Nutrición Hospitalaria
ISSN: 0212-1611
info@nutriciónhospitalaria.com
Grupo Aula Médica
España

Poroyko, V.; Morowitz, M.; Bell, T.; Ulanov, A.; Wang, M.; Donovan, S.; Bao, N.; Gu, S.; Hong, L.; Alverdy, J. C.; Bergelson, J.; Liu, D. C.

Diet creates metabolic niches in the "inmature gut" that shape microbial communities

Nutrición Hospitalaria, vol. 26, núm. 6, noviembre-diciembre, 2011, pp. 1283-1295
Grupo Aula Médica
Madrid, España

Nutrición Hospitalaria

Nutr Hosp. 2011;26(6):1283-1295
ISSN 0212-1611 • CODEN NUHOEQ
S.V.R. 318

Original

Diet creates metabolic niches in the "inmature gut" that shape microbial communities

V. Poroyko[1], M. Morowitz[2], T. Bell[3,5], A. Ulanov[4], M. Wang[4], S. Donovan[4], N. Bao[5], S. Gu[5], L. Hong[5], J. C. Alverdy[1], J. Bergelson[6] and D. C. Liu[1]

[1]Department of Surgery. The University of Chicago. Chicago. IL. USA. [2]University of Pittsburgh. Pittsburgh. PA. USA. [3]Department of Zoology. University of Oxford. Oxford. OX1 3PS. United Kingdom. [4]University of Illinois. Champaign. IL. USA. [5]The institution is Shanghai Children's Medical Center. Shanghai Jiaotong University. School of Medicine. Shanghai. China. [6]Department of Ecology & Evolution. The University of Chicago. Chicago. IL. USA.

Abstract

Although diet composition has been implicated as a major factor in the etiology of various gastrointestinal diseases, conclusive evidence remains elusive. This is particularly true in diseases such as necrotizing enterocolitis where breast milk as opposed to commercial formula appears to confer a "protective effect" to the "immature gut." Yet the mechanism by which this occurs continues to remain speculative. In the present study we hypothesize that the basic chemical composition of diet fundamentally selects for specific intestinal microbiota which may help explain disparate disease outcome and therapeutic direction. Complimentary animal and human studies were conducted on young piglets (21 d.)(n = 8) (IACUC protocols 08070 and 08015) and premature infants (adjusted gestational age 34-36 weeks) (n = 11) (IRB Protocol 15895A). In each study, cecal or stool contents from two groups (Breast milk-fed (BF) vs. Formula-fed (FF)) were analyzed by gas chromatography/mass spectrometry (GC/MS) and comprehensive metabolic profiles generated and compared. Concurrently, bacterial community structure was assayed and respective representative microbiota of the groups determined by 16S rRNA gene amplicon pyrosequencing. Statistical modeling and analysis was done using SIMCA-P+ and R software. GC/MS metabolomics identified clear differences between BF and FF groups in the intestinal environment of piglets and humans. Sugars, amino-sugars, fatty acids, especially unsaturated fatty acids, and sterols were identified as being among the most important metabolites for distinguishing between BF and FF groups. Joint analysis of microbiota and metabolomics pinpointed specific sets of metabolites ($p < 0.05$) associated with the dominant bacterial taxa. The chemical composition of diet appears

Correspondence: Donald C. Liu.
Mary Campau Ryerson Professor and Surgeon-in-Chief.
University of Chicago Comer Children's Hospital.
Vice-Chairman, Department of Surgery.
University of Chicago Pritzker School of Medicine.
MC 4062.
5839 South Maryland Avenue.
Chicago Illinois 60637-1470
E-mail: dliu@surgery.bsd.uchicago.edu

Recibido: 23-VI-2011.
Aceptado: 11-VII-2011.

LA DIETA CREA NICHOS METABÓLICOS EN EL "INTESTINO INMADURO" QUE DA FORMA A COMUNIDADES MICROBIANAS

Resumen

Aunque se ha implicado a la composición de la dieta como un factor principal en la etiología de varias enfermedades gastrointestinales, la evidencia concluyente sigue siendo esquiva. Esto es particularmente cierto en enfermedades como la enterocolitis necrosante en la que la leche materna, en contraposición de las fórmulas comerciales, parece conferir un "efecto protector" para el "intestino inmaduro" o el ecosistema intestinal juvenil del "intestino inmaduro", si bien el mecanismo por el que esto ocurre sigue siendo una especulación. La hipótesis de nuestro estudio es que la composición química básica de la dieta selecciona fundamentalmente microbióticos intestinales específicos que pueden explicar los resultados dispares de la enfermedad y tener implicaciones terapéuticas. Se realizaron estudios adicionales en animales y humanos en lechones (21 d.) (n = 8) (protocolos IACUC 08070 y 08015) y lactantes prematuros (edad gestacional ajustada de 34-36 semanas) (n = 11) (Protocolo IRB 15895A). En cada estudio, se analizaron los contenidos cecales y fecales de ambos grupos (alimentación materna (AM) y alimentación con fórmula (AF)) mediante cromatografía de gases/espectrometría de masas (CG/EM) y se generaron y compararon perfiles metabólicos completos. De forma concurrente, se probó la estructura de la comunidad bacteriana y se determinaron los representantes micriobióticos respectivos de los grupos mediante la pirosecuenciación del ARNr 16S del amplicón. Se realizó un modelo y análisis estadístico utilizando los programas informáticos SIMCA-P+ y R. La metabolómica de CG/EM identificó diferencias claras entre los grupos AM y AF en el ambiente intestinal de los lechones y los humanos. Se encontró que los azúcares, los amino-azúcares, los ácidos grasos, especialmente los insaturados, y los esteroles estaban entre los metabolitos más importantes para distinguir entre los grupos AM y AF. El análisis conjunto de la microbiótica y de la metabolómica resaltó conjuntos específicos de metabolitos ($p < 0,05$) asociados con los taxones bacterianos dominantes. La composición química de la dieta parece tener un papel significativo en la definición de los microbióticos del intestino inmaduro. El análisis en tándem de los perfiles microbianos y metabóli-

to have a significant role in defining the microbiota of the immature gut. Tandem analysis of intestinal microbial and metabolic profiles is potentially a powerful tool leading to better understanding of the role of diet in disease perhaps even leading to specific strategies to alter microbial behavior to improve clinical outcome.

(*Nutr Hosp.* 2011;26:1283-1295)

DOI:10.3305/nh.2011.26.6.5362

Key words: *Prematurity. Diet. Breast milk. Human microbiota. Intestinal environment. Metabolomics.*

cos es, potencialmente, una herramienta poderosa que permitirá mejorar el conocimiento del papel de la dieta en la enfermedad y quizás conduzca a estrategias específicas para alterar el comportamiento microbiano mejorando el pronóstico clínico.

(*Nutr Hosp.* 2011;26:1283-1295)

DOI:10.3305/nh.2011.26.6.5362

Palabras clave: *Prematuridad. Dieta. Microbiota. Ambiente intestinal. Metabolómica.*

Abbreviations

NICU: Neonatal Intensive Care Unit.
GC/MS: Gas Chromatography/Mass Spectrometry.
PCR: Polymerase Chain Reaction.
FF: Formula-Fed.
BF: Breast Milk-Fed.
PCA: Principal Component Analysis.
PLS-DA: Partial Least-Squares Discriminant Analysis.

