ORIGINAL ARTICLE

Gut bacteria–host metabolic interplay during conventionalisation of the mouse germfree colon

Sahar El Aidy^{1,2}, Muriel Derrien^{1,2,10}, Claire A Merrifield³, Florence Levenez⁴, Joël Doré⁴, Mark V Boekschoten⁵, Jan Dekker^{1,6}, Elaine Holmes³, Erwin G Zoetendal^{1,2}, Peter van Baarlen⁷, Sandrine P Claus⁸ and Michiel Kleerebezem^{1,2,7,9}

¹Top Institute Food and Nutrition, Wageningen, Wageningen, The Netherlands; ²Laboratory of Microbiology, Wageningen University, Dreijenplein, Wageningen, The Netherlands; ³Biomolecular Medicine, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK; ⁴INRA, UMR1319, F-78350, Jouy-en-Josas France; ⁵Nutrition, Metabolism and Genomics Group, Division of Human Nutrition, Wageningen University, Wageningen, The Netherlands; ⁶Department of Animal Sciences, Wageningen University, Wageningen, The Netherlands; ⁷Host-Microbe Interactomics, Wageningen University, Wageningen, The Netherlands; ⁸Department of Food and Nutritional Sciences, The University of Reading, Reading, UK and ⁹Health Department, NIZO food research, Ede, The Netherlands

The interplay between dietary nutrients, gut microbiota and mammalian host tissues of the gastrointestinal tract is recognised as highly relevant for host health. Combined transcriptome, metabonome and microbial profiling tools were employed to analyse the dynamic responses of germfree mouse colonic mucosa to colonisation by normal mouse microbiota (conventionalisation) at different time-points during 16 days. The colonising microbiota showed a shift from early (days 1 and 2) to later colonisers (days 8 and 16). The dynamic changes in the microbial community were rapidly reflected by the urine metabolic profiles (day 1) and at later stages (day 4 onward) by the colon mucosa transcriptome and metabolic profiles. Correlations of host transcriptomes, metabolite patterns and microbiota composition revealed associations between Bacilli and Proteobacteria, and differential expression of host genes involved in energy and anabolic metabolism. Differential gene expression correlated with *scyllo*- and *myo*-inositol, glutamine, glycine and alanine levels in colonic tissues during the time span of conventionalisation. Our combined time-resolved analyses may help to expand the understanding of host–microbe molecular interactions during the microbial establishment.

The ISME Journal (2013) 7, 743–755; doi:10.1038/ismej.2012.142; published online 22 November 2012

Subject Category: microbe-microbe and microbe-host interactions

Keywords: C57; BL6 J germfree mice; microbiota; transcriptome; metabonome

Introduction

The mammalian gastrointestinal tract is a home to an estimated 100 trillion microbial cells, representing the largest microbial community associated with the mammalian body (Savage, 1977; Lee and Mazmanian, 2010). This complex ecosystem provides a vast reservoir of metabolic capabilities that complement the metabolism of the host (Bäckhed *et al.*, 2004; Nicholson *et al.*, 2005; Turnbaugh *et al.*, 2009) and has a crucial role in the developmental and nutritional processes in the intestine (Kau *et al.*, 2011). The colon is the most prominent site of

microbial colonisation, with bacterial densities reaching up to 10^{12} per gram of content (O'Hara and Shanahan, 2006). Nevertheless, the effects of the intestinal microbiota can be observed in diverse regions of the host. 'Top-down' systems biology using metabolic profiling of conventional mice revealed large, systemic effects of the microbial community on absorption, storage and metabolism of dietary compounds (Martin *et al.*, 2007; Claus *et al.*, 2008). These, and other studies, have demonstrated that host metabolism is responsive to intestinal–microbial metabolism, which provides complementary pathways for resorption and assimilation of dietary ingredients and drugs (Nicholson *et al.*, 2005).

In recent years, a clear influence of changes in the human diet on gut microbiota has been shown, to the extent that there appears to have been a co-evolution between the diet, human and the composition of the human microbiota (Walter and Ley,

Correspondence: M Kleerebezem, Health Department; NIZO food research, PO Box 20, 6710 BA Ede, The Netherlands.
E-mail: michiel.kleerebezem@nizo.nl

¹⁰Present address: Danone Research, 91767 Palaiseau, France.
Received 27 April 2012; revised 17 September 2012; accepted 28 September 2012; published online 22 November 2012

2011). The microbiota is composed of well-established, resident bacteria that form long-term associations with the host, as well as transient bacteria that do not colonise the gastrointestinal tract permanently. This large and dynamic community undergoes dramatic changes after initial colonisation of the neonates (Hooper, 2004; Palmer *et al.*, 2007). During colonisation, all microbes compete and have to resist the host's defence systems (Ley *et al.*, 2006). Stable establishment of microbial groups requires cooperation in food networks, such as cross-feeding where metabolites from one organism act as a substrate for another (Duncan *et al.*, 2004; Fischbach and Sonnenburg, 2011).

In the intestinal–microbial ecosystem, the most prominent microbial activity is the fermentation of dietary or host-derived components, in particular, the conversion of non-digestible carbohydrates and host glycans into short chain fatty acids (SCFAs) (Bäckhed *et al.*, 2004). SCFAs activate a G-protein coupled receptor 43, which has an important role in immune modulation (Maslowski *et al.*, 2009), and has a key-role in the regulation of energy balance (Bjursell *et al.*, 2011). In addition to fermentation of dietary or host glycans, the microbiota synthesises essential vitamins such as vitamin K and certain B vitamins; these vitamins have to be supplemented to the feed of germfree animals (Hooper *et al.*, 2002). In fact, the co-evolution of the host–microbe interaction has enabled mammals to harvest nutrients from novel sources (Bäckhed *et al.*, 2005; Walter and Ley, 2011). In return, the gut microbiota are provided with a nutrient-rich niche that enables bacterial growth (Hooper *et al.*, 2002), and that has been proposed to provide high-affinity adhesion sites for specific microorganisms to accommodate their persistence in the intestine (Bäckhed *et al.*, 2005), contributing to colonisation resistance via antagonism against incoming pathogenic bacteria (Salysers and Pajeau, 1989; Hultgren *et al.*, 1993).

