

MATERIALS AND METHODS

Sample cohorts. Fecal samples, which were collected previously for the analysis of distal gut microbiota in healthy children and children diagnosed with IBS (Rigsbee et al 2012), were available for this project. Overall, 44 samples were used: 22 from healthy pre- and adolescent volunteers (designated kHLT for healthy kids; n=22; age range: 11-18 years, average: 12.6 years; 10 males, 12 females) and 22 samples collected from children newly diagnosed with diarrhea-predominant IBS (designated kIBS; n=22; age range: 8-18 years, average: 13.2 years; 10 males, 12 females). The study was approved by the Wright State University and Dayton Children's Hospital IRB committees. Diagnosis of the IBS and inclusion and exclusion criteria were described previously (Rigsbee et al 2012). Fresh fecal samples were (i) collected into sterile containers in the hospital during patient visit, (ii) homogenized immediately after collection, and (iii) aliquoted and frozen at -80°C. Fecal metabolites were extracted from stool aliquots shortly after collection.

Fecal water extract preparation. Similar to previous reports (Jacobs et al 2008, Le Gall et al 2011, Martin et al 2010, Pettersson et al 2008), fecal metabolites were analyzed in fecal water extracts prepared from each sample. A total of 250 mg of homogenized stool was suspended in 1.25 ml of sterile cold phosphate buffer (4.3 mM Na₂HPO₄•7H₂O, 1.5 mM KH₂PO₄, 2.7 mM KCl). The mixture was homogenized for 5 minutes and then centrifuged at 16,000 g for 5 minutes. The supernatant was collected and filtered through a GDX syringe filter (10.0 µm - 0.2 µm pore size). The filtrate was centrifuged again at 16,000 g for 15 minutes, and the supernatant was retained and stored at -80 °C for subsequent analyses.

The pH level of each stool sample was determined with BDH pH test strips. The water content of each fecal sample was measured by weighing 50 mg of the stool in a microcentrifuge tube, dehydrating the sample in a vacuum evaporator for 30 minutes, and reweighing the stool aliquot. The water content was defined as $1 - [\text{dry weight}] / [\text{wet weight}]$.

Proton nuclear magnetic resonance (NMR) analysis of fecal extracts. A 550 µl aliquot of the prepared fecal extract sample was transferred to a 5 mm NMR tube together with 150 µl of 9 mM trimethylsilylpropionic-2,2,3,3-d₄ acid (TSP) in D₂O. Proton (¹H) NMR spectra were acquired at 25 °C using a Varian INOVA operating at 600 MHz (14.1 Tesla). TSP served as a chemical shift reference and quantification standard, and D₂O provided a field-frequency lock for NMR acquisition. Water suppression was achieved using the first increment of a NOESY pulse sequence, which incorporated saturating irradiation (on-resonance for water) during the relaxation delay (2 s pre-saturation) and the mixing time (50 ms). Data were signal averaged over 400 transients using a 4.0 s acquisition time and interpulse delay of 11.05 s. Spectral data were pre-processed using Varian software that employed exponential multiplication (0.3 Hz line-broadening), Fourier transformation, and phase correction. Spectra were then baseline corrected (flattened) in MATLAB (The Mathworks, Inc.) using the Whittaker Smoother algorithm ($\lambda=200$) (Eilers 2003, Whittaker 1923). Further spectral processing included removal of the residual water signal, chemical shift referencing, and sum normalization. For multivariate data analyses, spectra were binned to reduce the dimensionality and mitigate peak misalignment, and signal intensities were autoscaled. A dynamic programming-based adaptive binning technique was employed (Anderson et al 2011) using a minimum and maximum distance between peaks in a single bin of 0.001 and 0.04 ppm, respectively. Bin boundaries were then manually adjusted to further mitigate

peak misalignment, and to keep known J-coupled multiplets (e.g., doublets, triplets, etc) within the same bin.

Quantification of specific metabolite resonances was accomplished using an interactive spectral deconvolution algorithm in MATLAB adapted from our previously described methods (Anderson et al 2011). The deconvolution tool fits a defined spectral region using a combination of tunable baseline shapes (spline, v-shaped, linear, or constant) and a Gauss-Lorentz peak-fitting function. All metabolite peak intensities were corrected for equivalent number of protons and normalized relative to the TSP signal intensity; signals were not corrected for partial T1 (relaxation) saturation. Since all extracts were prepared from an identical stool weight, relative comparisons between samples are fully valid, but absolute quantitative amounts of each metabolite are not reported. We used a combination of three sources to assign peaks to specific metabolites – (1) database of proton NMR peaks assigned to specific small compounds (such as Human Metabolome Database), (2) literature that defines specific peaks to belong to specific compounds, and (3) the above tentative assignments were confirmed by addition of the suspect compound (spiking) to a test extract sample, carrying out proton NMR spectrum acquisition, and identifying corresponding peaks.

