The Journal of Infectious Diseases

MAJOR ARTICLE

Human Gut Microbiota Predicts Susceptibility to *Vibrio cholerae* Infection

Firas S. Midani,[1,2,3,a] Ana A. Weil,[7,8,a] Fahima Chowdhury,[12] Yasmin A. Begum,[12] Ashraful I. Khan,[12] Meti D. Debela,[7] Heather K. Durand,[3] Aspen T. Reese,[3,4] Sai N. Nimmagadda,[5] Justin D. Silverman,[1,2,3,6] Crystal N. Ellis,[7,13] Edward T. Ryan,[7,8,11] Stephen B. Calderwood,[7,8,9] Jason B. Harris,[7,10] Firdausi Qadri,[12] Lawrence A. David,[1,2,3,5,b] and Regina C. LaRocque[7,8,b]

[1]Program in Computational Biology and Bioinformatics, [2]Center for Genomic and Computational Biology, [3]Department of Molecular Genetics and Microbiology, [4]Department of Biology, [5]Department of Biomedical Engineering, and [6]Medical Scientist Training Program, Duke University, Durham, North Carolina; [7]Division of Infectious Diseases, Massachusetts General Hospital, [8]Department of Medicine, [9]Department of Microbiology and Immunobiology, and [10]Department of Pediatrics, Harvard Medical School, and [11]Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts; [12]Mucosal Immunology and Vaccinology Laboratory, Infectious Diseases Division, International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka; and [13]Massachusetts College of Pharmacy and Health Sciences University, Boston

Background. Cholera is a public health problem worldwide, and the risk factors for infection are only partially understood.

Methods. We prospectively studied household contacts of patients with cholera to compare those who were infected to those who were not. We constructed predictive machine learning models of susceptibility, using baseline gut microbiota data. We identified bacterial taxa associated with susceptibility to *Vibrio cholerae* infection and tested these taxa for interactions with *V. cholerae* in vitro.

Results. We found that machine learning models based on gut microbiota, as well as models based on known clinical and epidemiological risk factors, predicted *V. cholerae* infection. A predictive gut microbiota of roughly 100 bacterial taxa discriminated between contacts who developed infection and those who did not. Susceptibility to cholera was associated with depleted levels of microbes from the phylum Bacteroidetes. By contrast, a microbe associated with cholera by our modeling framework, *Paracoccus aminovorans*, promoted the in vitro growth of *V. cholerae*. Gut microbiota structure, clinical outcome, and age were also linked.

Conclusion. These findings support the hypothesis that abnormal gut microbial communities are a host factor related to *V. cholerae* susceptibility.

Keywords. *Vibrio cholerae*; microbiome; machine learning.

Vibrio cholerae causes millions of cases of acute watery diarrhea every year, and our understanding of susceptibility to the disease remains incomplete [1]. Cholera occurs in both endemic and epidemic patterns. In both instances, multiple symptomatic and asymptomatic *V. cholerae* infections within a household are common [2, 3]. Transmission events within a household may occur through shared sources of contaminated food and water or through fecal-oral spread [4, 5]. Observational studies have identified host factors that correlate with susceptibility to *V. cholerae* infection, including young age, blood group O status, variants in genes of the innate immune system, and lack of preexisting immunity [2, 3, 6–8]. Nevertheless, these risk factors only partially explain the variation in clinical outcomes seen following exposure to *V. cholerae* [3, 6].

Human gut–associated bacterial communities (microbiota) may be another risk factor for enteric infections, including cholera. A case-control study of children in sub-Saharan Africa and southern Asia suggested that select bacterial taxa naturally occurring in the gut are protective against *Shigella*-induced diarrhea, and predeparture microbiota sampled in travelers showed that a specific microbial profile is associated with risk of *Campylobacter* infection [9, 10]. Animal studies have identified potential mechanisms by which commensal microbes might resist invading pathogens, including competition for nutrients or sites of adherence [11, 12] and stimulation of host epithelial cell defense [13, 14]. Cocolonization of gnotobiotic mice with *V. cholerae* and *Blautia obeum* (formerly named "*Ruminococcus obeum*") [15], a bacterial taxon enriched in human gut communities recovering from cholera, disrupts *Vibrio* virulence signaling pathways [16]. Further research on the relationship between human gut bacterial communities and *V. cholerae* is likely to advance our understanding of cholera pathogenesis and may lead to the development of novel interventions for disease prevention.

