

# Members of the human gut microbiota involved in recovery from *Vibrio cholerae* infection

Ansel Hsiao<sup>1</sup>, A. M. Shamsir Ahmed<sup>2,3</sup>, Sathish Subramanian<sup>1</sup>, Nicholas W. Griffin<sup>1</sup>, Lisa L. Drewry<sup>1</sup>, William A. Petri Jr<sup>4,5,6</sup>, Rashidul Haque<sup>3</sup>, Tahmeed Ahmed<sup>3</sup> & Jeffrey I. Gordon<sup>1</sup>

**Given the global burden of diarrhoeal diseases<sup>1</sup>, it is important to understand how members of the gut microbiota affect the risk for, course of, and recovery from disease in children and adults. The acute, voluminous diarrhoea caused by *Vibrio cholerae* represents a dramatic example of enteropathogen invasion and gut microbial community disruption. Here we conduct a detailed time-series metagenomic study of faecal microbiota collected during the acute diarrhoeal and recovery phases of cholera in a cohort of Bangladeshi adults living in an area with a high burden of disease<sup>2</sup>. We find that recovery is characterized by a pattern of accumulation of bacterial taxa that shows similarities to the pattern of assembly/maturation of the gut microbiota in healthy Bangladeshi children<sup>3</sup>. To define the underlying mechanisms, we introduce into gnotobiotic mice an artificial community composed of human gut bacterial species that directly correlate with recovery from cholera in adults and are indicative of normal microbiota maturation in healthy Bangladeshi children<sup>3</sup>. One of the species, *Ruminococcus obeum*, exhibits consistent increases in its relative abundance upon *V. cholerae* infection of the mice. Follow-up analyses, including mono- and co-colonization studies, establish that *R. obeum* restricts *V. cholerae* colonization, that *R. obeum luxS* (autoinducer-2 (AI-2) synthase) expression and AI-2 production increase significantly with *V. cholerae* invasion, and that *R. obeum* AI-2 causes quorum-sensing-mediated repression of several *V. cholerae* colonization factors. Co-colonization with *V. cholerae* mutants discloses that *R. obeum* AI-2 reduces *Vibrio* colonization/pathogenicity through a novel pathway that does not depend on the *V. cholerae* AI-2 sensor, LuxP. The approach described can be used to mine the gut microbiota of Bangladeshi or other populations for members that use autoinducers and/or other mechanisms to limit colonization with *V. cholerae*, or conceivably other enteropathogens.**

We used an approved protocol for recruiting Bangladeshi adults living in Dhaka Municipal Corporation area for this study. Of the 1,153 patients with acute diarrhoea who were screened, seven passed all entry criteria (Methods) and were enrolled (Supplementary Tables 1 and 2). Faecal samples collected at monthly intervals during the first 2 post-natal years from 50 healthy children living in the Mirpur area of Dhaka city, plus samples obtained at approximately 3-month intervals over a 1-year period from 12 healthy adult males also living Mirpur, allowed us to compare recovery of the microbiota from cholera with the normal process of assembly of the gut community in infants and children, and with unperturbed communities from healthy adult controls.

Using the standard treatment protocol of the International Centre for Diarrhoeal Disease Research, Bangladesh, study participants with acute cholera received a single oral dose of azithromycin and were given oral rehydration therapy for the duration of their hospital stay. Patients were discharged after their first solid stool. We divided the diarrhoeal period (from the first diarrhoeal stool after admission to the first solid stool) into four proportionately equal time bins: diarrhoeal phase 1 (D-Ph1)

to D-Ph4. Every diarrhoeal stool was collected from every participant. Faecal samples were also collected every day for the first week after discharge (recovery phase 1, R-Ph1), weekly during the next 3 weeks (R-Ph2), and monthly for the next 2 months (R-Ph3). For each individual, we selected a subset of samples from D-Ph1 to D-Ph3 (Methods), plus all samples from D-Ph4 to R-Ph3, for analysis of bacterial composition by sequencing PCR amplicons generated from variable region 4 (V4) of the 16S ribosomal RNA (rRNA) gene (Supplementary Information, Extended Data Fig. 1a and Supplementary Table 3). Reads sharing 97% nucleotide sequence identity were grouped into operational taxonomic units (97%-identity OTUs; Methods).

We identified a total of 1,733 97%-identity OTUs assigned to 343 different species after filtering and rarefaction (Methods). *V. cholerae* dominated the microbiota of the seven patients with cholera during D-Ph1 (mean maximum relative abundance 55.6%), declining markedly within hours after initiation of oral rehydration therapy. The microbiota then became dominated by either an unidentified *Streptococcus* species (maximum relative abundance 56.2–98.6%) or by *Fusobacterium* species (19.4–65.1% in patients B–E). In patient G, dominance of the community passed from a *Campylobacter* species (58.6% maximum) to a *Streptococcus* species (98.6% maximum) (Supplementary Table 4). Of the 343 species,  $47.9 \pm 6.6\%$  (mean  $\pm$  s.d.) were observed throughout both the diarrhoeal and recovery phases, suggesting that microbiota composition during the recovery phase may reflect an outgrowth from reservoirs of bacteria retained during disruption by diarrhoea (Extended Data Fig. 2a–d and Supplementary Information).

Indicator species analysis<sup>4</sup> (Methods) was used to identify 260 bacterial species consistently associated with the diarrhoeal or recovery phases across members of the study group, and in a separate analysis for each subject (Supplementary Table 5). The relative abundance of each of the discriminatory species in each faecal sample was compared with the mean weighted phylogenetic (UniFrac<sup>5</sup>) distance between that microbiota sample and all microbiota samples collected from the reference cohort of healthy Bangladeshi adults. The results revealed 219 species with significant indicator value assignments to diarrhoeal or recovery phases, and relative abundances with statistically significant Spearman's rank correlation values to community UniFrac distance to healthy control microbiota (Supplementary Table 6 and Extended Data Fig. 2d). Not surprisingly, the abundance of *V. cholerae* directly correlated with increased distance to a healthy microbiota. *Streptococcus* and *Fusobacterium* species, which bloomed during the early phases of diarrhoea, were also significantly and positively correlated with distance from a healthy adult microbiota. Increases in the relative abundances of species in the genera *Bacteroides*, *Prevotella*, *Ruminococcus/Blautia*, and *Faecalibacterium* (for example, *Bacteroides vulgatus*, *Prevotella copri*, *R. obeum*, and *Faecalibacterium prausnitzii*) were strongly correlated with a shift in community structure towards a healthy adult configuration (Extended Data Fig. 2d and Supplementary Table 6).

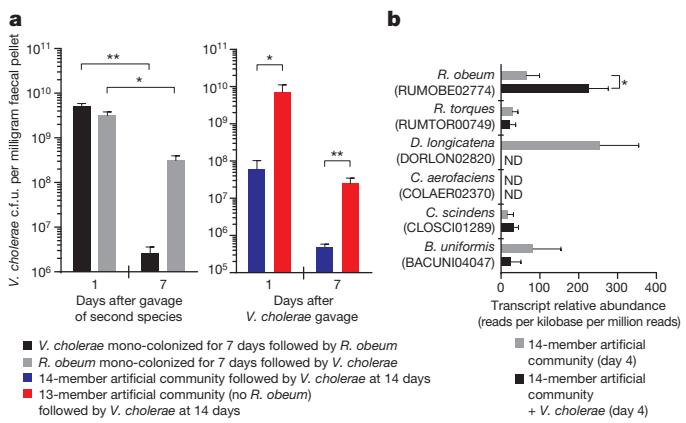
<sup>1</sup>Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St Louis, Missouri 63108, USA. <sup>2</sup>School of Population Health, The University of Queensland, Brisbane, Queensland 4006, Australia. <sup>3</sup>Centre for Nutrition and Food Security, International Centre for Diarrhoeal Disease Research, Dhaka 1212, Bangladesh. <sup>4</sup>Department of Medicine, University of Virginia School of Medicine, Charlottesville, Virginia 22908, USA. <sup>5</sup>Department of Microbiology, University of Virginia School of Medicine, Charlottesville, Virginia 22908, USA. <sup>6</sup>Department of Pathology, University of Virginia School of Medicine, Charlottesville, Virginia 22908, USA.

Previously we used Random Forests, a machine-learning algorithm, to identify a collection of age-discriminatory bacterial taxa that together define different stages in the postnatal assembly/maturation of the gut microbiota in healthy Bangladeshi children living in the same area as the adult patients with cholera<sup>3</sup>. Of those 60 most age-discriminatory 97%-identity OTUs representing 40 different species, 31 species were present in adult patients with cholera. Intriguingly, they followed a similar progression of changing representation during diarrhoea to recovery as they do during normal maturation of the healthy infant gut microbiota (Extended Data Fig. 2d). Twenty-seven of the 31 species were significantly associated with recovery from diarrhoea by indicator species analysis (see Supplementary Information and Extended Data Figs 3–5 for OTU-level and community-wide analyses). These 27 species, which serve as indicators and are potential mediators of restoration of the gut microbiota after cholera, guided construction of a gnotobiotic mouse model that examined the molecular mechanisms by which some of these taxa might affect *V. cholerae* infection and promote restoration.

We assembled an artificial community of 14 sequenced human gut bacterial species (Supplementary Table 7) that included (1) five species that directly correlated with gut microbiota recovery from cholera and with normal maturation of the infant gut microbiota (*R. obeum*, *Ruminococcus torques*, *F. prausnitzii*, *Dorea longicatena*, *Collinsella aerofaciens*), (2) six species significantly associated with recovery from cholera by indicator species analysis (*Bacteroides ovatus*, *Bacteroides vulgatus*, *Bacteroides caccae*, *Bacteroides uniformis*, *Parabacteroides distasonis*, *Eubacterium rectale*), and (3) three prominent members of the adult human gut microbiota that have known capacity to process dietary and host glycans (*Bacteroides cellulosilyticus*, *Bacteroides thetaiotaomicron*, *Clostridium scindens*<sup>6–8</sup>; as noted in Extended Data Fig. 6 and Supplementary Table 8, shotgun sequencing of diarrhoeal- and recovery-phase human faecal DNA samples revealed that genes encoding enzymes involved in carbohydrate metabolism were the largest category of identified genes specifying known enzymes that changed in relative abundance within the faecal microbiome during the course of cholera). One group of mice was directly inoculated with approximately  $10^9$  colony-forming units (c.f.u.) of *V. cholerae* at the same time they received the 14-member community to simulate the rapidly expanding *V. cholerae* population during diarrhoea ('D1invasion' group). A separate group was gavaged with the community alone and then invaded 14 days later with *V. cholerae* ('D14invasion' group) (Extended Data Fig. 1c).

*V. cholerae* levels remained at a high level in the D1invasion group over the first week (maximum 46.3% relative abundance), and then declined rapidly to low levels (<1%). Introduction of *V. cholerae* into the established 14-member community produced much lower levels of *V. cholerae* infection (range of mean abundances measured daily over the 3 days after gavage of the enteropathogen, 1.2–2.7%; Supplementary Table 9). Control experiments demonstrated that *V. cholerae* was able to colonize at high levels for at least 7 days when it was introduced alone into germ-free recipients ( $10^9$ – $10^{10}$  c.f.u. per milligram wet weight of faeces; Fig. 1a). Together, these data suggest that a member or members of the artificial human gut microbiota had the ability to restrict *V. cholerae* colonization.