There is an increasing evidence for an association between changes in proper microbial colonisation (dysbiosis) and development of human disease (Neish 2009). Dysbiosis of the microbial community has been associated with a variety of diseases including inflammatory bowel disease and colon cancer (Azcarate-Peril *et al.*, 2011) along with systemic diseases (Huycke and Gaskins 2004) such as obesity (Bäckhed *et al.* 2004; Bäckhed *et al.* 2007). In view of the increasing awareness of disease-associated shifts in intestinal microbiota communities, it is important to improve our understanding of the molecular basis and dynamics of homeostatic host–microbe interactions. To this end, we aimed to monitor the succession of microbial colonisation of germfree mice and the corresponding changes in the host–microbe metabolic relationships. Following conventionalisation, the time-resolved composition of colonic microbial communities was determined in parallel with colon mucosa transcriptomes, and ^1H NMR (hydrogen-1 nuclear

magnetic resonance) spectroscopic profiling of the colonic tissue and urine. Statistical modelling showed correlations between microbial diversity and host transcriptome and metabolome, and allowed us to reconstruct a comprehensive overview of the transient and more permanent alterations in the symbiotic host–microbe relationship.

Materials and methods

Animals, experimental design and sampling

All procedures were carried out according to the European guidelines for the care and use of laboratory animals and with permission 78–122 of the French Veterinary Services. Germfree and conventionalised mice (male, C57 BL/6J) were maintained in sterile conditions, on a commercial laboratory chow diet. After 2 weeks of acclimatisation and diet adaptation, a first set of germfree mice ($n=3$) were randomly assigned to killing by oral anaesthesia using isoflurane. The remaining germfree mice were conventionalised by oral gavage with 0.5 ml of mixed faecal suspension obtained from 0.2 g of freshly obtained faecal material of conventionally raised mice (C57 BL/6J) diluted 100-fold in Brain Heart infusion broth. Following conventionalisation, the mice were maintained in the isolator in standard cages with six mice per cage, until the moment of killing. Two independent biological experiments were performed using mice of different age after 2 weeks of acclimatisation and diet adaptation. The first and second experiments included 36 mice obtained in two biologically independent batches of 18 mice each ($n=3/\text{batch}/\text{day}$), aged 8 and 10 weeks, respectively. The colon from each mouse was removed and divided into 2-cm segments that were immediately stored in RNA later (Invitrogen, Bleiswijk, The Netherlands) at room temperature for 1 h prior to storage at -80°C for RNA isolation or snap frozen and stored at -80°C for metabolic profiling. Luminal content from colonic segments and caecum was removed by gentle squeezing, snap frozen and stored at -80°C for microbiota analysis and SCFAs high-performance liquid chromatography (HPLC) analyses, respectively. Urine samples (50–70 μl) were collected from the bladder directly after killing and stored at -80°C for ^1H NMR metabolites analysis.

Microbial profiling of colonic contents

Luminal contents from colon were analysed by Mouse Intestinal Tract Chip (MITChip), a diagnostic 16S rRNA array that consists of 3580 unique probes especially designed to profile murine gut microbiota as previously described and in analogy to the human intestinal tract chip (Rajilić-Stojanović *et al.*, 2009; Geurts *et al.*, 2011). Statistical analysis was performed between all conventionalisation days using the Kruskal–Wallis test executed in SPSS Statistics

17.0 (SPSS Inc., Chicago, IL, USA) (for more details see supplementary methods).

Microbial fermentation product analysis. Caecal content samples were analysed for SCFAs profiles, including the quantitative detection of acetate, butyrate, propionate as well as lactate and succinate using HPLC (The detector was a SpectraSYSTEM RI-150 Refractive Index). Samples of intestinal content (~0.1 g) were thoroughly mixed with four volumes of distilled water. Insoluble residue was removed by centrifugation (15 min at 13 000 g, 4 °C). The supernatant was mixed with the same volume of 1 M HClO₄ and mixed organic acid analyses by HPLC as previously described (Starrenburg and Hugenholtz, 1991) (for more details see supplementary methods).

Transcriptome analysis

High quality total RNA (RNA integrity number >8.0) was isolated from a 2-cm segment colon by extraction with TRIzol reagent, followed by DNase treatment and column purification. Samples were hybridised on Affymetrix GeneChip Mouse Gene 1.1 ST arrays. Complementary methods were used for the biological interpretation of the transcriptome data; gene clustering using multi-experimental viewer (Saeed *et al.*, 2006), overrepresentation analysis of gene ontology (GO) terms using temporal and location comparative analysis using STEM (Short Time-series Expression Miner) (Ernst and Bar-Joseph, 2006) and construction of biological interaction networks using Ingenuity Pathways Analysis are described in detail in the supplementary methods.

Metabolite profiling

Urine samples (25–30 µl) were added to 50 µl of 0.2 M phosphate buffer (pH 7.4) in D₂O plus 0.05% 3-(trimethylsilyl) propionate-2,3-d₄ before transferring to capillary tubes for analysis by ¹H NMR spectroscopy. Tissue samples were homogenised and extracted in acetonitrile/water (1:1), as previously described (Waters *et al.*, 2002). The supernatant containing the aqueous phase was collected, subsequently freeze-dried and dissolved in 600 µl of D₂O. Samples were centrifuged for 10 min at 15 000 g, and 500 µl of the supernatant plus 50 µl of water were transferred into 5 mm (outer diameter) NMR tubes for analysis by ¹H NMR spectroscopy. All NMR spectra were digitalised and imported to Matlab (version R2011a, MathWorks) for statistical analysis. All data were first visualised by Principal Component Analysis in order to identify potential outliers. Orthogonal partial least-square discriminant analysis (OPLS-DA) models (Trygg, Svante 2003) were then fitted between successive time-points in order to highlight discriminant metabolites. Principal Component Analysis, O-PLS, O-PLS-DA and Statistical Total Correlation Spectroscopy were performed using an in-house routines (Cloarec *et al.*, 2005).

