Increased sporulation underpins adaptation of *Clostridium difficile* strain 630 to a biologically–relevant faecal environment, with implications for pathogenicity

Nigel George Ternan¹, Nicola Diana Moore¹, Deborah Smyth¹, Gordon James McDougall², James William Allwood², Susan Verrall², Christopher Ian Richard Gill¹, James Stephen Gerard Dooley¹ & Geoff McMullan³

*Clostridium difficile* virulence is driven primarily by the processes of toxinogenesis and sporulation, however many *in vitro* experimental systems for studying *C. difficile* physiology have arguably limited relevance to the human colonic environment. We therefore created a more physiologically–relevant model of the colonic milieu to study gut pathogen biology, incorporating human faecal water (FW) into growth media and assessing the physiological effects of this on *C. difficile* strain 630. We identified a novel set of *C. difficile*–derived metabolites in culture supernatants, including hexanoyl– and pentanoyl–amino acid derivatives by LC-MS. Growth of *C. difficile* strain 630 in FW media resulted in increased cell length without altering growth rate and RNA sequencing identified 889 transcripts as differentially expressed (p < 0.001). Significantly, up to 300–fold increases in the expression of sporulation–associated genes were observed in FW media–grown cells, along with reductions in motility and toxin genes’ expression. Moreover, the expression of classical stress–response genes did not change, showing that *C. difficile* is well–adapted to this faecal milieu. Using our novel approach we have shown that interaction with FW causes fundamental changes in *C. difficile* biology that will lead to increased disease transmissibility.

The Gram–positive spore–forming anaerobe *Clostridium difficile* is recognised as one of the major causes of health–care associated infections¹ and exerts a negative and well–publicised impact on hospital morbidity and mortality rates. *C. difficile* infection (CDI) can develop when broad spectrum antibiotics are deployed to treat underlying infections: they disrupt the body’s natural colonic microbiota thus allowing development of CDI if spores or cells of this multidrug–resistant pathogen are also present. Colonisation of the host gastrointestinal tract depends on the germination of *C. difficile* spores, with subsequent growth of vegetative cells and the release of two large clostridial glycosylating toxins, toxin A and toxin B. These toxins are responsible for the inflammation and epithelial tissue damage that results in rapid loss of fluid and consequent diarrhoea. Clinical manifestations and severity of CDI vary from mild self-limiting diarrhoea to life-threatening pseudomembranous colitis and, in severe cases, death.

While it is well established that gut microbiome disruption by antibiotics can lead to the development of *C. difficile* infection (CDI), the mechanisms underlying *C. difficile* expansion after microbiota disturbance are only just emerging. Both dietary and microbiota compositional changes have been demonstrated to lead to alterations in the colonic environment that favour or suppress certain enteric pathogens such as *C. difficile* and indeed...
C. difficile virulence has been linked to the ability to both effectively utilise nutrients in the dysbiotic gut environment\textsuperscript{11} and to sporulate\textsuperscript{3}. During CDI, there is rapid expansion of the C. difficile vegetative cell population, with subsequent production of the two proven virulence factors, the toxins (A&B), and spores which serve as the transmissible elements\textsuperscript{12,13}. While a variety of model systems that are indispensable in the study of C. difficile pathogenesis have been described\textsuperscript{16}, studies with humans are, by contrast, limited to prospective or retrospective sampling and elucidation of C. difficile strain variants. While it could be argued that many of the experimental systems that exist to study C. difficile pathogenesis have rather limited relevance to the human gut, recent work with in vitro continuous flow bioreactors has elegantly demonstrated the increased competitive fitness of ribotype 027 C. difficile strains in a mixed microbiota model\textsuperscript{15} and shown that microbial communities representative of key features of the gut can be cultivated and manipulated successfully\textsuperscript{18}. Such in vitro models have also been used to investigate and model antibiotic exposure\textsuperscript{19}, intestinal biofilm development\textsuperscript{20} and genomic stability of C. difficile during simulated infection experiments\textsuperscript{21}. In other systems that seek to investigate changes within the human gut, faecal water (FW), the aqueous phase of human faeces\textsuperscript{22}, is an attractive means of linking changes in colonic contents with gut health outcomes\textsuperscript{23–27}. FW has been used as a biologically–relevant challenge agent in a range of gut studies\textsuperscript{28–30} as it contains a variety of unbound, soluble components including bile acids, fatty acids, amino acid residues and derivatives, (poly)phenols, and short–chain fatty acids\textsuperscript{31–33}. These metabolites are likely to modulate the function and composition of the microbiome. To allow physiologically–relevant modelling of C. difficile under controlled culture conditions, we have incorporated FW into growth media in order to better mimic the human gut environment. Using our novel approach we now demonstrate for the first time that expression of genes essential for pathogenesis are significantly differentially expressed in the human faecal water milieu.

Results and Discussion

C. difficile cell length increases, yet overall population growth is unaffected in FW media. To create a physiologically more relevant model of the colonic milieu to study gut pathogen biology, a pooled faecal water (FW) sample was produced as previously described\textsuperscript{32} from two male donors (age 40+/-2 years). We characterised the FW using LC–MS\textsuperscript{8} and demonstrated that it contained components identified in previous investigations\textsuperscript{32,34–36}. Some 30 FW components, many of which are known constituents of faeces (e.g. stercobilin and urobilinogen) were identified (Table 1), while others gave MS data and putative IDs consistent with previous analyses including a number of bile acid derivatives\textsuperscript{37} as well as some unidentified components.

To test the hypothesis that the presence of faecal water would change C. difficile physiology in a way more reflective of the in vivo environment, we compared growth in faecal water/BHIS growth media (“FW media”) with the BHIS control. FW media was not detrimental to growth of C. difficile strain 630 over 6 h (Fig. 1A), although we noted that bacterial cell length increased by almost 70% at 6 h in FW media (4.3 m versus 3.3 m, p = 0.015) (Fig. 1B). (Supplementary Data File 1, Tables S1 and S2)

C. difficile generates novel, amino-acid derived metabolites during growth in FW media. Having demonstrated that C. difficile population growth was not affected by the presence of FW components, we determined whether any of the FW components were utilised by C. difficile during growth, or if any bacterially–derived metabolites could be identified in culture supernatants. FW–derived components (n = 30) were not subject to further metabolism or degradation by C. difficile and were largely unchanged between the start and end of the 6 h incubation (Fig. 2A). The main trend was towards a slight increase in abundance during incubation (Fig. 2B, Table 2), possibly caused by release of FW components bound to BHI media constituents during growth of C. difficile.

We focused on metabolites whose abundance consistently increased during incubation, using the XCMS process with data checking to eliminate low abundance peaks, adducts and multiply charged ions (see Supplementary Data File 1, Figs S1–S6). Overall, there was an increase in certain components, most notably a set of amino acids esterified to hexanoic and pentanoic acids including glycine, lysine, alanine, tyrosine, methionine, valine, phenylalanine, tryptophan and leucine and/or isoleucine. No threonine, arginine, histidine, asparagine, aspar- tate, cysteine, glutamine, proline or serine hexanoyl derivatives were found. These putative hexanoyl amino-acid derivatives yielded characteristic MS\textsuperscript{2} fragments which suggested fragmentation at the amide bond to produce the M\textsuperscript{+–H} ion of the amino acid and a neutral loss of 98 atomic mass units (amu), which could be due to the aldehyde, hexenal (C\textsubscript{6}H\textsubscript{10}O). In a similar manner, fragmentation of pentanoyl derivatives produced the M\textsuperscript{+–H} ion of the amino acid and a neutral loss of 84 amu due to the aldehyde, pentenal, (C\textsubscript{6}H\textsubscript{10}O; 14 amu less than hexenal). Other common fragmentations can be assigned (i.e. neutral loss of 42 (C\textsubscript{3}H\textsubscript{6}) and 22 amu, neutral loss of 56 (C\textsubscript{4}H\textsubscript{8}) = butene), neutral loss of 70 (C\textsubscript{5}H\textsubscript{10} = pentene) and are consistent with such structures. Other neutral losses such as 44 and 46 amu, respectively, can be ascribed to CO\textsubscript{2} and formic acid from the amino acid components. These putative hexanoyl and pentanoyl amino acid derivatives have not previously been identified in cultures of C. difficile but this microorganism is known to produce isocaproic acid (also known as isohexanoic acid) during growth and the accumulation of this C6 fatty acid has previously been used as a diagnostic for C. difficile in stool samples\textsuperscript{38,39}. Investigations into biofuel production by Clostridia have shown that hexanoyl–coA is a key metabolite for the production of hexanol\textsuperscript{40} and the formation of these putative hexanoyl and pentanoyl amino-acid derivatives may be a consequence of growth of C. difficile in the amino acid–rich BHI media. Earlier work showing isocaproic acid accumulation\textsuperscript{38,39} used less sensitive GC–MS techniques that required sample derivatisation and thus would not have detected the amino acid derivatives we identified. Of all the metabolites whose abundance increased during growth of C. difficile (Table 2), only hexanoyl lysine (peak 5, m/z 243) was higher in the BHI media than in FW media (Supplementary Data File 1, Fig. S5). In fact, growth of C. difficile in the presence of FW did not reduce the levels of any of the other putatively–identified novel components (i.e. six components were substantially increased and four components marginally increased).
Table 1. Components identified in faecal water. Most of these peaks were present in media at the end of the incubation. Most were unchanged during growth of Clostridium difficile strain 630 but some were marginally increased. Databases: PB = PubChem (https://pubchem.ncbi.nlm.nih.gov/); ChEBI (http://www.ebi.ac.uk/chebi/); LM - LIPID MAPS (http://www.lipidmaps.org/data/structure/); HMDB (http://www.hmdb.ca/).