Changes in relative abundances of the 14 community members in faecal samples in response to *V. cholerae* were consistent for most species across the D1invasion and D14invasion mice (Supplementary Table 9). We focused on one member, *R. obeum*, because its relative abundance increased significantly after introduction of *V. cholerae* in both the D1invasion and D14invasion groups (Extended Data Fig. 7a and Supplementary Table 9) and because it is a prominent age-discriminatory taxon in the Random Forests model of gut microbiota maturation in healthy Bangladeshi children<sup>3</sup> (Extended Data Fig. 4b). Mice were mono-colonized with either *R. obeum* or *V. cholerae* for 7 days and then the other species was introduced (Extended Data Fig. 1d). When *R. obeum* was present, *V. cholerae* levels declined by 1–3 logs (Fig. 1a). Germ-free mice were also colonized with the defined 14-member community or the same community without *R. obeum* for 2 weeks, and *V. cholerae* was



**Figure 1 | *R. obeum* restricts *V. cholerae* colonization in adult gnotobiotic mice.** **a**, *V. cholerae* levels in the faeces of mice colonized with the indicated human gut bacterial species ( $n = 4$ –6 mice per group). **b**, Expression of *R. obeum luxS* AI-2 synthase in the 14-member community 4 days after introduction of  $10^9$  c.f.u. of *V. cholerae* or no pathogen ( $n = 5$  mice per group). Note that *D. longicatena* levels fall precipitously after *V. cholerae* invasion (Supplementary Table 9). Mean values  $\pm$  s.e.m. are shown. ND, not detected. \* $P < 0.05$ , \*\* $P < 0.01$  (unpaired Mann–Whitney *U*-test).

then introduced by gavage (Extended Data Fig. 1e). *V. cholerae* levels 1 day after gavage were 100-fold higher in the community that lacked *R. obeum*; these differences were sustained over time (50-fold higher after 7 days;  $P < 0.01$ , unpaired Mann–Whitney *U*-test; Fig. 1a).

Having established that *R. obeum* restricts *V. cholerae* colonization, we used microbial RNA sequencing (RNA-seq) of faecal RNAs to determine the effect of *R. obeum* on expression of known *V. cholerae* virulence factors in mono- and co-colonized mice. Co-colonization led to reduced expression of *tcpA* (a primary colonization factor in humans<sup>9,10</sup>), *rtxA* and *hlyA* (encode accessory toxins<sup>11,12</sup>), and *VC1447–VC1448* (RtxA transporters) (threefold to fivefold changes;  $P < 0.05$  compared with *V. cholerae* mono-colonized controls, Mann–Whitney *U*-test; see Supplementary Information and Supplementary Table 10 for other regulated genes that could impact colonization, plus Extended Data Fig. 8 for an ultra-performance liquid chromatography mass spectrometry (UPLC–MS) analysis of bile acids reported to effect *V. cholerae* gene regulation<sup>13</sup>).

Two quorum-sensing pathways are known to regulate *V. cholerae* colonization/virulence<sup>14–17</sup>: an intra-species mechanism involving cholera autoinducer-1, and an inter-species mechanism involving autoinducer-2 (refs 18, 19). Quorum sensing disrupts expression of *V. cholerae* virulence determinants through a signalling pathway that culminates in production of the LuxR-family regulator HapR<sup>15,16</sup>. Repression of quorum sensing in *V. cholerae* is important for virulence factor expression and infection<sup>20–22</sup>. The *luxS* gene encodes the S-ribosylhomocysteine lyase responsible for AI-2 synthesis. Homologues of *luxS* are widely distributed among bacteria<sup>18,19</sup>, including 8 of the 14 species in the artificial human gut community (Supplementary Table 11 and Extended Data Fig. 9). RNA-seq of the faecal meta-transcriptomes of D1invasion mice colonized with the 14-member artificial community plus *V. cholerae*, and mice harbouring the 14-member consortium without *V. cholerae*, revealed that of predicted *luxS* homologues in the community, only expression of *R. obeum luxS* (RUMOBE02774) increased significantly in response to *V. cholerae* ( $P < 0.05$ , Mann–Whitney *U*-test; Fig. 1b). Moreover, *R. obeum luxS* transcript levels directly correlated with *V. cholerae* levels (Extended Data Fig. 7c).

In addition to *luxS*, the *R. obeum* strain represented in the artificial community contains homologues of *lsrABC* that are responsible for import and phosphorylation of AI-2 in Gram-negative bacteria<sup>23</sup>, as well as homologues of two genes, *luxR* and *luxQ*, that play a role in AI-2 sensing and downstream signalling in other organisms<sup>24</sup>. Expression of all these *R. obeum* genes was detected *in vivo*, consistent with *R. obeum*

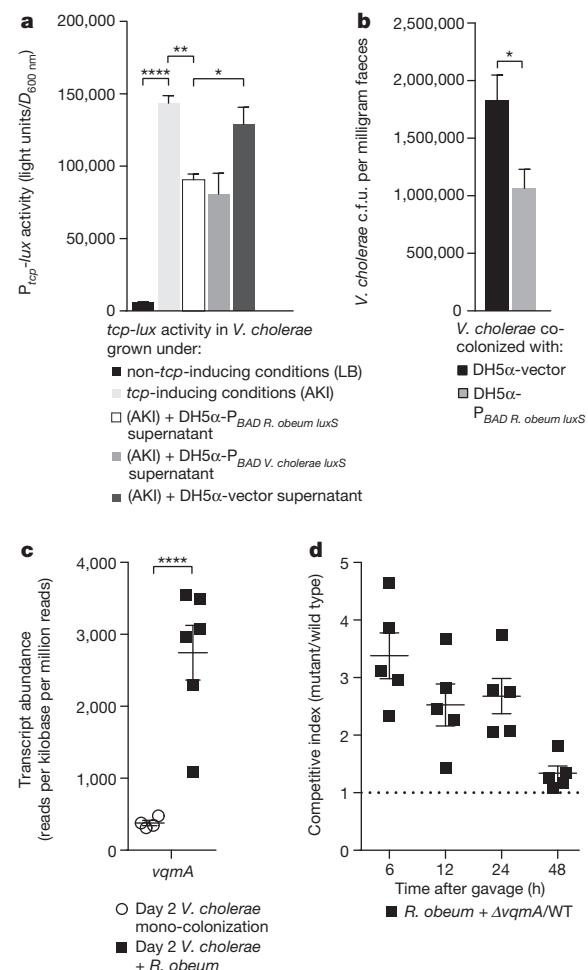
having a functional AI-2 signalling system (Extended Data Fig. 7b). (See Supplementary Information for results showing that *R. obeum* AI-2 production is stimulated by *V. cholerae* *in vitro* and in co-colonized animals (Extended Fig. 7d-f), plus (1) a genome-wide analysis of the effects of *V. cholerae* on *R. obeum* transcription in co-colonized mice (Supplementary Table 10c) and (2) a community-wide view of the transcriptional responses of the 14-member consortium to *V. cholerae* (Supplementary Table 12).)

Quorum sensing downregulates the *V. cholerae* *tcp* operon that encodes components of the toxin co-regulated pilus (TCP) biosynthesis pathway required for infection of humans<sup>9,10</sup>. To confirm that *R. obeum* LuxS could signal through AI-2 pathways, we cloned *R. obeum* and *V. cholerae* *luxS* downstream of the arabinose-inducible *P<sub>BAD</sub>* promoter in plasmids that were maintained in an *Escherichia coli* strain unable to produce its own AI-2 (DH5 $\alpha$ )<sup>25</sup>. High *tcp* expression can be induced in *V. cholerae* after slow growth in AKI medium without agitation followed by rapid growth under aerobic conditions<sup>26</sup>. Addition of culture supernatants harvested from the *E. coli* strains expressing *R. obeum* or *V. cholerae luxS* caused a two- to threefold reduction in *tcp* induction in *V. cholerae* ( $P < 0.05$ , unpaired Student's *t*-test; replicated in four independent experiments). Supernatants from a control *E. coli* strain with the plasmid vector lacking *luxS* had no effect (Fig. 2a). These findings are consistent with our *in vivo* RNA-seq results and provide direct evidence that *R. obeum* AI-2 regulates expression of *V. cholerae* virulence factor.

Germ-free mice were then colonized with *V. cholerae* and *E. coli* bearing either the *P<sub>BAD</sub>-R. obeum luxS* plasmid or the vector control. Mice that received *E. coli* expressing *R. obeum luxS* showed a significantly lower level of *V. cholerae* colonization 8 h after gavage than mice that received *E. coli* with vector alone (Fig. 2b; there was no statistically significant difference in levels of *E. coli* between the two groups (data not shown)). Together, these results establish a direct causal relationship between *R. obeum*-mediated restriction of *V. cholerae* colonization and *R. obeum* AI-2 synthesis.

Several *V. cholerae* mutants were used to determine whether known *V. cholerae* AI-2 signalling pathways are required for the observed effects of *R. obeum* on *V. cholerae* colonization. LuxP is critical for sensing AI-2 in *V. cholerae*. Co-colonization experiments in gnotobiotic mice revealed that levels of isogenic  $\Delta$ *luxP* or wild-type *luxP*<sup>+</sup> *V. cholerae* strains were not significantly different as a function of the presence of *R. obeum* (Extended Data Fig. 10a), suggesting that *R. obeum* modulates *V. cholerae* levels through other quorum-sensing regulatory genes. The *luxO* and *hapR* genes encode central regulators linking known *V. cholerae* quorum-signalling and virulence regulatory pathways. Deletion of *luxO* typically results in increased *hapR* expression<sup>15</sup>. However, our RNA-seq analysis had shown that both *luxO* and *hapR* are repressed in the presence of *R. obeum* (six- to sevenfold,  $P < 0.0001$ ; Mann-Whitney *U*-test), as are two important downstream activators of virulence repressed by HapR<sup>16</sup>, encoded by *aphA* and *aphB*. These findings provide additional evidence that *R. obeum* operates to regulate virulence through a novel regulatory pathway.

The quorum-sensing transcriptional regulator VqmA was upregulated more than 25-fold when *V. cholerae* was introduced into mice mono-colonized with *R. obeum* (Fig. 2c and Supplementary Table 10). When germ-free mice were gavaged with *R. obeum* and a mixture of  $\Delta$ *vqmA* (*lacZ*)<sup>27</sup> and wild-type *V. cholerae* (*lacZ*<sup>+</sup>) strains, the  $\Delta$ *vqmA* mutant exhibited an early competitive advantage (Fig. 2d), suggesting that *R. obeum* may be able to affect early colonization of *V. cholerae* through VqmA. VqmA is able to bind to and activate the *hapR* promoter directly<sup>27</sup>. Since RNA-seq showed that *hapR* activation did not occur in gnotobiotic mice despite high levels of *vqmA* expression (Extended Data Fig. 10b and Supplementary Table 10), we postulate that the role played by VqmA in *R. obeum* modulation of *Vibrio* virulence genes involves an uncharacterized mechanism rather than the known pathway passing through HapR.



**Figure 2 | *R. obeum* AI-2 reduces *V. cholerae* colonization and virulence gene expression.** **a**, *R. obeum* AI-2 produced in *E. coli* represses the *tcp* promoter in *V. cholerae* (triplicate assays; results representative of four independent experiments). **b**, Faecal *V. cholerae* levels in gnotobiotic mice 8 h after gavage with *V. cholerae* and an *E. coli* strain containing either the *P<sub>BAD</sub>-R. obeum luxS* plasmid or vector control. **c**, Faecal *vqmA* transcript abundance in mono- or co-colonized mice. **d**, Competitive index of  $\Delta$ *vqmA* versus wild-type *V. cholerae* during co-colonization with *R. obeum* ( $n = 5$  animals per group). Mean values  $\pm$  s.e.m. are shown. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.0001$  (unpaired two-tailed Student's *t*-test).

We have identified a set of bacterial species that strongly correlate with a process in which the perturbed gut bacterial community in adult patients with cholera is restored to a configuration found in healthy Bangladeshi adults. Several of these species are also associated with the normal assembly/maturation of the gut microbiota in Bangladeshi infants and children, raising the possibility that some of these taxa may be useful for ‘repair’ of the gut microbiota in individuals whose gut communities have been ‘wounded’ through a variety of insults, including enteropathogen infections. Translating these observations to a gnotobiotic mouse model containing an artificial human gut microbiota composed of recovery- and age-indicative taxa established that one of these species, *R. obeum*, reduces *V. cholerae* colonization. As an entrenched member of the gut microbiota in Bangladeshi individuals, *R. obeum* could function to increase median infectious dose ( $ID_{50}$ ) for *V. cholerae* in humans and thus help to determine whether exposure to a given dose of this enteropathogen results in diarrhoeal illness. The modest effects of *R. obeum* AI-2 on *V. cholerae* virulence gene expression in our adult gnotobiotic mouse model may reflect the possibility that we have only identified a small fraction of the microbiota’s full repertoire of virulence-suppressing mechanisms. Culture collections generated from the faecal microbiota

of Bangladeshi subjects are a logical starting point for ‘second-generation’ artificial communities containing *R. obeum* isolates that have evolved in this population, and for testing whether the observed effects of *R. obeum* generalize across many different strains from different populations. Moreover, the strategy described in this report could be used to mine the gut microbiota of Bangladeshi or other populations where diarrhoeal disease is endemic for additional species that use quorum-related and/or other mechanisms to limit colonization by *V. cholerae* and potentially other enteropathogens.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 October 2013; accepted 6 August 2014.