O2-PLS models were used to integrate metabonome, transcriptome and microbiota data sets as previously described (Li *et al.*, 2008). O2-PLS models were calculated for pair-wise data sets. Significant variables were then selected based on their correlation with the scores of the model ($P < 0.01$) (for detailed descriptions see supplementary material).

Accession numbers

The mouse microarray data set was deposited in NCBI Gene Expression Omnibus with accession number GSE32513.

Results

Dynamic establishment of microbial communities

The colonic microbiota composition of conventionalised mice was assessed for five time-points post-conventionalisation using the MITChip phylogenetic platform (Geurts *et al.*, 2011). The colonic microbiota of conventionalised mice proceeded through two major stages discriminated by low diversity of the microbiota during early stages (days 1–2) and increasing microbial diversity at later stages (days 8 and 16) of conventionalisation, ultimately reaching levels similar to those observed in conventional animals (El Aidy *et al.*, 2012). MITChip analysis confirmed that this early-to-late shift is characterised by early, rapid colonisers (days 1–2) belonging predominantly to the phyla Bacteroidetes, Firmicutes (notably the class Bacilli), Proteobacteria and Actinobacteria, whereas at later stages (days 4, 8 and 16 post-conventionalisation) specific subgroups of the Firmicutes, particularly the members of *Clostridium* clusters IV and XIVa increased in abundance, (Figures 1a and b). Days 1 and 2 post-conventionalisation were characterised by relatively high abundance of genera-like *Enterococcus* (*Enterococcus urinaeequi* *et rel.*), *Bacteroides fragilis* *et rel.*, *Prevotella*, and *Lactobacillus salivarius* (Supplementary Table 1). At day 4 post-conventionalisation, a transient abundance of Proteobacteria was detected, whereas days 8 and 16 post-conventionalisation were typically characterised by an increased abundance of genera such as *Dorea*, *Butyrivibrio crossotus* *et rel.*, and unclassified TM7 (Supplementary Table 1). Taken together, MITChip analysis identified changes in the composition of the microbial communities that had established in the colon of mice over time during conventionalisation. Moreover, the MITChip analysis identified the dynamic abundance of the microbial groups that comprised the climax community. This community appeared to be established from day 8 post-conventionalisation onward, and strongly resembled the inoculum (conventional) microbiota in terms of composition and abundance. The established microbial community was characterised by large numbers of anaerobes and distinct

Figure 1 Dynamics of the colonising microbiota in the colon. (a) Hierarchical clustering analysis of MITChip fingerprints generated from the inoculum and colonic samples collected from days 1–16 post-conventionalisation ($n=5-6$ mice/time-point). The highest phylogenetic assignments of probe specificity are provided at the right side of the gel-view dendrogram. (b) Dynamics of the relative contribution of different level 1 (class like) microbial groups to the total microbiota in the colon of mice at different time-points post-conventionalisation.

colonisation patterns of the major groups Bacteroidetes and *Clostridium* cluster IV and XIVa.

In addition to determination of the phylogenetic composition of the microbiota, gene-specific quantitative PCR was employed to quantify the relative abundance of the 16S rRNA gene and of two well-studied marker genes that represent the pathways of sulphate reduction (dissimilatory sulphite reductase (*dsr*) gene; Ben-Dov *et al.*, 2007) and methanogenesis (methyl coenzyme-M reductase (*mcrA*) gene; Steinberg and Regan, 2008). Expression levels of *dsr* and *mcrA* exemplify the succession of hydrogen-utilizing species (sulphate reducers and methanogens), which ensure efficient H_2 removal to maintain fermentation balance in the colon (Gibson *et al.*, 1993). Butyrate producers were detected by the most common gene involved in intestinal butyrate production; *butyryl-CoA-transferase* (Louis and Flint, 2007). The overall microbial community size as estimated by 16S rRNA gene copy number per gram of colonic content was stable at a level of $\sim \log_{10}-11.63 \pm 0.59$ during conventionalisation. The relative abundance of the *mcrA* gene also appeared to be

stable and stayed at low levels ($\log_{10}-4.82 \pm 0.82$ at day 1, $\log_{10}-5.2 \pm 0.26$ at day 2, $\log_{10}-5.15 \pm 0.76$ at day 4, $\log_{10}-3.82 \pm 1.18$ at day 8 and $\log_{10}-4.93 \pm 0.64$ at day 16) over the entire duration of the conventionalisation, suggesting that the methanogen population was among the early colonisers and had already reached its final population size early in the experiment. Although the relatively low absolute abundance of the *mcrA* gene may be an underestimation due to DNA isolation procedures that were not optimised for lysis of the methanogens (Dridi *et al.*, 2009), the stability of the gene's abundance in the microbiota is most probably an accurate reflection of the evolution of this microbial group over time. In contrast, the *dsr* gene appeared to be ~ 100 -fold more abundant in the ecosystem during later stages of the conventionalisation (Figure 2a), increasing from $\sim \log_{10}-2.54 \pm 0.2$ during days 1–2 to $\log_{10}-4.4 \pm 0.86$ at days 4, 8 and 16 post-conventionalisation. In addition, the *butyryl-CoA-transferase* gene abundance in the ecosystem increased more than 100-fold from $\sim \log_{10}-9.3 \pm 0.42$ at days 1–2 to $\log_{10}-11.87 \pm 0.78$ on days

Figure 2 Dynamics of specific functional gene abundance in the microbiota. (a) Quantification of butyrate producers and sulphate reducers expressed as mean $\pm$ s.d. log₁₀ number of *Butyryl-CoA-transferase* (●) and *dsr* (◆) genes/g content, respectively. Statistical analysis was performed using a one-way analysis of variance test executed in SPSS Statistics 17.0 (SPSS Inc., Chicago, IL, USA). Significant differences between time-points are indicated by distinct characters above the measurement groups ($P < 0.05$). (b) HPLC analysis of the large intestinal content for SCFAs including acetate, butyrate, propionate, lactate and succinate, ($n = 5$ –6 mice/time-point).