Considerable modulation of the C. difficile transcriptome results from growth in FW media.

Some 889 C. difficile strain 630 transcripts were differentially expressed (DE) (pad < 0.001, fold-change (FC) > 1.45) as a result of exposure to faecal water with 497 (56%) exhibiting increased expression and 392 (41%) being decreased (Supplementary Table S3). The largest numbers of DE genes were in categories “Similar to unknown proteins” (16%), “Transport binding proteins and lipoproteins” (14.4%), “Metabolism of amino acids and related molecules” (7.9%), “Transposon and IS function” (6.2%), “Sporulation” (6.8%), “Specific metabolic pathways” (5.7%), and regulation of RNA synthesis (4.8%) (Fig. 3). Orthogonal validation of expression changes using qRT-PCR showed good correlation (R^2 = 0.95) between RNAseq and qRT-PCR data (Fig. 4) (Supplementary Data File 1, Table S4) when applied to a number of motility and sporulation genes.

We noted that within the 889 DE genes, a slightly larger proportion exhibited increased expression in FW media with the exception of those in the categories of signal transduction, motility, genes associated with specific pathways, metabolism of co-factors, and detoxification (Fig. 5). Considerable expression changes in the transcriptional programme of C. difficile strain 630 were apparent in regard to genes associated with sporulation, protein synthesis and protein modification. Nine sporulation-associated genes exhibited transcriptional programme of C. difficile strain 630 but some were marginally increased. Databases: PB = PubChem (https://pubchem.ncbi.nlm.nih.gov/); ChEBI (http://www.ebi.ac.uk/chebi/); LM - LIPID MAPS (http://www.lipidmaps.org/data/structure/); HMDB (http://www.hmdb.ca/).
RNA metabolism, protein modification, adaptation to atypical conditions and those categorised as miscellaneous exhibited fold changes of no more than 20–fold. The least extensively–changed genes were those involved with metabolism of phosphorus, sulphur, lipids, and coenzymes (Fig. 6), indicating that these central metabolic pathways were relatively unperturbed in the presence of FW. We have previously demonstrated apparent robustness of \textit{C. difficile} central metabolic pathways under mild heat stress\textsuperscript{44–46}, however the extreme perturbations in sporulation, transport and conserved hypothetical protein–encoding genes led us to consider these biological processes, and their implications for the lifestyle of this important pathogen, in detail.

\textit{C. difficile} sporulation gene expression increases during growth in FW media. In the model Gram positive organism \textit{Bacillus subtilis}, and also in \textit{C. difficile}, sporulation is initiated by a two–component system with Spo0A and associated kinases\textsuperscript{47,48}, leading to the sequential, compartment–specific activation of the strictly conserved sporulation–specific sigma factors, $\sigma$H (early), $\sigma$F, $\sigma$E, $\sigma$G, and $\sigma$K (late)\textsuperscript{47,49}. Thereafter, however, differences have been shown in the order, activation, and function of the sigma factors in \textit{C. difficile}\textsuperscript{50}.

Genome–wide mutational analyses of sigma factor function in \textit{C. difficile} have revealed that while their transcriptional and functional sequence ($\text{sigE} & \text{sigF} – \text{early, sigG} & \text{sigK} – \text{late}$) is broadly conserved with the \textit{B. subtilis} model, there are differences in the \textit{C. difficile} developmental programme\textsuperscript{51}. Spores were not visible in cultures of sigma factor mutants, indicating their critical role in the various stages of sporulation and in the production of mature, heat-resistant spores\textsuperscript{42}. Intriguingly, and in contrast to \textit{B. subtilis}, the activity of $\sigma$E was partially independent of $\sigma$F; $\sigma$G or $\sigma$K did not require $\sigma$F or $\sigma$G, respectively and $\text{sigG}$ transcription was not dependent on $\sigma$F\textsuperscript{41,42}. Taken together, the published data suggests minimal intercompartmental communication and a weaker connection between forespore and mother cell\textsuperscript{50}, in addition to a looser association between gene expression and morphology in \textit{C. difficile}\textsuperscript{41,51}. While sporulation in \textit{C. difficile} is more akin to \textit{B. subtilis} than to other Clostridia, it at the same time represents a more ancestral, less tightly–controlled sporulation programme that facilitates a degree of population heterogeneity during infection\textsuperscript{52}. The recent work of Browne \textit{et al.}\textsuperscript{53} showed extensive sporulation ability within the human gut microbiota, with many taxa present in the spore form. Animal models of \textit{C. difficile} infection have revealed that in mouse, \textit{C. difficile} spores germinate by 6h post–infection, leading to pathogenic lesions and heat–resistant spores – comprising up to 20% of the total cfu – being detectable 24–36h post–challenge. Thus, while sporulation under infection conditions takes some time) with \textit{C. difficile} strain 630\textsuperscript{55},
the process happens as early as 6 h post–infection with the ribotype 027 strain, *C. difficile* R20291 [54]. Temporal examination of *C. difficile* 630 gene expression during infection in mouse [55] and pig ligated loop models [56] has shown that many genes associated with host adaptation, all stages of sporulation and a diversity of genes encoding “hypothetical proteins” were expressed at increased levels *in vivo*, indicating their importance in the infection process and the requirement for extensive remodelling of the transcriptome during infection. We thus hypothesised that FW would induce sporulation in *C. difficile* and sporulation genes were indeed some of the most

Figure 2. Metabolite peaks that changed during growth of *Clostridium difficile* strain 630 in faecal water media. (A) Metabolite peaks derived from faecal water (Identities in Supplementary Data File 1, Table S1). (B) Metabolite Peaks that changed during growth of *C. difficile* strain 630. Peaks A–R were noted to alter between the MS traces of the 0 h and 6 h time points while peaks 1–17 were noted as being increased during growth using the XCMS data processing method (Identities in Supplementary Data File 1, Table 2).
| Peak | RT (s) | Comparison of change in metabolite over time (0–6 hr) between FW & BHI | m/z [M–H]− | MS2* | Formula* [M–H]− | Putative identity | Database |
|------|--------|-------------------------------------------------|-------------|------|-----------------|-----------------|----------|
| 1    | 5.18   | ↑FW = ↑ BHI                                          | 144.0658   |      | C8H10N2O3 | N-butryl glycine | PC 88412 |
| 2    | 8.03   | ↑FW = ↑ BHI                                          | 149.0270   |      | C9H16N2O6 | hydroxyn (methylthio) butanoic acid | PC 11427 |
| 3    | 8.88 9.97 | ↑FW = ↑ BHI ↑FW = ↑ BHI | 158.0813  |      | C7H11NO3 | pentanoyl glycine isomers | NA |
| A    | 9.92   | ↑FW = ↑ BHI                                          | 227.1024   |      | C9H16N2O6 | pyrogulatyaline | PC 152416 |
| B    | 10.60  | ↑FW = ↑ BHI                                          | 259.0741   |      | C9H16O4 | pyreneccetic acid | PC 186770 |
| 4    | 10.66  | ↑FW = ↑ BHI                                          | 181.0494   |      | C7H12O3 | homovanillic acid | CS 1675 |
| 5    | 11.28  | ↑FW = ↑ BHI                                          | 243.1696   |      | C10H12N2O6 | hexanoyl lysine | CS 7992036 |
| C    | 11.35  | ↑FW = ↑ BHI                                          | 269.0761   |      | C9H15N2O6 | UK |
| 6    | 11.51  | ↑FW = ↑ BHI                                          | 222.0756   |      | C10H16N2O6 | N-acetyl tyrosine | PC 68310 |
| D    | 13.45  | ↑FW = ↑ BHI                                          | 131.0707   |      | C7H10O3 | hydroxycapric acid | PC 99824 |
| E    | 13.79  | ↑FW = ↑ BHI                                          | 131.0706   |      | C7H10O3 | hydroxycapric acid isomer | PC 99824 |
| 7    | 14.85  | ↑FW = ↑ BHI                                          | 172.0968   |      | C10H15N2O4 | pentanoyl alanine | CS 15621024 |
| F    | 15.20  | ↑FW = ↑ BHI                                          | 241.1182   |      | C9H15N2O4 | UK |
| 8    | 16.18  | ↑FW = ↑ BHI                                          | 172.0966   |      | C10H12N2O6 | hexanoyl glycine (N-caproylglycine) | CS 89859 |
| G    | 16.66  | ↑FW = ↑ BHI                                          | 165.0546   |      | C7H14O3 | 3-(4-hydroxyphenyl) propionic acid | PC 10394 |
| 9    | 16.75  | ↑FW = ↑ BHI                                          | 250.1065   |      | C9H16N2O6 | ethyl-N-acetyl tyrosine | CS 12729 |
| H    | 16.82  | ↑FW = ↑ BHI                                          | 273.0862   |      | C10H15N2O6 | hexahydrophenazine-1,6-dicarboxylic acid | ChEBI 132261 |
| 10   | 19.30  | ↑FW = ↑ BHI                                          | 186.1126   |      | C11H17N2O4 | hexanoyl alanine | NA |
| 11   | 19.62  | ↑FW = ↑ BHI                                          | 264.1221   |      | C9H17N2O4 | pentanoyl tyrosine | CS 5142226 |
| 12   | 20.12  | ↑FW = ↑ BHI                                          | 145.0861   |      | C10H14N2O3 | hydroxy heptanoic acid | PC 275049 |
| J    | 20.65  | ↑FW = ↑ BHI                                          | 598.2504   |      | C13H22O6 | UK |
| 13   | 22.11  | ↑FW = ↑ BHI                                          | 200.1277   |      | C11H17N2O4 | caproyltycine (2-oc tamamidoacetic acid) | CS 76040 |
| K    | 23.15  | ↑FW = ↑ BHI                                          | 385.1411   |      | C11H16N4O6 | UK |
| L    | 24.40  | ↑FW = ↑ BHI                                          | 278.1374   |      | C12H18N4O6 | hexanoyl tyrosine | CS 32674367 |
| 14   | 25.63  | ↑FW = ↑ BHI                                          | 273.1221   |      | C12H18N4O6 | pentanoyl tyrosine | NA |
| 15   | 26.67  | ↑FW = ↑ BHI                                          | 246.1151   |      | C10H15N2O5 | hexanoyl methionine | ChEBI 80768 |
| M    | 26.69  | ↑FW = ↑ BHI                                          | 214.1434   |      | C10H17N2O4 | hexanoyl valine | CS 24223294 |
| N    | 26.91  | ↑FW = ↑ BHI                                          | 740.4299   |      | C13H22O6 | UK |
| 16   | 27.47  | ↑FW = ↑ BHI                                          | 248.1273   |      | C11H17N2O4 | pentanoyl phenylalanine | CS 14753835 |
| 17   | 27.71  | ↑FW = ↑ BHI                                          | 287.1377   |      | C11H17N2O4 | pentanoyl triphenylalanine | CS 16818921 |
| O    | 29.43  | ↑FW = ↑ BHI                                          | 228.1588   |      | C12H18N4O6 | hexanoyl leucine/isoleucine | CS 20041531 |
| P    | 29.76  | ↑FW = ↑ BHI                                          | 301.1538   |      | C13H20N4O4 | hexanoyl triphenylalanine | CS 53673714 |
| Q    | 30.48  | ↑FW = ↑ BHI                                          | 262.1430   |      | C13H20N4O4 | hexanoyl phenylalanine | CS 3440214 |
| R    | 30.62  | ↑FW = ↑ BHI                                          | 262.1430   |      | C13H20N4O4 | hexanoyl phenylalanine | CS 3440214 |

