

## ORIGINAL ARTICLE

# *Clostridium difficile* heterogeneously impacts intestinal community architecture but drives stable metabolome responses

David Rojo<sup>1,7</sup>, María J Gosalbes<sup>2,3,7</sup>, Rafaela Ferrari<sup>2,3</sup>, Ana E Pérez-Cobas<sup>2,3</sup>, Ester Hernández<sup>4</sup>, Rosa Oltra<sup>5</sup>, Javier Buesa<sup>6</sup>, Amparo Latorre<sup>2,3</sup>, Coral Barbas<sup>1</sup>, Manuel Ferrer<sup>4,8</sup> and Andrés Moya<sup>2,3,8</sup>

<sup>1</sup>Centro de Metabolómica y Bioanálisis (CEMBIO), Facultad de Farmacia, Universidad CEU San Pablo, Campus Montepríncipe, Madrid, Spain; <sup>2</sup>Unidad Mixta de Investigación en Genómica y Salud de la Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO) and Instituto Cavanilles de Biodiversidad y Biología Evolutiva de la Universidad de Valencia, Valencia, Spain; <sup>3</sup>CIBER en Epidemiología y Salud Pública (CIBEResp), Madrid, Spain; <sup>4</sup>CSIC, Institute of Catalysis, Madrid, Spain; <sup>5</sup>Unidad Enfermedades Infecciosas, Servicio Medicina Interna, Hospital Clínico Universitario de Valencia-INCLIVA, Valencia, Spain and <sup>6</sup>Departamento de Microbiología, Facultad de Medicina, Universidad de Valencia and Hospital Clínico Universitario de Valencia-INCLIVA, Valencia, Spain

***Clostridium difficile*-associated diarrhoea (CDAD) is caused by *C. difficile* toxins A and B and represents a serious emerging health problem. Yet, its progression and functional consequences are unclear. We hypothesised that *C. difficile* can drive major measurable metabolic changes in the gut microbiota and that a relationship with the production or absence of toxins may be established. We tested this hypothesis by performing metabolic profiling on the gut microbiota of patients with *C. difficile* that produced ( $n=6$ ) or did not produce ( $n=4$ ) toxins and on non-colonised control patients ( $n=6$ ), all of whom were experiencing diarrhoea. We report a statistically significant separation ( $P$ -value  $< 0.05$ ) among the three groups, regardless of patient characteristics, duration of the disease, antibiotic therapy and medical history. This classification is associated with differences in the production of distinct molecules with presumptive global importance in the gut environment, disease progression and inflammation. Moreover, although severe impaired metabolite production and biological deficits were associated with the carriage of *C. difficile* that did not produce toxins, only previously unrecognised selective features, namely, choline- and acetylputrescine-deficient gut environments, characterised the carriage of toxin-producing *C. difficile*. Additional results showed that the changes induced by *C. difficile* become marked at the highest level of the functional hierarchy, namely the metabolic activity exemplified by the gut microbial metabolome regardless of heterogeneities that commonly appear below the functional level (gut bacterial composition). We discuss possible explanations for this effect and suggest that the changes imposed by CDAD are much more defined and predictable than previously thought.**

The ISME Journal (2015) 9, 2206–2220; doi:10.1038/ismej.2015.32; published online 10 March 2015

## Introduction

*Clostridium difficile* is a major nosocomial pathogen and is the main causative agent of antibiotic-associated diarrhoea (*C. difficile*-associated diarrhoea

(CDAD)) in hospitals worldwide (Kelly and LaMont, 2008; Britton and Young, 2014). Approximately 0.5–3 million individuals per year are infected by *C. difficile*, with 20–30% of them failing to be cured, depending on the production of toxins A and B, which lead to cellular death and tissue damage (Fordtran, 2006; Kuehne *et al.*, 2010). One of the abnormalities observed in patients with CDAD is antibiotic-induced dysbiosis in the gastrointestinal tract, which favours *C. difficile* germination and growth (Kuehne *et al.*, 2010; Knecht *et al.*, 2014; Theriot *et al.*, 2014). However, gut dysbiosis has also been observed in non-colonised patients (Pérez-Cobas *et al.*, 2013).

Standard treatment regimes and new medications exist, but the rate of CDAD recurrence remains high

Correspondence: M Ferrer, CSIC, Institute of Catalysis, Madrid 28049, Spain.

or A Moya, Unidad Mixta de Investigación en Genómica y Salud de la Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO), Valencia 46020, Spain.

E-mail: mferrer@icp.csic.es or andres.moya@uv.es

<sup>7</sup>These authors contributed equally to this work.

<sup>8</sup>These authors contributed equally to this work.

Received 7 August 2014; revised 12 January 2015; accepted 6 February 2015; published online 10 March 2015

(Borody *et al.*, 2014; Knecht *et al.*, 2014; Pérez-Cobas *et al.*, 2014). This relapse prevalence may be due to the limited knowledge of how *C. difficile* infection evolves and, importantly, which pathogen-induced changes occur in relation or not to the production of toxins A and B and our inability to counterbalance the as yet unknown functional declines in the gastrointestinal tract. In relation to this, *C. difficile* infection development during antibiotic therapy has not been associated with specific gut microbial composition but rather with the presence of a limited set of bacterial species (Knecht *et al.*, 2014), which may be involved in colonisation resistance against *C. difficile*. As one or two strains are sufficient to drive major changes in gastrointestinal and host metabolic profiles (Marcobal *et al.*, 2013), we considered the possibility that the development of CDAD may also be linked to specific microbial metabolic patterns, which may help identify metabolic associations with *C. difficile* infection.

This investigation moves beyond descriptive studies of community composition (for example, Rea *et al.*, 2011; Peterfreund *et al.*, 2012; Knecht *et al.*, 2014; Pérez-Cobas *et al.*, 2014) and presents the metabolic profiles of the gut microbiota in patients colonised with *C. difficile* that were positive or negative for the production of toxins and in non-colonised control patients. The study aimed to determine whether *C. difficile* has a role in shaping gastrointestinal microbial metabolism and, if so, whether the production of toxins has an influence. This question is of particular importance, as the role of the microflora throughout the body and particularly in the gastrointestinal tract has been recognised (Spor *et al.*, 2011). To reduce the possibility that our assay was affected by the occurrence of diarrhoea in the colonised patients, non-colonised patients with diarrhoea caused by bacteria other than *C. difficile* were investigated as controls. Results are presented that show that *C. difficile* and its toxins appear to produce different but stable metabolic profiles of the gut microbiota regardless of the patient medical characteristics. Biological significance and mechanistic plausibility for these associations are proposed.

## Materials and methods

### Chemicals and reagents

The following reagents were used: acetonitrile (high-performance liquid chromatography mass spectrometry (HPLC-MS) grade, Sigma-Aldrich, Taufkirchen, Germany), formic acid (MS grade, Sigma-Aldrich, Steinheim, Germany), L-methionine sulphone (Sigma-Aldrich, Taufkirchen, Germany) and sodium hydroxide (Panreac, Montcada I Reixac, Spain). The reference masses purine and hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazine were obtained from Agilent (Santa Clara, CA, USA) (atmospheric pressure ionization time-of-flight reference mass solution kit).

All solutions were prepared using MilliQ water (Millipore, Billerica, MA, USA).

### Patients and study cohorts

Faecal samples were collected by health personnel following standard procedures from patients with diarrhoea treated at the Hospital Clínico Universitario in Valencia (Valencia, Spain). The samples were kept at  $-80^{\circ}\text{C}$  until use. The diagnosis of *C. difficile* infection was performed in the Microbiology Service using the kit Techlab C. DIFF Quik Chek Complete (Alere Healthcare S.A, Barcelona, Spain), which simultaneously detects both glutamate dehydrogenase antigen and toxins A and B. We identified patients who carried toxin-producing *C. difficile* (referred to as p/p or toxin +), patients with *C. difficile* without toxin production (referred to as p/n or toxin -) and non-colonised control patients (referred to as n/n). Using 16S rRNA gene sequence analysis (see below), a number of bacteria were identified as putative diarrhoea-causing pathogens in the n/n cases, including species related to: (i) the *Pseudomonas* genus (3.4% of the total 16S ribosomal RNA (rRNA) sequences) in HCC18; (ii) the *Fusobacterium* (20.0%) and *Aggregatibacter* (9.3%) genera and the *Enterobacteriaceae* family (19.5%, including 4.7% related to the *Escherichia:Shigella* genus) in HCC19; (iii) the *Enterobacteraceae* family (2.4%) and the *Burkholderiales* order (14.2%, including 3.5% related to the *Sutterellaceae* family and 3.3% to the *Parasutterella* genus) in HCC21; (iv) the *Enterobacteriaceae* family (12.4%, including 2.17% related to the *Escherichia:Shigella* genus) in HCC16; and (v) the *Fusobacterium* genus (19.52%), the *Enterobacteriaceae* family (21.66%, including 18.36% related to the *Escherichia:Shigella* genus) and the *Aggregatibacter* genus (17.5%) in HCC23. No presumptive pathogen could be unambiguously identified in HCC6 because of its low diversity.

Information regarding the clinical manifestations, age, gender, medical history and medication, duration of infection and clinical-related factors for the patients was recorded (Supplementary Table 1). Briefly, the mean (s.d.) ages were  $71.0 \pm 15.0$  (case p/p),  $66.3 \pm 27.0$  (case p/n) and  $52.0 \pm 13.2$  (case n/n) years. All the individuals were affected by diarrhoea episodes for a range of 1–7 days (case p/p), 2–12 days (case p/n) and 5–60 days (case n/n) before sampling and were treated with antibiotics (Supplementary Table 1). Both male (p/p: 4; p/n: 4; n/n: 2) and female (p/p: 2; p/n: 0; n/n: 4) patients were included. The participants gave informed consent, and the study was approved by the independent ethics committee at the recruiting institution (Hospital Clínico Universitario).

### Stool samples

Fresh stool samples were collected from each subject, frozen immediately and stored at  $-80^{\circ}\text{C}$

until processing. Before analysis, stool samples were split into two samples (0.4 g each). One sample underwent differential centrifugation to enrich for microbial cells, from which metabolites were further extracted. The second sample, which was used for comparative purposes, was directly subjected to metabolite extraction.