8 and 16 post-conventionalisation, whereas the microbial ecosystem on day 4 post-conventionalisation appeared to contain an intermediate abundance of this gene (log₁₀–10.5 $\pm$ 0.96). These results are in accordance with phylogenetic analysis of the community where the relative populations of typical sulphate-reducing organisms such as *Bilophila* and *Desulfovibrio* and butyrate-producing organisms such as *Roseburia intestinalis* *et rel.*, *Subdoligranulum*, *Faecalibacterium prausnitzii* *et rel.* and *Butyrivibrio* spp. (Supplementary Table 1) increased during later stages of conventionalisation (days 4,

8 and 16 for the sulphate reducers and days 8 and 16 for butyrate producers).

Dynamics of microbial fermentation end-products

The phylogenetic analysis predicted a change in global fermentative capacities of the microbiota, with a shift towards increased production of butyrate at later time-points. To assess the fermentative capacities of the successive microbial communities, concentrations of the SCFAs, acetate, propionate, butyrate, as well as lactate and succinate were determined in caecal contents. Especially

lactate and succinate, together with relatively high amounts of acetate and propionate, dominated the microbial fermentation-profile measured during the early stages of conventionalisation (days 1–2 and to lesser extent day 4), whereas butyrate concentrations were below the detection limit ($0.11 \mu\text{mol mg}^{-1}$). The concentrations of fermentation metabolites drastically shifted at later stages of conventionalisation (days 8 and 16 post-conventionalisation) when lactate and succinate levels were decreased below the detection limit and acetate and propionate were significantly increased (Figure 2b). Moreover, butyrate was clearly detected at these later stages, which correlates very well with the increased abundance of butyrate producers (as shown above, Supplementary Figure 1).

Colon mucosa transcriptome profiling focusing on metabolic functions

Previously we established a time- and region-dependent modulation of gene expression profiles during conventionalisation, in which the colon

mucosa transcriptomes supported the establishment of a novel state of immune homeostasis within 16 days of conventionalisation (El Aidy *et al.*, 2012). Genes belonging to the GO category 'metabolic processes' showed time-dependent differential expression in the colon; this GO category was significantly enriched ($P < 0.001$). The genes belonging to this GO category were further analysed by Ingenuity Pathway Analysis software, which illustrated their involvement in a variety of metabolism-associated pathways (see below; Supplementary Figure 2). To precisely predict possible metabolic consequences of the changing metabolic gene expression profiles during conventionalisation, gene expression changes were projected onto Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic maps (www.genome.jp/kegg/) (Figure 3a). These projections showed a significant modulation of glycolytic, amino acid and nucleotide metabolic pathways on days 4, 8 and 16, in comparison to day 0 (germfree mice) and earlier days (1 and 2) post-conventionalisation. The significant induction of nucleotide synthesis and metabolic pathways were

Figure 3 Altered colonic metabolites during conventionalisation. (a) Metabolic pathway map including the genes that participate in certain metabolic pathways conversions. Genes are indicated in box symbols; each box is divided into five sub-boxes with colour codes, representing the changes observed at days 1, 2, 4, 8 and 16 post-conventionalisation, respectively. Direct and indirect interactions are depicted by solid and dashed arrows respectively, ($n = 6$ mice/time-point). Predicted altered metabolites are indicated with orange and purple arrows. (b) Heat map summarising the colonic metabolic variation during conventionalisation. A series of pair-wise OPLS-DA models were constructed for colon tissue. (+) refers to tentative assignment.

illustrated by the induction of thymidylate synthase (*Tyms*) and ribonucleotide reductase (*Rrm1* and *2*), which is an essential enzyme for the production of deoxyribonucleotides (Parker *et al.*, 1995). The colonic metabolism also underwent time-dependent changes in expression of genes involved in amino-acid metabolism, including the glutamine and glutamate-associated pathways. This was apparent from the induction of *Asns* (asparagine synthetase), and the repression of the *Glul* (glutamine synthetase), *Glud1* (mitochondrial glutamate dehydrogenase), and *Abat* (4-aminobutyrate aminotransferase or GABA transaminase), which converts 4-aminobutyrate into oxaloglutarate and glutamate. The strong downregulation of the genes encoding *Sord* (sorbitol dehydrogenase), *Sis* (sucrase isomaltase) and *Gbe1* (glucan (1,4- α -), branching enzyme 1), suggest that carbohydrate metabolism is repressed. Moreover, inositol and choline metabolism appeared to be repressed, as inferred from decreased expression of *Chpt1* (choline phosphotransferase 1) and *Chkb* (choline kinase beta). *Fut2* (fucosyltransferase 2) and *B3galt5*

(UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 5) involved in glycosphingolipid biosynthesis were upregulated from day 1 post-conventionalisation onward (Figure 3a). These transcriptome changes illustrate the dynamic changes in the expression of metabolic pathway genes in the colon mucosa upon conventionalisation, which encompass a broad area of intracellular (amino acid, glycolysis and nucleotide metabolism) as well as membrane (sphingolipid) and extracellular matrix (glycan biosynthesis) metabolic pathways.

Dynamics of local and systemic metabolic profiles

The projections of the differentially expressed genes on KEGG metabolic maps shown above, suggested that microbial colonisation would lead to changes in

the abundance of those colon tissue metabolites that serve as substrate or product of the enzymes encoded by the corresponding gene differentials (Figure 3a). Metabolic phenotyping using ^1H NMR, was performed on colonic tissue samples from germfree and conventionalised mice to investigate the impact of the colonizing gut microbiota on host colonic metabolism during conventionalisation. The time-resolved metabolite data sets were used to build OPLS-DA models, focusing on the differences between germfree and conventionalised animals over time. The analysis showed that the presence of specific microbial groups was not significantly reflected by the colonic mucosa metabolic profiles during the first two days of conventionalisation (days 1–2). During later time-points (days 8 and 16), the microbial colonisation was detectable in the mucosal metabolite profiles through significantly increasing levels of alanine, fumarate, glycine, uracil and methylmalonate† († = tentative assignment), and decreasing levels of glycerol, glucose and formate. Some, but not all, changes in these metabolic profiles appeared to be initiated at day 4 post-conventionalisation (Figure 3b). In addition, concentrations of several mucosal metabolites appeared to transiently respond to conventionalisation, including modulated levels of acetate, aspartate, glutamine and five other metabolites (Figure 3b).