Introduction

Colonization of the sterile gastrointestinal tract of human newborns begins at birth and continues until a complex microbial community is established months or years later.[1,2] These gut microbial communities have been implicated in a variety of human diseases, some of which are more frequent in children who are fed formula,[3,4,5,6] however the direct connection between diet and disease is poorly understood. Numerous studies have sought to distinguish the factors that control colonization of intestinal microbial populations.[1,2,7,8] Factors that appear to impact colonization patterns include gestational age at delivery[9], maternal microbial profiles,[10,11] host genotype,[12] and mode of delivery.[13,14] In premature newborns, colonization is likely affected further by deranged host physiology,[15] administration of broad-spectrum antibiotics,[16] and an underdeveloped immune system.[6] Numerous studies have demonstrated that several disease states (including neonatal necrotizing enterocolitis, neonatal sepsis, obesity, and diabetes) are observed more frequently in children who have received artificial formula than in children who have received human milk.[7] Interestingly, many of these diseases are also partially mediated by microbes. It remains unclear whether the increased disease incidence in formula-fed infants is itself related to the impact of diet on the establishment of the intestinal microbiota.

The importance of host diet remains controversial due to the high degree of variability across experiments and replicates within experiments. Strong effects of breast milk versus formula on the composition of the microbial communities have been demonstrated in some,[13,17] but not other, studies.[7,8,9] Furthermore, it has been suggested that a breast-milk diet promotes growth of a distinct microbial community consisting of "friendly" commensals,[2,18] although evidence of these long-held beliefs still remains elusive.[1,18,19] We believe this debate can best be resolved with an understanding of how diet influences the gut environment, and how the environment impacts the colonization of particular bacteria. Here, we use a tandem approach combining GC/MS metabolomics[3,20,21,22] with 16s rRNA gene PCR amplicon pyrosequencing to relate the intestinal ecosystems of breast milk-fed (BF) and formula Fed (FF) full term piglets and human premature infants. The relative simplicity of the microbial communities in the immature gut or juvenile intestinal ecosystem allows us to determine key metabolites in the gut that drive predictable changes in community composition.

Materials and methods

Ethics statement

Samples were collected from infants admitted to the neonatal intensive care unit of The University of Chicago Hospital (Chicago, IL) between January 2008 and March 2008. Samples were obtained with written, informed parental consent under a research protocol that was approved by the Institutional Review Board of the University of Chicago (Protocol 15895A). Animals were managed throughout the study in accordance with requirements of the Institutional Animal Care and Use Committee at the University of Illinois (IACUC protocols 08070 and 08015), in accordance with approved NIH guidelines.

Sample collection

Spontaneously evacuated fecal samples were collected from 11 infants; the adjusted gestational age of the subjects ranged from 34-36 weeks. Each study subject was receiving enteral nutrition only, had no major comorbidities, and was not receiving antibiotics for 5 days at time of sampling. Six of the infants were receiving premature infant formulas Abbott Nutrition (OH) (Similac NeoSure (DHA + ARA) with iron 22 cal. Infant

formula #59649, Similac Special Care (DHA + ARA) 20 cal. premature infant formula #52418, Similac Special Care (DHA + ARA) 24 cal. premature infant formula #51024) and 5 were receiving human milk. To get an overall view of the juvenile intestinal ecosystem, subjects were matched by adjusted gestational age, but the gender, birth weights, and age at sampling were as heterogeneous as possible. This study design was selected to emphasize the role of diet in gut microbial colonization and to exclude the influence of developmental stage. Samples were collected, frozen at -80ºC and stored until use. Preceding the analysis, samples were lyophilized using FreeZone 2,5L Freeze Dry System (Labconco, MO) and weighed.

Samples of breast milk were collected from healthy volunteer donors in sterile plastic containers and frozen at -20ºC until delivery to the laboratory, where they were stored at -80ºC. Samples of three brands of premature infant formula (see above) were stored until use in factory sealed containers according to manufacturer instructions.

Vaginally-delivered piglets were allowed to suckle for 48 h postnatal to obtain colostrum. Piglets were then randomized to be sow-reared (BF; n = 4) or formula fed (FF; n = 4). BF remained with the sow throughout the study, while FF piglets were transported to the animal facilities and housed individually in metabolism cages with 12 h light/dark cycle as previously described.[23] Room temperature was maintained between 30 to 32°C with supplemental heat provided through of radiant heaters suspended above the cages. The FF piglets were fed with a cow-milk based sow milk replacer (Advance Baby Pig LiquiWean; Milk Specialties, Dundee, IL). Formula was prepared fresh each morning at a concentration of 18.3% solids and was offered 14-times daily via a pump into a bowl at a rate of 360 mL/kg body weight. Piglets were weighed each morning and monitored three times per day for general health. Samples were collected on postnatal day 21. Piglets were first sedated with an intramuscular injection of Telazol (7 mg/kg body weight; Fort Dodge Animal Health, Fort Dodge, IA) and then euthanized by intracardiac administration of sodium pentobarbital (Fatal Plus: 72 mg/kg body weight; Vortech Pharmaceuticals, Dearborn, MI). The large intestine was excised and separated into cecum and colon at the cecocolic junction. Tissues and luminal contents from cecum were immediately collected into sterile tubes and snap frozen in liquid nitrogen. The samples were stored at -80°C until processing.

Fecal DNA preparation, 16S rRNA gene amplification and sequencing

DNA was extracted from 50 mg of lyophilized fecal material (human study) or frozen cecal content (animal study) using an established laboratory protocol[24] or QIAamp DNA Stool mini Kit (Qiagen) protocol with modifications[25] respectively. Variable regions 1-4 of the 16S rRNA gene were amplified using the 8-27F/788-806R primer pair. Primers contained multiplex identifiers and the recommended 454 adapter sequences. Amplification conditions are described in Poroyko et al 2010.[26] PCR products were purified using AMPure Kit (Agenicourt Bioscience Corp.) and sequenced on GS Titanium 70x75 picotiter plate from adaptor B according to the manufacturer's protocols for GS FLX (Roche Applied Science, Indianapolis, IN). The sequences with the length > 100 bp were then annotated to the level of genus using Ribosomal Database Project Classifier (http://rdp.cme.msu.edu/).[27] The sequences have been deposited in the Sequence Read Archive (human study submission number is SRA026959, samples: SRX033349 A, SRX033350 B, SRX033351 C, SRX033352 G, SRX033353 D, SRX033354 E, SRX033355 F, SRX033356 H, SRX033357 I, SRX033358 J, SRX033359 K; pig study submission number is SRA029819, milk diet samples: SRS172359-SRS172362, formula diet samples: SRS172363, SRS172364, SRS172367, SRS172371).

Metabolomic analysis

The fecal (10-15 mg DW) cecal (10-15 mg WW) and food samples (1 ml) were extracted and derivatized according to Roessner et al.[28] with minor changes. Samples were extracted with 1.5 ml of 70% methanol for 15 min at 60^{0}C following extraction with water at room temperature and chloroform at 37^{0}C. Polar (methanol + water) and nonpolar extracts were separated and dried under vacuum. Prior to GC/MS analysis, dried polar extracts were derivatized with 80 l Methoxyamine hydrochloride (20 mg ml^{-1}) for 60 min at 40^{0}C and then with 80 l MSTFA at 60^{0}C for 40 min. The GC/MS system (Agilent Inc, Palo Alto, CA, USA) consisted of an Agilent 6890N gas chromatograph, an Agilent 5973 mass selective detector, and a HP 7683B autosampler. Gas chromatography of the polar extracts was performed on a HP-5MS (30 m × 0.25 mm I.D. and 0.25 mm film thickness) capillary column (Agilent Inc, Palo Alto, CA, USA). The inlet and MS interface temperatures were 250^{0}C, and the ion source temperature was adjusted to 230^{0}C. An aliquot of 1 mL was injected with the split ratio of 5:1. The helium carrier gas was kept at a constant flow rate of 1.3 ml min^{-1}. The temperature program was an initial 5-min isothermal heating at 70^{0}C, followed by an oven temperature increase of 5^{0}C min^{-1} to 310^{0}C and a final 10 min at 310^{0}C. Nonpolar metabolites (free fatty acids) were converted to methyl esters and analyzed with the same instrument on a ZB-WAX capillary column (30 m × 0.25 mm I.D. and 0.25 mm film thickness) (Phenomenex, Torrance, CA, USA) with an injection port temperature of 260^{0}C, the interface set to 280^{0}C, and the ion source adjusted to 230^{0}C. The helium carrier gas was set at a constant flow rate of 2 ml min^{-1}. The temperature program was 5-min

isothermal heating at 140ºC, followed by an oven temperature increase of 10ºC min^{-1} to 265ºC for a final 25 min.