#### *Stool sample processing for enrichment of gut microbial cells*

Stool samples were directly subjected to differential centrifugation to enrich for microbial cells according to Pérez-Cobas *et al.* (2013), with minor modifications. Briefly, 0.4 g of faecal samples were resuspended in 1.2 ml of HPLC-grade H<sub>2</sub>O (1:3 w/v faeces to H<sub>2</sub>O) and immediately vortexed for 10 s; after re-suspension, the samples were placed in an ice water bath and sonicated for 120 s (at 15 W). Following sonication, the samples were then centrifuged at 2000 g at 4 °C for 2 min to remove faecal debris. The supernatant was then carefully transferred to a clean 2-ml eppendorf tube and then centrifuged at 13 000 g at 4 °C for 15 min to pellet the microbial cells. The pellet was suspended again in 1.2 ml of HPLC-grade H<sub>2</sub>O. The entire procedure was repeated for a total of three rounds. The final pellet was immediately subjected to metabolite extraction (see below). Note that this protocol eliminates most soluble faecal metabolites. In addition, targeting the V1–V2 hypervariable regions of the bacterial 16S rRNA and the human 18S rRNA genes revealed that the amount of eukaryotic cells in the derivative pellet obtained by differential centrifugation was below the detection limit, suggesting that the pellet contains mostly microbial cells.

#### *Metabolite extraction and preparation of metabolite solution for analysis*

Pellets containing microbial cells (obtained from 0.4 g stool material; see above) or stool samples (0.4 g) were further subjected to metabolite extraction by a two-step extraction method. The samples were first extracted with methanol (MeOH) by adding 1.2 ml of cold (–80 °C) HPLC-grade MeOH. The samples were then stored at –80 °C for 60 min. Next, the samples were again mixed by vortexing (for 10 s) and were sonicated for 20 s (in a Sonicator 3000; Misonix, Berlin, Germany) at 15 W in an ice cooler (–20 °C). This protocol was repeated twice, with a 5-min incubation at –20 °C between the cycles, and the final pellet was removed following centrifugation at 16 000 g for 20 min at 4 °C. Immediately after the MeOH extracts were obtained, the MeOH solution was transferred to a clean 20-ml penicillin vial and stored at –80 °C, whereas the remaining cell pellet was resuspended in 1.2 ml of cold (4 °C) HPLC-grade H<sub>2</sub>O and subjected to three cycles of sonication for 20 s (in a Sonicator 3000; Misonix) at 15 W in ice water; the samples were kept

on ice for 2 min between cycles. Finally, the supernatants were centrifuged at 16 000 g for 20 min at 4 °C and immediately transferred to a clean 20-ml penicillin vial and stored at –80 °C. Immediately after the MeOH and H<sub>2</sub>O extracts were obtained from the same microbial pellet, a mixture was prepared by combining equal volumes (1 ml) of each extract solution. Once prepared, the final metabolite solution was stored in 20-ml penicillin vials at –80 °C until use.

A total of 16 samples were metabolotyped using a combination of untargeted mass spectrometry (MS) and two different but complementary separation techniques (LC-ESI-QTOF-MS (liquid chromatography electrospray ionization quadrupole time-of-flight mass spectrometry) and CE-TOF-MS (capillary electrophoresis time-of-flight mass spectrometry)). Before analysis, 500 µl of each metabolite solution was centrifuged at 13 000 g at 4 °C for 10 min to separate any solid impurities. The supernatants were removed, filtered through 0.22-µm nylon syringe filters, and then, for the LC-MS analysis, 100 µl was transferred to the analytical vials. For the CE-MS analysis, 70 µl of filtered extract was evaporated to dryness using a Speedvac Concentrator (Thermo Fisher Scientific, Waltham, MA, USA), and then, each sample was reconstituted in 70 µl of MilliQ water with an internal standard (0.2 mM L-methionine sulphone) and 0.1 M formic acid.

#### *Preparation of quality controls (QCs) for metabolomics fingerprinting*

As the samples interact with the separation technique, as well as with the MS in LC-MS and CE-MS, QCs are crucial to ensure the analytical reproducibility. Indeed, QC samples must to be analysed throughout the analytical runs at periodic intervals of time to monitor the variation of the signal across time, as well as at the beginning of the sequence to stabilise the system (Dunn *et al.*, 2011). QC samples were independently prepared for LC-MS and CE-MS by pooling and mixing equal volumes of each metabolite solution. After gently vortexing, the mix was also filtered and subsequently transferred to an analytical vial.

#### *Metabolomics fingerprinting by LC-ESI-QTOF-MS*

The metabolic profile was achieved by a LC system consisting of a degasser, binary pump and autosampler (1290 Infinity, Agilent). A total of 0.5 µl of each sample was applied to a reversed-phase column (Zorbax Extend C<sub>18</sub> 50 × 2.1 mm, 3 µm; Agilent) that was kept at 60 °C during the analysis. The system was operated at a flow rate of 0.6 ml min<sup>–1</sup> with solvent A (water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% formic acid). The following gradient was used: 5% B (0–1 min), 5–80% B (1–7 min), 80–100% B (7–11.5 min) and 100–5% B (11.5–12 min). The flow was then held at 5% B to re-equilibrate the system for 3 min (15 min of total analysis time). Data were collected in positive and



negative electrospray ionization modes in separate runs on a quadrupole time-of-flight (Agilent 6550 iFunnel). The analyses were performed in both the positive and negative ion modes. The positive mode was operated in full scan mode from  $m/z$  50 to 1000. The capillary voltage was set at 3000 V with a scan rate of 1.0 spectra $s^{-1}$ . The gas temperature was 250 °C, the drying gas flow was 12 lmin $^{-1}$  and the nebuliser was used at 52 psi. The MS-TOF had the following parameters: fragmentor, 175 V; skimmer, 65 V; and octopole radio frequency voltage (OCT RF Vpp), 750 V. The negative mode was operated in full scan mode from  $m/z$  50 to 1100. The capillary voltage was 3000 V with a scan rate of 1.0 spectra $s^{-1}$ . The gas temperature was 250 °C, the drying gas flow was 12 lmin $^{-1}$  and the nebuliser was used at 52 psi. The MS-TOF parameters were as follows: fragmentor, 250 V; skimmer, 65 V; and OCT RF Vpp, 750 V. During the analyses, two reference masses were used in each mode: 121.0509 (detected  $m/z$  [C<sub>5</sub>H<sub>4</sub>N<sub>4</sub>+H]<sup>+</sup>) and 922.0098 (detected  $m/z$  [C<sub>18</sub>H<sub>18</sub>O<sub>6</sub>N<sub>3</sub>P<sub>3</sub>F<sub>24</sub>+H]<sup>+</sup>) in the positive mode and 112.9855 (detected  $m/z$  [C<sub>2</sub>O<sub>2</sub>F<sub>3</sub>-H]) and 1033.9881 (detected  $m/z$  [C<sub>18</sub>H<sub>18</sub>O<sub>6</sub>N<sub>3</sub>P<sub>3</sub>F<sub>24</sub>+TFA-H]<sup>-</sup>) in the negative mode. They were continuously infused into the system to allow for constant mass correction. The samples were analysed in a randomised run during which time they were kept in the autosampler at 4 °C. The analytical runs for both polarities were set up starting with the analysis of 10 QCs followed by the samples; a QC sample was injected in between each block of three samples until the end of the experiment.

#### Metabolomics fingerprinting by CE-TOF-MS

CE (7100 Agilent) coupled to TOF MS (6224 Agilent) was used. The CE mode was controlled by ChemStation software (B.04.03, Agilent), and the MS mode was controlled by Mass Hunter Workstation Data Analysis software (B.02.01, Agilent). The separation occurred in a fused-silica capillary (Agilent: total length, 100 cm; i.d., 50 µm). All separations were performed in normal polarity with a background electrolyte containing 0.8 M formic acid in 10% MeOH (v/v) at 20 °C. New capillaries were pre-conditioned with a flush of 1.0 M NaOH for 30 min, followed by MilliQ water for 30 min and the background electrolyte for 30 min. Before each analysis, the capillary was conditioned with a flush of background electrolyte for 5 min. The sheath liquid (6 µlmin $^{-1}$ ) was MeOH:H<sub>2</sub>O (1:1) containing 1.0 mM formic acid with two reference masses,  $m/z$  121.0509 ([C<sub>5</sub>H<sub>4</sub>N<sub>4</sub>+H]<sup>+</sup>) and  $m/z$  922.0098 ([C<sub>18</sub>H<sub>18</sub>O<sub>6</sub>N<sub>3</sub>P<sub>3</sub>F<sub>24</sub>+H]<sup>+</sup>), which allowed for correction and higher mass accuracy in the MS. The samples were hydrodynamically injected at 50 mBar for 50 s. The stacking was performed by applying the background electrolyte at 100 mBar for 10 s. The separation voltage was 30 kV, the internal pressure was 25 mBar and the analyses

were performed in 35 min. The MS parameters were as follows: fragmentor, 100 V; skimmer, 65 V; octopole, 750 V; drying gas temperature, 200 °C; flow rate, 10 lmin $^{-1}$  and capillary voltage, 3500 V. The data were acquired in positive mode with a full scan from  $m/z$  85 to 1000 at a rate of 1.41 scan $s^{-1}$ .

#### Metabolomics data treatment and statistical analysis

The resulting data files were cleaned of background noise and unrelated ions by the Molecular Feature Extraction tool of the Mass Hunter Qualitative Analysis software (B.05.00, Agilent). The alignment of the raw data was performed with MassProfiler Professional software (B.12.01, Agilent) and subsequent models were built by SIMCA-P+ software (12.0.1.0, Umetrics, Umea, Sweden). Next, the variables were filtered to select those that were present in at least 50% of the QCs and that had a coefficient of variation that was <30% across the QCs (Dunn *et al.*, 2011). Then, any missing values were replaced by the mean (if the variable was present in >2/3 of the samples per group) or by half of the minimum value (if the variable was present in between 1/3 and 2/3 of the samples per group). Zero was kept in the variables that were present in <1/3 of the samples per group. Finally, the groups were compared in pairs (p/p vs p/n, p/p vs n/n and p/n vs n/n) with the Mann–Whitney *U*-test followed by a Benjamini–Hochberg correction to minimise false positives (*P*-value ≤ 0.05 for both) (MATLAB 7.10.0.499, MathWorks, Natick, MA, USA). The resulting list of accurate masses that were significantly different among the groups were searched against the CEU Mass Mediator database (<http://biolab.uspceu.com/mediator>) (error ± 10 p.p.m.), and cluster analysis (<http://imagej.nih.gov/ij/>) was performed based on the relative abundance of each mass per sample. This procedure was performed independently for each analytical platform.