Microbiota composition dataset. The dataset of relative species and genus abundances measured in the same set of fecal samples was available from our previous study (Rigsbee et al 2012). Phylogenetic Microbiota Array was used to obtain quantitative phylotype and genus abundance values; the normalized phylotype signal values were adjusted for estimated cross-hybridization amount and unequal 16S rRNA gene copy number as described (Rigsbee et al 2011). Microarray results were corroborated with high-throughput sequencing, qPCR, and FISH. The full dataset of phylotype abundances is available at <http://www.nature.com/ajg/journal/v107/n11/supinfo/ajg2012287s1.html>. This dataset was used in all microbiota-metabolite comparative analyses.

Statistical data analyses. Analyses of proton NMR data included multivariate data analyses of the full binned spectra and of 19 measured fecal metabolites. To reduce sample-to-sample within-group variance, the normalized binned data were autoscaled to kIBS group mean (van den Berg et al 2006). To obtain average levels of metabolites among all samples within each sample type, weighted mean around median was calculated as we described previously (Agans et al 2011). The complete dataset of metabolite values is provided in supplementary materials (Table S1).

Mann-Whitney U test and general linear modeling (GLM) were run in SPSS (SPSS, Inc.). Principal components analysis (PCA), correspondence analysis (CA), canonical discriminant analysis (CDA), orthogonal projections to latent structures discriminant analysis (OPLS-DA) were run in MATLAB utilizing custom scripts (Shankar et al 2013). Statistical significance of group separation in the CA and CDA analyses was obtained through permutation of sample group assignment (Shankar et al 2013); PCA separation was tested based on the permutation of the Davies-Bouldin index measure (Davies and Bouldin 1979). Procrustes analysis was performed in MATLAB based on the previously described approach (McHardy et al 2013).

To assess putative associations between metabolite and microbiota profiles, non-parametric Spearman rank correlation was utilized. The pair-wise correlation matrix was computed on the metabolite-genus abundance dataset; to account for multiple hypothesis testing, R_s significance values were adjusted based on the Benjamini and Hochberg false discovery rate (FDR) method (Benjamini and Hochberg 1995). FDR was set to 10%. Hierarchical bi-clustering of the resulting correlation matrix was

used to visualize metabolite-genus associations. The network of significant correlations was visualized as a bipartite graph in NAVIGATOR (Brown et al 2009).

REFERENCES

Agans R, Rigsbee L, Kenche H, Michail S, Khamis HJ, Paliy O (2011). Distal gut microbiota of adolescent children is different from that of adults. *FEMS Microbiol Ecol* **77**: 404-412.

Anderson PE, Mahle DA, Doom TE, Reo NV, DelRaso NJ, Raymer ML (2011). Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data. *Metabolomics* **7**: 179-190.

Benjamini Y, Hochberg Y (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* **57**: 289-300.

Brown KR, Otasek D, Ali M, McGuffin MJ, Xie W, Devani B *et al* (2009). NAViGaTOR: Network Analysis, Visualization and Graphing Toronto. *Bioinformatics* **25**: 3327-3329.

Davies DL, Bouldin DW (1979). A cluster separation measure. *IEEE Trans Pattern Anal Mach Intell* **1**: 224-227.

Eilers PHC (2003). A perfect smoother. *Anal Chem* **75**: 3631-3636.

Jacobs DM, Deltimple N, van Velzen E, van Dorsten FA, Bingham M, Vaughan EE *et al* (2008). (1)H NMR metabolite profiling of feces as a tool to assess the impact of nutrition on the human microbiome. *NMR in biomedicine* **21**: 615-626.

Le Gall G, Noor SO, Ridgway K, Scovell L, Jamieson C, Johnson IT *et al* (2011). Metabolomics of Fecal Extracts Detects Altered Metabolic Activity of Gut Microbiota in Ulcerative Colitis and Irritable Bowel Syndrome. *Journal of proteome research* **10**.

Martin FP, Sprenger N, Montoliu I, Rezzi S, Kochhar S, Nicholson JK (2010). Dietary modulation of gut functional ecology studied by fecal metabolomics. *Journal of proteome research* **9**: 5284-5295.

McHardy IH, Goudarzi M, Tong M, Ruegger PM, Schwager E, Weger JR *et al* (2013). Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome* **1**.

Pettersson J, Karlsson PC, Choi YH, Verpoorte R, Rafter JJ, Bohlin L (2008). NMR metabolomic analysis of fecal water from subjects on a vegetarian diet. *Biol Pharm Bull* **31**: 1192-1198.

Rigsbee L, Agans R, Foy BD, Paliy O (2011). Optimizing the analysis of human intestinal microbiota with phylogenetic microarray. *FEMS Microbiol Ecol* **75**: 332-342.

Rigsbee L, Agans R, Shankar V, Kenche H, Khamis HJ, Michail S *et al* (2012). Quantitative profiling of gut microbiota of children with diarrhea-predominant Irritable Bowel Syndrome. *Am J Gastroenterol* **107**: 1740-1751.

Shankar V, Agans R, Holmes B, Raymer M, Paliy O (2013). Do gut microbial communities differ in pediatric IBS and health? *Gut Microbes* **4**: 347-352.

van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC genomics* **7**: 142.

Whittaker E (1923). On new method of graduation. *Proceedings of the Edinburgh Mathematical Society* **41**: 63-75.