Table 2: Identified metabolites that changed during growth of *Clostridium difficile* strain 630 in FW media. Databases - PB = PubChem (https://pubchem.ncbi.nlm.nih.gov/); CS = Chem Spider (http://www.chemspider.com/); ChEBI (http://www.ebi.ac.uk/chebi//) Peaks A – R are featured in Fig. 3. **Identified in faecal water** peak designation; †retention time; ‡arrows denote whether the compound increases or decreases during the growth period; <, > and + denote if the increase or decrease was more apparent in the + FW or BHI alone incubations; *MS*2 fragments in bold are the more intense, figures in brackets are neutral loss. All MS*2 fragments apart from bold or underlined are minor fragments; *predicted formula* based on m/z [M–H] values, UK = unknown.

differentially expressed in FW medium, with increases up to 300–fold. Of 60 DE genes in our dataset that had a predicted or known role in sporulation, only two exhibited decreased expression (CD2273, a ‘putative sporulation integral membrane protein’—possibly under εE control, and CD3669, a ‘putative exported protein’—part of the mature spore proteome). Of the remaining 58 genes, the expression of 22 of these was increased > 50–fold, indicating that FW is a potent inducer of sporulation genes in *C. difficile* strain 630. *C. difficile* sporulation has been extensively mapped, allowing definition of genes under control of SpotA (CD1214, increased by 2.5–fold)—the
master regulator of sporulation—and the identification of potential links between sporulation and other phenotypes41,42,52,57. Phosphorylation of the Spo0A protein initiates a sigma factor cascade that, acting in both mother cell and forespore, influences expression of the sporulation–specific sigma factors σF (CD0772, 9–fold up), σE (CD2643, 23–fold up), σG (CD2642, 40–fold up) and σK that control expression of early (Bacillus stage II and III) and late (Bacillus stage V and VI) sporulation genes 50. During sporulation, a septum results in asymmetric division of the bacterial cell and creates two unequally–sized compartments. The smaller—the forespore—develops into the spore, while the larger compartment prepares the forespore for dormancy42,51. Taken in the context of previous identification/analysis of sporulation–associated genes 48,53 our data indicates that, at the point of harvest, FW–grown C. difficile cells are physiologically at what in B. subtilis would be categorised as stage III of sporulation, i.e. the point at which engulfment of the forespore has occurred, but prior to cortex formation. Thus, genes including the spoIIIAA–spoIIAH operon (all >50–fold increased), in addition to spoIIIJ (oxaA1, 1.6–fold increased), spoIIID (56–fold increased) and sigG (40–fold increased) exhibited considerably increased expression in FW. The σE protein acts on a number of genes in the Clostridial sporulation cascade41,42 and we noted increased expression of σE–controlled genes including spoIIID, spoIVA (56–fold up), spoIIB (CD1433, 29–fold up). The peroxiredoxin and chitinase activities of CotE contribute to pathogenesis by facilitating degradation of gut mucus during infection58,59.

It has been demonstrated that decreased oppABC (CD0853–855, encoding an oligopeptide transporter) expression leads to earlier expression of sporulation–associated genes13 and the observation that oppABC expression was 50% lower in FW appears consistent with our other observations of FW–induced changes to the C. difficile transcriptome. Taken together, therefore, our gene expression data indicates that C. difficile cells are induced by FW components towards sporulation more rapidly than cells grown in BHIS media. The spores are the transmissible, resilient, and infectious form of the organism14 and thus our observation has clear implications for pathogenesis and transmission of the disease, in addition to being entirely consistent with observations by other researchers of extensive sporulation within the gut microbiota in vivo15,53–56.

A variety of C. difficile transport systems are differentially expressed in FW media. In previous work we showed that phosphotransferase (PTS) sugar transport systems were largely unperturbed by heat stress45. By contrast, our current investigation revealed considerable changes in transporter gene expression. The PTS is the major bacterial carbohydrate assimilation system for hexoses, hexitols and disaccharides and consists of two general components—enzyme I (EI), and the histidine phosphocarrier protein (HPr)—in addition to sugar specific permeases (enzymes II) in the cell membrane. In FW media, expression of the gene encoding the EI component (CD2755) common to, and essential for, all phosphotransferase systems in the cell, was increased by 1.58–fold. In contrast, expression of the HPr kinase/phosphorylase (CD3409) that phosphorylates...
the cytoplasmic phosphocarrier protein Hpr at Ser42, and which also leads to activation of the LacI family carbon catabolite repressor, ccpA, was 1.8-fold lower. Consistent with these observations, the gene encoding the IIABC component of the PTS system for uptake of beta-glucosides (bglF, CD0388) was increased, as was the downstream gene bglA (CD0389) encoding 6-phospho-beta-glucosidase, reflective of the likely increased availability of such glucoside substrates in the FW media. In addition, expression of the sorbitol specific IIB component, srlEa (CD0765), was increased as were genes encoding the fructose specific IIABC component, fruABC, as were genes encoding the IIA and IIB components of the glucose PTS transport system (CD2512, CD2510, respectively). Expression of the IIC and IID components of the mannose/fructose/sorbose transport system (CD3277, CD3276) were increased by 4- and 6.6-fold, respectively. Conversely, expression of PTS system components associated with uptake of xylosides (xyl and xyn operons, CD3064–CD3070) was lower in FW, while expression of the associated transcriptional regulator (xylose repressor, xylR, CD3066), which functions to reduce expression of genes for uptake and metabolism of xylose, was increased. These diverse perturbations in expression of carbohydrate transport–associated genes most likely underpin an adaptive response of C. difficile to additional carbon sources and other diet derived metabolites present in the FW. However, the PTS is also a signalling device which has been linked to chemotaxis and regulatory functions associated with C, N and P metabolism and to the virulence of C. difficile. The complex interplay between a variety of cellular systems (sugar transport, carbon catabolite repression, quorum sensing and amino acid metabolism), controls toxin production. It is known, for example, that butyrate stimulates toxin production but, in FW–grown cells we noted lower expression of 12 genes associated with carbohydrate fermentation to butyrate. The likely reduction in metabolic flux towards butyrate is consistent with the 4-fold lower expression of tcdA (CD0663) observed in FW (Fig. 4). Other genes encoded by the pathogenicity locus are not discussed here as our padj cutoff value precluded their inclusion in the list of statistically significant DE genes.

Figure 4. Comparison of qRT-PCR and RNAseq data for selected Clostridium difficile strain 630 genes. For each individual gene, expression changes determined by RNAseq (up-hatched columns) and by qRT-PCR (down-hatched columns) are shown relative to the BHIS control, and show good correlation between the two datasets ($R^2 = 0.97$). rpsJ, gyrA and adk were used as reference genes.

