecular monitoring of succession of bacterial communities in human neonates. *Appl Environ Microbiol.* 2002; 68:219–26. [PubMed: 11772630]
46. Rosenfeld M, Emerson J, Accurso F, Armstrong D, Castile R, Grimwood K, et al. Diagnostic accuracy of oropharyngeal cultures in infants and young children with cystic fibrosis. *Pediatr Pulmonol.* 1999; 28:321–8. [PubMed: 10536062]
47. Armstrong DS, Grimwood K, Carlin JB, Carzino R, Olinsky A, Phelan PD. Bronchoalveolar lavage or oropharyngeal cultures to identify lower respiratory pathogens in infants with cystic fibrosis. *Pediatr Pulmonol.* 1996; 21:267–75. [PubMed: 8726151]
48. Ramsey BW, Wentz KR, Smith AL, Richardson M, Williams-Warren J, Hedges DL, et al. Predictive value of oropharyngeal cultures for identifying lower airway bacteria in cystic fibrosis patients. *Am Rev Respir Dis.* 1991; 144:331–7. [PubMed: 1859056]
49. Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, et al. Analysis of the lung microbiome in the “healthy” smoker and in COPD. *PLoS One.* 2011; 6:e16384. [PubMed: 21364979]
50. van der Gast CJ, Walker AW, Stressmann FA, Rogers GB, Scott P, Daniels TW, et al. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J.* 2011; 5:780–91. [PubMed: 21151003]
51. Van den Abbeele P, Verstraete W, El Aidy S, Geirnaert A, Van de Wiele T. Prebiotics, faecal transplants and microbial network units to stimulate biodiversity of the human gut microbiome. *Microb Biotechnol.* 2013; 6:335–40. [PubMed: 23594389]
52. Gollwitzer ES, Marsland BJ. Microbiota abnormalities in inflammatory airway diseases - Potential for therapy. *Pharmacol Ther.* 2014; 141:32–9. [PubMed: 23969226]
53. Duytschaever G, Huys G, Bekaert M, Boulanger L, De Boeck K, Vandamme P. Cross-sectional and longitudinal comparisons of the predominant fecal microbiota compositions of a group of pediatric patients with cystic fibrosis and their healthy siblings. *Appl Environ Microbiol.* 2011; 77:8015–24. [PubMed: 21926193]
54. Omland O, Hjort C, Pedersen OF, Miller MR, Sigsgaard T. New-onset asthma and the effect of environment and occupation among farming and nonfarming rural subjects. *J Allergy Clin Immunol.* 2011; 128:761–5. [PubMed: 21752438]
55. Bottcher MF, Nordin EK, Sandin A, Midtvedt T, Bjorksten B. Microflora-associated characteristics in faeces from allergic and nonallergic infants. *Clin Exp Allergy.* 2000; 30:1590–6. [PubMed: 11069568]
56. Lynch SV, Bruce KD. The cystic fibrosis airway microbiome. *Cold Spring Harb Perspect Med.* 2013; 3:a009738. [PubMed: 23457293]
57. Rogers GB, Hart CA, Mason JR, Hughes M, Walshaw MJ, Bruce KD. Bacterial diversity in cases of lung infection in cystic fibrosis patients: 16S ribosomal DNA (rDNA) length heterogeneity PCR and 16S rDNA terminal restriction fragment length polymorphism profiling. *J Clin Microbiol.* 2003; 41:3548–58. [PubMed: 12904354]
58. Langille MG, Meehan CJ, Koenig JE, Dhanani AS, Rose RA, Howlett SE, et al. Microbial shifts in the aging mouse gut. *Microbiome.* 2014; 2(1):50. [PubMed: 25520805]
59. Sakamoto M, Kitahara M, Benno Y. *Parabacteroides johnsonii* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol.* 2007; 57(Pt 2):293–6. [PubMed: 17267966]