In both cases mass spectra were recorded in the m/z 25-800 scanning range. The spectra of all chromatogram peaks were compared with electron impact mass spectrum libraries NIST02 (NIST, MD, USA), WILEY7n (Palisade Corporation, NY, USA) and the custom library. To allow comparison between samples, all data were normalized to the internal standard (hentriacontanoic acid at 10 mg ml^{-1}) in each chromatogram and dry weight of each sample. The spectra of all chromatogram peaks were evaluated using the HP Chemstation (Agilent, Palo Alto, CA, USA) and AMDIS (NIST, Gaithersburg, MD, USA) programs.

Statistical analysis

Metabolite concentrations were measured on different scales, so GC/MS data were mean-centered and Pareto-scaled to facilitate comparisons across metabolites. If the concentration of metabolite *i* in sample *j* is m_{ij}, we calculated the mean $\bar{m}_{i\bullet}$ and standard deviation $s_{i\bullet}$ across the 11 individuals, and transformed the raw metabolite concentrations as:

$$\tilde{m}_{ij} = \frac{m_{ij} - \bar{X}\,\bar{m}_{i\bullet}}{\sqrt{s_{i\bullet}}}$$

We first inspected the GC/MS and sequencing data using principal components analysis (PCA) and linear discriminant analysis using R software v2.8.1 (http://www.r-project.org). Partial least squares to latent structures discriminant analysis (PLS-DA) was conducted using the SIMCA-P+ software (v. 12.0.0.0, Umetrics, Umea, Sweden).

Linear models and ANOVA analysis were conducted using R v2.9.2. Generalized linear models were constructed to evaluate correlations between metabolites and bacterial taxa that: (i) could be identified to at

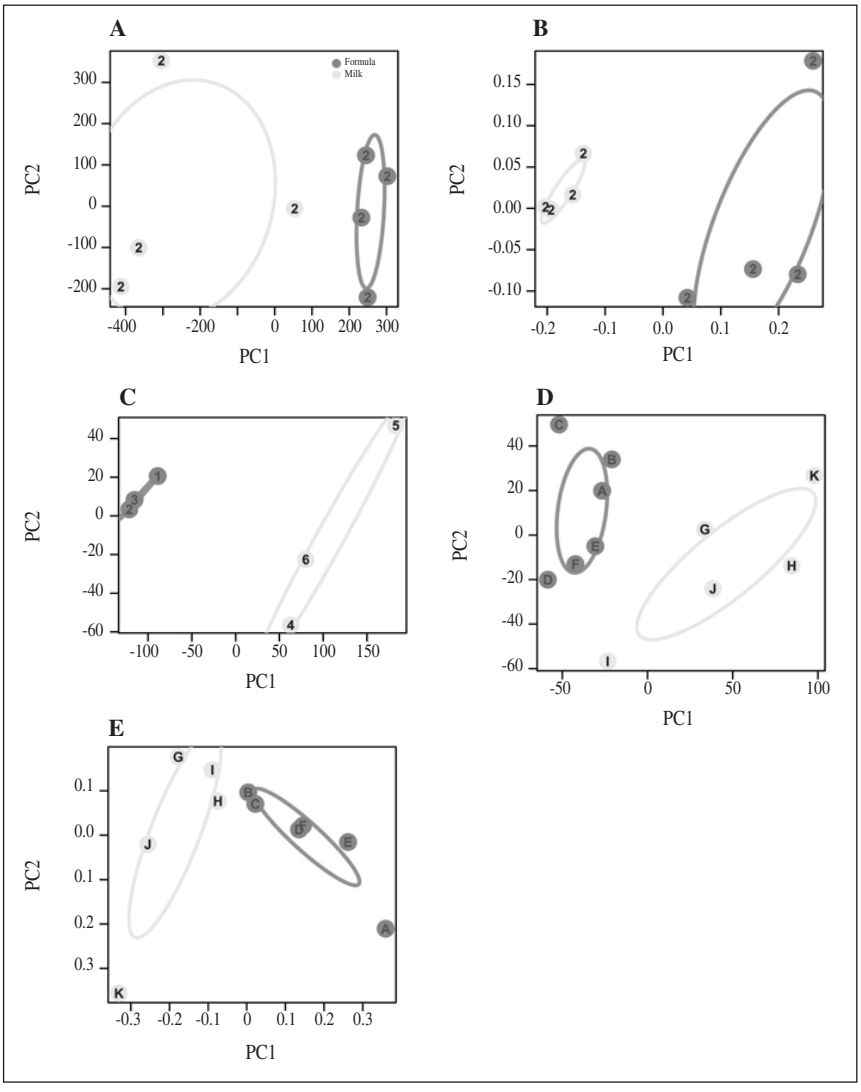

Fig. 1.—Specimens clustered according to (A) array of metabolites in the piglet cecal samples; (B) relative abundances of bacterial species at the genus level in pig cecal samples; (C) array of metabolites in the diet (human milk or formula); (D) array of metabolites in the human fecal samples; (E) relative abundances of bacterial species at the genus level in human fecal samples. The PCA is based on the same data as used for the generalized linear models (see Methods). Ellipses are 95% confidence intervals for each of the diet categories.

least the level of genus (ii) were detected in at least 2 individuals, and (iii) were detected with abundance greater than 20 sequences. This produced a list of 27 bacterial genera in piglets and 14 bacterial genera in human infants on which to conduct the analysis. The taxon abundances (counts) are dependent on the number of clones obtained from each individual. We therefore analyzed the relationship between each taxon and metabolite as a generalized linear model, where the observed counts of a particular taxon and the total number of individuals sampled are entered as the response variable, and the transformed metabolite concentration is the independent variable. Most of the models were overdispersed (residual deviance > residual degrees of freedom), so quasibinomial errors were used throughout. We constructed models for all pair wise combinations of bacterial genera and metabolites. The significance threshold was Bonferonni-adjusted for the total number of statistical test that were conducted in each study.