Published online 17 September 2014.

1. World Health Organization Cholera, 2013. *Wkly Epidemiol. Rec.* **89**, 345–356 (2014).
2. Chowdhury, F. et al. Impact of rapid urbanization on the rates of infection by *Vibrio cholerae* O1 and enterotoxigenic *Escherichia coli* in Dhaka, Bangladesh. *PLoS Negl. Trop. Dis.* **5**, e999 (2011).
3. Subramanian, S. et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510**, 417–421 (2014).
4. Dufrene, M. & Legendre, P. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecol. Monogr.* **67**, 345–366 (1997).
5. Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**, 8228–8235 (2005).
6. Martens, E. C. et al. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biol.* **9**, e1001221 (2011).
7. McNulty, N. P. et al. Effects of diet on resource utilization by a model human gut microbiota containing *Bacteroides cellulosilyticus* WH2, a symbiont with an extensive glycomicrobiome. *PLoS Biol.* **11**, e1001637 (2013).
8. McNulty, N. P. et al. The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. *Sci. Translat. Med.* **3**, 106ra106 (2011).
9. Taylor, R. K., Miller, V. L., Furlong, D. B. & Mekalanos, J. J. Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl Acad. Sci. USA* **84**, 2833–2837 (1987).
10. Herrington, D. A. et al. Toxin, toxin-coregulated pili, and the toxR regulon are essential for *Vibrio cholerae* pathogenesis in humans. *J. Exp. Med.* **168**, 1487–1492 (1988).
11. Olivier, V., Salzman, N. H. & Satchell, K. J. Prolonged colonization of mice by *Vibrio cholerae* El Tor O1 depends on accessory toxins. *Infect. Immun.* **75**, 5043–5051 (2007).
12. Olivier, V., Haines, G. K., III, Tan, Y. & Satchell, K. J. Hemolysin and the multifunctional autoprocessing RTX toxin are virulence factors during intestinal infection of mice with *Vibrio cholerae* El Tor O1 strains. *Infect. Immun.* **75**, 5035–5042 (2007).
13. Yang, M. et al. Bile salt-induced intermolecular disulfide bond formation activates *Vibrio cholerae* virulence. *Proc. Natl Acad. Sci. USA* **110**, 2348–2353 (2013).
14. Miller, M. B., Skorupski, K., Lenz, D. H., Taylor, R. K. & Bassler, B. L. Parallel quorum sensing systems converge to regulate virulence in *Vibrio cholerae*. *Cell* **110**, 303–314 (2002).
15. Zhu, J. et al. Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae*. *Proc. Natl Acad. Sci. USA* **99**, 3129–3134 (2002).
16. Kovacikova, G. & Skorupski, K. Regulation of virulence gene expression in *Vibrio cholerae* by quorum sensing: HapR functions at the aphaA promoter. *Mol. Microbiol.* **46**, 1135–1147 (2002).
17. Higgins, D. A. et al. The major *Vibrio cholerae* autoinducer and its role in virulence factor production. *Nature* **450**, 883–886 (2007).
18. Pereira, C. S., Thompson, J. A. & Xavier, K. B. AI-2-mediated signalling in bacteria. *FEMS Microbiol. Rev.* **37**, 156–181 (2013).
19. Sun, J., Daniel, R., Wagner-Dobler, I. & Zeng, A. P. Is autoinducer-2 a universal signal for interspecies communication: a comparative genomic and phylogenetic analysis of the synthesis and signal transduction pathways. *BMC Evol. Biol.* **4**, 36 (2004).
20. Duan, F. & March, J. C. Engineered bacterial communication prevents *Vibrio cholerae* virulence in an infant mouse model. *Proc. Natl Acad. Sci. USA* **107**, 11260–11264 (2010).
21. Liu, Z. et al. Mucosal penetration primes *Vibrio cholerae* for host colonization by repressing quorum sensing. *Proc. Natl Acad. Sci. USA* **105**, 9769–9774 (2008).
22. Liu, Z., Stirling, F. R. & Zhu, J. Temporal quorum-sensing induction regulates *Vibrio cholerae* biofilm architecture. *Infect. Immun.* **75**, 122–126 (2007).
23. Taga, M. E., Semmelhack, J. L. & Bassler, B. L. The LuxS-dependent autoinducer AI-2 controls the expression of an ABC transporter that functions in AI-2 uptake in *Salmonella typhimurium*. *Mol. Microbiol.* **42**, 777–793 (2001).
24. Bassler, B. L., Wright, M. & Silverman, M. R. Multiple signalling systems controlling expression of luminescence in *Vibrio harveyi*: sequence and function of genes encoding a second sensory pathway. *Mol. Microbiol.* **13**, 273–286 (1994).
25. Surette, M. G., Miller, M. B. & Bassler, B. L. Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: a new family of genes responsible for autoinducer production. *Proc. Natl Acad. Sci. USA* **96**, 1639–1644 (1999).
26. Iwanaga, M. et al. Culture conditions for stimulating cholera toxin production by *Vibrio cholerae* O1 El Tor. *Microbiol. Immunol.* **30**, 1075–1083 (1986).
27. Liu, Z., Hsiao, A., Joelson, A. & Zhu, J. The transcriptional regulator VqmA increases expression of the quorum-sensing activator HapR in *Vibrio cholerae*. *J. Bacteriol.* **188**, 2446–2453 (2006).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank S. Wagoner, J. Hoisington-López, M. Meier, J. Cheng, D. O’Donnell, and M. Karlsson for technical support, J. Zhu for providing strains of *V. cholerae* and *Vibrio harveyi*, and W.-L. Ng for providing *AluxP*. This work was supported in part by a grant from the Bill & Melinda Gates Foundation. The singleton birth cohort of Bangladeshi children was supported by a grant from the National Institutes of Health (AI 43596). The post-doctoral fellowship stipend of A.H. was funded in part by NIH training grants (T32DK077653, T32AI007172) and by the Crohn’s and Colitis Foundation of America. The International Centre for Diarrhoeal Disease Research, Bangladesh, acknowledges the following donors, which provided unrestricted support: the Australian Agency for International Development, the Government of Bangladesh, the Canadian International Development Agency, the Swedish International Development Cooperation Agency, and the Department for International Development, UK.

**Author Contributions** A.H. and J.I.G. designed the metagenomic and gnotobiotic mouse study; A.M.S.A., R.H., and T.A. designed and implemented the clinical study, participated in patient recruitment, sample collection, sample preservation and clinical evaluations; R.H. and W.A.P. participated in recruitment of and sample collection from healthy Bangladeshi controls; A.H. generated the 16S rRNA, AI-2, RNA-seq, shotgun microbial community DNA sequencing, and *V. cholerae* colonization data. S.S. generated 16S rRNA data from extended sampling of the Bangladeshi singleton birth cohort. L.L.D. performed 16S rRNA sequencing of the additional samples from patients C and E and helped generate the colonization data in *in vivo* competition experiments involving isogenic wild-type,  $\Delta vqmA$  and  $\Delta luxP$  strains of *V. cholerae* C6706; A.H., S.S., N.W.G., and J.I.G. analysed the data; A.H. and J.I.G. wrote the paper.

**Author Information** All 16S rRNA, shotgun sequencing, and RNA-seq data sets generated from faecal samples have been deposited in the European Nucleotide Archive in raw format before post-processing and data analysis under accession number PRJEB6358. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.I.G. (jgordon@wustl.edu).

## METHODS

**Human studies.** Subject recruitment. Protocols for recruitment, enrollment, and consent, procedures for sampling the faecal microbiota of healthy Bangladeshi adults and children, and the faecal microbiota of adults during and after cholera infection, plus the subsequent de-identification of these samples, were approved by the Human Studies Committees of the International Centre for Diarrhoeal Disease Research, Bangladesh, and Washington University School of Medicine in St. Louis.

Enrollment into the adult cholera study was based on the following criteria: residency in the Dhaka Municipal Corporation area, a positive stool test for *V. cholerae* as judged by dark-field microscopy, diarrhoea for no more than 24 h before enrollment, and a permanent address that allowed follow-up faecal sampling after discharge from Dhaka Hospital (International Centre for Diarrhoeal Disease Research, Bangladesh). Non-prescription antibiotic usage is prevalent in Bangladesh<sup>28,29</sup>. Since a history of previous antibiotic consumption could be a confounder when interpreting the effects of cholera on the gut microbiota, we excluded individuals if they had received antibiotics in the 7 days preceding admission to the hospital. Since this was an observational study with no experimental treatment arm, blinding for study inclusion was not necessary. See Supplementary Table 1 for the number of individuals screened for inclusion in the study, the number of potential subjects excluded from the study and the reasons for their exclusion, and the number of subjects enrolled who satisfied all criteria for inclusion.

The healthy adults were fathers in a cohort of healthy twins, triplets, and their parents living in Mirpur that is described in ref. 3. Fathers were sampled every 3 months during the first 2 years of their offspring's postnatal life. Histories of diarrhoea and antibiotic use were not available for these fathers. However, histories of diarrhoea and antibiotic use in their healthy children were known: 46 of the 49 paternal faecal samples used were obtained during periods when none of their children had diarrhoea; 36 of these 49 samples were collected at a time when there had been no antibiotic use by their children in the preceding 7 days.

DNA extraction from human faecal samples, sequencing, and analysis. All diarrhoeal stools were collected from each participant (one sterilized bowl per sample), frozen immediately at  $-80^{\circ}\text{C}$ , then subjected to the same bead beating and phenol chloroform extraction procedure for DNA purification that was applied to the formed frozen faecal samples collected from these individuals during the recovery phases (and previously to a wide range of samples collected from individuals representing different ages, cultural traditions, geographical locations, and physiological and disease states<sup>3,30</sup>).

DNA was isolated from all frozen faecal samples from D-Ph1 to D-Ph4, from the period of frequent sampling during the first week following discharge (recovery phase 1; R-Ph1), the period of less frequent sampling during weeks 2–3 (R-Ph2), and from weeks 4 to 12 of recovery (R-Ph3) ( $n = 1,053$  samples in total). For analyses involving healthy adult and child control groups, samples were excluded from our analysis where antibiotic use or diarrhoea was known to have occurred in the 7 days before sample collection.

For each participant in the cholera study, we selected one sample with high DNA yield ( $\geq 2\ \mu\text{g}$ ) from each 2-hour period during D-Ph1 to D-Ph3. An additional  $7 \pm 2$  samples (mean  $\pm$  s.d.) that had been collected during the approximately 5-h period before the rate of diarrhoea began to decrease at the beginning of D-Ph3 were included. All faecal samples collected after this time point (that is, from the remainder of D-Ph3 to R-Ph3), were also included in our analysis ( $n = 19.7 \pm 7.4$  total samples (mean  $\pm$  s.d.) per individual in the diarrhoeal phase, and  $14 \pm 3.3$  total samples per individual in the recovery phase). Two patients (C and E) were chosen for additional sequencing of all their diarrhoeal samples ( $n = 100$  and 50, respectively; see Supplementary Table 3b).

The V4 region of bacterial 16S rRNA genes represented in each selected faecal microbiota sample was amplified by PCR using primers containing sample-specific barcode identifiers. Amplicons were purified, pooled, and paired-end sequenced with an Illumina MiSeq instrument (250 nucleotide paired-end reads;  $86,315 \pm 2,043$  (mean  $\pm$  s.e.m.) assembled reads per sample; see Supplementary Table 3). Healthy control samples were analysed using the same sequencing platform and chemistry ( $n = 293$  total samples).