#### DNA extraction, amplification and sequencing of the bacterial 16S rRNA gene

Before DNA extraction, the samples were resuspended in saline phosphate buffer (PBS)/RNAlater (Ambion, Austin, TX, USA) and centrifuged at low speed (2,000 r.p.m., 2 min) to remove faecal debris as much as possible. Then, the genomic DNA was extracted from supernatants using QIAamp DNA Stool Kit (Qiagen, West Sussex, UK) according to the manufacturer's instructions. The total DNA was quantified with a Nanodrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). A region of the 16S rRNA gene (V1, V2) was amplified by PCR technology for each sample. For this, the universal primers E8F (5'-AGAGTTTGATC MTGGCTCAG-3') and 357R (5'-TGCTGCCTCCCGT AGGAGT-3') were utilised using the sample-specific Multiplex Identifier (MID) for pyrosequencing according to 454 standard protocols. A pool of amplicons was

sequenced on a 454 Genome Sequencer FLX Titanium platform (Roche Diagnostics, Branford, CT, USA). All sequences have been deposited in the European Bioinformatics Institute database under accession number PRJEB7637.

#### 16S rRNA gene analysis

The sequences generated from pyrosequencing were analysed with the MOTHUR software package (Schloss *et al.*, 2009). The reads with a Phred quality score below 20 ( $Q_{20}$ ) and with a short length (<250 bp) were removed. Thereafter, the remaining sequences were checked for potential chimeras applying the align.seqs and chimera.slayer tools incorporated in the MOTHUR package (Schloss *et al.*, 2009). Taxonomic information of the 16S rDNA sequences was obtained by comparison with the Ribosomal Database Project-II (RDP) (Cole *et al.*, 2009). For each sequence, we recovered and used for further analysis the assignment obtained with a confidence level (bootstrap cut-off) >80%, leaving the assignment at the last well-identified level and the consecutive levels as unclassified (uc). We obtained an average of 4200 taxonomically assigned reads per sample.

Canonical correspondence analysis was performed to determine the relation between the microbiota composition of the three groups (p/p, p/n and n/n) and external variables (presence of *C. difficile* and toxin production). To statistically assess the effect of these factors on the bacterial composition, a multivariate analysis of variance based on dissimilarity tests (Adonis) was applied, as implemented in the package Vegan, R software (<http://cran.r-project.org>) (Oksanen *et al.*, 2011; R Development Core Team, 2011). The differences in bacterial composition between samples were addressed comparing groups using LefSe algorithm (Segata *et al.*, 2011).

#### Biochemical tests: transformation of putrescine to N-acetylputrescine

Protein extracts from each of the patients were obtained as previously described (Hernández *et al.*, 2013), using bacterial pellets separated from faecal material, as described above. Note that the bacterial cells were from the same batch used for metabolite extraction. The extent of transformation of putrescine to N-acetylputrescine was quantified at 30 °C for a total time of 60 min, in a solution containing 1 ml of 50 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulphonic acid), pH 7.0, 0.15 mg ml<sup>-1</sup> putrescine, 0.15 mg ml<sup>-1</sup> acetyl-coenzyme A (CoA) and 0.1 mg ml<sup>-1</sup> total protein. Reactions without proteins and substrates were used as controls; they were subsequently treated in the same manner as samples. The transformation efficiency was followed by the formation of the reaction product CoA using a CoA detection kit in a microplate reader (Synergy HT Multi-Mode Microplate Reader—BioTek,

Winooski, VT, USA). All of the chemicals and the kit were supplied by Fluka-Aldrich-Sigma Chemical Co. (St Louis, MO, USA). Before analysis (CoA determination), the samples (triplicates) were diluted 1:3 with MilliQ water (dilution factor 4) and were further analysed in 96-well plates.

## Results

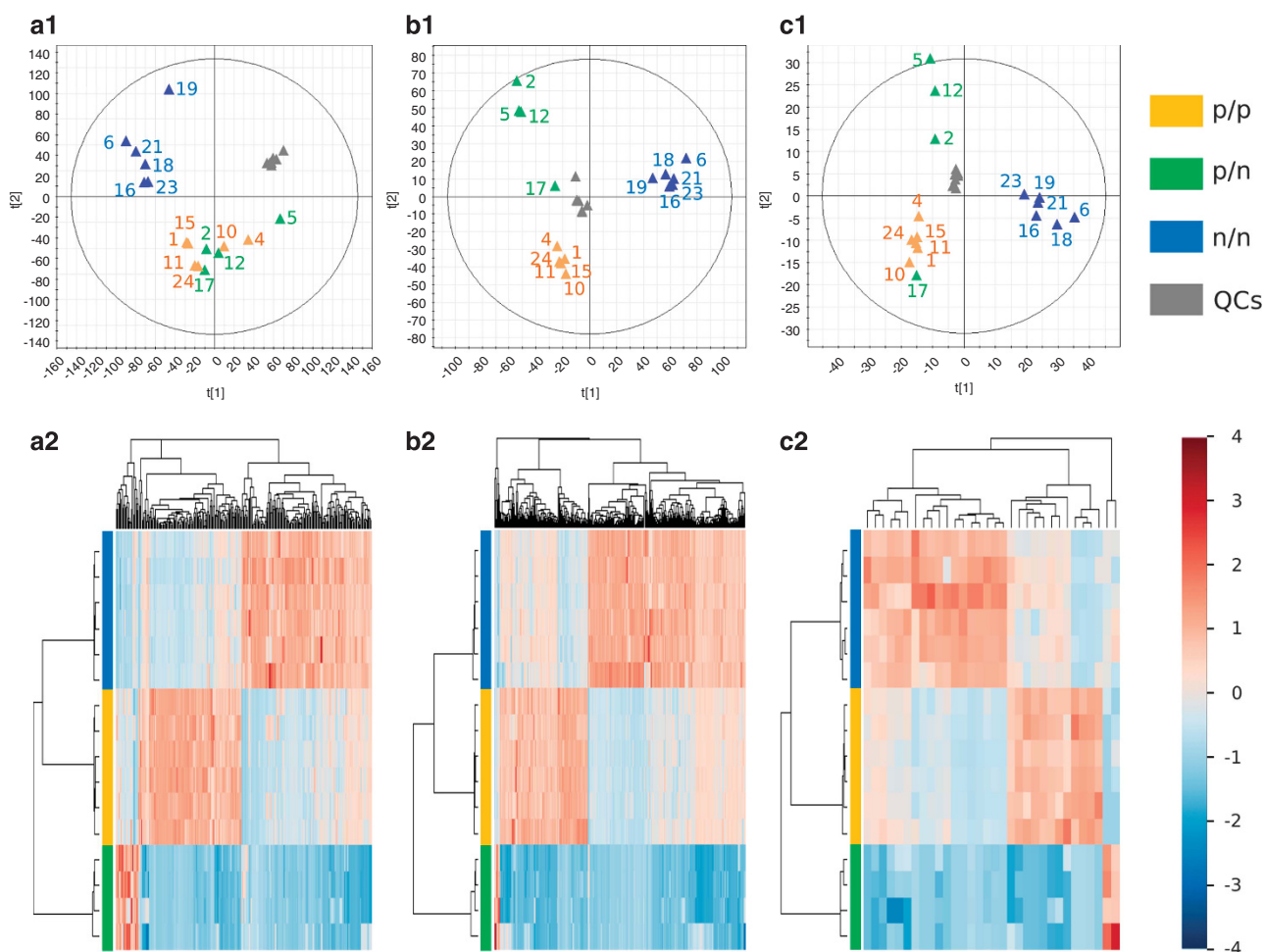
#### CDAD is linked to marked, stable and measurable microbial metabolism

In this study, we focused on the isolation of metabolites from microbes isolated from fresh stool material and then on the metabolic profiling of the corresponding gut microbiota by a combination of MS with LC and CE separations. Of 115 228 masses, a total of 725 (LC-MS positive mode: 300; LC-MS negative mode: 393; CE-MS: 32) masses passed the statistical criteria (Supplementary Table 2;  $P$ -value <0.05; Mann–Whitney  $U$ -test) and were determined to be responsible for the real biological separation of the three groups of samples (Figure 1), namely, patients with *C. difficile* that were positive ( $n=6$ ) or negative ( $n=4$ ) for the production of toxins and non-colonised control patients ( $n=6$ ). As shown in Figure 1, our data sets were highly similar for all subjects within each of the three groups, regardless of age, gender and other patient characteristics, the presence and duration of disease (including one or more additional disorders), antibiotic treatment (type and duration) and medical history (Supplementary Table 1). This finding suggests the prevalence of the presence of pathogens in the gastrointestinal tract as a dominant factor that swiftly regulates gastrointestinal microbial metabolism compared with other environmental factors. Accordingly, infection by a pathogen may become a driving factor determining the changes in the microbial metabolism, the level of which may be pathogen dependent, as revealed by the separate grouping of the metabolome profiles shown in Figure 1.

Notably, the separation of the three groups was not observed when metabolic profiling was performed using metabolite solutions from stool samples without pre-treatment, that is, when examining metabolites that were directly extracted from total faecal material (Supplementary Figure 1). The total fluid from the faecal material may contain a complex mixture of metabolites derived from the diet, the host and intestinal bacteria, which may explain the heterogeneity of the metabolome profiles. Metabolites from intestinal bacteria are therefore good indicators of gut microbial function under CDAD, and they may be more effective than faecal fluids as a readout of CDAD-induced alterations.