Results and discussion

A group of 8 vaginally-delivered piglets were randomized to be milk (n = 4) or formula-fed (n = 4). At the postnatal day 21, animals were euthanized and samples of cecal content were collected; and subjected to GC/MS metabolic and 16S rRNA gene based bacteria community profiling. The metabolic profiles of cecal samples from BF and FF piglets were clearly distinguished in PCA analysis (fig. 1A). We used partial least-squares discriminant analysis (PLS-DA) to determine the array of descriptive metabolites associated with each diet. PLS-DA maximizes separation of treatment classes (here BF and FF) and suggests candidate response variables (here the relative concentrations of metabolites) that distinguish the classes. In PLS-DA, Q^2Y describes the predictive ability of the model where predictions are achieved as Q^2Y approaches unity. The PLS-DA model ($Q^2Y = 0.97$), applied to the set of 243 cecal metabolites, identified 59 compounds discriminating BF and FF cecal environmental conditions (table I). The sugars, amino-sugars, fatty acids, especially unsaturated fatty acids, and sterols were identified as being among the important metabolites for distinguishing between BF and FF groups (table II). When PCA was applied to the relative abundances of gut microbes, the first two principal components explained 92.6% of the variation in the data, and the samples were clearly separated according to the diet type (fig. 1B). The BF specific genera *Prevotella* and *Oscillibacter,*

Table I
Distribution of metabolites detected in samples

Classes of metabolites	*Piglet cecal contents*			*Infant diet*			*Infant intestinal contents*		
	FF-specific	*BF-specific*	*Common*	*Formula-specific*	*Human milk-specific*	*Common*	*FF-specific*	*BF-specific*	*Common*
Organic acids	4	8	44	10*	6	25	11	8	13
Amino Acids	0	2	24	0	11	14	2	2	17
Sugar Alcohols	1	2	7	0	1	10	8	3	2
Fatty Alcohols	1	1	6	0	0	9	0	0	0
Amines	0	2	7	0	1	3	0	0	6
Nucleosides/Purines/Pirimidines	2	2	9	1	0	8	0	3	1
Alkanes	0	1	2	0	0	2	0	0	1
Vitamines	2	0	3	3	0	0	1	0	1
Phosphates	1	0	2	1	0	0	0	0	2
Sugars/sugar phosphates	2	5	17	9	6	13	2	9	7
Monosaccharides	2	1	15	5	4	9	1	5	6
Disaccharides:	0	2	2	2	2	1	0	2	1
Trisaccharides:	0	0	0	1	0	1	0	1	0
Amino sugars	0	2	0	1	0	1	0	2	1
Fatty acids	4	16	37	5	6	40	12	16	28
Saturated	4	8	33	4	3	30	7	8	18
Unsaturated	0	8	4	1	3	10	5	8	10
Sterols and Steroids	2	5	9	0	2	3	0	0	3
Other	0	0	0	1	1	4	0	0	3

*Numbers are counts of detected metabolites within each class of metabolites.

Table II

Discriminative compounds detected by GC/MS within samples or cecal contents of piglets receiving milk and formula diets

Metabolite	*Diet specific**	*Metabolite*	*Diet specific**
Organic acids		**Phosphates**	
2,3-hydroxypropyl 2-aminoethylphosphoric acid	BF	Monomethylphosphate	FF
2-Hydroxyglutaric acid	BF		
Aminomalonic acid	BF	**Vitamins**	
α-Tocopherolacetic acid	FF	γ-Tocopherol	FF
Galactonic acid	FF	α-Tocopherol	FF
Glucoheptonic acid	BF		
Glycolic acid	FF	**Fatty acids**	
Threonic acid	FF	10-Heptadecenoic acid	BF
		11-Eicosenoic acid	BF
Amino acids		11-Octadecenoic acid	BF
4-hydroxyproline	BF	15-Tetracosenoic acid	BF
Homocysteine	BF	5-hydroxypentanoic acid	BF
		9,12-Octadecadienoic acid	BF
Sugars *Monosaccharides:*		9,15-Octadecadienoic acid	BF
Fructose-6-P	FF	9-hexadecenoic acid	BF
Mannose-6-P	FF	9-octadecenoic acid	BF
Rhamnose (deoxy sugar)	BF	Docosanoic acid	FF
		Hexadecanoic acid	BF
Sugars *Disaccharides:*		Nonadecanoic acid	FF
Lactose	BF	Octadecanoic acid, hexadecyl ester	BF
Trehalose	BF	Tetracosanoic acid	FF
		Tetradecanoic acid	BF
Aminosugars		Tricosanoic acid	FF
N-Acetyl glucosamine	BF		
N-Acetylglucosylamine	BF	**Fatty alcohols**	
		Octadecanol	BF
Sugar alcohols		Tetradecanol	FF
Galactitol	FF		
Glycerol	BF	**Short chain fatty acid**	
Glycerol-2-P	BF	2-Aminobutanoic acid	BF
		2-methyl-Propanoic acid	BF
Amines		3-Methyl-3-hydroxybutanoic acid	BF
Putrescine	BF	Butanoic acid	BF
Spermidine	BF		
		Sterols and steroids	
Nucleosides		24-Ethyl-.delta.(22)-coprostenol	FF
Adenosine	BF	Cholesterol	BF
Guanosine	BF	Coprostan-3-ol	BF
Thymine	FF	Stigmastan-3-ol	BF
9H-Purin-2-amine	FF	β-Sitosterol	FF
		Lanosta-8,24-dien-3-beta-ol	BF
Alkenes		Lanosterol	BF
1-Cyanohexadecane	BF		

*According to PLS-DA model R2X(cum) = 0.759; R2Y(cum) = 0.999; Q2Y = 0.968.
BF: milk fed; FF: formula fed.

and the FF specific genera *Bacteroidetes* and *Parabacteroidetes* were discriminative for cecal microbial populations as determined by PLS-DA ($Q^2Y = 0.93$). We speculated that differences in the metabolite concentrations within the gut lumen, created by breast milk or formula, select for particular bacterial taxa. To investigate this possibility, the relationships between intestinal environmental metabolite concentrations and the abundances of the most abundant bacterial taxa was examined. Clustering the significant relationships allowed us to visualize the degree to which bacteria were associated with the same or different metabolites and therefore to infer which metabolites inhibited or enhanced the abundance of each taxon. The analysis of significant p-values associated with the relationship between metabolites and bacterial taxa (fig. 2) has shown breast milk to be complementary for the genera *Oscillibacter, Sporacetigenium, Clostridium, Prevotella* and *Anaerovibrio,* and formula to be complementary for the genera *Roseburia, Acidaminococcus, Bacteroides, Parabacteroides* and *Alistipes*. Hence, by pinpointing the suite of metabolites that were corre-