Sequences were assembled, then de-multiplexed and analysed using the QIIME software package<sup>31</sup> and custom Perl scripts. For analysis of diarrhoeal and recovery phase samples, rarefaction was performed to 49,000 reads per sample. For analyses including samples from healthy adults and children, samples were rarefied to 7,900 reads per sample. Reads sharing 97% nucleotide sequence identity were grouped into operational taxonomic units (97%-identity OTUs). To ensure that we retained less abundant bacterial taxa in our analysis of the faecal samples of patients with cholera, a 97%-identity OTU was called 'distinct and reliable' if it appeared at 0.1% relative abundance in at least one faecal sample. Taxonomic assignments of OTUs to species level were made using the Ribosomal Database Project version 2.4 classifier<sup>32</sup> and a manually curated Greengenes database<sup>33</sup>.

Indicator species analysis<sup>4</sup> was used to classify bacterial species as highly associated with either diarrhoeal phases or recovery. This approach is used in studies of macroecosystems to identify species that associate with different environmental groupings; it assigns for each species an indicator value that is a product of two components: (1) the species' specificity, which is the probability that a sample in which the species is found came from a given group; and (2) the species' fidelity, which is the proportion of samples from a given group that contains the species. We performed indicator species analysis in the set of 236 faecal specimens, selected from the seven patients according to the subsampling scheme described above, to identify bacterial species consistently associated with the diarrhoeal or recovery phases across members of the study group; statistical significance was defined using permutation tests in which permutations were constrained within subjects. We also conducted a separate indicator species analysis for each subject, using each individual's replicate diarrhoeal and recovery phase samples as the groupings.

For analyses of variation across communities, we used UniFrac<sup>5</sup>, a metric that measures the overall degree of phylogenetic similarity of any two communities based on the degree to which they share branch length on a bacterial tree of life; low pairwise UniFrac distance values indicate that communities are more similar to one another. UniFrac distances were calculated using the QIIME software package<sup>31</sup>.

The gut microbiomes of study participants were characterized by paired-end 2  $\times$  250 nucleotide shotgun sequencing of faecal DNA using an Illumina MiSeq instrument (mean 216,698 reads per sample; Supplementary Table 3). Paired sequences were assembled into single reads using the SHERA software package<sup>34</sup>, and annotated by mapping to version 58 of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database<sup>35</sup> using UBLAST<sup>36</sup>.

**Gnotobiotic mouse experiments.** All experiments involving animals used protocols approved by the Washington University Animal Studies Committee. Germ-free male C57BL/6J mice were maintained in flexible plastic film gnotobiotic isolators and fed an autoclaved, low-fat, plant polysaccharide-rich mouse chow (B&K, catalogue number 7378000, Zeigler Bros) *ad libitum*. Mice were 5–8 weeks old at time of gavage. The number of mice used in each experiment is reported in the text, relevant figure legends, and summarized in Extended Data Fig. 1.

Bacterial strains and plasmids. Supplementary Table 7 lists the sequenced human gut-derived bacterial strains used to generate the artificial communities and their sources. Since all Bangladeshi faecal samples were devoted to DNA extraction, we were unable to utilize strains that originated from culture collections generated from study participants' faecal biospecimens. Thus, the strains incorporated into the artificial community were from public repositories, represented multiple individuals, and were typically not accompanied by information about donor health status or living conditions.

A  $P_{tcp}$ -*lux* reporter strain was constructed by introducing  $P_{tcp}$ -*lux* (pJZ376) into *V. cholerae* C6706 via conjugation from SM10λpir.  $P_{BAD}$ -*luxS* expression vectors were produced by first amplifying the *luxS* sequences of *V. cholerae* C6706 and *R. obium* ATCC2917 using PCR and the primers described in Supplementary Table 13. Amplicons were then cloned into pBAD202 (TOPO TA Expression Kit; Life Technologies), and introduced into *E. coli* DH5α by electroporation.

All cultures of *V. cholerae* C6706, the isogenic  $\Delta luxS$  mutant (MM883), and *E. coli* strains containing *luxS* expression vectors were grown aerobically in Luria Broth (LB) medium with appropriate antibiotics (Supplementary Table 13). All members of the 14-member artificial human gut microbiota, including *R. obium* ATCC29174, were propagated anaerobically in MegaMedium<sup>37</sup>.

Colonization of gnotobiotic mice. All animal experiments involved administration of known consortia of bacterial species; as such, no blinding to group allocation was performed. The order of administration of microbial species to given groups of recipient mice was intentionally varied, as described in Extended Data Fig. 1c–e.

Mono-colonized animals received either 200  $\mu\text{l}$  of overnight cultures of *R. obium* strain ATCC29174 or *V. cholerae* strain C6706. All *V. cholerae* colonization studies in mice used the current pandemic El Tor biotype (strain C6706). Mice receiving the defined 13- or 14-member communities of sequenced human bacterial symbionts were gavaged with 200  $\mu\text{l}$  of an equivalent mixture of bacteria assembled from overnight monocultures of each strain ( $D_{600\text{nm}} \approx 0.4$  per strain; grown in MegaMedium). In the case of mice that received mixtures of *V. cholerae* and *E. coli* strains with *R. obium luxS*-expressing plasmids (or vector controls), the *E. coli* strains were first grown overnight in LB medium containing 50  $\mu\text{g ml}^{-1}$  kanamycin. Two millilitres of the culture were removed and cell pellets were obtained by centrifugation, washed three times with 2 ml LB medium to remove antibiotics, and re-suspended in 6 ml LB medium containing 0.1% arabinose. The suspension of *E. coli* cells was then incubated at  $37^{\circ}\text{C}$  for 90 min, and mixed with *V. cholerae* C6706 such that each mouse was gavaged with  $\sim 50\ \mu\text{l}$  and  $\sim 2.5\ \mu\text{l}$  of overnight cultures of each organism, respectively. All gavages involving *V. cholerae* were preceded by a gavage of 100  $\mu\text{l}$  sterile 1 M sodium bicarbonate to neutralize gastric pH. Colonization levels of *V. cholerae* were determined by serial dilution plating of faecal homogenates on selective medium.

Competitive index assays were performed with mice gavaged with 50 µl aliquots of cultures of mutant and wild-type *V. cholerae* C6706 strains that had been grown to  $D_{600\text{nm}} = 0.3$ . For experiments involving competitive index calculations as a function of the presence of *R. obeum*, 100 µl of an overnight *R. obeum* culture was co-inoculated with the mixture of *V. cholerae* strains. Faecal samples from recipient gnotobiotic mice were subjected to dilution plating and aerobic growth on LB agar with the LacZ substrate Xgal; blue–white screening was used to determine colonization levels of the individual *V. cholerae* strains.

Community profiling by shotgun sequencing (COPRO-seq). Shotgun sequencing of faecal community DNA was used to define the relative abundance of species in the artificial communities; experimental and computational tools for COPRO-seq have been described previously<sup>8</sup>.

Microbial RNA-seq analysis of faecal samples collected from mice colonized with the 14-member artificial community with and without *V. cholerae*. Faecal samples were collected from colonized gnotobiotic mice and immediately snap-frozen in liquid nitrogen. RNA was extracted using bead-beating in phenol/chloroform/isoamyl alcohol followed by further purification using MEGAclear (Life Technologies). Purified RNA was depleted of 16S rRNA, 5S rRNA, and transfer RNA as previously described<sup>8</sup> or by using a RiboZero kit (Epicentre). Complementary DNA (cDNA) libraries were generated and sequenced (50 nucleotide unidirectional reads; Illumina GA-IIx, HiSeq 2000 or MiSeq instruments; see Supplementary Table 3). Reads were mapped to the genomes of members of the artificial community using Bowtie<sup>38</sup>.

To profile transcriptional responses to *V. cholerae*, all cDNA reads that mapped to the genomes of the 14 consortium members were binned based on enzyme classification level annotations from KEGG. ShotgunFunctionalizeR<sup>39</sup> was then used compare the faecal meta-transcriptomes of 'D14invasion' animals sampled 4 days after gavage of the 14-member community to the faecal meta-transcriptomes of D1invasion mice sampled 4 days after gavage of the 14-member community plus *V. cholerae*. A mean twofold or greater difference in expression between the conditions, with an adjusted *P* value less than 0.0001 (ShotgunFunctionalizeR) was considered significant. This approach of binning to enzyme classifications mitigates issues with low-abundance transcripts being insufficiently profiled owing to limitations in sequencing depth<sup>8</sup>.

Owing to the higher sequencing depth achieved for *R. obeum* and *V. cholerae* in mono- and co-colonization experiments, reads were mapped to reference genomes using Bowtie, and changes at the single transcript level were analysed using DESeq<sup>40</sup> (Supplementary Table 11). Transcripts that satisfied the criteria of (1) having greater than twofold differential expression after DESeq normalization, (2) an adjusted *P* value less than 0.05, and (3) a minimum mean count value more than 10 were retained.

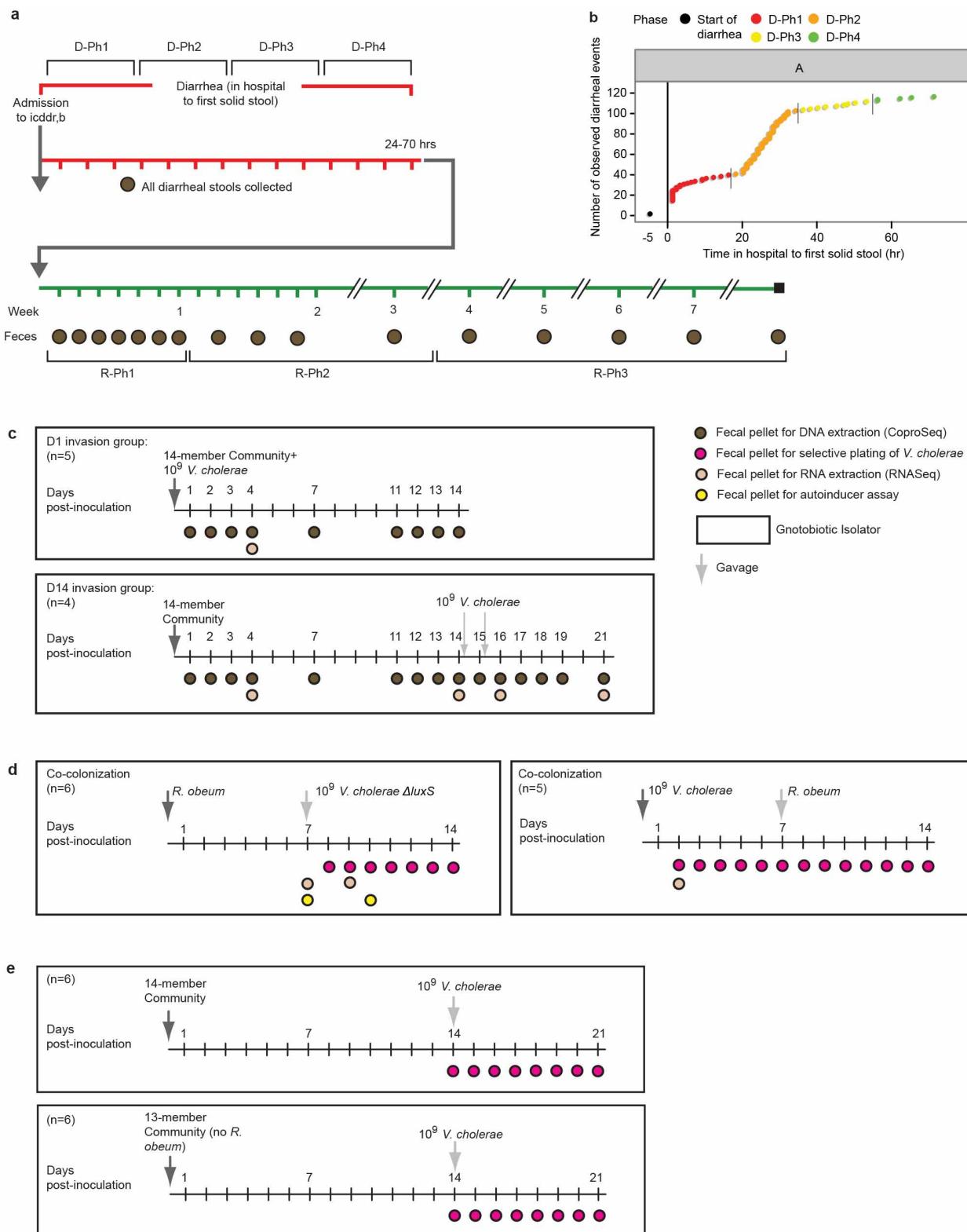
**AI-2 assays.** Previously frozen faecal pellets from gnotobiotic mice were re-suspended in AB medium<sup>24</sup> by agitation with a rotary bead-beater (25 mg faecal

pellet per millilitre of medium). AI-2 assays were performed using the *V. harveyi* BB170 bioassay strain<sup>24</sup>, with reported results representative of at least two independent experiments, each with five technical repeats. *V. harveyi* BB170 cultures were grown aerobically overnight in AB medium, and diluted 1:500 in this medium for use in the AI-2 bioassay<sup>24</sup>. Luminescence was measured using a BioTek Synergy 2 instrument after 4 h of growth at 30 °C with agitation (300 r.p.m. using a rotatory incubator).