Analogously, metabolic phenotyping using ^1H NMR, was employed on urine samples as a proxy for systemic metabolic changes. Contrary to the tissue metabolite data, urine metabolite profiles appeared to reflect the consequences of microbial colonisation during the early time-points (days 1–2), where the excretion of metabolites such as creatine and formate were increased on day 1 and days 1–2 post-conventionalisation, respectively (Figure 4a). Whereas concentrations of the urine metabolites 2-hydroxy-3-methylvalerate, trimethylamine, phenylacetylglutamine, TMAO, B-aminobutyrate and B-hydroxybutyrate

Figure 4 (a) Heat map summarising the urine metabolic variation during conventionalisation. A series of pair-wise OPLS-DA models were constructed for urine metabolite profiles with significant modulation during the days (d) 1–16 post-conventionalisation period, compared to germfree. (†) refers to tentative assignment. Purple arrows refer to metabolites as predicted from the projections of differentially expressed metabolic genes on KEGG maps. Orange asterisks refer to metabolites that have been previously reported to be exclusively produced by microbial metabolism. (b) Correlation heat map between urine metabolites and the bacterial taxonomical level from MITChip data (level 2–genera like) showing the association of early colonisers and urine metabolites.

trimethylamine *N*-oxide (TMAO), β -aminoisobutyrate and β -hydroxybutyrate $\dagger$, appeared to fluctuate (Figure 4a), the metabolites tryptophan and phenylacetylglutamine were consistently detected at a higher level during conventionalisation as compared with germfree animals.

Microbiota–metabolite–transcriptome correlation mining

The identified changes in the local and systemic metabolic profiles appeared to coincide with the transcriptome shifts and the succession of the microbial communities that established during the conventionalisation period. To test this notion, statistical modelling (using O2-PLS methods) was employed to detect significant correlations between microbial groups and the mucosal transcriptome, the mucosal and urine metabolic profiles. O2-PLS regression-models (Trygg, Svante 2003) between ^1H NMR urine spectra and level-2 (genera-like) MITChip absolute abundance scores were performed to search for statistically and biologically meaningful correlations between microbial taxa in the colon, colon transcriptome and concentrations of specific metabolites. The MITChip probe intensities that were best predicted by the variation in the ^1H NMR data were assigned to level-2 groups within the phylum Bacteroidetes. The closest relative isolates of the identified genera included, *Prevotella ruminicola et rel.*, *Alistipes*, *Rikenella* and *unclassified Porphyromonadaceae*. O2-PLS modelling found multiple correlations between these microbial groups and the urine metabolites, fumaric acid, 2-oxo-glutaric acid and malic acid $\dagger$ (Figure 4b).

A similar modelling strategy was performed to search for correlations between colonic tissue ^1H NMR metabolome and transcriptome. During conventionalisation, the metabolites glutamate, alanine and glycine were positively correlated with a set of induced genes that have a role in multiple metabolic pathways, including nucleotide metabolism, and O- and N-glycan biosynthesis and degradation. In contrast, the tissue concentrations of *scyllo*- and *myo*-inositol were positively correlated with a set of repressed genes that also had roles in multiple metabolic pathways, including phosphoglycerolipid metabolism, sialylated glycan biosynthesis and degradation and glycine and serine metabolism (Supplementary Figure 3).

Correlation analysis of changes in colonic tissue metabolites, host metabolic gene expression and microbiota taxonomical assignments illustrated that phylum/class microbial groups correlated to subsets of genes within the mucosal transcriptome data. More specifically, a positive correlation between Bacilli with a subset of the genes of which expression was positively correlated with *scyllo*- and *myo*-inositol concentrations was identified. Analogously, a positive correlation between alpha- and epsilon-Proteobacteria and a subset of the genes of which the

expression correlated with tissue concentrations of glutamate, alanine and glycine was detected (Figure 5). The O2-PLS correlation analysis found strong correlations between changes in the microbial colonisers of the colon, subsets of genes from the tissue transcriptomes and concentrations of specific tissue metabolites. These statistical correlations can be considered to be of biological relevance as the differentially expressed genes and the correlating metabolites belong to coherent metabolic pathways (Figure 5, Supplementary Figure 3).

Discussion

Several studies have shown that interactions between colonic microbes and germfree mice, lead to basal changes in host metabolism (Bäckhed *et al.*, 2004; Nicholson *et al.*, 2005; Turnbaugh *et al.*, 2009). However, these changes have hardly been investigated at a time-resolved molecular level. Here an ecosystems biological approach was employed to investigate the interactions between colonising microbes and colonic transcriptome and metabolome, as well as systemic changes in metabolome based on urine metabolites (Figure 6).

Detailed analysis of the composition of the colonising microbiota demonstrated that the establishment of microbiota in the large intestine of mouse is characterised by two major phases. The early, transient phase (days 1 and 2) showed low microbial diversity with typical formate, lactate and succinate production (as measured in the caecum), while the later, more stable phase showed a microbial community of higher diversity in terms of composition that resembled the inoculum. The latter phase was characterised by expansion of sulphate reducers such as *Desulfovibrio spp.* and strictly anaerobic species belonging to the major Clostridia clusters such as *Clostridium* cluster IV and XIVa, which contain butyrate producers. This bacterial expansion was confirmed by the quantitative PCR detection of the genes associated with these bacterial groups (*dsr* and *butyryl-CoA-transferase*, respectively). At day 4 post-conventionalisation, transient colonisation by some Proteobacteria was detected. This microbiota colonisation succession showed statistically significant and biologically relevant correlations with alterations in the gene-expression patterns and metabolite profiles from the colonic tissue. These colonic mucosa correlations were especially significant from day 4 post-conventionalisation onward. In contrast, statistical correlations between systemic metabolite concentrations that were measured in urine samples could be detected at the earliest time-points during conventionalisation (that is, days 1–2), suggesting a rapid and transient systemic metabolome-change by altered metabolite absorption from the gut. This implies that already 1 day after microbial colonisation of the intestine, the presence of the microbiota

Figure 5 Microbiota–metabolite–transcriptome correlation. Correlation heat maps showing: Model (1) correlation computed between colonic tissue transcriptome (shown in Supplementary Figure 3A), Bacilli and *scyllo*- and *myo*-inositol tissue metabolites (a) (positive correlation). Model (2) correlation computed between colonic tissue transcriptome (shown in Supplementary Figure 3B), glutamate, alanine and glycine tissue metabolites and epsilonproteobacteria, (b) and alphaproteobacteria (c) (positive correlation).

led to altered systemic metabolites profiles that were most likely because of altered luminal metabolite profiles and their absorption by the intestine. Remarkably, such absorption changes did not appear to induce changes in gene expression profiles in the colon mucosa, suggesting passive absorption or the presence of the respective transporters in the epithelium of germfree mice.