Figure 5. Differentially–expressed genes in the Clostridium difficile strain 630 faecal water transcriptome by functional category. Orange – increased expression; blue – decreased expression, $p < 0.001$. Analysis was conducted only on functional categories within which > 5 genes were differentially expressed.
A number of ABC transporter–encoding genes were DE in FW, including some associated with transport of sugar phosphates, vitamins, oligopeptides, amino acids and also transporters associated with multidrug efflux mechanisms. The most downregulated gene (270–fold lower in FW) in our dataset was an ABC transporter ‘substrate-binding lipoprotein’ (CD0873), recently identified as an adhesin that enables C. difficile to bind Caco–2 cells. Our data suggests that, in FW at the point of cell harvest and possibly at later stages in the infection cycle in a host, C. difficile exhibits reduced binding to epithelial cells, consistent with increased sporulation and lowered motility. This physiological state could thus facilitate evacuation of the bacterial population from the host. Ten different lantibiotic/multidrug ABC transporters were also DE. Four exhibited increased expression–CD0161 (4.73–fold); CD1349/50 (5.3, 2.9–fold); CD2210/11 (3.7, 2.5–fold); CD2406/7/8 (all 2.58–fold up). The precise function of such transporters has not yet been defined, and consequently genomic annotations are a “general function prediction” only. Nonetheless, within the intestine, C. difficile and other gut pathogens must contend with innate host defences including cationic antimicrobial peptides (CAMPs, e.g., nisin) produced by both host and indigenous microbiota. McBride and Sonenshine have shown that proteins encoded by CD1349/CD1350 are involved in resistance to CAMPs, proposing the designation cprABC-cprK for CD1349 to CD1352. In FW–grown C. difficile cells, expression of CD1349 and CD1350 (cprA, cprB, encoding the ATP–binding protein and the permease respectively) were increased by 5.3– and 2.9–fold, consistent with our hypothesis that increased levels of host or microbiota–derived antimicrobial peptides present in FW lead to increased expression of this specific mechanism.

Expression of motility genes is decreased in FW media. Bacterial flagella are self–assembling molecular machines, with flagella and type-IV pili comprising motility devices essential for the pathogenesis of certain bacteria. Including a variety of motile enteropathogens. Flagellar biosynthesis is a highly–ordered process in which hierarchal control of gene expression ensures that synthesis of late–stage components is repressed until assembly of earlier components is complete. Thus, only when the basal body and motor machinery is in place, do late–stage genes, including flagellin (fliC) become expressed. Expression of motility genes in C. difficile is regulated by a sigma–28 factor encoded by CD0266 (fliA, or sigD) whose expression was 1.5–fold decreased in FW–grown cells. El Meouche et al. (2013) demonstrated that sigD acts as a positive regulator of both flagella and toxin gene expression in C. difficile. Decreased expression of sigD could therefore be partially responsible for the reduced expression of motility genes. In FW–grown C. difficile, expression of genes located broadly in the F3 loci (CD0245–CD0271), including those encoding components of the basal body, motor, hook and rod, exhibited 1.5– to 2–fold decreased expression. Gene expression in the F1 locus (CD0226–CD0240) was reduced between 2– and 5.7–fold, as would be expected if these gene products were not required until assembly of the basal body was complete. In addition, expression of fliC (CD0239) was decreased by 4.4–fold to ~20% of the level in the control, an observation corroborated by our qRT-PCR data (Fig. 4). flgN (CD0230) expression was reduced by 5.7–fold and we noted likewise that expression of genes in the interflagellar F2 locus (CD0240–CD0244) decreased by just over three–fold in FW. Levels of transcript (Supplementary Table S3, base mean values), for class I flagellar genes were higher than those for the class II genes, with genes in the interflagellar locus
expressed at a yet lower level still, which is logical from an assembly perspective. Flagellar operon gene expression, and thus motility of *C. difficile*, decreases in FW concomitantly with increased expression of sporulation–associated genes. The precise role of flagella in *C. difficile* pathogenesis is still unclear however, depending in many cases upon the strain tested. Decreased FliC expression under clinically–relevant heat stress may enable better adherence to epithelial cells, a hypothesis supported by the work of Dingle et al. who assessed *flic* and *fliD* disruption mutants, concluding that flagellar motility *per se* did not contribute to adherence to epithelial cells *in vitro*. Indeed, they argued that flagella were either not necessary for virulence, or that repression of motility could be a pathogenic mechanism. In a *C. difficile* dnaK mutant, lack of motility was underpinned by a 4–fold decrease in *flic* expression with the mutant also exhibiting significantly enhanced biofilm–forming ability. Other mutational studies have also shown that non–flagellated *C. difficile* cells exhibit lower levels of toxin production in addition to increased sporulation as a result of the pleiotropic role of FliC in *C. difficile* gene regulation.

In addition to changes in flagellar operon gene expression, we noted a 2– to 3–fold increase in expression of some genes in the secondary type IV pilus (TFP) locus. This increased expression of genes encoding a type–IV pilin, an associated type–II secretion system protein, and a pilus assembly ATPase (CD3294–6) suggests that pilus–driven motility may possibly be more important in a FW milieu and in certain stages of the infection process. Regardless of the role of TFP during infection, bacterial flagella are known to promote intestinal lesions via host inflammatory responses: *C. difficile* FliC protein recognizes TLR5 and consequently activates the NF–κB and the MAPK signalling pathways that elicit synthesis of pro–inflammatory cytokines. Such host receptors are not present in our experiment and decreased expression of flagellar loci may represent interplay between a putative motility phenotype and an adhesion, or indeed sporulation, phenotype. Nonetheless, flagella are energetically extravagant structures and, in the challenging environment of the gut, it makes strategic sense for motility in a semi–solid milieu to be driven by less resource–intensive structures such as type–IV pili.

Many Genes encoding Conserved Hypothetical Proteins are differentially expressed in FW media. The largest number of DE genes (n = 141) were placed in the 'Similar to unknown proteins' category and 19 of these had expression changes of >20–fold in FW–grown cells. None of the proteins encoded by these genes had predicted signal peptides and with the exception of CD1726 and CD3522, were all predicted to lie transcriptionally downstream of CD2344 in the same operon structure (e.g. CD1884, CD2657, CD2374 and CD2375 were also reported by Janoir et al.). In addition, DE genes including CD1929, CD2150A & CD3522) were also DE in FW. Dembek et al. reported that a large proportion of *C. difficile* conserved hypothetical protein–encoding genes was CD2344, which has been identified as a putative succinate transporter with a role in sporulation. The most downregulated conserved hypothetical protein–encoding gene was CD2344, indicating their potential function. Nonetheless, literature and database interrogation allowed us to link many to the cytoplasmic membrane and the majority of these 141 proteins possessed no conserved domains that might indicate their potential function. Nonetheless, literature and database interrogation allowed us to link many to a role in sporulation. The most downregulated conserved hypothetical protein–encoding gene was CD2344, which has been identified as a putative succinate transporter with a role in *C. difficile* gut colonisation. We have also shown here that expression of a variety of other genes in the succinate to butanoate fermentation pathway–which lie transcriptionally downstream of CD2344 in the same operon structure (e.g. cat1, sucD, abfD, and cat2, hbd)–were decreased by ~4–fold in FW. A number of genes reported to be regulated by sporulation–specific sigma factors 41–43, including rE (CD3580 & CD1065), σG (CD2808 & CD3275) and σE (CD1063A-C, CD2150A & CD3522) were also DE in FW. Dembek et al. reported that a large proportion of *C. difficile* spore transcripts encoded proteins of unknown function and proposed that these were indicative of the difference between the transcriptional programme of vegetative cells and spores. Three such genes exhibited increased expression in FW–CD351A (71–fold), CD2374 (30–fold) and CD2229 (36–fold). In addition, DE genes including CD1929, CD1884, CD2657, CD2374 and CD2375 were also reported by Janoir et al. to be expressed at higher levels in stationary phase *C. difficile* 630 cells at 14 h and 38 h–i.e. where the sporulation process would be well–established.

Conclusions

We set out to establish a new means of investigating gut pathogen biology *in vitro*. LC–MS metabolomic analysis of FW allowed us to identify 30 individual components including urobilinogen, steroblin and several cholic acid derivatives. Having established that the FW was–metabolomically–consistent with previous reports, we demonstrated that in the presence of FW, growth of *C. difficile* strain 630 was largely unaffected, save for an increase in cell length that our transcriptome data indicates is most likely a prelude to sporulation. A primary question was whether *C. difficile* strain 630 could utilise components of FW. Our analysis showed that while FW metabolites were not further metabolised during growth, a set of previously unknown *C. difficile*–derived hexanoyl– and pentanoyl– derivatives of amino acids were produced. These metabolites are not only novel biomarkers for the presence of this pathogen, but also reflect previously unrecognised metabolic capabilities within *C. difficile* strain 630. RNA sequencing showed clearly that the primary transcriptomic response of *C. difficile* strain 630 to FW was an acceleration of the sporulation cascade. FW–grown cells exhibited increases of up to 300–fold in the expression of sporulation–associated genes, with concomitant decreases in motility and toxin gene expression. These changes are reflective of the interplay between FW components and the expression of sensor kinase/response regulator systems, and transcription factors, many of which exhibited increased expression in FW–grown cells. Interestingly, none of the classical stress–response genes were differentially expressed, supporting the rationale that *C. difficile* adapts easily to a faecal milieu. The considerable modulation of a variety of transport systems is consistent with the addition of FW components to the growth media. Overall, therefore, our *ex vivo* FW model represents a new and unique means of assessing the response of *C. difficile* strain 630 to gut metabolites allowing us to describe, for the first time, the faecal milieu–associated physiological changes in this important pathogen.