For *in vitro* measurements of *R. obeum* AI-2 production, a 100 µl aliquot from an overnight monoculture of the bacterium grown in MegaMedium without glucose was diluted 1:20 in fresh MegaMedium without glucose. In addition, cells pelleted from 100 µl of an overnight culture of *V. cholerae* ΔluxS (MM883 (ref. 14)) grown in LB medium were added to *R. obeum* that had also been diluted 1:20 in MegaMedium without glucose. The resulting mono- and co-cultures were incubated anaerobically at 37 °C for 16 h. Cells were pelleted by centrifugation, and supernatants were harvested and then added to *V. harveyi* BB170 cultures for AI-2 bioassay.

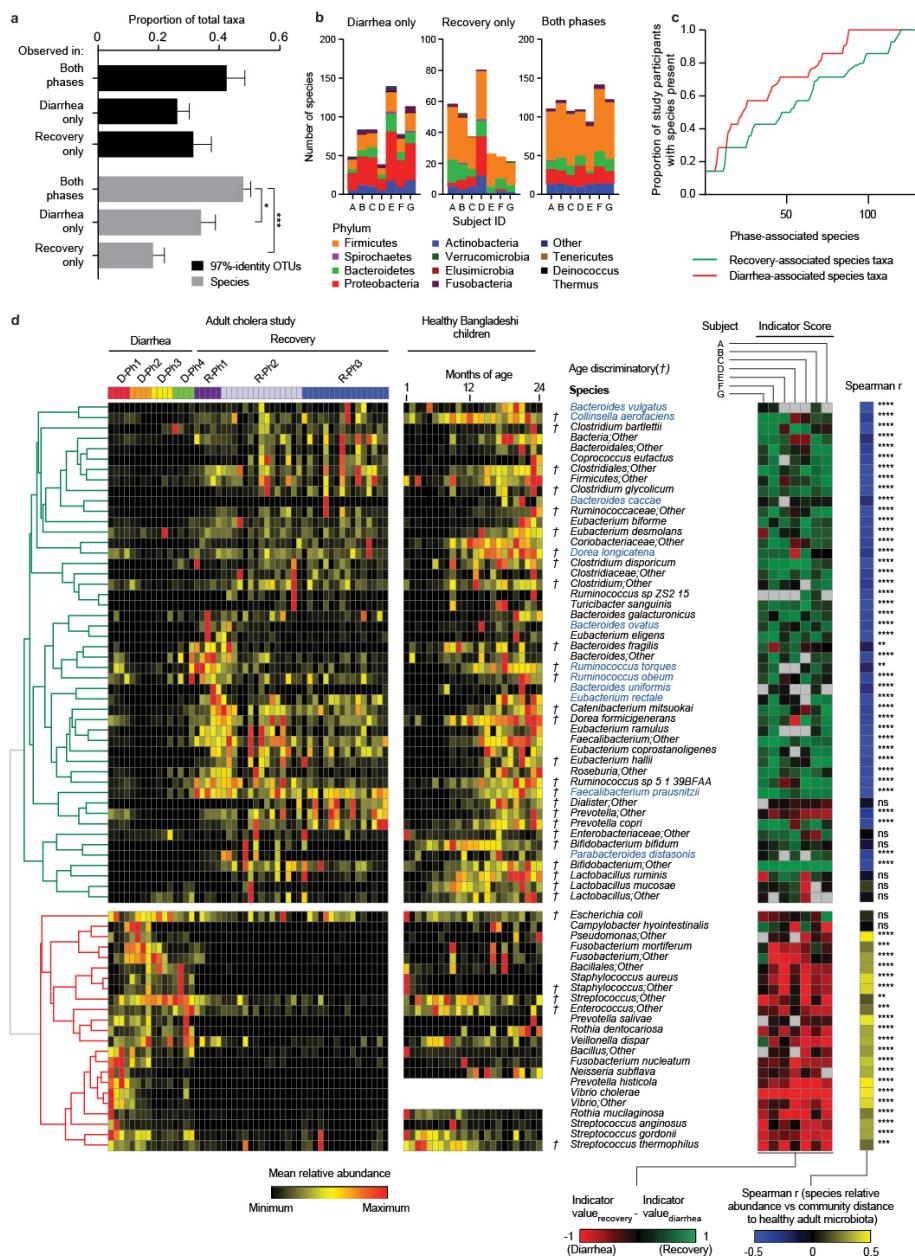
**UPLC–MS.** Procedures for UPLC–MS of bile acids have been described in ref. 37.

28. Faiz, M. A. & Basher, A. Antimicrobial resistance: Bangladesh experience. *Reg. Health Forum* **15**, 1–18 (2011).
29. Morgan, D. J., Okeke, I. N., Laxminarayan, R., Perencevich, E. N. & Weisberg, S. Non-prescription antimicrobial use worldwide: a systematic review. *Lancet Infect Dis.* **11**, 692–701 (2011).
30. Yatsunenko, T. et al. Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012).
31. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336 (2010).
32. Cole, J. R. et al. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Res.* **35**, D169–D172 (2007).
33. DeSantis, T. Z. et al. Greengenes, a chimeric-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
34. Rodrigue, S. et al. Unlocking short read sequencing for metagenomics. *PLoS ONE* **5**, e11840 (2010).
35. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
36. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
37. Ridaura, V. K. et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* **341**, 1241214 (2013).
38. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
39. Kristiansson, E., Hugenholtz, P. & Dalevi, D. ShotgunFunctionalizeR: an R-package for functional comparison of metagenomes. *Bioinformatics* **25**, 2737–2738 (2009).
40. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).



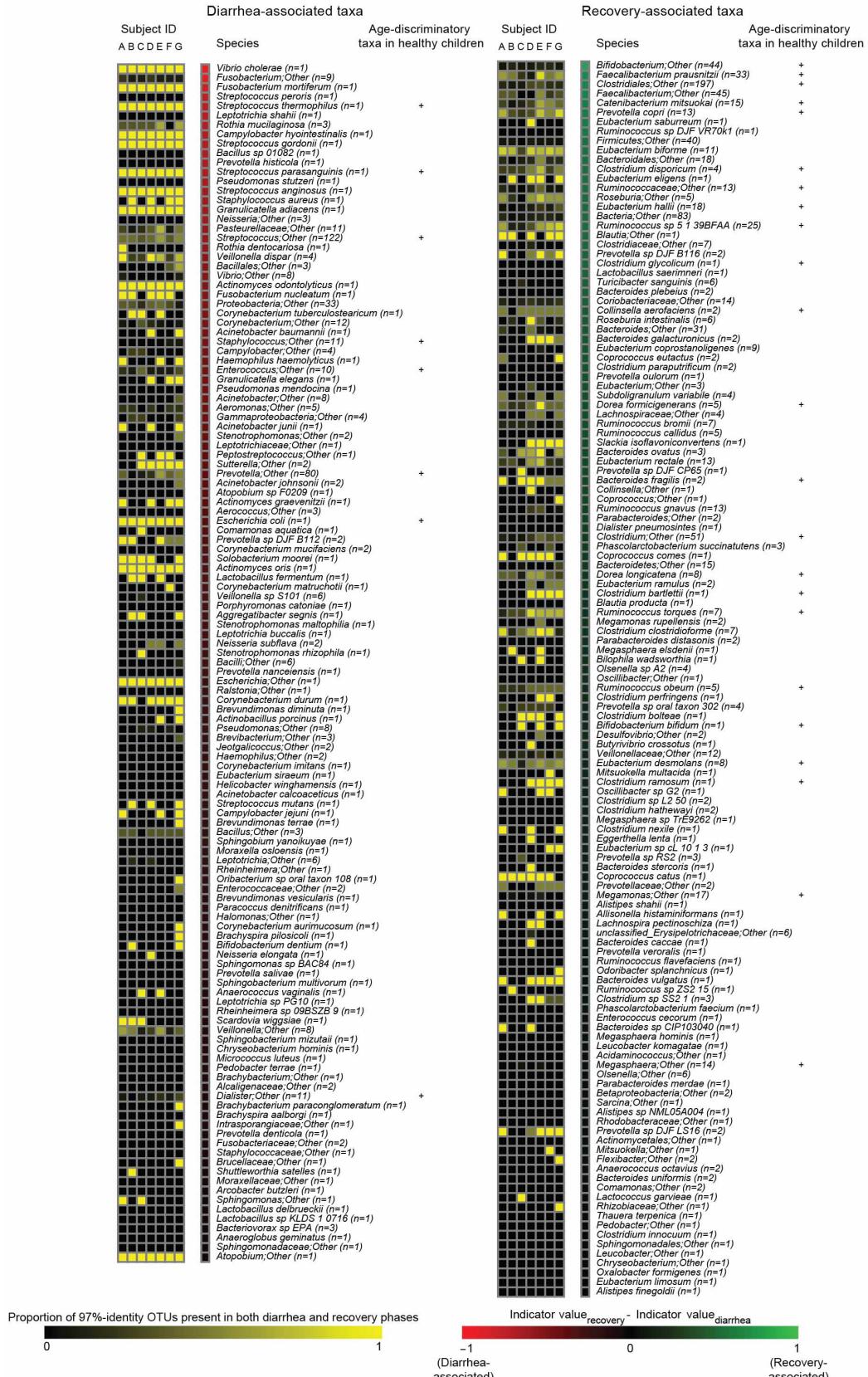
**Extended Data Figure 1 | Experimental designs for clinical study and gnotobiotic mouse experiments.** **a**, Sampling schedule for human cholera study. **b**, Frequency of diarrhoeal episodes over time for a representative participant (patient A). Initial time (black circle) represents beginning of

diarrhoea. The long vertical line marks enrollment into the study. Colours and short vertical lines denote boundaries of study phases defined in **a**. **c–e**, Gnotobiotic mouse experimental design. The number (*n*) of animals in each treatment group is shown.