#### Overall impact of CDAD in gut microbial metabolism

The abundance level of only 725 out of 115 228 (or 0.63% of the total) mass features was found to be statistically ( $P$ -value <0.05) different between the



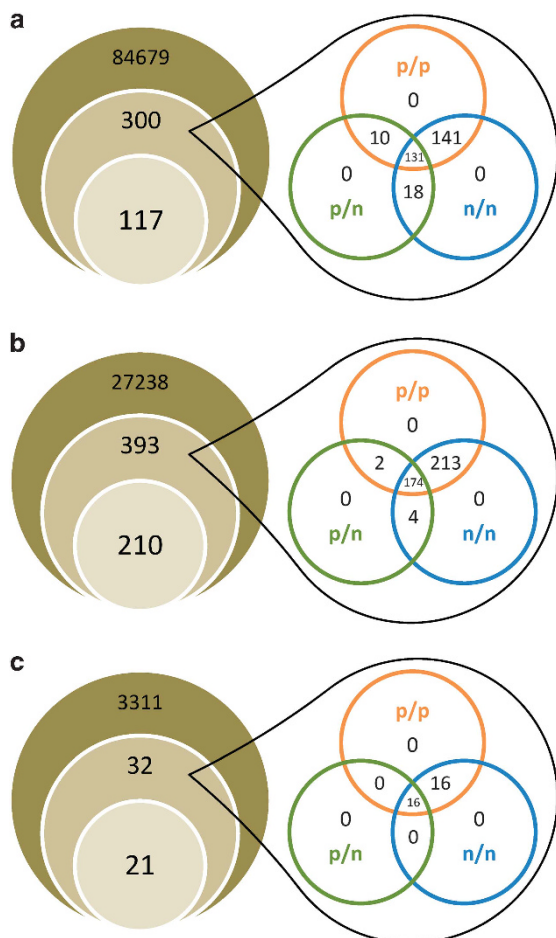
**Figure 1** Composition of the intestinal microbial metabolite profiles. (1) Principal components analysis (PCA) plot for models built with the entire set of raw data: (a1) LC-MS (+), two components,  $R^2=0.171$ ,  $Q^2=-0.013$ ; (b1) LC-MS (-), two components,  $R^2=0.238$ ,  $Q^2=0.068$ ; (c1) CE-MS, two components,  $R^2=0.290$ ,  $Q^2=0.124$ . (2) Two-way hierarchical clustering analysis performed with the matrices of the significant compounds in which the abundances were scaled by mean-centring and dividing by the standard deviation of each variable: (a2) LC-MS (+); (b2) LC-MS (-); (c2) CE-MS. The following codes were established: code p/p for patients with *C. difficile* producing toxins, code p/n for patients with *C. difficile* not producing toxins and code n/n for non-colonised control patients. The robustness of the sample clustering is based on three main reasons. First, the PCA is a non-supervised model, which means that distribution of the samples plotted in Figure 1 are due only to the directions of the principal components that maximise the explained percentage of variance ( $R^2$ ). Second, raw data are plotted, including all possible sources of variability besides the biological sources. Therefore, the robustness of the analytical procedure was demonstrated by the tight clustering of the QCs in the non-supervised PCA models (see the Materials and Methods section for details), proving that separation among the groups was due to real biological variability and not due to analytical variance (random). Third, in both techniques (LC-MS and CE-MS), the samples are almost all clustered. The slight differences among them are related to the complementarity between platforms because each of them is devoted to a certain part of the metabolome that includes a wide set of chemical functions.

three groups. This finding indicates that the impact of CDAD caused by either toxin - or toxin + *C. difficile* on the gut microbial metabolite-wide flux distribution and the metabolism itself is moderately low. However, only ~56% (or 321 out of 725) of those masses were shared (Figure 2), suggesting high metabolic heterogeneity among the groups, which is consistent with the different natures of the diseases. Further examination of the masses responsible for the separation of the three groups of samples confirmed that the p/p (toxin + *C. difficile*) and n/n (non-colonised controls) groups shared a higher number of common metabolites. This finding is shown in the Venn diagram in Figure 2, which demonstrates that ~51% of the total masses that

passed the statistical criteria (725) were shared by both groups, whereas only ~1.7% and ~3.0% of the masses were shared between the p/p and p/n and the p/n and n/n groups, respectively.

To identify the specific effects associated with CDAD and toxins, empirical formulas were assigned to the subset of statistically significant features, which were obtained after examining the metabolic profiles of the gut microbiota, with a maximum error of 10 p.p.m. using the CEU Mass Mediator database (<http://biolab.uspceu.com/mediator>). Supplementary Table 2 contains all of the putatively identified masses features per technique (LC-MS-positive mode: 117; LC-MS-negative mode: 210; CE-MS: 21).





**Figure 2** Number of masses that passed each data treatment step and the distribution of the significant masses per comparison. The data are colour-coded as follows: dark brown, after alignment; medium brown, significant; light brown, tentatively identified. (a) LC-MS (+). (b) LC-MS (-). (c) CE-MS. Venn diagrams indicate the number of shared statistically significant metabolites per technique and sample group.

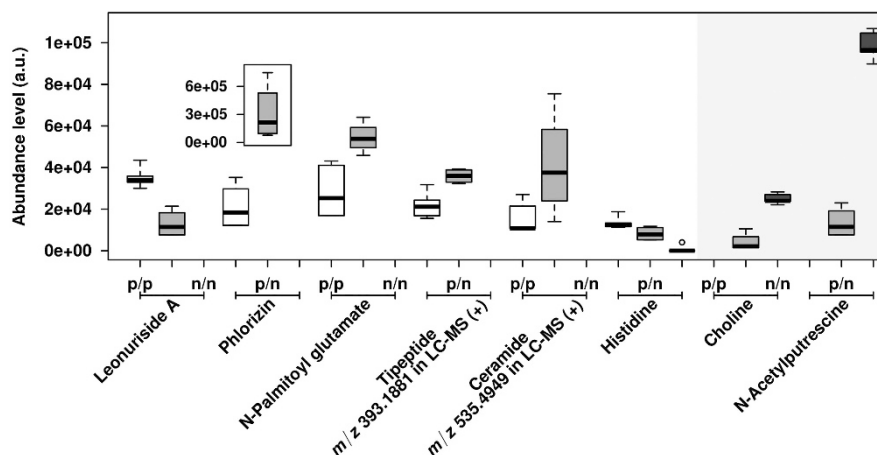
#### *CDAD associated with the production of molecules presumptively involved in disease progression*

Among the set of metabolites tentatively identified (Supplementary Table 2), we first observed common similarities between patients with toxin - and toxin + *C. difficile*. Briefly, six putatively identified metabolites (Leonuriside A, *N*-palmitoyl glutamic acid, phlorizin, two tripeptides and ceramide) were *C. difficile* specific (Figure 3), as they were found in patients with CDAD regardless of toxin production, but not in non-colonised control patients ( $P$ -value  $\leq 0.0152$ ). These chemicals are implicated in the regulation of chronic intestinal inflammation, cell growth and signal transduction (Qiu *et al.*, 1996; Lauzon *et al.*, 2003; Tan *et al.*, 2010; Kurek *et al.*, 2013). They also function as binding sites for bacteria and their toxins to prevent the translocation of pathogens (Kurek *et al.*, 2013). Accordingly, the presumptive role of these CDAD-specific chemicals in the progression of disease cannot be ruled out.

However, further experimental evidence is needed to confirm the cause and effect relationship between these altered metabolites and the underlying CDAD disease. In any case, these metabolites were associated with subjects carrying both toxin - and toxin + *C. difficile*, irrespective of the presence or absence of different antibiotic therapies and of patient characteristics. However, the metabolites were absent in non-colonised control patients who were also experiencing diarrhoea, suggesting that their presence is due to CDAD rather than due to diarrhoea or environmental factors. In addition, the presence of *C. difficile* was also associated with a significantly higher production of histidine (12.2-fold;  $P$ -value  $\leq 0.0095$ ) that was enhanced in the presence of toxins (19.8-fold;  $P$ -value  $\leq 0.0021$ ) compared with that in non-colonised control patients (Figure 3).

#### *Carriage of toxin-negative C. difficile associated with impaired metabolite production*

Patients with *C. difficile* that was negative for the production of toxins showed increased metabolic deficiencies, namely, the inability to produce 72 metabolites (Supplementary Table 2;  $P$ -value from 0.0095 to 0.0382), the presence and concentration of which have been shown to influence human health. They include (i) two fatty acids, 22 lipids (including gluco-, glycer- and glycerophospho-lipids) and 10 bile acid derivatives, low levels of which have been reported to promote cell lipotoxicity and apoptosis and have an overall cathartic effect on the colon (Yoon *et al.*, 2002; Senkal *et al.*, 2011; Swann *et al.*, 2011; Longato *et al.*, 2012); (ii) five *N*-acyl amino acids or polyamides (including arachidoyl glycine, *N*-stearoyl proline, *N*-oleoyl (iso)leucine, *N*-stearoyl tyrosine and *N*-palmitoyl threonine), the absence of which promotes dysfunction in the regulation of host temperature, locomotion and inflammation (Tan *et al.*, 2010); (iii) ferroxamine and five metabolites implicated in porphyrin and iron metabolism, the deficiency of which decreases the concentration of beneficial gut microbiota and induces iron deficiency anaemia (Dostal *et al.*, 2014); (iv) 12 presumptive secondary metabolites and bioactive peptides, such as methionine enkephalin and a number of tripeptides, the metabolism failure of which has been shown to lead to failure in immune and neuroactive ligand-receptor interactions (Yoshimasa *et al.*, 1982; Salzet and Tasiemski, 2001); (v) inosine, pseudouridine and hypoxanthine, nucleoside and purine derivatives, the depletion of which may negatively influence the initiation of translation and nucleic acid synthesis in human gut microbiota and in gut mucosal defence (Grimble, 1994); and (vi) six ceramide/sphingolipid derivatives, creatinine, *N*-acetylhistamine, glyoxylate and succinate-ceramide and creatinine deficits in the gut have been associated with liver and renal dysfunctions (Peral *et al.*, 2002; Longato *et al.*,



**Figure 3** Box plot of the metabolite abundance levels in the three groups of patients examined. The abundance levels of key statistically significant metabolites are given in arbitrary units (a.u.) for the metabolites associated with patients with asymptomatic and toxigenic carriage of CDAD. The data were extracted from Supplementary Table 2. *P*-values between the groups of samples, calculated by a Mann–Whitney *U*-test followed by a Benjamini–Hochberg correction to minimise false positives (the mean  $\pm$  s.e.m.), range from 0.0021 to 0.0381. The codes are as in Figure 1. Note that in many cases, the complete absence of metabolites was found, and accordingly, the metabolite level is absent in the Figure (value: zero).