Clear correlation was found between the abundance of *Prevotella ruminicola et rel.* with the urine

concentrations of the metabolites, fumaric acid and 2-oxo-glutaric acid. This correlation may have reflected a temporarily decreased utilisation of tricarboxylic acid cycle intermediates by the host (Figure 4b). Alternatively, these metabolites could have been derived from the bacterial metabolism, which would be in agreement with the capacity of the correlated bacterial groups to produce oxaloglutarate by reductive carboxylation of succinate (Henderson, 1980). Succinate was detected at high

Figure 6 A schematic model summarising how different microbial taxa and their metabolites influence local (tissue metabolites and transcriptomes) and systemic (urine metabolites) metabolism in the time-resolved ecosystems biology approach applied in this study. Blue arrows indicate the O2-PLS based statistical correlations between microbiota, transcriptome and metabolites. Induced pathways and higher level metabolites are indicated in red and repressed pathways and lower level metabolites are indicated in green (in comparison to the germfree mice).

levels in the caecal lumen from 1 to 4 days post-conventionalisation, which could be due to the abundance of Bacteroidetes groups, known producers of succinate (Scheifinger and Wolin, 1973), during the early phase of conventionalisation. The early phase of conventionalisation of the colon was also characterised by accumulation of lactate, which can be formed by intestinal lactic acid bacteria or by a variety of other microorganisms in the gut ecosystem (Barcenilla *et al.*, 2000). It is of relevance that the results have shown that the initial microbial ecosystem that establishes during the first days of conventionalisation (days 1–2) appears to be relatively ineffective in the extraction of energy from dietary materials, which is exemplified by the accumulation of ‘high-energy fermentation metabolites’ such as lactate and succinate. At later days, the subsequent establishment of typical secondary fermenters in the microbial ecosystem such as the members of the *Clostridium* clusters IV, XIVa and sulphate reducers led to the depletion of lactate and succinate in the large intestinal lumen. The increasing concentrations of typical secondary metabolites

such as acetate, but notably propionate and butyrate in the large intestinal lumen, known to be produced by different bacterial groups including the Clostridia and Proteobacteria (Falony *et al.*, 2006; Flint *et al.*, 2008; Marquet *et al.*, 2009), provide support for the secondary conversion of lactate and succinate by these bacteria. These later stage developments in the microbiota community are paralleled by prominent changes in the colon–mucosa transcriptome and metabolite profiles. SCFAs are recognised by G-protein coupled receptor (GPCR) 43, a gene encoding a GPCR (Brown *et al.*, 2003) that has been proposed to constitute a molecular link between diet, microbiota and immune responses (Maslowski *et al.*, 2009). The microbiota-accommodating homeostasis that we previously reported to establish during the later phase of colonisation (days 8–16) (El Aidy *et al.*, 2012), was also associated with the appearance of butyrate in the large intestine where it has been proposed to serve as the primary energy source for colonocytes (Donohoe *et al.*, 2011), and has an important role in the regulation of fatty acid oxidation (Vanhoutvin *et al.*, 2009). The

mitochondrial hydroxy-3-methylglutaryl-CoA synthase (*Hmgcs2*), which is involved in fatty acid oxidation (Scheppach *et al.*, 1995) has been shown to be regulated by butyrate production by the intestinal microbiota (Cherbuy *et al.*, 2004). The temporarily decreased expression of *Hmgcs2* at day 4 post-conventionalisation but not at later time-points (Figure 3) may illustrate a transient impairment in β -oxidation, suggesting a metabolic shift towards anabolic metabolism via amino acid and nucleotide metabolism.

The decreasing concentrations of inositol metabolites was strongly correlated with the decreased expression of genes encoding enzymes involved in the metabolism of inositol, choline and related metabolic pathway involved in the assimilation of sarcosine, glycine and betaine. These modulations of the inositol and related metabolic pathways could also be associated with the intestinal colonisation by bacteria that can metabolise choline into methylamines (Kiene, 1998). The choline metabolism was reflected by the increased levels of its derivative-metabolites, trimethylamine (TMA) and TMAO that were identified in the urine samples of conventionalised mice, which is in agreement with previous reports (Nicholls *et al.*, 2003). Interestingly, TMA and subsequent TMAO production by gut microbiota has recently been reported to promote atherosclerosis and cardiovascular diseases (Wang *et al.*, 2011). This further supports the importance of improved understanding of the microbe–host metabolic interaction in view of its link to diseases of the host.

This study also identified a statistical correlation between the transiently increased abundance of alpha- and epsilon-Proteobacteria, and alterations in tissue and luminal metabolite levels (increased levels of glutamate, alanine, lactate and glycine, decreased levels of glucose) as well as specific changes in expression of metabolic pathways genes (increased expression of genes involved in the rate limiting steps of glycolysis, amino acid and nucleotide metabolic pathways) (Figures 5b and c). This association appears to be of biological relevance in that it suggests that the colonic metabolism is shifting towards increased energy production via glycolysis and assimilation, via anabolic metabolism, of novel cell components starting at day 4 onward. This proposed assimilation and production of novel cell components correlates well with the morphological changes that we previously observed in the colonic tissue during later time-points of conventionalisation (El Aidy *et al.*, 2012).