Materials and Methods

**Chemicals and Glassware.** All chemicals and reagents of Analar grade or better were purchased from Sigma–Aldrich (Poole, UK) unless stated. Brain Heart Infusion (BHI) agar and broth and yeast extract were purchased from Oxoid (Basingstoke, UK). All molecular biology reagents were purchased from Invitrogen (Renfrewshire, UK) save for qPCR reagents, which were obtained from Roche Diagnostics (Hertfordshire, UK)
and random primers, which were obtained from Promega (Southampton, UK). Lysing Matrix A tubes were from MP Biomedicals (Cambridge, UK) and all glassware was cleaned with 1% Virkon (Antec Intl. Ltd., UK) overnight prior to steeping in 2% Decon (Decon Labs Ltd., UK) for 4 h prior to use.

Preparation of faecal water for inclusion in BHIS media. The Ulster University Research Ethics Committee exempted this study from review because donors were not involved in any intervention; the samples received were not collected by means of intervention and were used solely for preparation of a bacterial growth media. Written consent was obtained for provision of the donor faecal samples. Fresh faecal samples were provided by two apparently healthy individuals (2 males, age range 38–42 years, who had not taken antibiotics within the previous three months). Stool samples were collected from donors and stored at 4 °C for up to 2 h before processing. Faecal water was produced as described in Gill et al.16. In brief, the faecal sample was weighed, mixed 1:1 wt/vol with ice cold 0.01 M phosphate buffered saline (PBS) then homogenised in a Seward 600 stomacher (2 × 2 min cycle). The resultant faecal slurry was centrifuged at 50,000 × g for 2 h at 4 °C in the 70.1Ti rotor of a Beckman L8-M centrifuge and the supernatant removed and filter-sterilised (0.22 μm filter) on ice before being aliquoted for storage at −70 °C.

Growth of Clostridium difficile strain 630. Clostridium difficile 630 (ATCC no. BAA-1382D-5) was grown under anaerobic conditions in a Don Whiteley MACS MG500 anaerobic cabinet (Don Whiteley Scientific Ltd, Yorkshire, UK) using a single gas mix (BOC, UK) of 80% N₂, 10% CO₂, and 10% H₂, at 37 °C. Standard growth media was brain heart infusion broth supplemented with 5 g L⁻¹ yeast extract and 1 g L⁻¹ L-cysteine (BHIS). For media containing faecal water, 2-fold concentrated BHIS (50 mL) was prepared and a 50 mL aliquot of filter-sterilised FW added to this aseptically post-autoclaving, resulting in 1 × BHIS containing 50% FW (“FW media”). Control media was prepared from 2-fold concentrated BHIS to which was added an equal volume of sterile PBS. Three starter cultures of C. difficile 630 were set up in 20 mL of FW media in glass universal bottles and each was inoculated with a single colony of freshly grown C. difficile 630 from a BHIS agar plate. Starter cultures were incubated overnight and used to inoculate fresh media, in triplicate, at 5% (vol/vol). Growth was recorded hourly as attenuation at 650 nm (D₆₅₀nm)44 against an uninoculated BHIS and FW media references. Multiple cell pellets were collected by centrifugation from all six cultures at mid-log phase of growth (D₆₅₀nm = 0.6). Culture supernatants were removed to fresh tubes and both cell pellets and supernatants were placed briefly in liquid nitrogen before immediate transfer to −70 °C until required.

RNA extraction and quality control. RNA extraction was via a Qia-gen RNeasy® Mini kit with the addition of a mechanical lysis step as described previously45. RNA was checked for absence of DNA contamination by PCR with gyrA, rpsJ, and adk primers (Table 3) followed by agarose gel electrophoresis and imaging under UV light. A Nanodrop™ 1000 spectrophotometer (Thermo Scientific) was used to quantify the amount of RNA in the samples and integrity of total RNA was then determined using an RNA 6000 Nano Assay kit with an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA) instrument as per the manufacturer’s instructions. Only RNA samples with RIN > 9.0 were used in subsequent procedures.

Transcriptome sequencing. RNA sequencing (RNAseq) and initial bioinformatics analysis was performed at Deepseq (University of Nottingham, UK). RNA samples were shipped to Deepseq on dry ice and upon receipt, total RNA quality was once more assessed using the Agilent RNA 6000 Nano Kit (Agilent Technologies, 5067–1511) on the Agilent 2100 Bioanalyzer. The total RNA concentration was measured using the Qubit RNA BR assay kit (Life technologies, Q10210). A 1 μg amount of Total RNA was used for rRNA depletion using the Ribo-Zero rRNA Removal Kit (Gram-Positive Bacteria) (Illumina, MRZGP126). Illumina stranded whole transcriptome sequencing libraries were prepared using NEBNext Ultra Directional RNA library prep kit for Illumina (NEB, E7420S). The standard protocol for use with Ribosome Depleted RNA was followed except that, after second strand synthesis, the samples were precipitated with 1 μL (20 ng μL⁻¹) glycerol and 1/10 vol 3 M sodium acetate. Pellets were washed once with 80% ethanol, followed by 70% ethanol and after air-drying, pellets were resuspended in 58 μL of water. The standard protocol for use with Ribosome Depleted RNA was resumed for the remaining steps, except libraries were size selected using Agencourt AMPure XP beads at a 1.5 x ratio to retain the smaller sized fraction (~150 bp). The NEB Next Multiplex Oligos for Illumina Primer set 1 (NEB, E7335S) was used to generate barcoded multiplex libraries. Library QC was performed using bioanalyser HS kit (Agilent technologies, 5067–4626) and libraries were quantified using qPCR (Kapa Biosystems, KK4824). Libraries were pooled at desired concentrations, denatured and loaded for sequencing according to the manufacturer’s instructions. Sequencing was performed over 3 runs on the Illumina MiSeq sequencing platform to generate 2 × 75 bp reads.

For differential gene expression analysis the sequencing reads were mapped onto the annotated C. difficile strain 630 reference genome (http://www.ncbi.nlm.nih.gov/nuccore/115249003) with appropriate alignment software. The aligned ﬁles were then processed for tag counts per location mapped or normalised tag counts (RPKM) and differential gene expression analysis. The DeepSeq Filtering Pipeline for Read Mapping was used to filter reads with low sequencing score, in addition to reads aligned to adaptor sequences. Reads from the sequencer were QC checked using FASTQC, then trimmed and filtered for low quality bases and adaptor sequences, and QC checked once more. Reads that passed this filter were mapped onto the reference genome in the context of known gene exon coordinates using the bwa mapping tool (http://bio-bwa.sourceforge.net/). Read alignments were recorded in a BAM formatted alignment file (named *.bam), and companion BAM index file (named *.bam.bai). Read alignments, both primary and unique, were then ﬁltered further according to their mapping quality score (MAPQ). For gene expression, MAPQ20 uniquely aligned reads were used to generate
counts per gene using 'htseq-count', which determines the number of uniquely aligned reads per gene (http://www-huber.embl.de/users/anders/HTSeq/doc/count.html).

These counts were used as the input for the DESeq program\textsuperscript{44,45}. DESeq models the distribution of the counts data in each sample and then compares the distributions to determine differentially expressed genes, with significantly differentially expressed genes having an adjusted p value < 0.05. The program implements a single analytical approach and when RNA–seq samples with biological replicates are available, as is the case here, DESeq analyses the variance between them in order to better model the expression values of individual genes within the group of replicates.

The data discussed herein has been deposited in NCBI's Gene Expression Omnibus\textsuperscript{86} repository (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112422) with accession number GSE112422.

### Data processing

Transcriptome sequence data was obtained from DeepSeq as a summary MS Excel file containing a list of genes with cognate base mean values for BHI medium (BHI, base mean A) and faecal water medium (FW, base mean B) growth conditions, in addition to p value, p-adjusted value and the ratio of FW/BHI containing a list of genes with cognate base mean values for BHI medium (BHI, base mean A) and faecal water medium (FW, base mean B) growth conditions, in addition to p value, p-adjusted value and the ratio of FW/BHI.

The NCBI C. difficile strain 630 genome (http://www.ncbi.nlm.nih.gov/nuccore/115249003) was used as a starting point for addition of the C. difficile strain 630 locus annotations\textsuperscript{46,47}, in addition to protein name and Sublist functional category\textsuperscript{48,49}. This process was carried out essentially as in our previous work\textsuperscript{44,45} using the NCBI CDD database, BioCyc pathway tools and metacyc visual pathways software\textsuperscript{42,43}, combined with literature searching to arrive at a functional role/categorisation and to identify predicted co–regulated genes and operon structures (for complete list of DE genes see Supplementary Table S3).