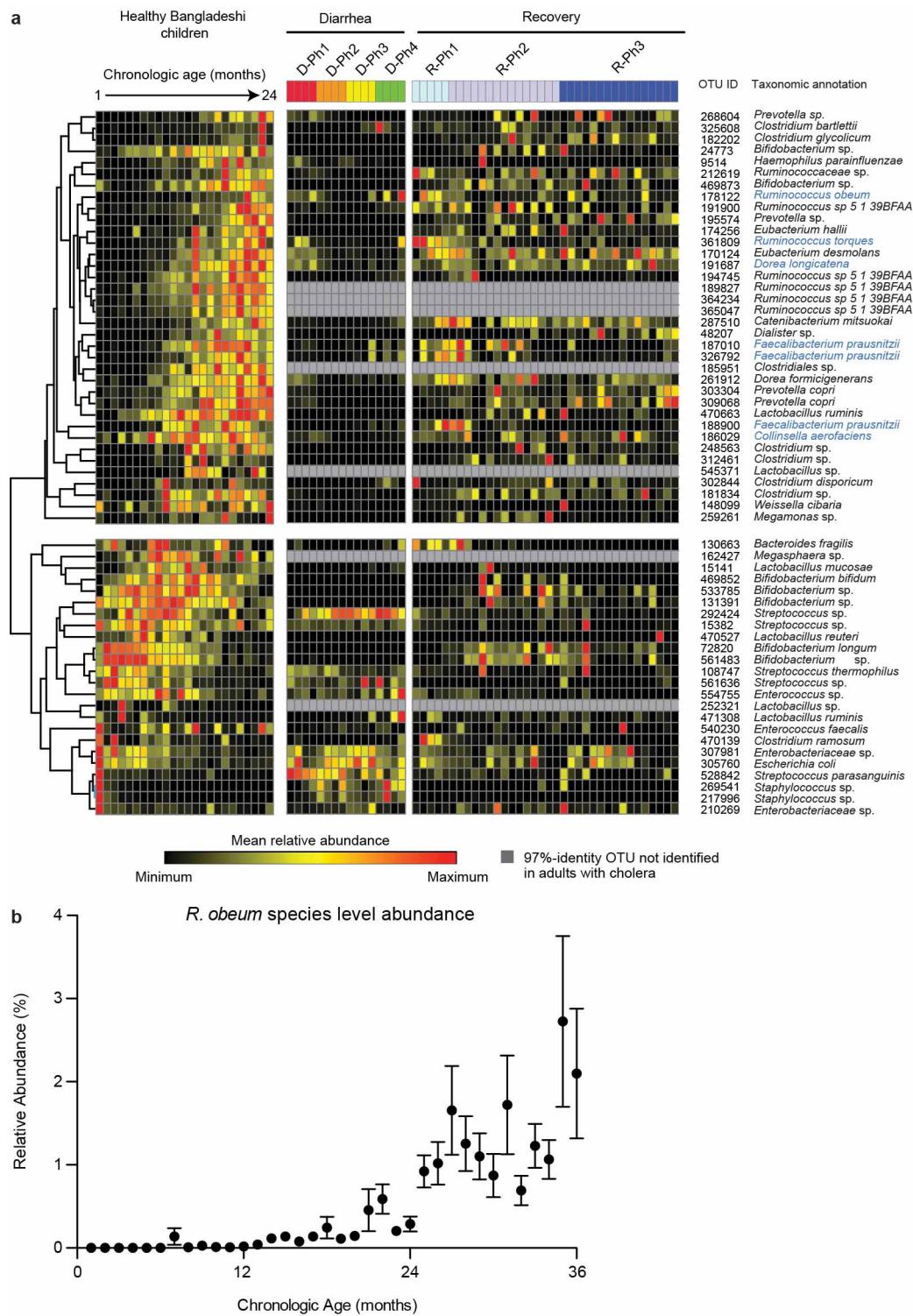
**Extended Data Figure 2 | Bacterial taxa associated with diarrhoeal and recovery phase.** **a**, Proportion of bacterial species-level taxa that were observed in both diarrhoeal and recovery phases, in D-Ph1 to D-Ph4 only, and in R-Ph1 to R-Ph3 only. Mean values  $\pm$  s.e.m. are plotted.  $*P < 0.05$ ,  $***P < 0.001$  (unpaired Mann-Whitney  $U$ -test). **b**, Phylum-level analysis. Mean values are plotted. **c**, Proportion of study participants having bacterial taxa associated by indicator species analysis with the diarrhoeal or recovery phase. The  $x$  axis shows species associated with each phase, ranked by proportion of subjects harbouring that species. For each species, ‘representation in study participants’ is the average presence/absence of all 97%-identity OTUs with that species taxonomic assignment. The OTU table was rarefied to 49,000 reads per sample. **d**, Bacterial species identified by indicator analysis as indicative of diarrhoea or recovery phases in adult patients with cholera, and species identified by Random Forests analysis as discriminatory for different stages in the maturation of the gut microbiota of healthy Bangladeshi infants/children aged 1–24 months (denoted by the symbol †). The heat map in the left-hand portion of the panel shows mean relative abundances of species across all individuals during D-Ph1 to D-Ph4, with each phase subdivided into four equal time bins. For recovery time points, columns represent the mean relative abundances for each sampling time point during R-Ph1 to R-Ph3. Mean relative abundance values are also presented for these same species in the faecal microbiota of 50 healthy Bangladeshi children sampled from 1 to 2 years of age at monthly

intervals. Unsupervised hierarchical clustering used relative abundances of species in the faecal microbiota of the patients with cholera. The green portion of the tree encompasses species that are more abundant during recovery whereas the red portion encompasses species that are more abundant during diarrhoea. Indicator scores are presented in the right-hand portion of the panel, with ‘score’ for a given taxon defined as its indicator value for recovery minus its indicator value for diarrhoea ( $-1$ , highly diarrhoea-associated;  $+1$ , highly recovery-associated). Spearman’s rank correlation coefficients of mean relative abundances of species by sample in the cholera study versus the mean sample-weighted UniFrac distance to healthy adult faecal microbiota are shown at the extreme right together with the statistical significance of correlations after Benjamini-Hochberg false discovery rate correction for multiple hypothesis testing (NS, not significant;  $*P < 0.05$ ;  $**P < 0.01$ ;  $***P < 0.001$ ). Higher coefficients indicate increasing divergence from a healthy configuration with higher relative abundance of a given species. Species shown satisfied two or more of the following criteria: (1) presence among the list of the top 40 age-discriminatory species in the Random-Forests-based model of gut microbiota maturation in healthy infants and children; (2) indicator value score greater than  $0.7$ ; (3) significant correlation (Spearman’s  $r$ ) between relative abundance in the faecal microbiota of patients with cholera and UniFrac distance to healthy adult faecal microbiota; and (4) inclusion in the artificial 14-member human gut community (species name highlighted in blue).



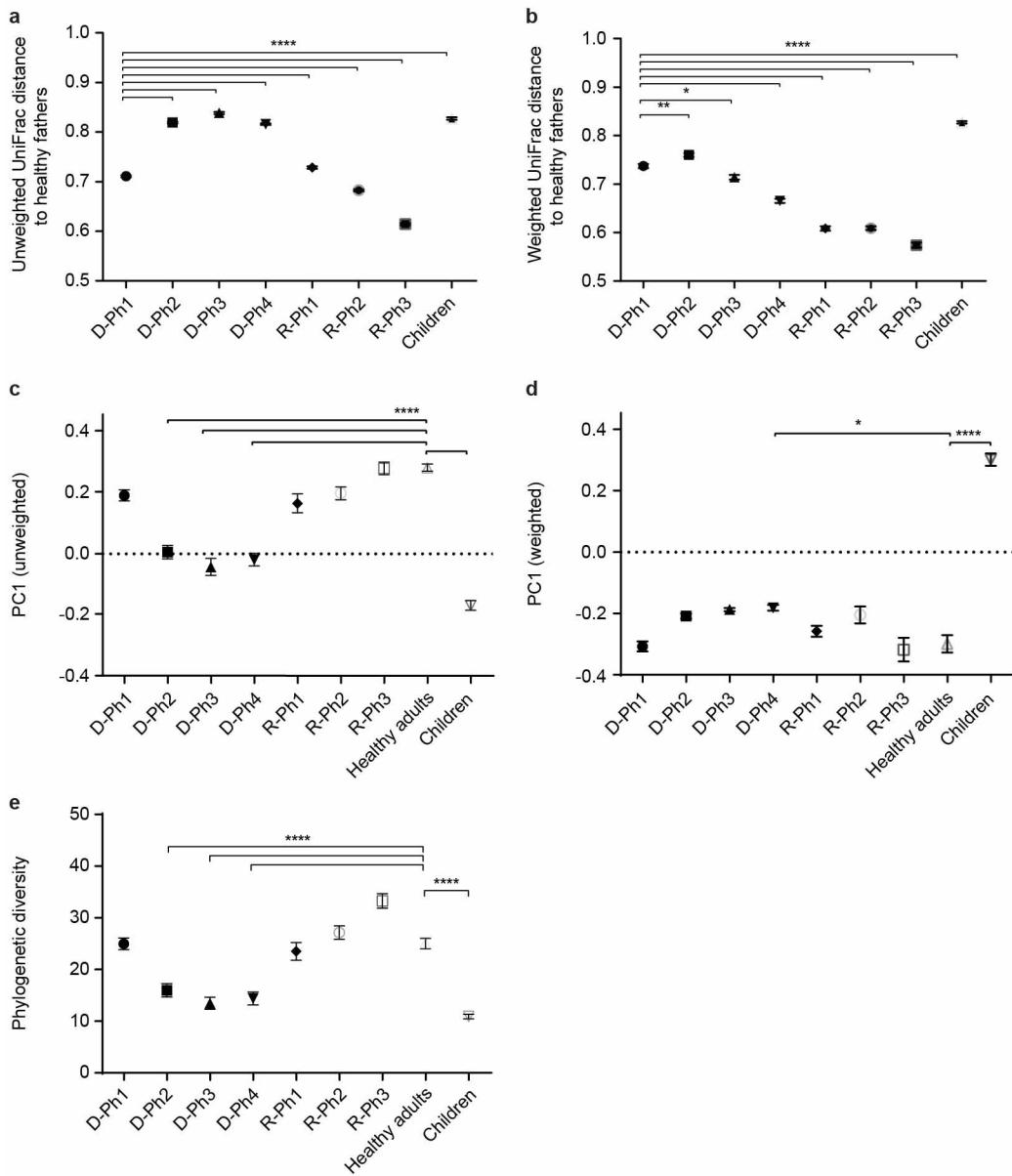
**Extended Data Figure 3 | The 97%-identity OTUs observed in both diarrhoeal and recovery phases.** The proportion of 97%-identity OTUs with a given species-level taxonomic assignment that were present in both diarrhoeal and recovery phases is shown for each individual in the study. The number of 97%-identity OTUs with a given species assignment is shown in parentheses. Species are ordered based on their ‘indicator scores’ (defined as indicator

value<sub>recovery</sub> minus indicator value<sub>diarrhoea</sub>). Age-discriminatory bacterial species incorporated into a Random-Forests-based model for defining relative microbiota maturity and microbiota-for-age *z*-scores<sup>3</sup> in healthy Bangladeshi infants and children are marked with a ‘+’ symbol. The 97%-identity OTUs were derived from data sets generated from all samples from adult patients with cholera; the OTU table was rarefied to 49,000 reads per sample.