To test the role of the gut microbiota in susceptibility to *V. cholerae* infection, we prospectively evaluated household contacts of patients with cholera in Dhaka, Bangladesh. This cohort was uniquely designed to enable sampling of individuals

Received 25 September 2017; editorial decision 26 March 2018; accepted 10 April 2018; published online April 12, 2018.

Presented in part: IDWeek 2015, Philadelphia, Pennsylvania, 5 October 2015. Abstract #49899; American Society for Microbiology, Boston, Massachusetts, 20 June 2016. Abstract #611; International Symposium on Microbial Ecology, Montreal, Canada, 25 August 2016. Abstract #218B; Keystone Conference Microbiome in Health and Disease, Keystone, Colorado, 9 February 2017.

[a]F. S. M. and A. A. W. are joint first authors.

[b]L. A. D. and R. C. L. are joint last authors.

Correspondence: L. A. David, PhD, Department of Molecular Genetics and Microbiology, Duke University, Box 3382, 101 Science Dr, Durham, NC 27708 (lawrence.david@duke.edu).

The Journal of Infectious Diseases® 2018;218:645–53

DOI: 10.1093/infdis/jiy192

prior to infection, and we have previously used it to define clinical and epidemiological risk factors for cholera [2, 3, 6, 17]. Here, we report that the baseline microbiota of household contacts can be used to predict susceptibility to *V. cholerae* infection at least as well as previously identified host risk factors. We also demonstrate an association between gut microbiota structure, contact age, and susceptibility. Last, we describe an experimentally validated interaction between *Paracoccus aminovorans* and *V. cholerae* that was predicted by our modeling framework. Our results illustrate a role for gut microbiota in predicting *V. cholerae* susceptibility in humans.

MATERIALS AND METHODS

Sample Collection, Classification of Outcomes, and 16S rRNA Analysis

We enrolled households with an index cholera case hospitalized at the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b). We included index cases with a stool culture positive for *V. cholerae* O1 as the sole pathogen. We enrolled household contacts within 6 hours of presentation of the index case; households were located in informal urban settlements with limited access to clean water and sanitation. We excluded individuals aged <2 or >60 years, those who resided outside Dhaka, and those with comorbid conditions. We collected demographic information from and characterized the blood group of all contacts, obtained rectal swab specimens for *V. cholerae* culture and 16S sequencing, measured vibriocidal titers, and took symptom histories over 30 days of follow-up (Figure 1). Contacts were considered infected if any rectal swab culture during the observation period was positive for *V. cholerae* or if they experienced diarrhea (defined as ≥3 loose stools over 24 hours) and developed a 4-fold increase in vibriocidal titer [2, 3]. We profiled the gut microbiota from rectal swab samples by extracting DNA, amplifying for the V4 region of the 16S ribosomal RNA, and sequencing the gene library by using Illumina MiSeq. Sequences were filtered for quality, reassigned to the sample of origin, clustered into operational taxonomic units (OTUs) at 97% similarity by using UCLUST via QIIME, and then assigned a taxonomy by mapping against Greengenes 16S reference database [18]. Further information on cohort enrollment, computational scripts, and microbiological and serological assays are described in the Supplementary Methods. Nucleotide sequences are available from the European Nucleotide Archive under accession number PRJEB17860. The study was approved by the Ethical Review Committee of the icddr,b and the institutional review boards of Massachusetts General Hospital and the Duke University Health System. Participants or their guardians provided written informed consent.

Predictive Statistical Modeling of Outcomes

We used multivariate logistic regression to evaluate the clinical and epidemiologic host factors known to influence susceptibility to cholera (Stata, College Station, TX). Models were constructed using scikit-learn 0.17 in Python. In the hold-out model, we partitioned the data into a training set of 48 samples and a testing set of 28 samples. We used a support vector machine (SVM) model that learned patterns of relative abundance of OTUs and distinguished infected from uninfected contacts. We coupled the SVM to a recursive feature elimination (RFE) algorithm, which simplifies models and increases accuracy by removing uninformative bacterial taxa [19]. For the