### Reverse transcription and qPCR

As in our previous work\textsuperscript{44,45} differential gene expression data was corroborated using qRT-PCR, on aliquots of the same RNA samples that were sent for sequencing. cDNA was prepared from 500 ng aliquots of the extracted RNA samples and 50 ng of random hexamer primer (Promega, WI, USA) with a SuperScript II Reverse Transcriptase kit (Invitrogen, Renfrewshire, UK). Successful reverse transcription and generation of cDNA was confirmed by PCR using rpsJ, gyrA or adk primers (Table 3), as compared to the "minus RT" controls. Quantitative PCR (qPCR) was performed on a LightCycler480 instrument using a Master SYBR Green 1 kit (Roche Diagnostics, UK). Standard curves were prepared by creating a 5–fold serial dilution (1, 1/5, 1/25, 1/125, 1/625, 1/3125, and 1/15625) of the pooled cDNA samples from all cultures with nuclease-free water. qPCR target run reactions were set up in technical triplicates. Bulk mastermix containing 5 µL of 2–fold concentrated master mix, 1 µL each of forward and reverse primer (at a concentration of 10 µM), 2 µL of nuclease free H₂O and 1 µL of a 1–in–10 dilution of the relevant cDNA template was prepared and 10 µL aliquots of this added to the plate. qPCR cycling conditions comprised an initial denaturation stage of 95 °C for 5 min followed by 40 cycles of 95 °C for 10 s, 57 °C for 10 s and 72 °C for 10 s. Melting curve analysis of target runs, in addition to "no template" and "no reverse transcriptase" controls confirmed the specificity of amplification.

Roche Re-Quant software (Roche Diagnostics, UK) was used to generate a Cq value for each sample using the second derivative maximum method. Cq values were transferred to Excel and the arithmetic mean of technical replicates was determined. These values were then log transformed to relative quantities (RQ) using the
information gained from standard curves previously constructed for each primer pair, thus ensuring PCR efficiencies were calculated accurately for each gene. All target gene RQs were normalised against the geometric mean of the reference gene RQs by dividing the former by the latter (target/housekeeping) to generate normalised relative quantity (NRQ) values. Control sample NRQ values (BHI5) were subsequently used as a calibrator and corrected to 1 with all experimental NRQ values (FW media) being expressed as a relative expression ratio to the calibrator for each gene. Values were subsequently expressed as fold change ratios relative to the BHI control (Supplementary Data File 1, Table S4).

**Liquid Chromatography-Mass Spectrometry (LC–MS) analysis of culture supernatant samples.**

*LC–MS Analysis.* Culture media samples were frozen and transported to the Hutton Institute on dry ice where they were stored at −70°C prior to analysis. After thawing on ice, samples (1 mL) were vortexed then transferred to 2 mL microcentrifuge tubes and centrifuged at 10,000 × g for 10 min at 5°C in a refrigerated microfuge. A sub-sample (0.5 mL) was removed and placed in a 0.45 μm PTFE filter vial (Thomson Instrument Company, supplied by Bioprocess Engineering Services Ltd, Kent, UK) prior to analysis using the LTQ-Orbitrap XL LC–MS system. Samples were analysed using a LC system consisting of a quaternary pump (Thermo Fisher Scientific, Acello 600) and a PDA detector (Thermo Fisher Scientific, Acello) coupled to an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Hemel Hempstead, U.K.). Duplicate 10 μL samples were injected in part-loop mode onto a 2 × 150 mm (4 μm) Synergy Hydro-RP 80Å column fitted with a C18 4 × 2 mm Security Guard cartridge (Phenomenex Ltd, Macclesfield, UK). Autosampler and column temperatures were maintained at 6°C and 30°C, respectively. The samples were analysed at a flow rate of 0.3 mL/min using a binary mobile phase of (A) 0.1% aqueous formic acid and (B) 0.1% formic acid in acetonitrile/water (1:1, vol/vol) with the following gradient: 0–4 min, 5% B; 4–22 min, 5–50% B; 22–32 min, 50–100% B. Following each analysis, the column was equilibrated for 5 min under starting solvent conditions. Mass detection was carried out using an LTQ Orbitrap XL mass spectrometer in negative ESI mode. Two scan events were employed; full-scan analysis (130–2000 m/z) in profile peak mode was followed by data–dependent MS/MS in centroid peak mode of the three most intense m/z peaks. Scan speeds of 0.1 s and 0.4 s were applied in the LTQ and FT-MS respectively. The Automatic Gain Control was set to 1 × 106 and 5 × 105 for the LTQ and FT-MS, respectively. Prior to the analytical run, the LTQ and FT-MS were tuned to optimise conditions for the detection of ions in the mid detection range of m/z 80–2000, as well as being calibrated with the manufacturer’s recommended calibration mixture. ESI conditions were as follows: spray voltage −3.5 kV (ESI-); sheath gas 60; aux gas 30; capillary voltage −35 V (ESI-); tube lens voltage −100 V (ESI-); capillary temperature 380°C. For the first three min of analysis, the eluent flow was directed to waste. All predicted formula data presented are accurate at ±0.005 ppm.

*Raw LC–MS data processing.* The raw LC–MS data files were first converted into an mzML centroid format using the Proteowizard MSConvert software package. Each mzML based three-dimensional data matrix (intensity × m/z × time) for each per sample was converted (or deconvoluted) into a vector of peak responses, where a peak response is defined as the sum of intensities over a window of specified mass and time range (e.g. m/z = 102.1 ± 0.01 and time = 130 ± 10 s) using the freely available XCMS software (http://massspec.scripps.edu/xcms/scms.php). A full description of the data deconvolution method performed within XC–MS is available59. In the current work, the band width setting was adjusted from 10 to 20 to accommodate the wider peak widths that result from HPLC as compared to UHPLC. The XC–MS deconvolution produced an MS Excel based X by Y matrix output as peak areas for detected peaks.

**Statistical analysis of LC–MS data.** The data from XCMS was loaded into SIMCA-P 12.0.1.0 software (Umetrics, available at https://umetrics.com) and principal components analysis (PCA) was carried out. PCA, using univariate scaling, clearly showed that the FW samples separated from the BHI-only samples on score 1, which explained 52% of the variation of the dataset. The beginning and end samples were clearly separated in scores 3 and 4 of the PCA, which explained 10% and 5%, respectively, of the variation (Supplementary Data File 1, Fig. S1). Following this robust PCA, a further discriminant analysis (optimized partial least squares, OPLS-DA) was performed with two classifications (“start” and “end” of incubation), resulting in a model that described ~9% of the variation with a Q2 (cum) value of 0.851 (Supplementary Data File 1, Fig. S2).

Using the loadings plots from this OPLS–DA plot (Supplementary Data File 1, Fig. S3), the m/z signals that drove the separation for the “end of incubation” could be extracted into an Excel file (Supplementary Data File 1, Table S5). Returning to the original XCMS data, the abundance of these components before and after incubation (for both biological and technical replicates) was plotted as peak areas (Supplementary Data File 1, Figs S4, S5). Over 120 “potential up-at-end” components were selected by this process and the graphs were quality checked to select only those with a clear distinction between before and after peak areas (e.g. with no overlap between before and after replicates). A final step of manual peak checking was carried out to check MS peak quality and to exclude peaks of very low abundance which often yielded no MS² data.

The XCMS data selected each m/z peak (along with any type of adduct ion(s) present) and the PutMedID set of workflows within the Taverna Workbench 1.7.2 software package was applied to predict putative metabolite identities using a library of known plant metabolites obtained from the Plant Metabolic Network PlantCyc database (http://www.plantcyc.org). In many cases, however, the putative identifications were not supported by subsequent examination (e.g.) of MS² data. Therefore, further manual putative peak assignment was carried out by comparing the predicted molecular formulae and MS² data with various databases and literature (Supplementary Data File 1, and Tables 1, 2).
Equipment and settings. Images shown in Figs 1, 3, 4 and 5 were produced using MS Excel, individually exported in PDF and imported into GiMP 2.8 for construction and final labelling. MS traces comprising Fig. 2A,B were exported from the resident MS Xcalibur software (Thermo Fisher Scientific, Hemel Hempstead, U.K.) into a Word document. After conversion into Microsoft Office drawing objects, the traces were edited to incorporate peak labels in Microsoft Word, exported as.jpg files and imported into GiMP 2.8 for construction and final labelling of Fig. 2.

Images in Supplementary data file 1 (Figs S1, S3) were made using the graphics package inherent in SIMCA then copied into a Word document. Figs S4 and S5 were made using the graphic package in GENESTAT, then copied into a Word document. The MS traces comprising Fig. S6 were exported from the resident MS Xcalibur software (Thermo Fisher Scientific, Hemel Hempstead, U.K.) into a Word document. After conversion into Microsoft Office drawing objects, the traces were edited to incorporate peak labels in Microsoft Word. All Supplementary Figures S1–S6 were saved as a single pdf file.

Ethical Approval. The Ulster University Research Ethics Committee exempted this study from review because donors were not involved in any intervention; the samples received were not collected by means of intervention and were used solely for preparation of a bacterial growth media. Written consent was obtained for provision of the donor faecal samples.

Data Availability. All data generated or analysed during this study are included in this published article (and in Supplementary Information files). The RNAseq datasets generated and analysed in this work are available in the NCBI GEO repository (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112422) with accession number GSE112422.