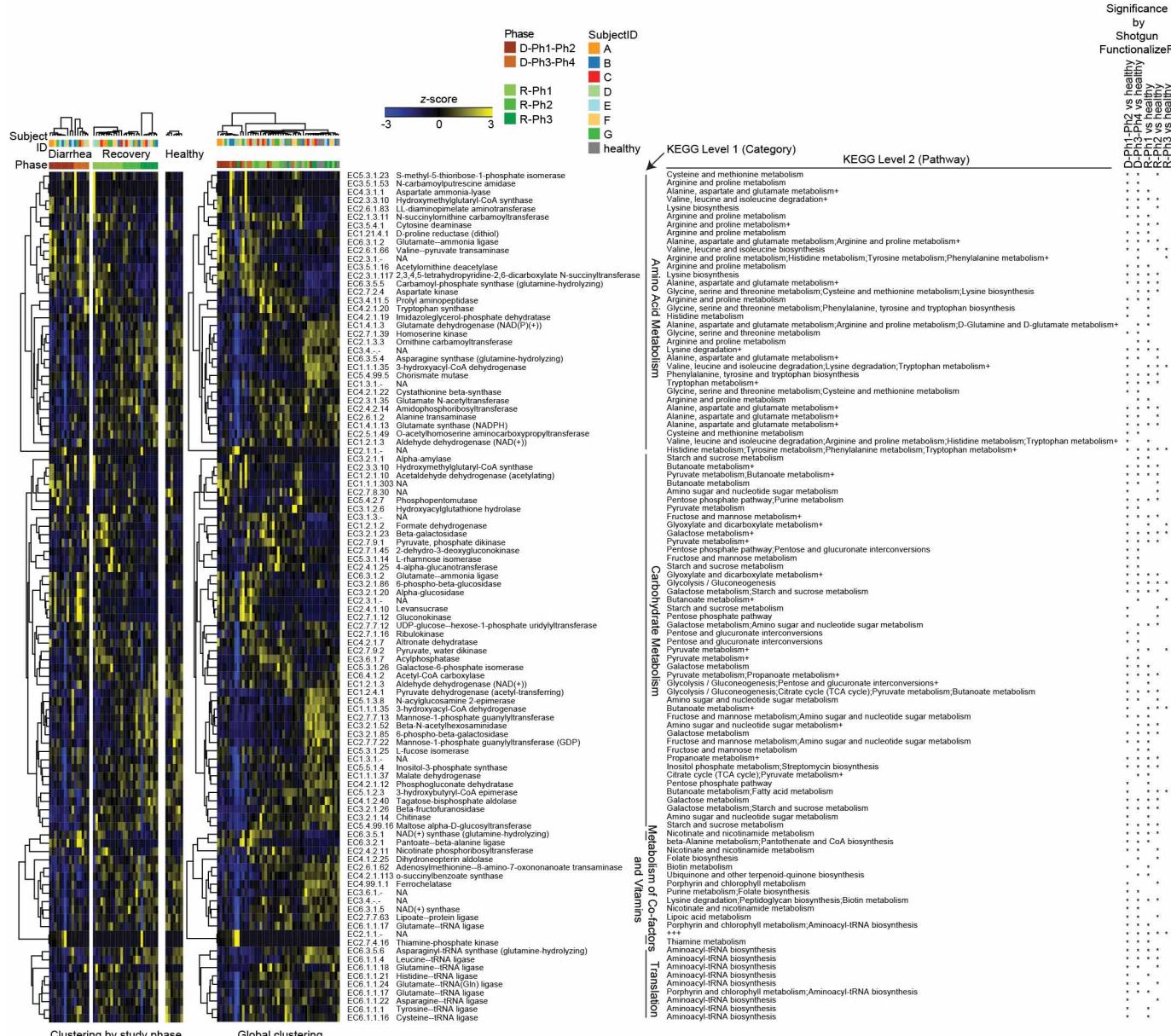
**Extended Data Figure 4 | Pattern of appearance of age-discriminatory 97%-identity OTUs in the faecal microbiota of patients with cholera mirrors the normal age-dependent pattern in the faecal microbiota of healthy Bangladeshi infants and children.** **a**, Left portion of the panel shows hierarchical clustering of relative abundance values for each of the top 60 most age-discriminatory 97%-identity OTUs in a Random-Forests-based model of normal maturation of the microbiota in healthy Bangladeshi infants/children (importance scores for the age-discriminatory taxa defined by Random Forests analysis are reported in ref. 3; these 60 97%-identity OTUs can be

grouped into 40 species-level taxa). Right portion of the panel presents the mean relative abundances of these OTUs in samples obtained from patients with cholera during D-Ph1 to D-Ph4, and R-Ph1 to R-Ph3. The 97%-identity OTUs corresponding to species included in the artificial community that was introduced into gnotobiotic mice are highlighted in blue. **b**, Relative abundance of *R. obaeum* strains in the faecal microbiota of healthy Bangladeshi children sampled monthly through the first 3 years of life. Mean values  $\pm$  s.e.m. are plotted.



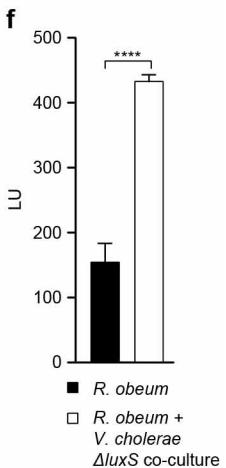
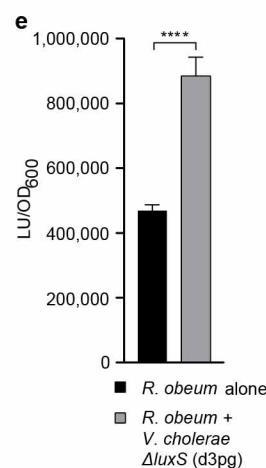
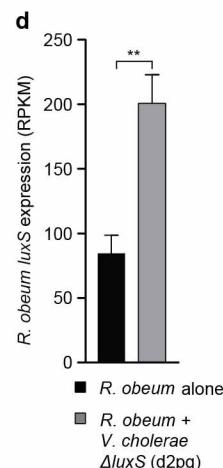
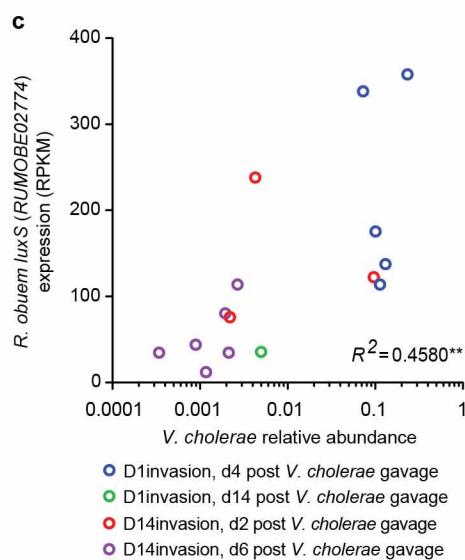
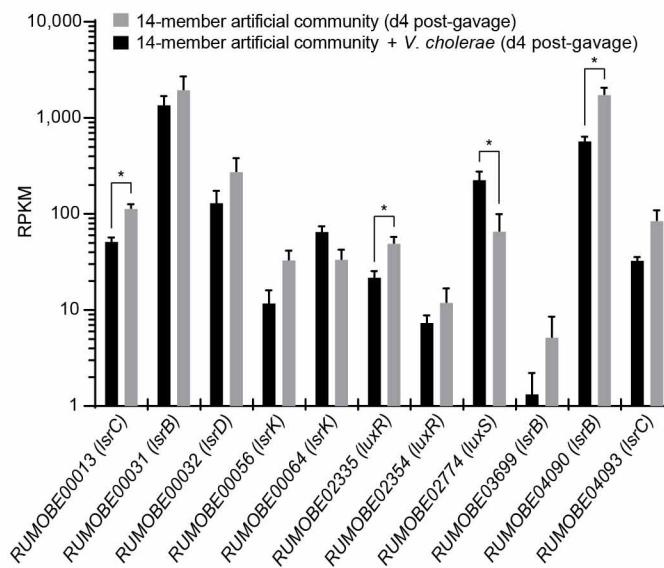
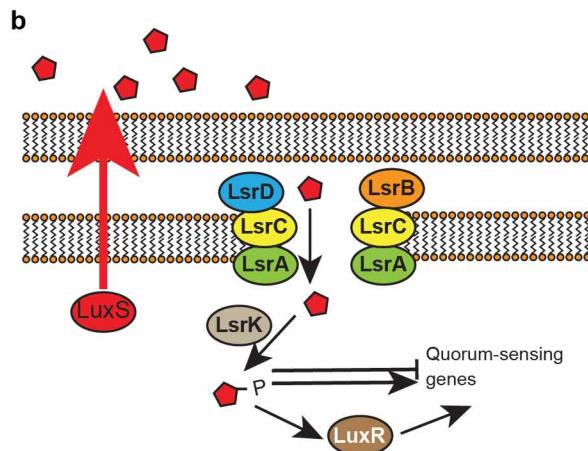
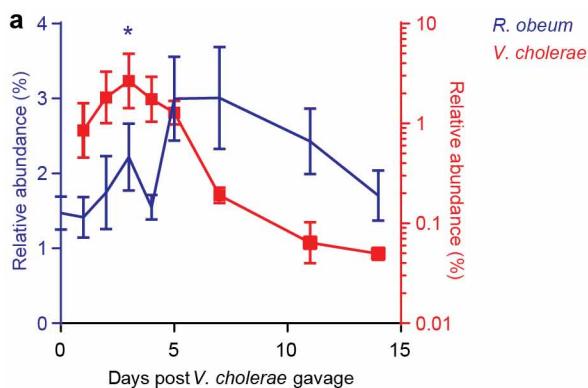
**Extended Data Figure 5 | Pattern of recovery of the gut microbiota in patients with cholera.** **a, b**, Mean unweighted (a) and weighted (b) UniFrac distances to healthy adult controls at each of the defined phases of diarrhoea and recovery. **c, d**, Principal coordinates analysis of UniFrac distances between gut microbiota samples. Location along the principal axis of variation (PC1) shows how acute diarrhoeal communities first resemble those of healthy Bangladeshi children sampled during the first 2 years of life, then evolve their

phylogenetic configurations during the recovery phase towards those of healthy Bangladeshi adults. PC1 accounts for 34.3% variation for weighted and 17.7% variation for unweighted UniFrac values. **e**, Alpha diversity (whole-tree phylogenetic diversity) measurements of faecal microbial communities through all study phases. Mean values  $\pm$  s.e.m. are plotted. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.0001$  (Kruskal–Wallis analysis of variance followed by multiple comparisons test).



**Extended Data Figure 6 | Proportional representation of genes encoding enzymes (classified according to Enzyme Commission number identifiers) in faecal microbiomes sampled during the diarrhoeal and recovery phases of cholera.** Shotgun sequencing of faecal community DNA was performed (MiSeq 2000 instrument; 2 × 250 bp paired-end reads; 341,701 ± 145,681 reads (mean ± s.d. per sample)). Read pairs were assembled (SHERA software package<sup>34</sup>). Read counts were collapsed based on their assignment to Enzyme Commission (EC) number identifiers. The significance of differences in EC abundances compared with faecal microbiomes in healthy adult Bangladeshi controls was defined using ShotgunFunctionalizeR<sup>39</sup>. Unsupervised hierarchical clustering identifies groups of ECs that characterize the faecal microbiomes of patients with cholera at varying diarrhoeal and recovery phases. The heatmap on the left shows the results of EC-based clustering by phase (diarrhoea/recovery). An asterisk on the extreme right of the figure indicates that differences in EC abundance observed across the specified study phases were statistically significant (adjusted  $P < 0.00001$ ,

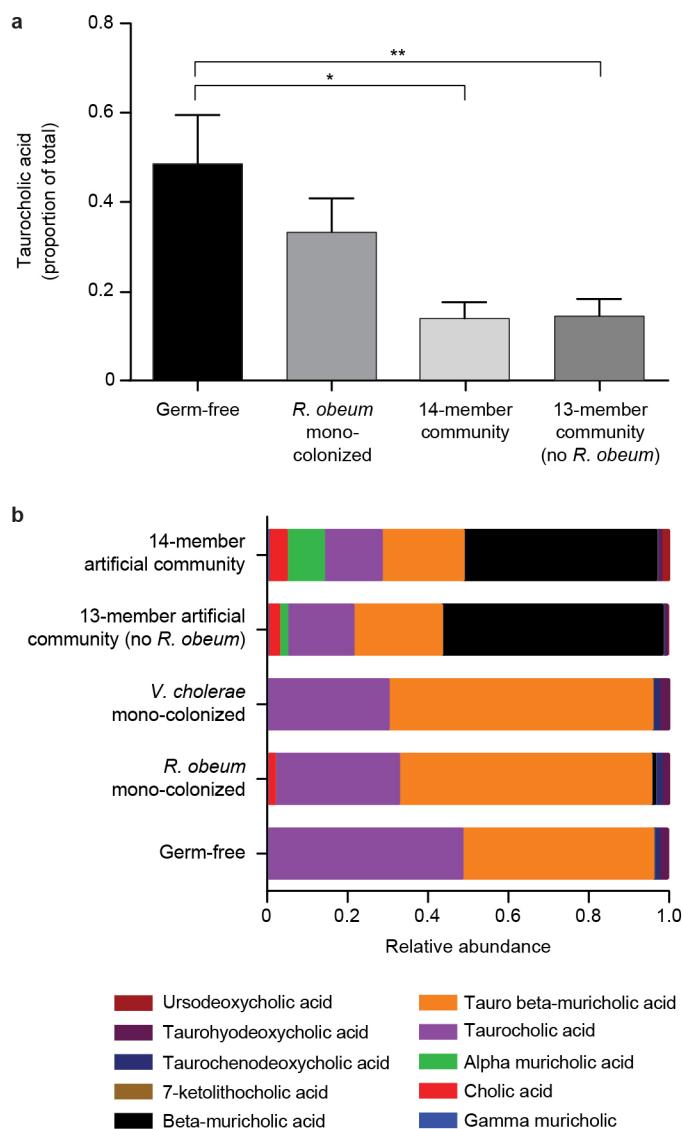
ShotgunFunctionalizeR). The heatmap on the right presents the results of a global clustering of all time-points and study phases. Genes encoding 102 ECs were identified with (1) at least 0.1% average relative abundance across the study and (2) significant differences in their representation relative to healthy microbiomes in at least one comparison (adjusted  $P < 0.00001$  based on ShotgunFunctionalizeR). In each of the heatmaps, z-scores for each EC across all samples are plotted. ECs are grouped by KEGG level 1 assignment and further annotated based on their KEGG Pathway assignments. A '+' symbol indicates that the EC has additional KEGG level 2 annotations (see Supplementary Table 8 for a list of all assignable functional annotations). Note that the majority of the 46 ECs that were more prominently represented in faecal microbiomes during diarrhoeal phases in study participants are related to carbohydrate metabolism. The faecal microbiomes of patients during recovery are enriched for genes involved in vitamin and cofactor metabolism (Supplementary Table 8).