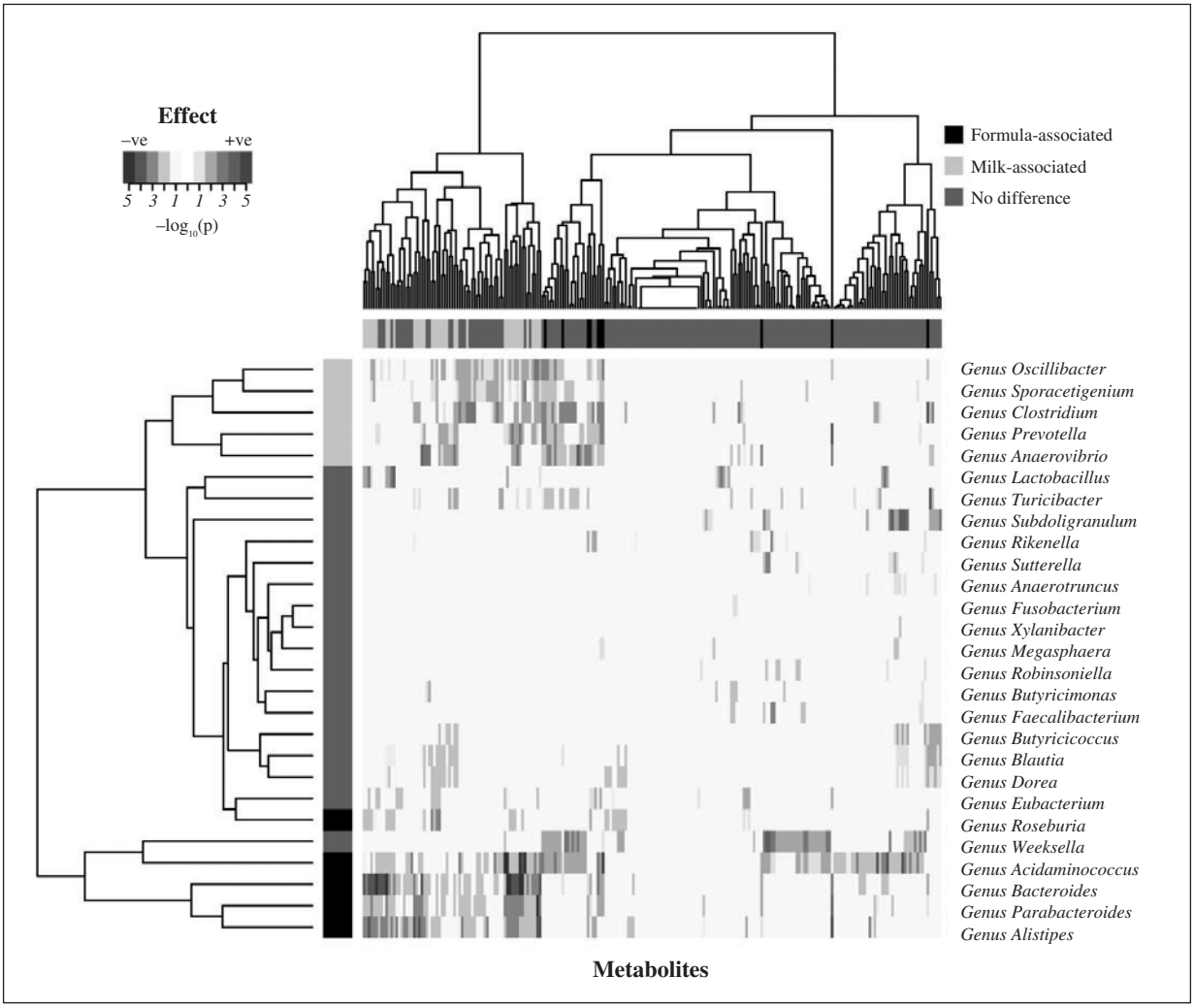

Fig. 2.—Heatmap of significant p-values associated with the relationship between metabolites and bacterial taxa in piglet cecal content. Analysis based on the intestinal contents collected from premature infants. Data are –log$_{10}$-transformed such that darker blue coloration denotes relatively more significant negative correlations and darker red coloration denotes relatively more significant positive correlations. All p-values < 0.05 are shown. Marginal cluster dendrograms are calculated from Euclidean distances. Greyscale sidebars denote whether metabolites or bacterial counts are significantly higher in formula, milk, or if there is no significant difference (p < 0.05). Significance is assessed using analysis-of-variance (metabolite data) or generalized linear models using binomial errors (count data).

lated with abundance, it was possible to identify the ecological niche of the most abundant taxa.

In order to test how the above environmental hypothesis complies with the juvenile intestinal ecosystem representative of infirmary settings we conducted analogous study on 11 NICU impatient premature infants. Fecal samples were obtained from 5 BF premature infants and 6 FF infants (table III), and subjected to metabolic profiling. The same metabolic profiling was applied to 3 samples of commercially available brands of premature formulas and 3 human milk samples donated by healthy volunteers. The GC/MS analysis identified a total of 257 metabolites, of which 105 were common to all datasets, 92 were diet specific (milk of formula) and 60 were found only among intestinal contents (fecal samples) (table I).

Principal component analysis (PCA) of diet derived metabolites demonstrated pronounced chemical differences between human milk and infant formula. The first two principal components were sufficient to explain 97% of the total variation, and there was clearly good separation of the diets over these principal components (fig. 1C). Linear discriminant analysis confirmed that the metabolite profiles enable samples to be categorized according to diet class, demonstrating that the substrates entering the infant digestive systems were substantially distinct.

As for the diets, PCA analysis of the fecal samples from BF and FF infants were clearly distinguished, with no overlap in the confidence intervals around the diet classes (fig. 1D). Linear discriminant analysis of these fecal metabolite profiles confirmed that each of

Table III
Clinical characteristics of human study subjects

Sample ID	*Sex*	*Mode of delivery*	*Antibiotics administration*	*Age at sample collection (days)*	*Adjusted gestational age (weeks)*	*Weight at sample collection (kg)*	*Diet*
A	Fem.	C	A/G Day 1-8	13	35	2	FF
B	Fem.	S	A/G Day 1-2	16	35	1.77	FF
C	Fem.	C	A/G Day 1-2, V/G Day 10-12	46	35	2	FF
D	Mal.	C	A/G Day 1-2	42	36	2.7	FF
E	Mal.	C	A/G Day 1-7	36	35	2.365	FF
F	Mal.	C	A/G Day 1-2; V/G Day 9-10	23	36	1.9	FF
G	Mal.	C	A/G Day 1-2	25	35	1.25	BF
H	Fem.	S	A/G Day 1-2	14	36	2.04	BF
I	Mal.	S	A/G Day 1-2	10	35	2.175	BF
J	Fem.	S	A/G Day 1-2	22	34	1.87	BF
K	Fem.	S	A/G Day 1-2	23	34	1.82	BF

C: Caesarean section; S: Spontaneous vaginal delivery; A/G: Ampicillin/gentamicin; V/G: Vancomycin/gentamicin; FF: Formula fed; BF: Breast milk-fed.

the infants could be correctly categorized according to their diet class. For the diet-derived metabolites, we found PLSDA model $Q^2Y = 0.96$, indicating excellent discrimination and predictive ability.

Several reasons are preventing in depth comparison of metabolites acquired from human stool samples and piglet cecal contents; among them the deeply anaerobic environment of the animal cecum, the noninvasive nature of the human study, and the differences of animal and human diets. However, as it was previously shown in our animal experiment, the PLS-DA model identified sugars as being among the important metabolites for distinguishing between FF and BF groups (tables I, IV). Sixteen mono-, di- and trisaccharides were observed, of which 8 were specific to the BF group and one (galactofuranose) was specific to the formula diet. Increased carbohydrate supply appears to suppress microbial virulence genes in many situations.[29,30,31] It is thus noteworthy that the environmental metabolomic approach identified a preponderance of galactofuranose (Gal*f*) in the FF samples, as Gal*f* has been shown to be essential for microbial pathogenicity[32] and formula fed premature infants are known to be especially susceptible to microbe-mediated complications[7]. Another group of metabolites, in contrary, was strongly associated with the FF group: of the 13

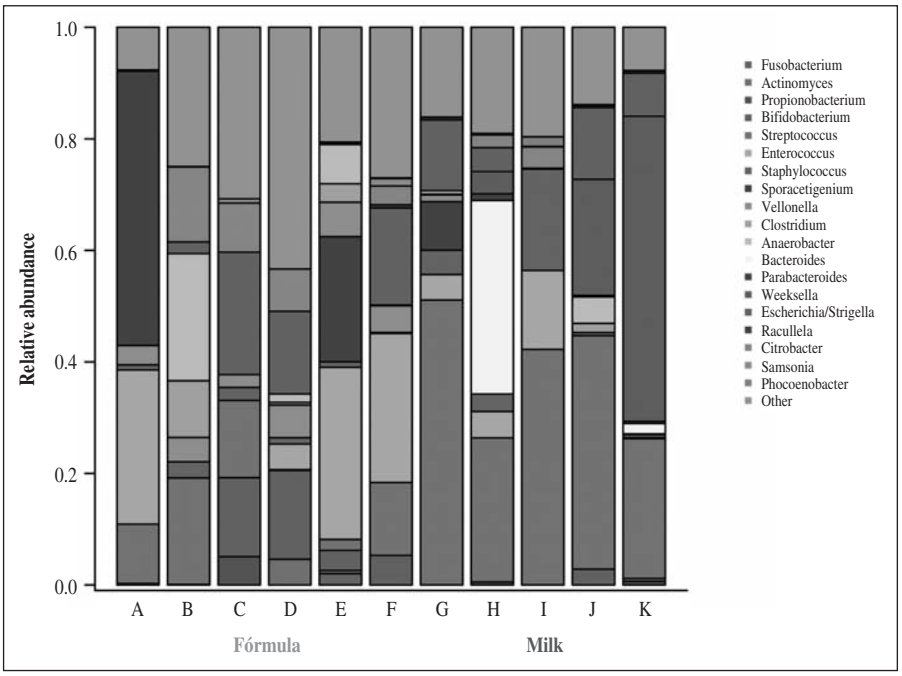

Fig. 3.—Distribution of bacterial genera in fecal samples of human milk- and formula-fed premature infants. The "Other" category combines all sequences that could not be identified at the genus level, or had counts fewer than 20, or that were contained in only a single individual