In conclusion, the time-resolved ecosystems biology approach applied in this study showed how different microbial taxa and their metabolites influence local (tissue metabolites and transcriptomes) and systemic (urine metabolites) metabolism in time. We inferred from these reciprocal metabolic changes in mice and microbiota that the changes in host transcriptomes and metabolites are exemplary for the establishment of a novel homeostasis that

accommodates the microbiota. Based on the changes in gene expression and metabolite concentrations, homeostasis was reached in the colon of the mice after 16 days post-conventionalisation (Figure 6). The transient phases prior to this novel state of homeostasis appeared to include a transient state of dysbiosis apparent from the inefficiency of early colonising communities to effectively extract energy from the dietary components, leading to the accumulation of ‘high-energy metabolites’ and impaired fatty acid oxidation. After this initial, transient state, establishment of later colonisers depleted the high-energy metabolites and generated second-stage fermentation SCFAs that can be efficiently metabolised by colonic epithelia. We propose that the dynamic molecular interactions between the microbiota and their hosts presented here revealed metabolic pathways and processes that contributed to the molecular definition of symbiotic, homeostatic interrelations between intestinal microbiota and the colon mucosa.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank several members of the team of Dr Joël Doré (INRA, Jouy-en-Josas) for assistance with animal killing and sampling, J Jansen and M Grootte Bromhaar (Division of Human Nutrition, Wageningen University) for excellent microarray hybridisation, P de Groot (Division of Human Nutrition, Wageningen University) for performing microarray quality control and primary data processing. This work was funded from the top institute of food and nutrition.

Author contributions

SEA and MK conceived and designed the experiments. SEA, MD, FL and CAM performed the experiments. SEA, MD, CAM, MB and SPC analysed the data. JD, JDE, EGZ and EH contributed material/analysis tools. SEA, PvB and MK wrote the paper.

References

- Azcárate-Peril MA, Sikes M, Bruno-Bárcena JM. (2011). The intestinal microbiota, gastrointestinal environment and colorectal cancer: a putative role for probiotics in prevention of colorectal cancer? *Am J Physiol Gastrointest Liver Physiol* **301**: G401–G424.
- Barcenilla A, Pryde SE, Martin JC, Duncan SH, Stewart CS, Henderson C *et al.* (2000). Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl Environ Microbiol* **66**: 1654–1661.
- Ben-Dov E, Brenner A, Kushmaro A. (2007). Quantification of sulfate-reducing bacteria in industrial wastewater, by real-time polymerase chain reaction (PCR) using *dsrA* and *apsA* genes. *Microb Ecol* **54**: 439–451.

- Bjursell M, Admyre T, Göransson M, Marley AE, Smith DM, Oscarsson J *et al.* (2011). Improved glucose control and reduced body fat mass in free fatty acid receptor 2-deficient mice fed a high-fat diet. *Am J Physiol Endocrinol Metab* **300**: E211–E220.
- Brown AJ, Goldsworthy SM, Barnes AA, Eilert MM, Tcheang L, Daniels D *et al.* (2003). The orphan G protein-coupled receptors GPR41 and GPR43 are activated by propionate and other short chain carboxylic acids. *J Biol Chem* **278**: 11312–11319.
- Bäckhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A *et al.* (2004). The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci USA* **101**: 15718–15723.
- Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JL (2005). Host-bacterial mutualism in the human intestine. *Science* **307**: 1915–1920.
- Bäckhed F, Manchester JK, Semenkovich CF, Gordon JL (2007). Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc Natl Acad Sci USA* **104**: 979–984.
- Cherbuy C, Andrieux C, Honvo-Houeto E, Thomas M, Ide C, Druet N. (2004). Expression of mitochondrial HMGCoA synthase and glutaminase in the colonic mucosa is modulated by bacterial species. *Eur J Biochem* **271**: 87–95.
- Claus SP, Tsang TM, Wang Y, Cloarec O, Skordi E, Martin FP. (2008). Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes. *Mol Syst Biol* **4**: 219.
- Cloarec O, Dumas M-E, Craig A, Barton RH, Lindon JC, Nicholson JK *et al.* (2005). Evaluation of the O-PLS model limitations caused by chemical shift variability and improved visualization of biomarker changes in 1H NMR spectroscopic metabonomic studies. *Anal Chem* **77**: 517–526.
- Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunger MK *et al.* (2011). The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* **13**: 517–526.
- Dridi B, Henry M, El Khéchine A, Raoult D, Drancourt M. High prevalence of *Methanobrevibacter smithii* and *Methanospaera stadtmanae* detected in the human gut using an improved DNA detection protocol. *PLoS One* (2009). **4**: e7063.
- Duncan SH, Holtrop G, Lobley GE, Calder AG, Stewart CS, Flint HJ. (2004). Contribution of acetate to butyrate formation by human faecal bacteria. *Br J Nutr* **91**: 915–923.
- EL Aidy S, van Baarlen P, Derrien M, Lindenbergh-Kortleve DJ, Hooiveld G, Levenez F *et al.* (2012). Temporal and spatial interplay of microbiota and intestinal mucosa drive establishment of immune homeostasis in conventionalised mice. *Mucosal Immunol* **5**: 567–579.
- Ernst J, Bar-Joseph Z. (2006). STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics* **7**: 191.
- Falony G, Vlachou A, Verbrugghe K, De Vuyst L. (2006). Cross-feeding between *Bifidobacterium longum* BB536 and acetate-converting, butyrate-producing colon bacteria during growth on oligofructose. *Appl Environ Microbiol* **72**: 7835–7841.
- Fischbach MA, Sonnenburg JL. (2011). Eating for two: how metabolism establishes interspecies interactions in the gut. *Cell Host Microbe* **10**: 336–347.
- Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA. (2008). Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol* **6**: 121–131.
- Geurts L, Lazarevic V, Derrien M, Everard A, Van Roye M, Knauf C *et al.* (2011). Altered gut microbiota and endocannabinoid system tone in obese and diabetic leptin-resistant mice: impact on apelin regulation in adipose tissue. *Front Microbiol* **2**: 149.
- Gibson GR, Macfarlane S, Macfarlane GT. (1993). Metabolic interactions involving sulfate-reducing and methanogenic bacteria in the human large-intestine. *Fems Microbiol Ecol* **12**: 117–125.
- Henderson C. (1980). The influence of extracellular hydrogen on the metabolism of *Bacteroides ruminalis*, *anaerobium lipolyticum* and *selenomonas ruminantium*. *J Gen Microbiol* **119**: 485–491.
- Hooper LV, Midtvedt T, Gordon JL. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr* **22**: 283–307.
- Hooper LV. (2004). Bacterial contributions to mammalian gut development. *Trends Microbiol* **12**: 129–134.
- Hultgren SJ, Abraham S, Caparon M, Falk P St, Geme JW, Normark S. (1993). Pilus and nonpilus bacterial adhesins: assembly and function in cell recognition. *Cell* **73**: 887–901.
- Huycke MM, Gaskins HR. (2004). Commensal bacteria, redox stress, and colorectal cancer: mechanisms and models. *Exp Biol Med* **229**: 586–597.
- Kau AL, Ahern PP, Griffin NW, Goodman AL, Gordon JL. (2011). Human nutrition, the gut microbiome and the immune system. *Nature* **474**: 327–336.
- Kiene RP. (1998). Uptake of choline and its conversion to glycine betaine by bacteria in estuarine waters. *Appl Environ Microbiol* **64**: 1045–1051.
- Lee YK, Mazmanian SK. (2010). Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science* **330**: 1768–1773.
- Ley RE, Peterson DA, Gordon JL. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**: 837–848.
- Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H *et al.* (2008). Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci USA* **105**: 2117–2122.
- Louis P, Flint HJ. (2007). Development of a semiquantitative degenerate real-time PCR-based assay for estimation of numbers of butyryl-coenzyme A (CoA) CoA transferase genes in complex bacterial samples. *Appl Environ Microbiol* **73**: 2009–2012.
- Marquet P, Duncan SH, Chassard C, Bernalier-Donadille A, Flint HJ. (2009). Lactate has the potential to promote hydrogen sulphide formation in the human colon. *Fems Microbiol Lett* **299**: 128–134.
- Martin FP, Wang Y, Sprenger N, Holmes E, Lindon JC, Kochhar S *et al.* (2007). Effects of probiotic *Lactobacillus Paracasei* treatment on the host gut tissue metabolic profiles probed via magic-angle-spinning NMR spectroscopy. *J Proteome Res* **6**: 1471–1481.
- Maslowski KM, Vieira AT, Ng A, Kranich J, Sierro F, Yu D *et al.* (2009). Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* **461**: 1282–U1119.
- Neish AS. (2009). Microbes in gastrointestinal health and disease. *Gastroenterology* **136**: 65–80.