60. Garrido D, Dallas DC, Mills DA. Consumption of human milk glycoconjugates by infant-associated *Bifidobacteria*: mechanisms and implications. *Microbiology*. 2013; 159:649–64. [PubMed: 23460033]
61. Fukuda S, Toh H, Hase K, Oshima K, Nakanishi Y, Yoshimura K, et al. *Bifidobacteria* can protect from enteropathogenic infection through production of acetate. *Nature*. 2011; 469:543–7. [PubMed: 21270894]
62. Jost T, Lacroix C, Braegger CP, Chassard C. New insights in gut microbiota establishment in healthy breast fed neonates. *PLoS One*. 2012; 7:e44595. [PubMed: 22957008]
63. Backhed F. Host responses to the human microbiome. *Nutr Rev*. 2012; 70(Suppl 1):S14–7. [PubMed: 22861802]
64. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. *Cell*. 2012; 148:1258–70. [PubMed: 22424233]
65. Collado MC, Cernada M, Bauerl C, Vento M, Perez-Martinez G. Microbial ecology and host-microbiota interactions during early life stages. *Gut Microbes*. 2012; 3:352–65. [PubMed: 22743759]
66. Hansen CH, Nielsen DS, Kverka M, Zakostelska Z, Klimesova K, Hudcovic T, et al. Patterns of early gut colonization shape future immune responses of the host. *PLoS One*. 2012; 7:e34043. [PubMed: 22479515]
67. Lee YK, Mazmanian SK. Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science*. 2010; 330:1768–73. [PubMed: 21205662]
68. Ivanov II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, et al. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*. 2009; 139:485–98. [PubMed: 19836068]
69. Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell*. 2005 Jul 15; 122:107–18. [PubMed: 16009137]
70. Kamada N, Seo SU, Chen GY, Nunez G. Role of the gut microbiota in immunity and inflammatory disease. *Nat Rev Immunol*. 2013; 13:321–35. [PubMed: 23618829]
71. Luoto R, Kalliomaki M, Laitinen K, Delzenne NM, Cani PD, Salminen S, et al. Initial dietary and microbiological environments deviate in normal-weight compared to overweight children at 10 years of age. *J Pediatr Gastroenterol Nutr*. 2011; 52:90–5. [PubMed: 21150648]
72. Ajslev TA, Andersen CS, Gamborg M, Sorensen TI, Jess T. Childhood overweight after establishment of the gut microbiota: the role of delivery mode, pre-pregnancy weight and early administration of antibiotics. *Int J Obes*. 2011; 35:522–9.
73. Bisgaard H, Li N, Bonnelykke K, Chawes BL, Skov T, Paludan-Muller G, et al. Reduced diversity of the intestinal microbiota during infancy is associated with increased risk of allergic disease at school age. *J Allergy Clin Immunol*. 2011; 128:646–52. e1–5. [PubMed: 21782228]
74. Lynch SV, Goldfarb KC, Wild YK, Kong W, De Lisle RC, Brodie EL. Cystic fibrosis transmembrane conductance regulator knockout mice exhibit aberrant gastrointestinal microbiota. *Gut Microbes*. 2013; 4:41–7. [PubMed: 23060053]
75. Schippa S, Iebba V, Santangelo F, Gagliardi A, De Biase RV, Stamato A, et al. Cystic fibrosis transmembrane conductance regulator (CFTR) allelic variants relate to shifts in faecal microbiota of cystic fibrosis patients. *PLoS One*. 2013; 8:e61176. [PubMed: 23613805]
76. Scanlan PD, Buckling A, Kong W, Wild Y, Lynch SV, Harrison F. Gut dysbiosis in cystic fibrosis. *J Cyst Fibros*. 2012; 11:454–5. [PubMed: 22538067]
77. Tanaka A, Seki M, Yamahira S, Noguchi H, Kosai K, Toba M, et al. *Lactobacillus pentosus* strain b240 suppresses pneumonia induced by *Streptococcus pneumoniae* in mice. *Lett Appl Microbiol*. 2011; 53:35–43. [PubMed: 21554343]
78. Thomas DJ, Husmann RJ, Villamar M, Winship TR, Buck RH, Zuckermann FA. *Lactobacillus rhamnosus* HN001 attenuates allergy development in a pig model. *PLoS One*. 2011; 6:e16577. [PubMed: 21386995]
79. Youn HN, Lee DH, Lee YN, Park JK, Yuk SS, Yang SY, et al. Intranasal administration of live *Lactobacillus* species facilitates protection against influenza virus infection in mice. *Antiviral Res*. 2012; 93:138–43. [PubMed: 22120759]

Fig. 1. Heatmap of intestinal and respiratory relative microbial abundance

Heatmap based on the hierarchical clustering solution (Euclidean distance metric and average linkage) of the intestinal (n=66) and respiratory (n=54) microbiome samples. Rows represent samples (subjects and respective time-points at which the samples were collected), columns represent genera, and the values in the heatmap represent logged relative microbial abundance, with warmer colors representing greater relative abundance. Black-white shading indicates age at which the sample was collected, with black earlier in life (near time zero) and gray to white indicating the oldest timepoints at which samples were collected (nearing 34 months). Box A: Core microbiome crossing respiratory and intestinal samples across all time points. Box B: microbes more abundantly represented in the respiratory microbiome predominately at older timepoints.