Table IV
Metabolites detected by GC/MS within dietary and/or fecal samples

Metabolites	Specific for		Detected in		Metabolites	Specific for		Detected in	
	*Diet	**Feces	Diet	Feces		*Diet	**Feces	Diet	Feces
Organic acids					*Amino Acids*				
2-Keto-L-gluconic acid	F		+	+	Phenylalanine	M		+	+
2-methyl-2-hydroxyglutaric acid	M		+	-	Proline	M		+	+
2-methylbenzoic acid	F		+	+	Pyroglutamic acid	M		+	+
2-methylmalic acid		FF	-	+	Threonine	M		+	+
4-hydroxybenzenepropanoic acid		BF	-	+	Tyrosine	M	BF	+	+
α-Ketoglutaric acid	M	BF	+	+	Valine	M		+	+
Alpha-Tocopherolacetic acid	F		+	-	*Sugar Alcohols*				
Aminomalonic acid		FF	-	+	Erythritol		FF	+	+
Arabino-hexonic acid		FF	-	+	Glycerol	M	BF	+	+
Aspartic acid	M	BF	+	+	Inositol, muco-		FF	-	+
Azelaic acid	M		+	-	Inositol, myo-		FF	+	+
Fumaric acid	F	FF	+	+	Inositol, scyllo-		FF	+	+
Galactonic acid		FF	+	+	Inositol-1-P		BF	-	+
Galacturonic acid		FF	-	+	Pentitol		FF	-	+
Glucuronic acid		BF	+	+	Pinitol		FF	-	+
Glutaric acid	M	FF	+	+	Sorbitol		FF	+	+
Glyceric acid	F	FF	+	+	Threitol		BF	-	+
Glycolic acid	F	FF	+	+	Xylitol		FF	+	+
Gulonic acid		BF	-	+	*Amines*				
Lactic acid		BF	+	+	Ethanolamine	M		+	+
Malic acid		BF	+	+	*Nucleosides*				
p-Coumaric acid		FF	-	+	Adenosine		BF	+	+
Pimelic acid	M		+	-					
Quinic acid		FF	-	+	*Purines*				
Ribonic acid	F		+	+	Hypoxanthine	F		+	-
Succinic acid	F	BF	+	+					
Threonic acid	F		+	+	*Pyrimidines*				
Xylonic acid-1,5-lactone	F		+	-	Cytosine		BF	+	+
					Uracil		BF	+	+
					Cholesta-5,24-diene-3-beta-ol	M		+	-
Amino Acids					Cholesterol	M		+	+
Glutamic acid	M		+	+					
Glycine	M		+	+	*Vitamines*				
Isoleucine	M		+	+	Orotic acid (vitamin B13)	F		+	-
Leucine	M		+	+	Panthotenic acid (vitamin B5)	F		+	+
Lysine		FF	+	+	Tocopherol (Vitamin E)	F	FF	+	+
N-Acetylglutamic acid	M	FF	+	+	*Phosphates*				
Ornithine		BF	+	+	Phosphoric acid	F		+	+

Table IV *(continuing)*
Metabolites detected by GC/MS within dietary and/or fecal samples

Metabolites	Specific for		Detected in		Metabolites	Specific for		Detected in	
	*Diet	**Feces	Diet	Feces		*Diet	**Feces	Diet	Feces
Sugars/sugar phosphates					*Saturated Fatty acids:*				
1,6-Anhydroglucose		FF	+	+	2,3-dihydroxypropyloctadecanoic acid	F		+	-
1-Methyl-beta-D-galactopyranoside	F		+	-	2,4,5-Trihydroxypentanoic acid		FF	+	+
2-O-Glycerol-alfa-d-galactopyranoside		BF	-	+	3,4-Dihydroxybutanoic acid	F	FF	+	+
					3-methyl-2-hydroxypentanoic acid		BF	-	+
Monosaccharides					4-methyl-2-hydroxypentanoic acid		BF	-	+
Arabinose (aldopentose)	F		+	+	Hexanoic acid	M		+	-
Xylose (aldopentose)	M		+	+	Hydroxypentanoic acid		BF	-	+
Fructose (ketohexose)	F		+	+	Nonadecanoic acid		BF	+	+
Fucose (deoxy sugar)	M		+	+	Octanoic acid triglyceride	F		+	-
Galactofuranose		FF	+	+	Pentacosanoic acid		FF	-	+
Galactopyranose		BF	+	+					
Galactose (aldohexose)	F	BF	+	+	*Carboxylic acids*				
Glucose (aldohexose)	F	BF	+	+	Decanoic acid		FF	-	+
Mannose (aldohexose)		BF	+	+	Docosanoic acid		FF	+	+
Sorbose (ketohexose)	F		+	-	Dodecanoic acid	M	FF	+	+
Rhamnose (deoxy sugar)	M	BF	+	+	Octanoic acid		FF	+	+
Sedoheptulose (ketoheptose)	M		+	-					
					Unsaturated Fatty acids				
Disaccharides					10-Nonadecenoic acid		BF	-	+
Lactose	M	BF	+	+	11-Hexadecenoic acid, (Z)-		BF	-	+
Maltose	F	BF	+	+	9-Octadecenoic acid nitrile	F		+	-
Melibiose	M		+	-	11,14,17-Eicosatrienoic acid		BF	+	+
Sucrose	F		+	-	9,12,15-Octadecatrienoic acid, (Z,Z,Z)-		FF	+	+
					9-tetradecenoic acid (Z)		FF	-	+
Trisaccharides					13,16-Docosadienoic acid		BF	-	+
Maltotriose		BF	+	+	9,12-Octadecadienoic acid (E,E)-		BF	-	+
Melezitose	F		+	-	9,12-Octadecadienoic acid (Z,Z)-	M	FF	+	+
					9-Hexadecenoic acid	M	BF	+	+
Amino sugars					11-Eicosenoic acid		BF	+	+
N-Acetyl glucosamine	F	BF	+	+	15-Tetracosenoic acid, (Z)-		BF	-	+
N-Acetyl mannosamine		BF	+	+	9-Octadecenoic acid (Z)-	M	FF	+	+
					9-Octadecenoic acid, (E)-		FF	-	+
Saturated Fatty acids									
12-methyltetradecanoic acid		BF	-	+					
12-methyltridecanoic acid		BF	-	+					
14-methylhexadecanoic acid		BF	-	+	*Other*				
1-Monohexadecanoylglycerol		BF	+	+	1-Methyl-4-hydroxy-1H-imidazol-2-amine	F		+	-
1-Monooctadecanoylglycerol	F		+	-	Urea	M		+	+
2,3-dihydroxypropylhexadecanoic acid	M		+	-					