References

1. Rupnik, M., Wilcox, M. H. & Gerdng, D. N. Clostridium difficile infection: new developments in epidemiology and pathogenesis. Nat. Rev. Microbiol. 7, 526–536 (2009).
2. Burke, K. E. & Lamont, J. T. Clostridium difficile infection: a worldwide disease. Gut Liver. 8, 1–6 (2014).
3. Jarrad, A. M., Karoli, T., Blaskovich, M. A., Lyra, D. & Cooper, M. A. Clostridium difficile drug pipeline: challenges in discovery and development of new agents. J. Med. Chem. 58, 5164–5185 (2015).
4. Kuehne, S. A. et al. The role of toxin A and toxin B in Clostridium difficile infection. Nature. 467, 711–713 (2010).
5. Carter, G. P., Rood, J. I. & Lyra, D. The role of toxin A and toxin B in the virulence of Clostridium difficile. Trends Microbiol. 20, 21–29 (2012).
6. Baines, S. D. & Wilcox, M. H. Antimicrobial resistance and reduced susceptibility in Clostridium difficile: potential consequences for induction, treatment, and recurrence of C. difficile infection. Antibiotics. 4, 267–298 (2015).
7. Ng, K. M. et al. Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. Nature. 502, 96–99 (2013).
8. Moore, J. H. et al. Defined nutrient diets alter susceptibility to Clostridium difficile-associated disease in a murine model. PLoS ONE. 10, e0131829 (2015).
9. Zackular, J. P. et al. Dietary zinc alters the microbiota and decreases resistance to Clostridium difficile infection. Nat. Med. 22, 1330–1334 (2016).
10. Hrycykowi, A. J. et al. Microbiota-accessible carbohydrates suppress Clostridium difficile infection in a murine model. Nat Microbiol. 3, 662–669 (2018).
11. Collins, J. et al. Dietary trehalose enhances virulence of epidemic Clostridium difficile. Nature. 553, 291–294 (2018).
12. Ferreyra, J. A. et al. Gut microbiota-produced succinate promotes C. difficile infection after antibiotic treatment or motility disturbance. Cell Host Microbe. 16, 770–777 (2014).
13. Edwards, A. N., Nawrocki, K. L. & McBride, S. M. Conserved oligopeptide permeases modulate sporulation initiation in Clostridium difficile. Infect. Immun. 82, 4276–4291 (2014).
14. Deakin, L. J. et al. The Clostridium difficile spotA gene is a persistence and transmission factor. Infect. Immun. 80, 2704–2711 (2012).
15. Koenigsknecht, M. J. et al. Dynamics and establishment of Clostridium difficile infection in the murine gastrointestinal tract. Infect. Immun. 83, 934–941 (2015).
16. Best, E. L., Freeman, J. & Wilcox, M. H. Models for the study of Clostridium difficile infection. Gut Microbes. 3, 145–167 (2012).
17. Robinson, C. D., Auchtung, J. M., Collins, J. & Britton, R. A. Epidemic Clostridium difficile strains demonstrate increased competitive fitness compared to nonepidemic isolates. Infect. Immun. 82, 2815–2825 (2014).
18. Auchtung, T. M., Robinson, C. D. & Britton, R. A. Cultivation of stable, reproducible microbial communities from different fecal donors using mini-bioreactor arrays (MBRAs). Microbiome. 3, 42 (2015).
19. Crowther, G. S. et al. Efficacy of vancomycin extended-dosing regimens for treatment of simulated Clostridium difficile infection within an in vitro human gut model. J. Antimicrob. Chemother. 71, 986–991 (2016).
20. Crowther, G. S., Wilcox, M. H. & Chilton, C. H. An in vitro model of the human colon: studies of intestinal biofilms and Clostridium difficile infection. Methods Mol. Biol. 1476, 223–234 (2016).
21. Eyer, D. W. et al. Short-term genome stability of serial Clostridium difficile ribotype 027 isolates in an experimental gut model and recurrent human disease. PLoS ONE. 8, e63540 (2013).
22. Rafter, J. J. et al. Cellular toxicity of fecal water depends on diet. Am. J. Clin. Nutr. 45, 559–563 (1987).
23. Pearson, J. R., Gill, C. I. & Rowland, I. R. Diet, fecal water, and colon cancer–development of a biomarker. Nutr. Res. 67, 509–526 (2009).
24. Daniela, E. et al. Fecal water genotoxicity in healthy free–living young Italian people. Food Chem. Toxicol. 64, 104–109 (2014).
25. Verbeke, K. A. et al. Towards microbial fermentation metabolites as markers for health benefits of prebiotics. Nutr. Res. Rev. 28, 42–66 (2015).
26. Windey, K. et al. Wheat bran extract alters colonic fermentation and microbial composition, but does not affect faecal water toxicity: a randomised controlled trial in healthy subjects. Br. J. Nutr. 113, 225–238 (2015).
27. Eid, N. et al. Impact of palm date consumption on microbiota growth and large intestinal health: a randomised, controlled, crossover, human intervention study. Br. J. Nutr. 114, 1226–1236 (2015).
28. Boyd, L. A. et al. Assessment of the anti-genotoxic, anti-proliferative, and anti-metastatic potential of crude watercress extract in human colon cancer cells. Nutr. Cancer. 55, 232–241 (2006).
29. Brown, E. M. et al. Persistence of anticancer activity in berry extracts after simulated gastrointestinal digestion and colonic fermentation. PLoS ONE. 7, e49740 (2012).
30. Nowak, A., Sliżewska, K., Blasiak, J. & Libudzisz, Z. The influence of Lactobacillus casei DN 114 001 on the activity of faecal enzymes and genotoxicity of faecal water in the presence of heterocyclic aromatic amines. Anaerobe. 30, 129–136 (2014).
31. Monléon, D. et al. Metabolite profiling of fecal water extracts from human colorectal cancer. NMR Biomed. 22, 342–348 (2009).
32. Gill, C. I. R. et al. Profiling of phenols in human fecal water after raspberry supplementation. J. Agric. Food Chem. 58, 10389–10395 (2010).

33. Claesson, M. J. et al. Gut microbiota composition correlates with diet and health in the elderly. Nature. 488, 178–184 (2012).

34. Gao, X. et al. Metabolite analysis of human fecal water by gas chromatography/mass spectrometry with ethyl chloroformate derivatization. Anal. Biochem. 393, 163–175 (2009).

35. Gao, X., Puertas-Guillot, E. & Sibédo, J. L. Development of a quantitative metabolomic approach to study clinical human fecal water metabolome based on trimethylsilylation derivatization and GC/MS analysis. Anal. Chem. 82, 6447–6456 (2010).

36. Jiménez-Girón, A., Muñoz-González, I., Martín-Ivarez, P. J., Moreno-Arribas, M. V. & Bartolomé, B. Towards theecal metabolome derived from moderate red wine intake. Metabolites. 4, 1101–1118 (2014).

37. McDougall, G. J. et al. Nontargeted LC-MSn profiling of compounds in ileal fluids that decrease after raspberry intake identifies consistent alterations in bile acid composition. J. Nat. Prod. 79, 2606–2615 (2016).

38. Brooks, J. R., Nunez-Montiel, O. L., Wycoff, R. J. & Moss, C. W. Frequency pulsed electron capture gas-liquid chromatographic analysis of metabolites produced by Clostridium difficile in broth enriched with amino acids. J. Clin. Microbiol. 20, 539–548 (1984).

39. Madan, E. & Slikf, M. Stool capric acid for screening of Clostridium difficile. Am. J. Clin. Pathol. 89, 525–527 (1988).

40. Tracy, B. P., Jones, S. W., Fast, A. G., Indurthi, D. C. & Papoutsakis, E. T. Clostridium: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. Curr. Opin. Biotechnol. 23, 364–381 (2012).

41. Sajut, L. et al. Genome-wide analysis of cell type-specific gene transcription during spore formation in Clostridium difficile. PLoS Genet. 9, e1003756 (2013).

42. Finlay, K. A. et al. Global analysis of the sporation pathway of Clostridium difficile. PLoS Genet. 9, e1003660 (2013).

43. Pishdadian, K., Finlay, K. A. & Shen, A. SpoIIID-mediated regulation of ε function during Clostridium difficile sporulation. Mol. Microbiol. 95, 189–208 (2015).

44. Jain, S., Graham, C., Graham, R. L. J., McMullan, G. & Ternan, N. G. Quantitative proteomic analysis of the heat stress response in Clostridium difficile strain 630. J. Proteome Res. 10, 3880–3890 (2011).

45. Ternan, N. G., Jain, S., Srivastava, M. & McMullan, G. Comparative transcriptional analysis of clinically relevant heat stress response in Clostridium difficile strain 630 using a Gel/MS workflow with emPAI quantitation. PLoS ONE 9, e88960 (2014).

46. Underwood, S. et al. Characterization of the sporation initiation pathway of Clostridium difficile and its role in toxin production. J. Bacteriol. 191, 7296–7305 (2009).

47. Pettit, L. J. et al. Functional genomics reveals that Clostridium difficile Spo0A coordinates sporation, virulence and metabolism. BMC Genomics. 15, 160 (2014).

48. Al-Himai, M. A., Jones, S. W. & Papoutsakis, E. T. Of Clostridium acetobutylicum is the first known sporulation-specific sigma factor with two developmentally separated roles, one early and one late in sporation. J. Bacteriol. 196, 287–299 (2014).