Fig. 2. Changes in relative abundance of bacterial genera prior to initial *P. aeruginosa* colonization

Linear mixed effects model coefficient-based estimates of changes in logged normalized microbial abundance are plotted for each bacterial genus. Values below the line represent a decrease in abundance of the indicated bacterial genus over time and points above the line represent bacteria that increased in abundance over time (highlighted in red = significant change after correcting for multiple comparisons). **(A)** Changes in relative abundance of intestinal bacteria prior to *P. aeruginosa* colonization n=66 samples from 12 subjects. **(B)** Changes in relative abundance of respiratory bacteria prior to to *P. aeruginosa* colonization in n=54 samples from 13 subjects.

Fig. 3, online. Changes in relative abundance of intestinal and respiratory bacteria prior to first CF exacerbation

Linear mixed effects model coefficient-based estimates of changes in logged normalized microbial abundance are plotted for each bacterial genus. Values below the line represent a decrease in abundance of the indicated bacterial genus over time and points above the line represent bacteria that increase in abundance over time. (A) Changes in relative abundance of intestinal bacteria prior to CF exacerbation. (B) Changes in relative abundance of respiratory bacteria prior to CF exacerbation. Note that after correcting for multiple comparisons there were no significant changes observed.

Fig. 4. Comparison of subjects with and without CF exacerbation in the first six months of life according to microbiome composition using PCA

Principal components are based on average microbiome profiles over time points prior to first CF exacerbation for subjects experiencing CF exacerbation in the first 6 months of life and over the first 6 months of life for subjects without CF exacerbation in the first 6 months of life. Principal component 1 (PC1) is plotted against principal component 2 (PC2).

Subjects experiencing pulmonary exacerbation in the first six months of life (filled circles) are compared with subjects not experiencing pulmonary exacerbation in the first six months of life (open circles). **(A)** Comprison of 12 subjects based on n=23 intestinal samples; **(B)** Comprison of 11 subjects based on n=19 respiratory samples.

Fig. 5, online. Comparison of subjects with and without *Pseudomonas* colonization status in the first six months of life according to microbiome composition using PCA

Principal components are based on average microbiome profiles over time points prior to first *P. aeruginosa* colonization for subjects colonized by *P. aeruginosa* in the first 6 months of life and over the first 6 months of life for subjects without *P. aeruginosa* colonization in the first 6 months of life. Principal component 1 (PC1) is plotted against principal component 2 (PC2). Subjects experiencing *P. aeruginosa* colonization in the first six months of life (filled circles) are compared with subjects not experiencing *P. aeruginosa* colonization in the first six months of life (open circles). **(A)** Comparison of 12 subjects based on n=25 intestinal samples (p=0.14); **(B)** Comparison of 11 subjects based on n=19 respiratory samples (p=0.40).

Fig. 6. Time-to-event analysis of the relationship between breast milk exposure in infancy and time to first CF exacerbation

The proportion of subjects who have not had first CF exacerbation is plotted against chronological age (in months) for 10 breast fed subjects (broken lines) and 3 exclusively formula fed subjects (solid lines). Infants with exposure to breast milk experienced overall longer times to first CF exacerbation (log rank test: $p=0.06$).

Fig. 7, online. Time-to-event analysis of the relationship between intestinal or respiratory microbial diversity (assessed using Simpson's diversity index; SDI) and time to CF related outcome, adjusted for breast milk exposure

(A) Proportion of subjects without CF exacerbation versus age at varying levels of intestinal SDI. Models are based on 46 stool samples collected from 12 subjects prior to their first CF exacerbation. Log rank test $p=0.12$; (B) Proportion of subjects without *P. aeruginosa* colonization versus age at varying levels of intestinal SDI. Models are based on 59 stool samples collected from 12 subjects prior to their first *P. aeruginosa* colonization. Log rank test $p=0.62$; (C) Proportion of subjects without CF exacerbation versus age at varying levels of respiratory SDI. Models are estimated based on 37 oropharyngeal samples collected from 13 subjects prior to their first CF exacerbation. Log rank test $p=0.20$; (D) Proportion of subjects without *P. aeruginosa* colonization versus age at varying levels of respiratory SDI. Models are estimated based on 49 oropharyngeal samples collected from 13 subjects prior to their first *P. aeruginosa* colonization. Log rank test $p=0.85$.

Table 1

Clinical characteristics of study subjects

Clinical variable	Mean (SD)	N (%)
Ever breast fed		10 (77)
Ever given antibiotics		12 (92)
Age first antibiotics (months)	9.0 (4.6)	
Ever colonized with <i>P. aeruginosa</i>		6 (46)
Age first <i>P. aeruginosa</i> colonization (months)	9.6 (6.7)	
Ever had CF exacerbation		9 (69)
Age first CF exacerbation (months)	8.8 (3.9)	

SD: standard deviation