**Extended Data Figure 7 | *R. obeum* encodes a functional AI-2 system, and *R. obeum* AI-2 production is stimulated by the presence of *V. cholerae*.**

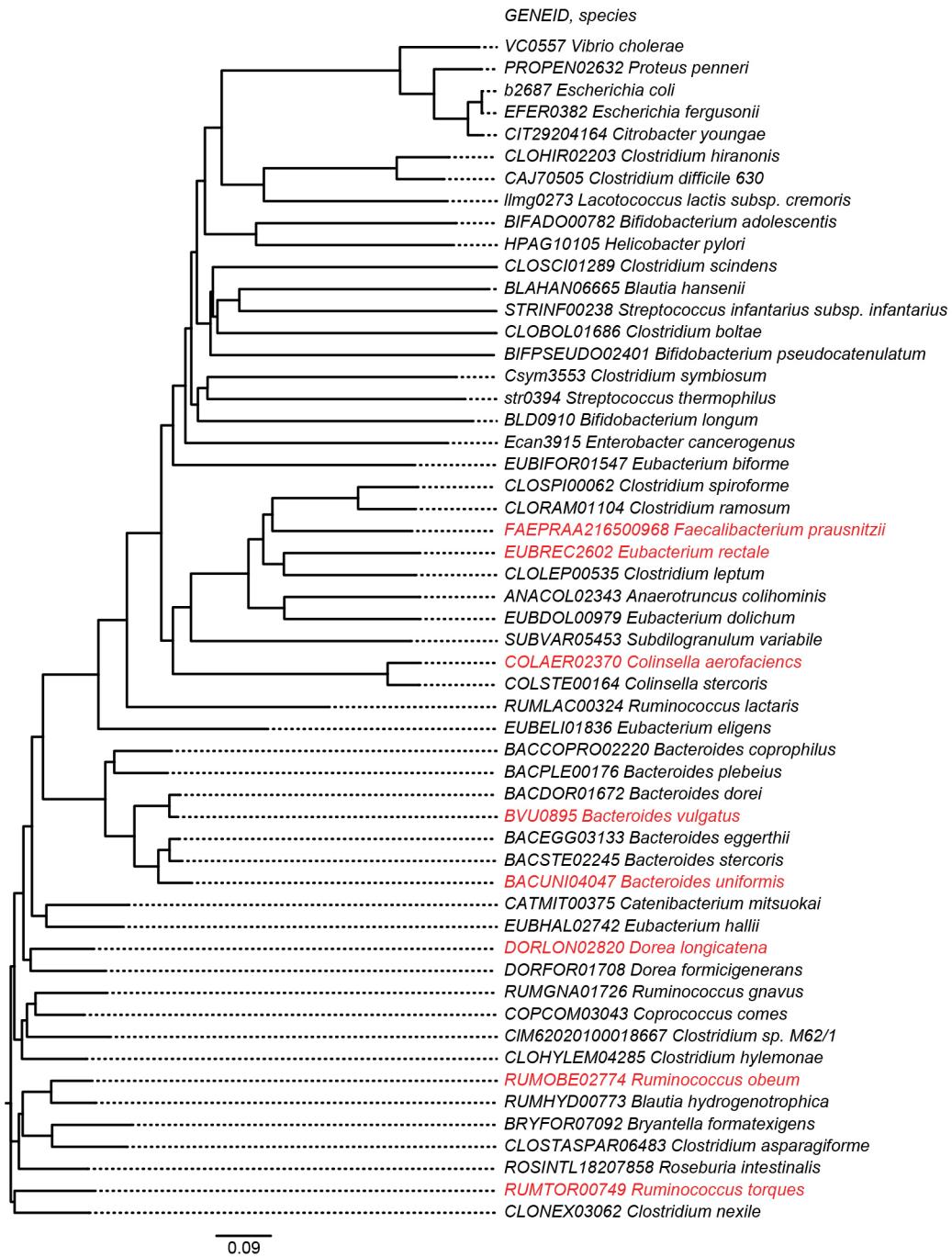
**a**, Relative abundances of *R. obeum* and *V. cholerae* in the faecal microbiota after introduction of *V. cholerae* into mice harbouring the artificial 14-member human gut community (D14invasion group, see Extended Data Figure 1c). ‘Days post *V. cholerae* gavage’ refers to the second of two daily gavages of  $10^9$  c.f.u. *V. cholerae* into animals that had been colonized 14 days earlier with the 14-member community. Mean values  $\pm$  s.e.m. are shown ( $n = 4$  or 5 mice,  $*P < 0.05$ , unpaired Student’s *t*-test). **b**, Left panel shows AI-2 signalling pathway components represented in the *R. obeum* genome. Right panel plots changes in expression of these components as defined by microbial RNA-seq of faecal samples obtained (1) 4 days after colonization of mice with the 14-member community and (2) 4 days after gavage of mice with the 14-member community together with  $10^9$  c.f.u. of *V. cholerae* ( $n = 4$ –6 animals per group; one faecal sample analysed per animal). Mean values  $\pm$  s.e.m. are shown.  $*P < 0.05$  (Mann–Whitney *U*-test). **c**, RNA-seq of faecal samples collected at the time points and treatment groups indicated reveals that *R. obeum luxS* transcription is directly correlated to *V. cholerae* abundance in the context of the 14-member community.  $**P < 0.01$  (*F* test). **d**, *R. obeum luxS* expression. Mice were colonized first with *R. obeum* for 7 day. Faecal samples were collected for microbial RNA-seq analysis 1 day before gavage of  $10^9$  c.f.u. of a *V. cholerae*

$\Delta luxS$  mutant, and then 2 days post-gavage (d2pg). Mean values for relative *R. obeum luxS* transcript levels ( $\pm$  s.e.m.) are shown ( $n = 5$  or 6 animals per group per experiment,  $n = 3$  independent experiments;  $**P < 0.01$  unpaired Mann–Whitney *U*-test). **e**, AI-2 levels in faecal samples, taken 1 day before and 3 days after gavage of the *V. cholerae*  $\Delta luxS$  strain, from the same mice as those analysed in **a**. AI-2 levels were measured based on induction of bioluminescence in *V. harveyi* BB170 using the same mass of input faecal sample for all assays. Mean values  $\pm$  s.e.m. are shown;  $****P < 0.0001$  (unpaired Mann–Whitney *U*-test). **f**, *R. obeum* produces AI-2 when co-cultured with *V. cholerae* *in vitro*. Aliquots of the supernatant from cultures containing *R. obeum* alone, or *R. obeum* plus the *V. cholerae*  $\Delta luxS$  mutant, were assayed for their ability to induce *V. harveyi* bioluminescence. Mean values  $\pm$  s.e.m. are presented ( $n = 4$  independent experiments). LU, light units; RPKM, reads per kilobase per million reads.  $****P < 0.0001$  (unpaired Mann–Whitney *U*-test). Note that (1) the number of *R. obeum* c.f.u. present in the samples obtained from mono-cultures of the organism was similar to the number in co-culture, as measured by selective plating, and (2) the *V. cholerae*  $\Delta luxS$  mutant cultured alone produced levels of AI-2 signal that were not significantly different from that of *R. obeum* in mono-culture (data not shown).



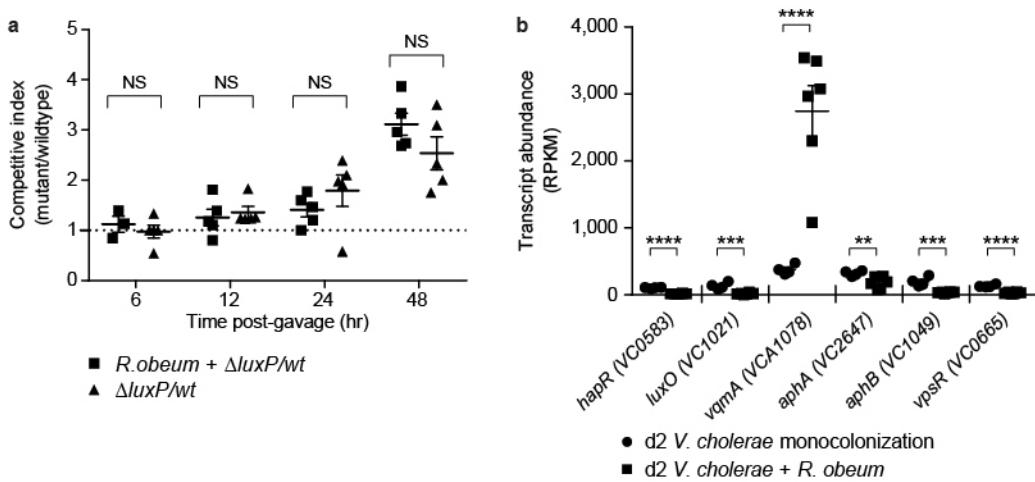
**Extended Data Figure 8 | UPLC-MS analysis of faecal bile acid profiles in gnotobiotic mice.** Targeted UPLC-MS used methanol extracts of faecal pellets obtained from age- and gender-matched germ-free C57BL/6J mice and gnotobiotic mice colonized for 3 days with *R. obeum* alone, for 7 days with the 14-member community ('D1invasion group'), and for 3 days with the

13-member community that lacked *R. obeum* ( $n = 4\text{--}6$  mice per treatment group; one faecal sample analysed per animal). **a**, Faecal levels of taurocholic acid. Mean values  $\pm$  s.e.m. are plotted.  $*P < 0.05$ ,  $**P < 0.01$ , Mann–Whitney *U*-test. **b**, Mean relative abundance of ten bile acid species in faecal samples obtained from the mice shown in **a**.



**Extended Data Figure 9 | Phylogenetic tree of luxS genes present in human gut bacterial symbionts and enteropathogens.** The tree was constructed from amino-acid sequence alignments using Clustal X. Red type indicates that the

homologue is represented in the genomes of members of the 14-member artificial human gut bacterial community.



**Extended Data Figure 10 | In vivo tests of the effects of known quorum-sensing components on *R. obeum*-mediated reductions in *V. cholerae* colonization.** **a**, Competitive index of  $\Delta luxP$  versus wild-type C6706 *V. cholerae* when colonized with or without *R. obeum* ( $n = 4$ –6 animals per group). Horizontal bars, mean values. Data from individual animals are shown using the indicated symbols. **b**, Transcript abundance (reads per kilobase per million reads) for selected quorum-sensing and virulence gene regulators in

*V. cholerae*. Microbial RNA-seq was performed on faecal samples collected 2 days after mono-colonization of germ-free mice with *V. cholerae* (circles), or 2 days after *V. cholerae* was introduced into mice that had been mono-colonized for 7 days with *R. obeum* (squares) ( $n = 5$  animals per group; NS, not significant ( $P \geq 0.05$ ); \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  (unpaired two-tailed Student's *t*-test)).