2012). Notably, the production of the above metabolites was observed in patients carrying toxin + *C. difficile* strains (Supplementary Table 2).

The negative effects provoked by the carriage of toxin – *C. difficile* towards the above metabolites have not been reported to be associated with other diseases. Only creatinine and hypoxanthine have been statistically associated with aqueous faecal extracts of patients with irritable bowel syndrome (IBS) (Saric *et al.*, 2008). However, the elevated concentrations of those molecules in the corresponding faecal samples contrast with the complete depletion observed herein. This suggests that different diseases (for example, IBS and CDAD) promote different metabolic consequences. The lower bile acid diversity observed in the microbiota of patients with *C. difficile* that was negative for the production of toxins has also been associated with germ-free and antibiotic-treated rats (Swann *et al.*, 2011) and with antibiotic-treated mice with *C. difficile* infections (Theriot *et al.*, 2014). Interestingly, the low diversity observed herein was highly similar for all subjects with carriage of toxin – *C. difficile* regardless of the absence of antibiotic therapy (for example, HCC12) and treatments with different antibiotics (imipenem, levofloxacin, cefuroxime, meropenem, co-trimoxazol, metronidazole and/or vancomycin), whose dosage and duration also differed (in HCC2, HCC5 and HCC17). This finding, together with the different chemical nature of the bile acids detected compared with previous reports (Swann *et al.*, 2011; Theriot *et al.*, 2014), suggests that carriage of toxin – *C. difficile*, and not the antibiotic treatment, is the dominant factor driving the observed metabolic differences. The fact that secondary bile acids are most likely products of microbial metabolism

suggests that in this particular case, it is plausible that the chemical composition of faecal fluids (Swann *et al.*, 2011; Theriot *et al.*, 2014) and gut microbial metabolites (in this study) may be comparable.

#### *Carriage of toxin-positive C. difficile strains but not toxin-negative strains induces selective effects*

Opposite to what was observed in patients with *C. difficile* strains that were negative for the production of toxins, which showed a higher number of metabolic deficiencies relative to non-colonised controls, the production of toxins was mostly associated with selective biological deficiencies, namely, choline- and acetylputrescine-deficient microbiota, as both chemicals were absent in patients with carriage of toxin + *C. difficile* ( $P$ -value  $\leq 0.0095$ ) (Figure 3; Supplementary Table 2). Both chemicals were found at significantly lower levels in patients with carriage of toxin – *C. difficile* than in non-colonised control patients (6.1-fold for choline and 7.4-fold for acetylputrescine;  $P$ -value  $\leq 0.0095$ ), which suggests that the negative effect of *C. difficile* on the adsorption and/or metabolism of both metabolites is accentuated when toxins are produced by the bacteria. Note that deficiencies in the microbial metabolism of both compounds have been associated with the development of human diseases, including a predisposition to impaired glucose homeostasis, non-alcoholic fatty liver disease (Dumas *et al.*, 2006), inflammation and intestinal mucosa stability (Kibe *et al.*, 2014), immune system and intestinal maturation (Gómez-Gallego *et al.*, 2012). Recently, a high level of *N*-acetylputrescine was found in the faecal extracts of patients with ulcerative colitis and IBS (Le Gall



*et al.*, 2011), in clear contrast to its absence in the patients with toxin + *C. difficile* herein observed. This finding again suggests that different diseases (for example, ulcerative colitis, IBS and CDAD) may be associated with not only distinct gut metabolic profiles but also specific measurable metabolic changes in the gut microbiota. The fact that *N*-acetylputrescine is the result of the metabolic activity of gut microbiota (Murray *et al.*, 1993; Le Gall *et al.*, 2011; Kibe *et al.*, 2014) and that the microbial intracellular concentration of this chemical may be a good estimator of its corresponding concentration in faecal fluids through the action of specific transporters (Kibe *et al.*, 2014) may agree with this hypothesis. In the case of choline, such direct evidence could not be established, as choline is not a direct product of microbial activity. We speculate that the reduction of choline herein observed could be linked either to problems associated with transport inside the microbial cells (Kiene, 1998) or the reduction of bacteria that are more active in its metabolism. Whatever the case, the possibility that the production of toxins could select gut microbiota in humans that are unfavourable to the assimilation of choline as well as the metabolism of acetyl polyamines such as *N*-acetylputrescine may be suggested.

Although choline is an essential nutrient for higher organisms, including humans, it is also used by bacteria to produce intermediates that are used as carbon sources (Craciun and Balskus, 2012) and in the biosynthesis of cell lipids such as phosphatidylcholine (Martínez-Morales *et al.*, 2003). Phosphatidylcholine is a membrane-forming phospholipid that is present in approximately 10% of all bacteria (Sohlenkamp *et al.*, 2003). Thus, it is plausible that the carriage of toxin-producing *C. difficile* may potentially induce gut microbiota deficient in phosphatidylcholine-containing bacteria. This idea was confirmed by examining the composition of bacterial groups known to contain phosphatidylcholine (Sohlenkamp *et al.*, 2003) and by showing statistically significant differences ( $P$ -value < 0.05) between their average taxon abundance in patients with toxin + *C. difficile* (~0.11% of the total 16S rRNA sequences) and in patients with toxin - *C. difficile* (~1.46%) or in non-colonised controls (~7.56%). *N*-acetylputrescine, along with other acetyl polyamines, results from the breakdown of amino acids and their further metabolism (Chae *et al.*, 2012). We reasoned that the decreased activity of putrescine-modifying enzymes may result in reduced production of *N*-acetylputrescine inside gut bacterial cells. This theory was further confirmed biochemically by activity tests in which the extension of the biochemical production of *N*-acetylputrescine from its reaction substrate putrescine (plus acetyl-CoA) was evaluated using microbial protein extracts obtained as described previously (Hernández *et al.*, 2013). The results revealed that

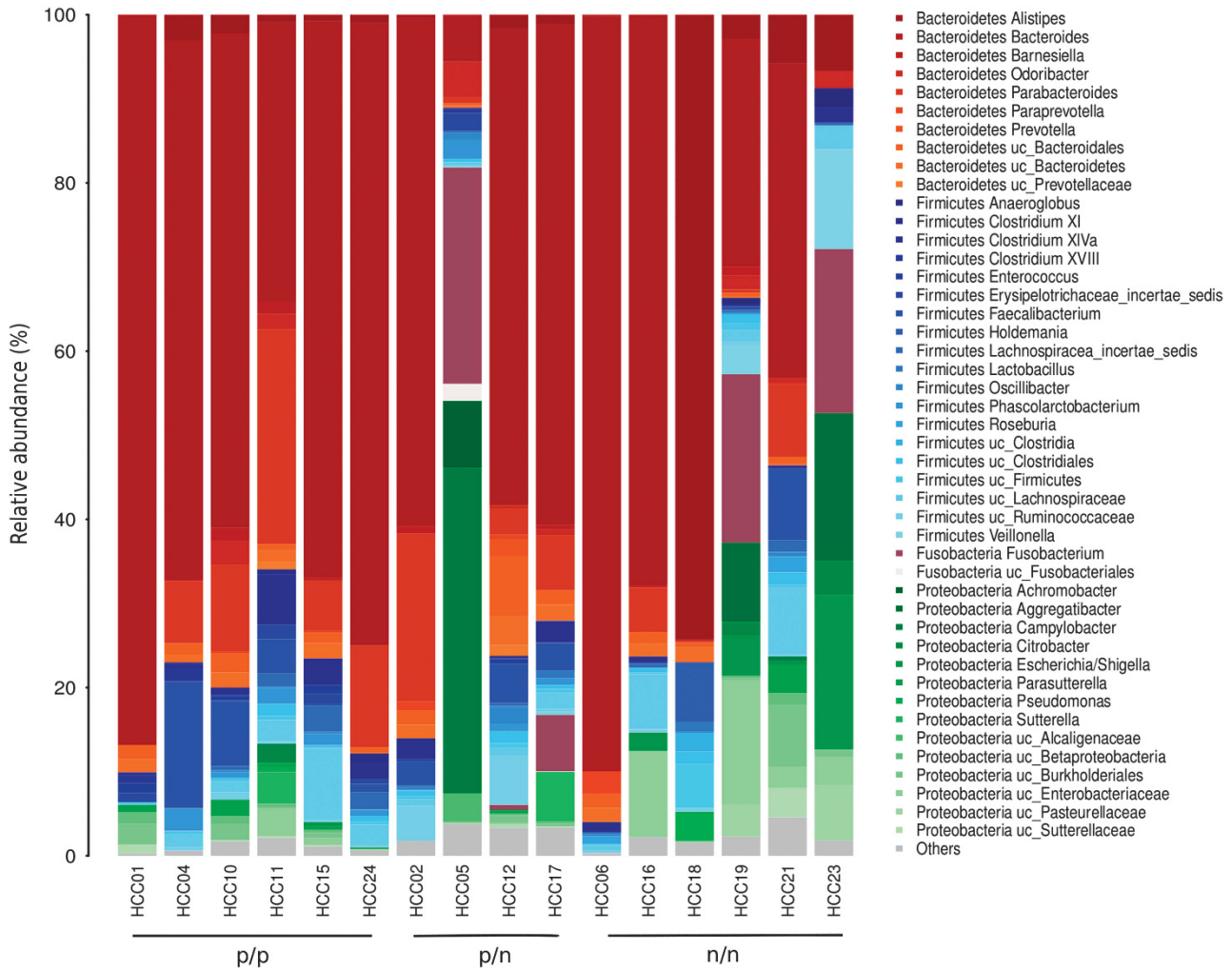
the carriage of toxigenic *C. difficile* is associated with microbial protein extracts ( $n=6$ ) that are not able to metabolise putrescine, whereas it was transformed to *N*-acetylputrescine and CoA in patients ( $n=4$ ) with toxin - *C. difficile* ( $25.0\% \pm 4.0\%$  conversion to CoA) and in non-colonised controls ( $82.5 \pm 8.4\%$  conversion to CoA;  $n=6$ ) at the end of the assay.