- Nicholls AW, Mortishire-Smith RJ, Nicholson JK. (2003). NMR spectroscopic-based metabonomic studies of urinary metabolite variation in acclimatizing germ-free rats. *Chem Res Toxicol* **16**: 1395–1404.
- Nicholson JK, Holmes E, Wilson ID. (2005). Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Microbiol* **3**: 431–438.
- O'Hara AM, Shanahan F. (2006). The gut flora as a forgotten organ. *EMBO Rep* **7**: 688–693.
- Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. (2007). Development of the human infant intestinal microbiota. *Plos Biol* **5**: 1556–1573.
- Parker NJ, Begley CG, Fox RM. (1995). Human gene for the large subunit of ribonucleotide reductase (RRM1)-Functional-analysis of the promoter. *Genomics* **27**: 280–285.
- Rajilić-Stojanović M, Heilig HG, Molenaar D, Kajander K, Surakka A, Smidt H *et al.* (2009). Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environ Microbiol* **11**: 1736–1751.
- Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA *et al.* (2006). TM4 microarray software suite. *Methods Enzymol* **411**: 134–193.
- Salyers AA, Pajeau M. (1989). Competitiveness of different polysaccharide utilization mutants of bacteroides-Thetaiotaomicron in the intestinal tracts of germfree-mice. *Appl Environ Microbiol* **55**: 2572–2578.
- Savage DC. (1977). Microbial ecology of gastrointestinal-tract. *Annu Rev Microbiol* **31**: 107–133.
- Scheifinger CC, Wolin MJ. (1973). Propionate formation from cellulose and soluble sugars by combined cultures of Bacteroides succinogenes and Selenomonas ruminantium. *Appl Microbiol* **26**: 789–795.
- Scheppach W, Bartram HP, Richter F. (1995). Role of short-chain fatty acids in the prevention of colorectal cancer. *Eur J Cancer* **31A**: 1077–1080.
- Starrenburg MJ, Hugenholtz J. (1991). Citrate fermentation by lactococcus and leuconostoc spp. *Appl Environ Microbiol* **57**: 3535–3540.
- Steinberg LM, Regan JM. (2008). Phylogenetic comparison of the methanogenic communities from an acidic, oligotrophic fen and an anaerobic digester treating municipal wastewater sludge. *Appl Environ Microbiol* **74**: 6663–6671.
- Trygg J, Svante W. (2003). O2-PLS, a two-block (X-Y) latent variable regression (LVR) method with an integral OSC filter. *J Chemom* **17**: 53–64.
- Turnbaugh PJ, Hamady M, Yatsunencko T, Cantarel BL, Duncan A, Ley RE *et al.* (2009). A core gut microbiome in obese and lean twins. *Nature* **457**: 480–U487.
- Vanhoutvin SA, Troost FJ, Hamer HM, Lindsey PJ, Koek GH, Jonkers DM *et al.* (2009). Butyrate-induced transcriptional changes in human colonic mucosa. *Plos One* **4**: e6759.
- Walter J, Ley R. (2011). The human gut microbiome: ecology and recent evolutionary changes. *Annu Rev Microbiol* **65**: 411–429.
- Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B *et al.* (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**: 57–U82.
- Waters NJ, Holmes E, Waterfield CJ, Farrant RD, Nicholson JK. (2002). NMR and pattern recognition studies on liver extracts and intact livers from rats treated with alpha-naphthylisothiocyanate. *Biochem Pharmacol* **64**: 67–77.

Supplementary Information accompanies the paper on The ISME Journal website (<http://www.nature.com/ismej>)