M: Human milk; BF: Breast milk-fed; F: Premature infant formula; FF: Premature formula-fed.
*According to PLS-DA model (R2X = 0.76; R2Y = 0.99; Q2Y = 0.96); **According to PLS-DA model (R2X = 0.73, R2Y = 0.96, Q2Y = 0.92)

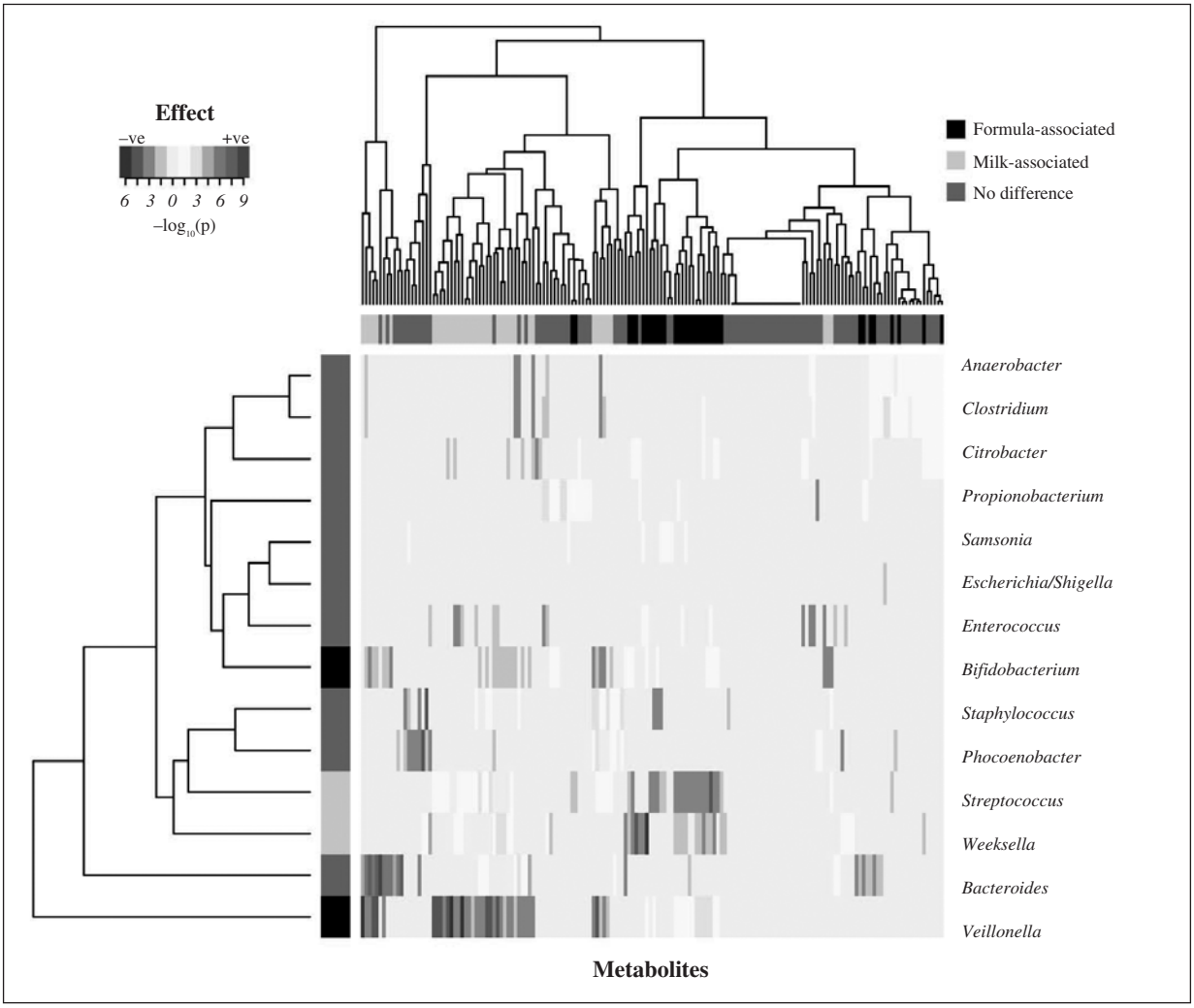

Fig. 4.—Heatmap of significant p-values associated with the relationship between metabolites and bacterial taxa in human stool samples. (For explanation see Figure 2 legend).

observed polyols or sugar alcohols, 8 (61.5%) were linked to the FF group, whereas only 3 (23%) are specific for BF and 2 (15%) (galactitol and ribitol) shared by both groups. Fatty acids represented the largest group of compounds identified in this study and fourth most homogenous. 56 total fatty acids were identified, of which 28 (50%) were seen in both groups, 12 (~21%) were specific to FF, and 16 (28%) to BF infants. Interestingly, the diversity of both types of fatty acids (saturated and unsaturated) was slightly greater among BF samples. Overall, the metabolic profiling demonstrates that the accumulation of chemical substrates within the lumen of the premature human intestine is a diet-dependent process. The question now is whether these different chemical environments select for particular bacterial communities.

Bacterial community profiling demonstrated low overall taxonomic diversity and high inter-individual variability (fig. 3), which is consistent with published reports.[1,7,24] Of the 10 bacterial phyla that have been observed within the human intestine,[33] only 5 *(Fusobacteria, Bacteroidetes, Proteobacteria, Actinobacteria, Firmicutes)* were detected in the infants studied. When PCA was applied to the relative abundances of gut microbes, the first two principal components explained 53% of the variation in the data. The PCA produced neat sample clustering matching diet type (fig. 1D) as seen for metabolites previously (fig. 1C). However PLS-DA created much lower predictive ability ($Q^2Y = 0.17$) compared to the data from the animal model. There was some suggestion that *Weeksella* and *Parabacteroides* were associated with the BF group, while *Veilonella* was associated with the FF group, but the low predictive ability prevents firm conclusions at this point. In general, the lower predictive ability of PLS-DA model supports the conclusion that bacterial colonization of the premature intestine is influenced by multiple factors, with diet explaining some of the variability. It is important to note that the premature infants in our study were delivered vaginally or by C-section, with unequal representation between the groups (table III). Although all subjects received antibiotics (table III) before start of our sur-

veys, it possible that the mode of delivery is still influenced our results.[1,34] Nevertheless, the analysis of significant associations between metabolites and bacterial taxa (fig. 4) determined metabolic/ecological niches of the dominant bacterial taxa within the diet categories. For example, milk provides a set of substrates positively correlated with the presence of the genera *Streptococcus* and *Weeksella* whereas formula provides a set of substrates positively correlated with the presence of the genera *Bifidobacterium* and *Veillonella*. Hence the clear separation of BF and FF groups, identical to the separation observed in a controlled for the mode of delivery piglet study, makes a strong statement toward the predominant role of diet in the establishment of juvenile intestinal ecosystem.

We propose that these data can be used to infer potential ecological interactions depending upon whether the metabolic niches of bacterial isolates overlap in the heat-map (interaction) or do not overlap (no interaction). Such considerations of species- and strain-specific metabolomes allow for the development of models of resource competition[35] between prebiotic-associated bacteria, pathogens and toxin-producing bacteria within naturally occurring microbial communities. These considerations may also suggest gut environments which alter bacterial behavior or pathogenesis. For example, a growing body of literature has revealed tight molecular links between carbohydrate utilization and expression of microbial virulence genes in gram positive bacteria, such as suppression of virulence factors by the catabolite control protein A in *Streptococcus pneumoniae*[29] and *S. pyogenes*.[30,31] Importantly, here, the close association between dominant bacterial taxa and particular diet associated metabolites confirms a strong role of diet in shaping bacterial communities even though large stochastic forces make it difficult to decipher consistent differences between the bacterial community compositions of infants in different treatment groups.