#### *Bacterial composition of faecal microbiota in the different groups of patients*

Figure 4 shows the bacterial composition of the faecal microbiota. The *Bacteroides* genus is the most abundant taxon in the three groups (p/p: 63.1%; p/n: 49.5%; n/n: 33%). The p/p and p/n groups showed a higher presence of *Parabacteroides* (9.5% and 7.9%, respectively), *Faecalibacterium* (5.9% and 3%, respectively) and the cluster *Clostridium* XIVa (1.9% and 1.5%, respectively). We found the cluster *Clostridium* XI (which includes *C. difficile*) in both groups but at very low abundance (p/p: 0.35%; p/n: 0.06%). The *Fusobacterium* genus belongs to the *Fusobacteria* phylum, which has been described to contain potential pathogens; it contains Gram-negative bacteria rather abundant in the p/n and n/n groups (7.1% and 10.9%, respectively). HCC5 also showed a high prevalence of the *Campylobacter* genus (38.7%).

Canonical correspondence analyses were performed to elucidate the relationship between the faecal microbiota composition of the different patient groups and external factors, such as the presence of *C. difficile* and toxin production. We found that the presence of *C. difficile* significantly affects the microbiota composition (Adonis,  $P$ -value = 0.05) (Figure 5). However, toxin production (Adonis,  $P$ -value = 0.612) did not support the bacterial composition differences that we detected using the metabolomics fingerprint of the gut microbiota (Figure 1).

To identify differences in microbial composition that may be related to the presence of *Clostridium* or to toxin production, we used the LEfSe biomarker discovery tool. We compared samples that were positive for the presence of *Clostridium* (the p/p and p/n groups) with those that were negative in the test (n/n group) (Figure 6a). In the *Clostridium*-positive groups, we found significant over-representation of taxa that mainly belonged to the *Firmicutes* phylum, such as *Phascolarctobacterium* ( $P$ -value = 0.0022), *Enterococcus* ( $P$ -value = 0.0054), *Clostridium* XI cluster ( $P$ -value = 0.0022), *Flavonifractor* ( $P$ -value = 0.032) and *Erysipelotrichaceae incertae sedis* ( $P$ -value = 0.0088). However, the *Blautia* ( $P$ -value = 0.046) and *Holdemania* ( $P$ -value = 0.017) genera and the *Enterobacteriaceae* ( $P$ -value = 0.047) and *Veillonellaceae* ( $P$ -value = 0.017) families were under-represented in these samples. Enrichment of the *Enterococcus* genus and of the *Clostridium* XI cluster and depletion of the *Blautia* genus and *Enterobacteriaceae* family have been described in other



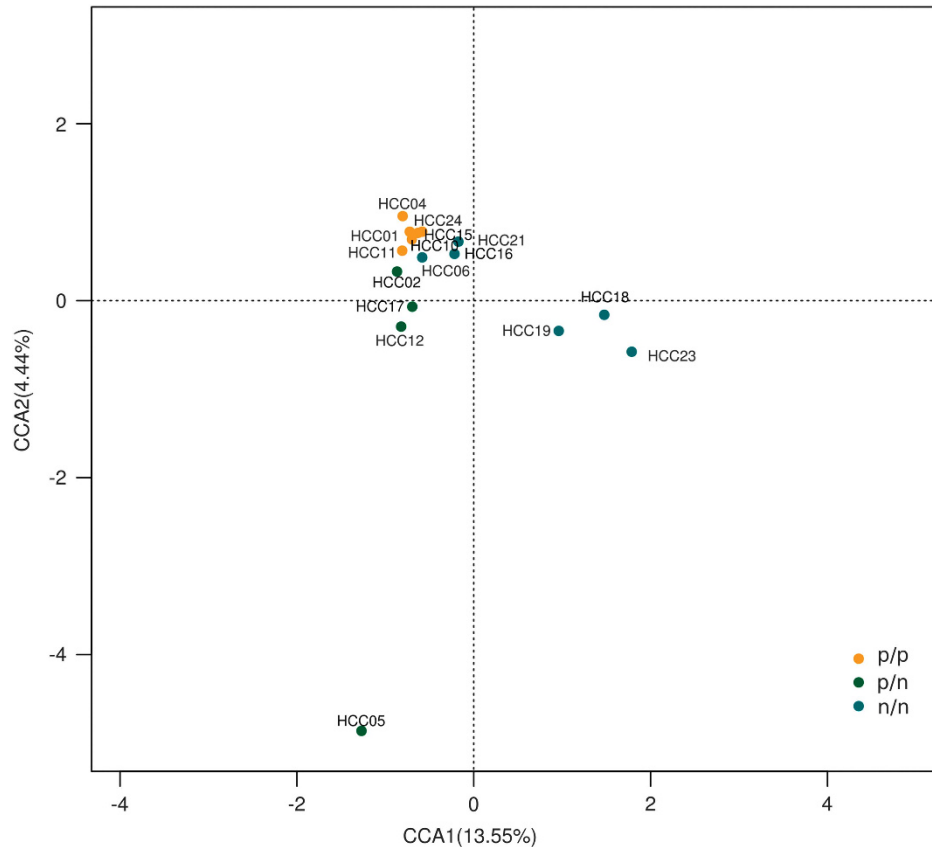
**Figure 4** Faecal microbiota composition of the samples. The composition of each sample is based on the RDP taxonomic assignment of the 16S rRNA sequences. The phylum and genus levels are shown for the most abundant bacterial groups (>1%). The three sample groups (p/p, p/n and n/n) are indicated.

faecal samples from patients infected by *C. difficile* (Pérez-Cobas *et al.*, 2014). The differences in bacterial composition between the toxin-positive (p/p) group and the toxin-negative (p/n and n/n) groups are shown in Figure 6b. The p/p group showed a high abundance of the *Enterococcus* genus ( $P$ -value = 0.042), the *Clostridium* XI cluster ( $P$ -value = 0.00068) and *Erysipelotrichaceae incertae sedis* ( $P$ -value = 0.0142) and a depletion of the *Fusobacterium* ( $P$ -value = 0.0024), *Prevotella* ( $P$ -value = 0.025), *Desulfovibrio* ( $P$ -value = 0.048) and *Dialister* ( $P$ -value = 0.048) genera.

## Discussion

The data generated in this study demonstrated that a single pathogen, such as *C. difficile* at very low abundance ( $\leq 0.35\%$  total 16S rRNA), can swiftly regulate the gastrointestinal microbial metabolism. This was shown by demonstrating that the presence

of *C. difficile*, regardless of patient characteristics and environmental factors, is associated with a consistent microbial metabolome landscape and that this profile was markedly different from that shown in patients suffering from diarrhoea caused by other presumptive pathogens. In addition, the fact that the mass signals were highly similar in all of the non-*C. difficile*-colonised control patients, regardless of the presence of putative different pathogens (for example, *Escherichia:Shigella*, *Aggregatibacter* or *Fusobacterium*, as detected by 16S rRNA), further suggests not only the prevalence of *C. difficile* (and thus, of CDAD) compared with other pathogens as a stronger driver of intestinal dysbiosis but also that pathogenic *Escherichia:Shigella*, *Aggregatibacter* or *Fusobacterium*, to cite some, most likely induced similar microbial metabolome patterns. This may open an interesting research window, namely, ranking the impact of different pathogens in the gut environment. This finding agrees with recent investigations that suggest that only a limited number of



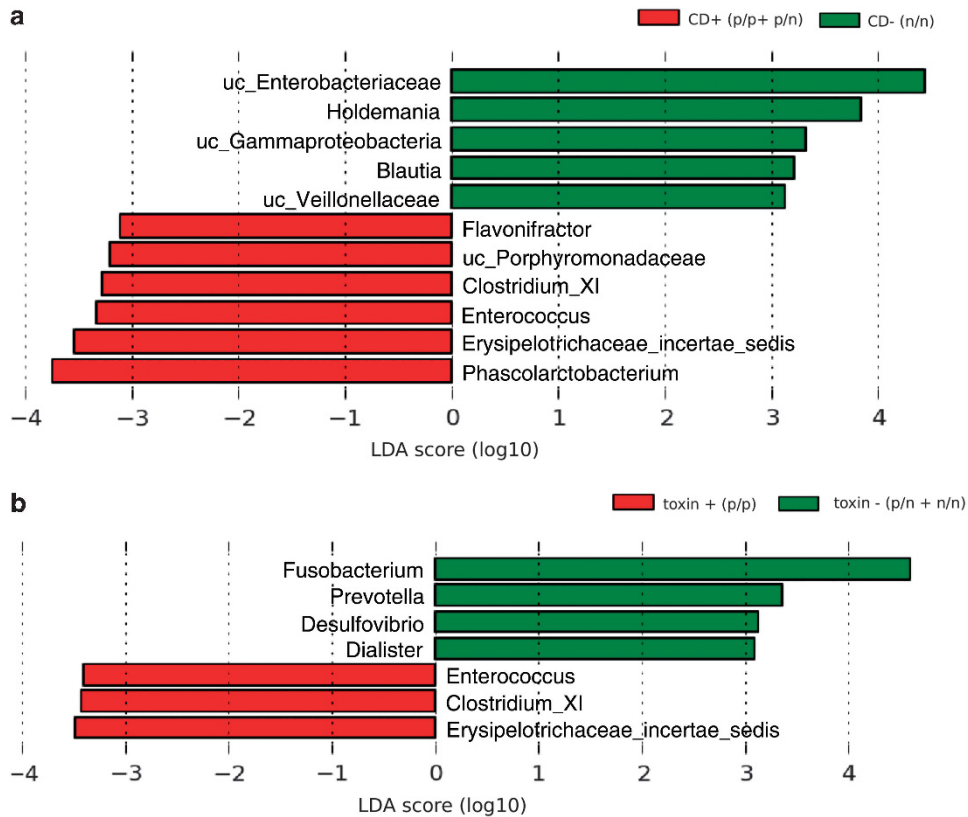
**Figure 5** Canonical correspondence analysis (CCA) based on taxa abundance and composition of gut microbiota. The data are colour coded as in Figure 1. For the sample codes, see Supplementary Table 1.

bacterial species can drive major changes in the gut microbiota and, thus, influence the metabolic profile (Marcobal *et al.*, 2013). By examining two variables, the presence of *C. difficile* (toxin -) and the presence of the toxin (toxin +), we observed that the effect of the toxin is demonstrated by the separation of the two classes. We reasoned that the changes imposed by *C. difficile* may become marked at the highest level of the functional hierarchy, that is, when considering the chemical (or metabolite) landscape of the gut microbiota representing the final metabolic outputs of biological systems (that is, the activity of the gut microbiota). This claim is supported by the demonstration that metabolic profile assessment of the gut microbiota may be better than 16S rRNA analysis for distinguishing different diseases, herein exemplified by CDAD and diarrhoea caused by pathogens other than *C. difficile*. In addition, the present work demonstrates that assessments of such metabolic profiles may allow the discrimination not only of different pathogenic diseases but also of different statuses (for example, toxin -/toxin +) of a disease. Therefore, a much more defined predictive behaviour emerged by investigating the metabolic patterns of gut microbiota (functional robustness) rather than by analysing the bacterial composition. This finding may open new research avenues related

to investigating how the heterogeneities that appear below the functional level (for example, microbiota composition based on 16S rRNA gene survey) that are commonly associated with patients with CDAD (Kuehne *et al.*, 2010; Theriot *et al.*, 2014; Knecht *et al.*, 2014), including the samples herein examined (Figures 4 and 5), can eventually result in distinctive metabolic patterns (Figure 1). This finding should also be of practical importance because although metabolites are downstream of gene expression, they also may modify gene expression.