49. Finlay, K. A. & Shen, A. Diverse mechanisms regulate sporation sigma factor activity in the Firmicutes. Curr. Opin. Microbiol. 24, 88–95 (2015).

50. Pereira, F. C. et al. The spore differentiation pathway in the enteric pathogen Clostridium difficile. PLoS Genet. 9, e1003782 (2013).

51. Sajut, L., Pereira, F. C.,Henriques, A. O. & Martin-Verastraeta, I. The regulatory network controlling spore formation in Clostridium difficile. FEMS Microbiol. Lett. 358, 1–10 (2014).

52. Browne, H. P. et al. Culturing of ‘unculturable’ human microbiota reveals novel taxa and extensive sporation. Nature. 533, 543–546 (2016).

53. Kansau, I. et al. Deciphering adaptation strategies of the epidemic Clostridium difficile 027 strain during infection through in vivo transcriptional analysis. PLoS ONE 11, e0158204 (2016).

54. Janoir, C. et al. Adaptive strategies and pathogenesis of Clostridium difficile from in vivo transcriptomics. Infect. Immun. 81, 3757–3769 (2013).

55. Scarfa, J. et al. Clostridium difficile transcriptome analysis using pig ligated loop model reveals modulation of pathways not modulated in vitro. Journal Infect. Dis. 203, 1613–1620 (2011).

56. Lawley, T. D. et al. Proteomic and detailed characterization of highly infectious Clostridium difficile 630 spores. J. Bacteriol. 191, 5377–5386 (2009).

57. Permpoonpanthana, P. et al. Functional characterization of Clostridium difficile spore coat proteins. J. Bacteriol. 195, 1492–1503 (2013).

58. Paredes-Sabja, D., Shen, A. & Sorg, J. A. Clostridium difficile spore biology: sporation, germination, and spore structural proteins. Trends Microbiol. 22, 406–416 (2014).

59. Martin-Verastraeta, I., Peltier, J. & Dupuy, B. The regulatory networks that control Clostridium difficile toxin synthesis. Toxins. 8, pii: E153 (2016).

60. Williamson, G. & Clifford, M. N. Role of the small intestine, colon and microbiota in determining the metabolic fate of polyphenols. Biochem. Pharmacol. 139, 24–39 (2017).

61. Antunes, A., Martin-Verastraeta, I. & Dupuy, B. CcpA-mediated repression of Clostridium difficile toxin gene expression. Mol. Microbiol. 79, 882–899 (2011).

62. Deutscher, J. et al. The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein–protein interactions. Microbiol. Mol. Biol. Rev. 78, 231–256 (2014).

63. Suárez, J. M., Edwards, A. N. & McBride, S. M. The Clostridium difficile cpr locus is regulated by a noncontiguous two-component system and response to type A and B antibiotics. J. Bacteriol. 195, 2621–2631 (2013).

64. Kovacs-Simon, A. et al. Lipoprotein CD0873 is a novel adhesin of Clostridium difficile. J. Infect. Dis. 210, 274–284 (2014).

65. Cole, J. N. & Nizet, V. Bacterial evasion of host antimicrobial peptide defenses. Microbiol. Spectr. 4. https://doi.org/10.1128/microbiolspec.VMBF-0006-2015 (2016).

66. McBride, S. M. & Sonenshein, A. L. Identification of a genetic locus responsible for antimicrobial peptide resistance in Clostridium difficile. Infect. Immun. 79, 167–176 (2011).

67. Evans, L. D., Hughes, C. & Fraser, G. M. Building a flagellum outside the bacterial cell. Trends Microbiol. 22, 566–572 (2014).

68. Baker, M. A. How biophysics may help us understand the flagellar motor of bacteria which cause infections. Adv. Exp. Med. Biol. 915, 231–243 (2016).

69. Erhardt, M. Strategies to block bacterial pathogenesis by interference with motility and chemotaxis. Curr. Top. Microbiol. Immunol. 398, 185–205 (2016).

70. Deligianni, E. et al. Pseudomonas aeruginosa cystic fibrosis isolates of similar RAPD genotype exhibit diversity in biofilm forming ability in vitro. BMC Microbiol. 10, 38 (2010).

71. Dingle, T. C., Mulvey, G. L. & Armstrong, G. D. Mutagenic analysis of the Clostridium difficile flagellar proteins, FlcA and FlID, and their contribution to virulence in hamsters. Infect. Immun. 79, 4061–4067 (2011).

72. Palen, M. J. & Matsu, N. J. From The Origin of Species to the origin of bacterial flagella. Nat. Rev. Microbiol. 4, 784–790 (2006).

73. El Meouchi, S. et al. Characterization of the SigD regulon of C. difficile and its positive control of toxin production through the regulation of tcpR. PLoS ONE 8, e83748 (2013).

74. Jain, S. et al. Inactivation of the dnaK gene in Clostridium difficile 630 Δerm yields a temperature–sensitive phenotype and increases biofilm–forming ability. Sci. Rep. 7, 17522 (2017).
76. Aubry, A. et al. Modulation of toxin production by the flagellar regulon in Clostridium difficile. Infect. Immun. 80, 3521–3532 (2012).
77. Barketi-Klia, A. et al. The flagellin FlIC of Clostridium difficile is responsible for pleiotropic gene regulation during in vivo infection. PLoS ONE. 9, e96876 (2014).
78. Bathe, J. & Kansau, I. Intestinal epithelial cell response to Clostridium difficile flagella. Methods Mol. Biol. 1476, 103–116 (2016).
79. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods. 8, 785–786 (2011).
80. Bendtsen, J. D., Kiemer, L., Fausboll, A. & Brunak, S. Non-classical protein secretion in bacteria. BMC Microbiol. 5, 58 (2005).
81. Yu, N. Y. et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. 26, 1608–1615 (2010).
82. Karp, P. D. et al. Pathway Tools version 19.0: integrated software for pathway/genome informatics and systems biology. Brief. Bioinform. 17, 877–890 (2016).
83. Dembek, M. Whole-genome analysis of sporulation and germination in Clostridium difficile. https://spiral.imperial.ac.uk/handle/10044/1/38630 (2014). Accessed 2 Nov 2017.
84. Anders, A. & Huber, W. Differential expression of RNA-Seq data at the gene level--the DESeq package. Eur. Mol. Biol. Lab (2013).
85. Anders, A. & Huber, W. Differential expression analysis for sequence count data. BMC Genomic 1, R106 (2010).
86. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30, 207–210 (2002).
87. Sebaihia, M. et al. The multidrug-resistant human pathogen Clostridium difficile: a highly mobile, mosaic genome. Nat. Genet. 38, 779–786 (2006).
88. Monot, M. et al. Reannotation of the genome sequence of Clostridium difficile strain 630. J. Med. Microbiol. 60, 1193–1199 (2011).
89. Moszer, I., Jones, L. M., Mooreira, S., Fabry, C. & Danchin, A. SubtiList: the reference database for the Bacillus subtilis genome. Nucleic Acids Res. 30, 62–65 (2002).
90. Lechat, P., Hummel, L., Rousseau, S. & Moszer, I. GenoList: an integrated environment for comparative analysis of microbial genomes. Nucleic Acids Res. 36(Database issue), D469–474 (2008).
91. Marchler-Bauer, A. et al. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43(Database issue), D222–226 (2015).
92. Caspi, R. et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. 44, (D1): D471–480 (2016).
93. Dunn, W. B. et al. Metabolic profiling of serum using ultra performance liquid chromatography and the LTQ-Orbitrap mass spectrometry system. J. Chromatogr. B 871, 288–298 (2008).
94. Brown, M. et al. Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. Analyst. 134, 1322–1332 (2009).
95. Goodert, J. J. et al. Fecal metabolomics: assay performance and association with colorectal cancer. Carcinogenesis. 35, 2089–2096 (2014).
96. Metcalf, D., Sharif, S. & Weese, J. S. Evaluation of candidate reference genes in Clostridium difficile for gene expression normalization. Anaerobe. 16, 439–443 (2010).
97. Denève, C., Deloménie, C., Barc, M. C., Collignon, A. & Janoir, C. Antibiotics involved in Clostridium difficile–associated disease increase colonization factor gene expression. J. Med. Microbiol. 57, 732–738 (2008).

Acknowledgements

NDM was supported by a Department of Employment and Learning (Northern Ireland) Ph.D. studentship (2013–2016). DS was supported by a Department of Employment and Learning (Northern Ireland) Ph.D. studentship (2011–2015). NGT, GMcM, CIRG and JSGD were supported by strategic funding from the Biomedical Sciences Research Institute, University of Ulster. GMcD, SV and JW A were part–supported by the Rural & Environment Science & Analytical Services Division of the Scottish Government. The funders had no input into study design or data analysis.

Author Contributions

N.G.T., C.I.R.G. and G.McM. conceived and designed the experiments. N.D.M., D.S., N.G.T., C.I.R.G., G.McD., S.V., J.W.A., J.S.G.D. and G.Mc.M. performed the experiments or analysis. N.G.T., N.D.M., C.I.R.G. and G.Mc.D. prepared the figures and tables. N.G.T., G.Mc.M., C.I.R.G., and G.Mc.D. wrote the paper. All authors read and approved the final version.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-35050-x.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018