In this report, a metabolomic approach was successful in identifying the metabolites specific to the intestinal contents of BF and FF newborns in both, animal and human, studies. By supplying the juvenile gut with an array of specific substrates available for bacterial consumption, breast milk and infant formula clearly serve as distinct bacterial growth environments and select on different bacterial communities. Furthermore, the metabolomics approach to studying intestinal ecosystems generates testable hypotheses that may help explain the increased incidence of microbe-mediated complications in FF premature infants.

Acknowledgments

We thank the metabolomics center of UIUC Urbana-Champaign IL for high quality professional support, and the healthy volunteer donors that supplied breast milk for metabolomic analysis.

References

1. Palmer C, Bik EM, DiGiulio DB, David A. Relman DA, Brown PO. Development of the human infant intestinal microbiota. *PLoS Biol* 2007; 5: e177.
2. Mackie RI, Sghir A, Gaskins HR. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am J Clin Nutr* 1999; 69: 1035-1045.
3. Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H, et al. Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci USA* 2008; 105: 2117-2122.
4. Turnbaugh PJ, Gordon JI. The core gut microbiome, energy balance and obesity. *J Physiol* 2009; 587: 4153-4158.
5. Claud EC, Walker WA. Hypothesis: Inappropriate colonization of the premature intestine can cause neonatal necrotizing enterocolitis. *FASEB J* 2001; 15: 1398-1403.
6. Lin PW, Stoll BJ. Necrotising enterocolitis. *Lancet* 2006; 368: 1271-1283.
7. Magne F, Abély M, Boyer F, Morville P, Pochart P, Suau A. Low species diversity and high interindividual variability in faeces of preterm infants as revealed by sequences of 16S rRNA genes and PCR-temporal temperature gradient gel electrophoresis profiles. *FEMS Microbiol Ecol* 2006; 57: 128-138.
8. Schwiertz A, Gruhl B, Löbnitz M, Michel P, Radke M, Blaut M. Development of the intestinal bacterial composition in hospitalized preterm infants in comparison with breast-fed, full-term infants. *Pediatr Res* 2003; 54: 393-399.
9. Butel MJ, Suau A, Campeotto F, Magne F, Aires J, Ferraris L et al. Conditions of bifidobacterial colonization in preterm infants: A prospective analysis. *J Pediatr Gastroenterol Nutr* 2007; 44: 577-582.
10. Mandar R, Mikelsaar M. Transmission of mother's microflora to the newborn at birth. *Biol Neonate* 1996; 69: 30-35.
11. Grönlund MM, Gueimonde M, Laitinen K, Kociubinski G, Grönroos T, Salminen S et al. Maternal breast-milk and intestinal bifidobacteria guide the compositional development of the bifidobacterium microbiota in infants at risk of allergic disease. *Clin Exp Allergy* 2007; 37: 1764-1772.
12. Stewart JA, Chadwick VS, Murray A. Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children. *J Med Microbiol* 2005; 54: 1239-1242.
13. 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
14. Hällström M, Eerola E, Vuento R, Janas M, Tammela O. Effects of mode of delivery and necrotising enterocolitis on the intestinal microflora in preterm infants. *Eur J Clin Microbiol Infect Dis* 2004; 23: 463-470.
15. Alverdy JC, Laughlin RS, Wu L. Influence of the critically ill state on host-pathogen interactions within the intestine: Gut-derived sepsis redefined. *Crit Care Med* 2003; 31: 598-607.
16. Dethlefsen L, Huse S, Sogin ML, Relman DA. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol* 2008; 6: e280.
17. Harmsen HJ, Wildeboer-Veloo AC, Raangs GC, Wagendorp AA, Klijn N, Bindels JG et al. Analysis of intestinal flora development in breast-fed and formula-fed infants by using molecular identification and detection methods. *J Pediatr Gastroenterol Nutr* 2000; 30: 61-67.
18. Penders J, Vink C, Driessen C, London N, Thijs C, Stobberingh EE. Quantification of bifidobacterium spp., escherichia coli and clostridium difficile in faecal samples of breast-fed and formula-fed infants by real-time PCR. FEMS Microbiol. Lett. 2005;243:141-147.
19. Lundequist B, Nord CE, Winberg J. The composition of the faecal microflora in breastfed and bottle fed infants from birth to eight weeks. *Acta Paediatr Scand* 1985; 74: 45-51.
20. Saric J, Wang Y, Li J, Coen M, Utzinger J, Marchesi JR et al. Species variation in the fecal metabolome gives insight into differential gastrointestinal function. *J Proteome Res* 2008; 7: 352-360.
21. Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC et al. Metabolomics analysis reveals large effects of gut

microflora on mammalian blood metabolites. *Proc Natl Acad Sci USA* 2009; 106: 3698-3703.
22. Jansson J, Willing B, Lucio M, Fekete A, Dicksved J, Halfvarson J et al. Metabolomics reveals metabolic biomarkers of Crohn's disease. *PLoS One* 2009; 4: e6386.
23. Houle VM, Schroeder EA, Odle J, Donovan SM. Small intestinal disaccharidase activity and ileal villus height are increased in piglets consuming formula containing recombinant human insulin-like growth factor-I. *Pediatr Res* 1997; 42 (1): 78-86.
24. Wang Y, Hoenig JD, Malin KJ, Qamar S, Petrof EO, Sun J et al. 16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis. *ISME J* 2009; 3: 944-954.
25. Zoetendal EG, Heilig HG, Klaassens ES, Booijink CC, Kleerebezem M, Smidt H et al. Isolation of DNA from bacterial samples of the human gastrointestinal tract. *Nature Protocols* 2006; 1 (2): 870-873.
26. Poroyko V, White JR, Wang M, Donovan S, Alverdy J, Liu DC et al. Gut microbial gene expression in mother-fed and formula-fed piglets. *PLoS One* 2010; 5 (8): e12459.
27. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* 2007; 73 (16): 5261-5267.
28. Roessner U, Willmitzer L, Fernie AR. High-Resolution Metabolic Phenotyping of Genetically and Environmentally Diverse Potato Tuber Systems. Identification of Phenocopies. *Plant Physiology* 2001; 127: 749-764.
29. Giammarinaro P, Paton JC. Role of RegM, a homologue of the catabolite repressor protein CcpA, in the virulence of streptococcus pneumoniae. *Infect Immun* 2002; 70: 5454-5461.
30. Almengor AC, Kinkel TL, Day SJ, McIver KS. The catabolite control protein CcpA binds to pmga and influences expression of the virulence regulator mga in the group A streptococcus. *J Bacteriol* 2007; 189: 8405-8416.
31. Shelburne SA 3rd, Keith D, Horstmann N, Sumby P, Davenport MT, Graviss EA et al. A direct link between carbohydrate utilization and virulence in the major human pathogen group A streptococcus. *Proc Natl Acad Sci USA* 2008; 105: 1698-1703.
32. Pedersen LL, Turco SJ. Galactofuranose metabolism: A potential target for antimicrobial chemotherapy. *Cell Mol Life Sci* 2003; 60: 259-266.
33. Yang X, Xie L, Li Y, Wei C. More than 9,000,000 unique genes in human gut bacterial community: Estimating gene numbers inside a human body. *PLoS ONE* 2009; 4: e6074.
34. 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 USA* 2010; 107 (26): 11971-11975.
35. Fabich AJ, Jones SA, Chowdhury FZ, Cernosek A, Anderson A, Smalley D et al. Comparison of Carbon Nutrition for Pathogenic and Commensal Escherichia coli Strains in the Mouse Intestine. *Infect Immun* 2008; 76: 1143-1152.