The results presented in this study further revealed that the chemical compositions of gut microbiota in the three groups of examined samples were significantly different and that such separation occurred regardless of age, body mass index, disease duration, dietary intake, lifestyle-related factors, or medical and antibiotic uptake history of the patients. This finding, in turn, may suggest the prevalence of pathogens compared with other environmental factors (for example, age, antibiotic uptake and medical history) as a dominant factor that swiftly regulates gastrointestinal microbial metabolism. This could explain why age or different antibiotic uptakes may not promote measurable metabolic differences in patients with pathogen-caused diarrhoea and that the differences within patients in each of the three





**Figure 6** Taxonomic biomarkers. (a) LefSe analysis comparing *C. difficile* + in red, including the p/p and p/n groups, and *C. difficile* - in blue (n/n). (b) LefSe analysis comparing toxin + samples in red (p/p) and toxin - samples in blue (p/n and n/n). LDA scores (log<sub>10</sub>) for the most prevalent taxa in *C. difficile* - or toxin + are represented in the positive scale, whereas negative LDA scores indicate those taxa that are enriched in *C. difficile* + or toxin +.

groups investigated were small regardless of patient characteristics and medical history.

The effects of various pathophysiologies in the human gut metabolome have been previously examined (Saric *et al.*, 2008; Le Gall *et al.*, 2011; Marcobal *et al.*, 2013; Weir *et al.*, 2013; Bondia-Pons *et al.*, 2014). However, no clear associations between faecal fluid metabolome patterns and individual pathophysiologies (for example, weight gain or the presence of a disease) have been previously observed; this was mainly due to large inter-individual variation. Such inter-individual variation was also observed in the present study when examining the metabolic profiles of whole stool samples (Supplementary Figure 1). However, it was not observed when analysing the metabolic profiles of the gut microbiota (Figure 1), as the mass signals within grouped subjects using this approach were highly similar regardless of age, disease duration, dietary intake, lifestyle-related factors or medical history. One of the major differences between our study and previous studies examining total faecal material is that herein we focused on isolating metabolites from microbes isolated from stool material followed by metabolic profiling of gut microbiota rather than on examining total faecal fluids, which are known to contain a complex mixture of metabolites provided from the diet, the host and intestinal bacteria. Such complex

mixtures are commonly investigated in metabolomics studies (Saric *et al.*, 2008; Le Gall *et al.*, 2011; Marcobal *et al.*, 2013; Weir *et al.*, 2013; Walker *et al.*, 2014). Metabolites from intestinal bacteria (rather than dietary or host metabolites) are required to maintain and repair the large intestine and to support human health (Kibe *et al.*, 2014). Therefore, any metabolites that are directly produced or absorbed (from environmental inputs or the host) by gut microbes, and not those present in complex whole faecal fluids, may be relevant and useful indicators not only for investigating gut physiology but also for determining the role of such metabolites in pathophysiologies and human health. Note, however, that our study relates to metabolite levels inside gut bacterial cells, which may be interpreted differently from metabolite levels in plasma or, to some extent, in faecal fluids; therefore, a direct relation may not be established (see additional comments in Supplementary Discussion). In addition, we are aware that the different extraction methods, analytical platforms and biological fluids used in this work and in previous investigations preclude any clear-cut conclusion about the differences between this report and previous reports. Whatever the case, we believe that this investigation will provide information that can be directly linked to complementary microbial data (for example, 16S

rRNA gene profiles). Such links are otherwise difficult to achieve when non-microbial metabolites (from the environment or host), which are commonly considered when working with whole faecal material, are investigated.

We are aware that our study reveals metabolic signatures that are associated with CDAD and toxins and that it is important to find a plausible mechanism for these associations. Indeed, based on the data generated in this study, it is plausible that differences in the production of key metabolites with major biological relevance or significance, without substantial loss of metabolic function, may be responsible for the association of toxin – or toxin + carriage with marked, stable and measurable metabolic changes in the gut microbiota. In agreement with this hypothesis, we observed deficiencies in chemical species known to be of global importance for the host, such as those regulating host chronic intestinal inflammation, temperature, locomotion, inflammation, iron deficiency, glucose homeostasis, liver and renal dysfunction, intestinal mucosa stability, immune system, intestinal maturation, lipotoxicity and apoptosis. We also observed deficiencies in chemical species that are important for gut microbial ecosystems (for instance, impaired microbial cell growth, signal transduction and initiation of translation and nucleic acid synthesis). It is also notable that, in many cases, minor differences (for example, at the population level) have been demonstrated to induce physiological changes. As an example, it has recently been demonstrated that one or two strains are sufficient to drive major changes in gastrointestinal and host (mouse) metabolic profiles, where up to  $10^{12}$  microbial cells or >500 species may coexist (Marcobal *et al.*, 2013). That said, in this study, we also demonstrated for the first time that carriage of toxin – *C. difficile* is associated with major biological deficits relative to the carriage of toxin + *C. difficile* strains. The toxin – *C. difficile* strains most likely provoked selective effects in specific, cell-critical systems, such as an inability to adsorb and/or metabolise choline and acetylputrescine.

As demonstrated, only 725 of 115 228 (or 0.63% of the total) mass features were found to significantly differ among the three groups. It is therefore important to evaluate whether such subtle differences can be considered to fall within a common range. It is notable, however, that no report to date has described the metabolomic profiling either of bacterial faecal extracts or of faecal fluids from patients with CDAD; therefore, little is known about whether the observed differences among the groups of samples are within a common range. In the case of other pathophysiology, when using metabolite profiling of faecal fluids, it was found that: (i) only 18 faecal metabolites allowed discrimination between ulcerative colitis or IBS and healthy control patients (Le Gall *et al.*, 2011); (ii) 99 metabolites allowed discrimination of humanised and

gnotobiotic mice, although they possess quite distinct microbiota (85% of the microbial genera and species are different) (Marcobal *et al.*, 2013); (iii) 43 metabolites were found to differ when comparing the human, mouse and rat faecal metabolomes (Saric *et al.*, 2008); and (iv) 22 metabolites allow the discrimination of patients with colorectal cancer from healthy adults (Weir *et al.*, 2013). In a different context, in the case of subjects that are discordant for weight, few examples exist in the literature that have examined faecal fluid metabolomes. Thus, it is noteworthy that a recent report examining the caecal metabolome revealed that only 65 out of a total of 10 515 mass signals (or 0.7% of the total) were significantly associated with a high-fat diet (Walker *et al.*, 2014). In a different study, only 22 faecal metabolites were differentially produced in monozygotic twin pairs that were discordant for weight (Bondia-Pons *et al.*, 2014). Therefore, based on bibliographic records in the specialised literature, the subtle differences associated with *C. difficile* that were positive or negative for the production of toxins and with non-colonised control patients reported in this study can be considered to be within (or even higher than) the common range reported for other pathophysiology.

Finally, a question may arise as to the causes of such metabolite alterations. The results support the notion that, for the majority of metabolites, the observed differences are most likely due to deficiencies in the metabolite-modifying enzyme (experimentally demonstrated here for putrescine) and/or differences in the abundance of bacterial groups known to transport or metabolise metabolites (demonstrated here by examining bacteria that metabolise choline-like derivatives). Before drawing any clear conclusion about the exact causes that induce such changes and to ascertain whether the differences are a result of deficiencies in metabolite absorption (as suggested here for choline) into bacterial cells, different microbial compositions (as demonstrated here for phosphatidylcholine) or different intracellular metabolite production (as demonstrated here for putrescine), further experimental analysis is needed. In any case, additional questions remain of how environmental or host inputs (metabolites) are transported into gut microbial cells and how intracellular metabolites produced by bacterial cells can be secreted into the gut environment. Here, we should consider that bacterial cells contain multiple transport systems. As an example, flavonoids such as phlorizin, which was detected in this study, are likely to move via vesicle-mediated transport or through the membrane-bound transporters of the ATP binding cassette or multidrug and toxic extrusion compound families (Hassan and Mathesius, 2012). In the case of choline, its transport into cells is believed to be dependent on the proton motive force (Kiene, 1998), and bacteria can transport putrescine derivatives into the cytoplasm with specific transporters (Kibe *et al.*, 2014). Notably, the

presence of a metabolite of non-microbial origin inside a bacterial cell suggests that the bacterium has the ability to absorb it, although any further metabolic steps require experimental evidence.

The degree to which socioeconomic, ethnic and geographic factors and other patient and medical (for example, antibiotic treatment) factors may affect the results of the metabolic profiles of the gut microbiota should also be investigated further in a larger cohort of subjects. In addition, the mechanistic basis for the metabolic associations observed herein and the implications of the biological deficits that were presumptively identified will require further investigation.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgements

The present investigation was funded by the Spanish Ministry of Economy and Competitiveness and the Federal Ministry of Education and Research (BMBF) within the ERA NET PathoGenoMics2 program, grant number 0315441A. This work was further funded by grants BFU2008-04501-E, SAF2009-13032-C02-01, SAF2012-31187, CSD2007-00005 and BIO2011-25012 from the Spanish Ministry of Economy and Competitiveness, PrometeoII/2'14/065 from Generalitat Valenciana (Spain) and AGL2006-11697/ALI. We gratefully acknowledge the financial support provided by the European Regional Development Fund (ERDF). We thank Rafael Bargiela for his excellent support in relation to the preparation of Figure 3.

## References

- Bondia-Pons I, Maukonen J, Mattila I, Rissanen A, Saarela M, Kaprio J *et al.* (2014). Metabolome and fecal microbiota in monozygotic twin pairs discordant for weight: a Big Mac challenge. *FASEB J pii: fj.14-250167*.
- Borody TJ, Peattle D, Kapur A. (2014). Could fecal microbiota transplantation cure all *Clostridium difficile* infections? *Future Microbiol* **9**: 1–3.
- Britton RA, Young VB. (2014). Role of the intestinal microbiota in resistance to colonization by *Clostridium difficile*. *Gastroenterology* **146**: 1547–1553.
- Chae YK, Kim SH, Ellinger JJ, Markley JL. (2012). Tracing metabolite footsteps of *Escherichia coli* along the time course of recombinant protein expression by two-dimensional NMR spectroscopy. *Bull Korean Chem Soc* **33**: 4041–4046.
- Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al.* (2009). The ribosomal database project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* **37**: 141–145.
- Craciun S, Balskus EP. (2012). Microbial conversion of choline to trimethylamine requires a glyceryl radical enzyme. *Proc Natl Acad Sci USA* **109**: 21307–21312.
- Dostal A, Lacroix C, Pham VT, Zimmermann MB, Del'homme C, Bernalier-Donadille A *et al.* (2014). Iron supplementation promotes gut microbiota metabolic activity but not colitis markers in human gut microbiota-associated rats. *Br J Nutr* **111**: 2135–2145.
- Dumas ME, Barton RH, Toye A, Cloarec O, Blancher C, Rothwell A *et al.* (2006). Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proc Natl Acad Sci USA* **103**: 12511–12516.
- Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N *et al.* (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat Protoc* **6**: 1060–1083.
- Fordtran JS. (2006). Colitis due to *Clostridium difficile* toxins: underdiagnosed, highly virulent, and nosocomial. *Proc (Bayl Univ Med Cent)* **19**: 3–12.
- Gómez-Gallego C, Collado MC, Ilo T, Jaakkola UM, Bernal MJ, Periago MJ *et al.* (2012). Infant formula supplemented with polyamines alters the intestinal microbiota in neonatal BALB/cOlaHsd mice. *J Nutr Biochem* **23**: 1508–1513.
- Grimble GK. (1994). Dietary nucleotides and gut mucosal defence. *Gut* **35**: S46–S51.
- Hassan S, Mathesius U. (2012). The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interactions. *J Exp Bot* **63**: 3429–3444.
- Hernández E, Bargiela R, Diez MS, Friedrichs A, Pérez-Cobas AE, Gosalbes MJ *et al.* (2013). Functional consequences of microbial shifts in the human gastrointestinal tract linked to antibiotic treatment and obesity. *Gut Microbes* **4**: 306–315.
- Kelly CP, LaMont LT. (2008). *Clostridium difficile* – more difficult than ever. *N Engl J Med* **359**: 1932–1940.
- Kibe R, Kurihara S, Sakai Y, Suzuki H, Ooga T, Sawaki E *et al.* (2014). Upregulation of colonic luminal polyamines produced by intestinal microbiota delays senescence in mice. *Sci Rep* **4**: 4548.
- Kiene RP. (1998). Uptake of choline and its conversion to glycine betaine by bacteria in estuarine waters. *Appl Environ Microbiol* **64**: 1045–1051.
- Knecht H, Neulinger SC, Heinsen FA, Knecht C, Schilhabel A, Schmitz RA *et al.* (2014). Effects of  $\beta$ -lactam antibiotics and fluoroquinolones on human gut microbiota in relation to *Clostridium difficile* associated diarrhea. *PLoS One* **9**: e89417.
- Kuehne SA, Cartman ST, Heap JT, Kelly ML, Cockayne A, Minton NP. (2010). The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature* **467**: 711–713.
- Kurek K, Łukaszuk B, Piotrowska DM, Wiesiołek P, Chabowska AM, Zendzian-Piotrowska M. (2013). Metabolism, physiological role, and clinical implications of sphingolipids in gastrointestinal tract. *Biomed Res Int* **2013**: 908907.
- Lauzon CR, Potter SE, Prokopy RJ. (2003). Degradation and detoxification of dihydrochalcone Phlorizin by *Enterobacter agglomerans*, a bacterium associated with the apple pest, *Rhagoletis pomonella* (Walsh) (Diptera: Tephritidae). *Environ Entomol* **32**: 953–962.
- 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. *J Proteome Res* **10**: 4208–4218.
- Longato L, Tong M, Wands JR, de la Monte SM. (2012). High fat diet induced hepatic steatosis and insulin resistance: role of dysregulated ceramide metabolism. *Hepato Res* **42**: 412–427.
- Marcobal A, Kashyap PC, Nelson TA, Aronov PA, Donia MS, Spormann A *et al.* (2013). A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *ISME J* **7**: 1933–1943.
- Martínez-Morales F, Schobert M, López-Lara IM, Geiger O. (2003). Pathways for phosphatidylcholine biosynthesis in bacteria. *Microbiology* **149**: 3461–3471.
- Murray KE, Shaw KJ, Adams RF, Conway PL. (1993). Presence of N-acyl and acetoxy derivatives of putrescine and cadaverine in the human gut. *Gut* **34**: 489–493.
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB *et al.* (2011). Vegan: community ecology package. *R package version 1*: 17–19; <http://CRAN.R-project.org/package=vegan>.
- Peral MJ, García-Delgado M, Calonge ML, Durán JM, De La Horra MC, Wallimann T *et al.* (2002). Human, rat and chicken small intestinal Na<sup>+</sup> – Cl<sup>-</sup> -creatine transporter: functional, molecular characterization and localization. *J Physiol* **545**: 133–144.
- Pérez-Cobas AE, Gosalbes MJ, Friedrichs A, Knecht H, Artacho A, Eismann K *et al.* (2013). Gut microbiota disturbance during antibiotic therapy: a multi-omic approach. *Gut* **62**: 1591–1601.
- Pérez-Cobas AE, Artacho A, Ott SJ, Moya A, Gosalbes MJ, Latorre A. (2014). Structural and functional changes in the gut microbiota associated to *Clostridium difficile* infection. *Front Microbiol* **5**: 335. doi:10.3389/fmicb.2014.00335.
- Peterfreund GL, Vandivier LE, Sinha R, Marozsan AJ, Olson WC, Zhu J *et al.* (2012). Succession in the gut microbiome following antibiotic and antibody therapies for *Clostridium difficile*. *PLoS One* **7**: e46966.
- Qiu B, Pothoulakis C, Castagliuolo I, Nikulasson Z, LaMont JT. (1996). Nitric oxide inhibits rat intestinal secretion by *Clostridium difficile* toxin A but not *Vibrio cholerae* enterotoxin. *Gastroenterology* **111**: 409–418.
- R Development Core Team (2011), R: A language and environment for statistical computing. R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Rea MC, Dobson A, O'Sullivan O, Crispie F, Fouhy F, Cotter PD *et al.* (2011). Effect of broad- and narrow-spectrum antimicrobials on *Clostridium difficile* and microbial diversity in a model of the distal colon. *Proc Natl Acad Sci USA* **108**: 4639–4644.
- Salzet M, Tasiemski A. (2001). Involvement of pro-enkephalin-derived peptides in immunity. *Dev Comp Immunol* **25**: 177–185.
- Saric J, Wang Y, Li J, Coen M, Utzinger J, Marchesi JR *et al.* (2008). Species variation in the fecal metabolome gives insight into differential gastrointestinal function. *J Proteome Res* **7**: 352–360.
- Senkal CE, Ponnusamy S, Manevich Y, Meyers-Needham M, Saddoughi SA, Mukhopadhyay A *et al.* (2011). Alteration of ceramide synthase 6/C16-ceramide induces activating transcription factor 6-mediated endoplasmic reticulum (ER) stress and apoptosis via perturbation of cellular Ca<sup>2+</sup> and ER/Golgi membrane network. *J Biol Chem* **286**: 42446–42458.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.
- Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS *et al.* (2011). Metagenomic biomarker discovery and explanation. *Genome Biol* **12**: R60.
- Spor A, Koren O, Ley R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat Rev Microbiol* **9**: 279–290.
- Sohlenkamp C, López-Lara IM, Geiger O. (2003). Biosynthesis of phosphatidylcholine in bacteria. *Prog Lipid Res* **42**: 115–162.
- Swann JR, Want EJ, Geier FM, Spagou K, Wilson ID, Sidaway JE *et al.* (2011). Systemic gut microbial modulation of bile acid metabolism in host tissue compartments. *Proc Natl Acad Sci USA* **108**: 4523–4530.
- Tan B, O'Dell DK, Yu YW, Monn MF, Hughes HV, Burstein S *et al.* (2010). Identification of endogenous acyl amino acids based on a targeted lipidomics approach. *J Lipid Res* **51**: 112–119.
- Theriot CM, Koenigsnecht MJ, Carlson PE Jr, Hatton GE, Nelson AM, Li B *et al.* (2014). Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to *Clostridium difficile* infection. *Nat Commun* **5**: 3114.
- Walker A, Pfitzner B, Neschen S, Kahle M, Harir M, Lucio M *et al.* (2014). Distinct signatures of host-microbial meta-metabolome and gut microbiome in two C57BL/6 strains under high-fat diet. *ISME J* **8**: 2380–2396.
- Weir TL, Manter DK, Sheflin AM, Barnett BA, Heuberger AL, Ryan EP. (2013). Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. *Plos One* **8**: e70803.
- Yoon G, Kim KO, Lee J, Kwon D, Shin JS, Kim SJ *et al.* (2002). Ceramide increases Fas-mediated apoptosis in glioblastoma cells through FLIP down-regulation. *J Neurooncol* **60**: 135–141.
- Yoshimasa T, Nakao K, Ohtsuki H, Li S, Imura H. (1982). Methionine-enkephalin and leucine-enkephalin in human sympathoadrenal system and pheochromocytoma. *J Clin Invest* **69**: 643–650.

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