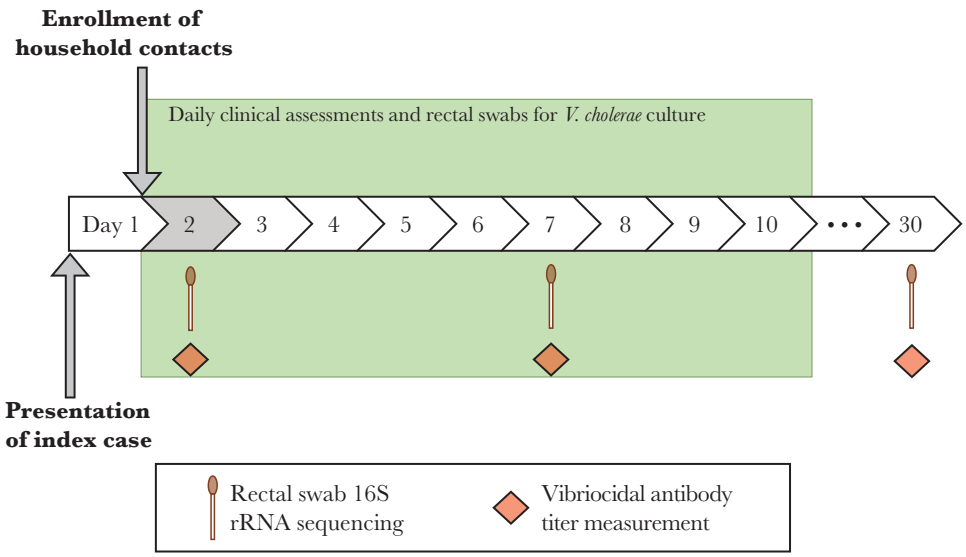

Figure 1. Study design for prospective observation of household contacts of patients with cholera. Index cases were hospitalized on day 1 with symptomatic *Vibrio cholerae* O1 infection, and their household contacts were enrolled on day 2. Daily clinical assessments and rectal swab sampling of household contacts for *V. cholerae* culture were conducted during the observation period on days 2–10 and day 30. On the day of enrollment, day 7, and day 30, vibriocidal antibody titers were measured, and rectal swab specimens for 16S ribosomal RNA (rRNA) gene sequencing were obtained.

combined model, an SVM was applied to age, blood group O status, vibriocidal antibody titer, and the OTUs selected by the microbiota-based model. We followed an identical procedure in the cross-validation model, using 30 replicates of 10-fold cross-validation.

Spent Supernatant Culture Experiments

Spent culture supernatant (SCS) of *P. aminovorans, Vibrio harveyi*, and *V. cholerae* was obtained by culturing in Luria-Bertani broth to an optical density of 1.5 and then removing cells by centrifugation and filter sterilization (pore diameter, 0.22 μm; Millipore). A total of 50 μL of overnight *V. cholerae* culture was inoculated into 4 mL of *P. aminovorans* SCS, *V. harveyi* SCS, *V. cholerae* SCS, and fresh broth and cultured overnight. Growth of *V. cholerae* was measured using spectrophotometry and colony counts. All experiments were conducted in replicates of 5 and replicated in brain heart infusion medium and minimal medium. *Blautia obeum, Prevotella buccalis, Bacteroides ovatus*, and *Bacteroides uniformis* were grown using an anaerobic environment (5% H_2 and 10% CO_2 balanced with N_2) in a vinyl anaerobic chamber (Coy Products) at 37°C. SCS from the stationary phase of growth in Gifu broth (Gibco) was used to conduct identical *V. cholerae* growth experiments as described above, using anaerobically cultured *V. cholerae* and fresh broth as controls. Further information on bacterial strains, controls, media, and experimental validation are listed in the Supplementary Methods. GraphPad Prism 7 (GraphPad Software) was used for analysis.

RESULTS

V. cholerae Infection in Household Contacts of Patients With Cholera

We prospectively evaluated 124 contacts in 66 households that included an index patient with cholera. Contacts were enrolled in 2 temporal cohorts (February 2012–December 2012 and February 2013–May 2014). No cholera cases were enrolled in January 2013, during the season of low cholera incidence. We excluded 48 contacts with missing data; with microbiologic, genomic, or clinical evidence of *V. cholerae* infection upon enrollment; or with recent antibiotic use (Supplementary Methods).

A total of 76 contacts formed our cohort for predicting cholera susceptibility. Of these, 22 (29%) developed *V. cholerae* infection during the follow-up period, and 54 (71%) were uninfected (Table 1). The average age of contacts was 26 years (range, 4–60 years), and 63% (48 of 76) were women. Additional demographic characteristics are listed in Supplementary Table 1. Of the 22 infected contacts, 10 were symptomatic, and 12 were asymptomatic. Symptomatic cholera led to significant changes in gut microbial community structure that persisted through the 30-day follow-up period, as we have previously observed (Supplementary Figure 1) [18]. However, beta-diversity patterns indicate that OTUs that were lost or gained in the course of infection were compensated for by shifts in the abundance of related taxa. These changes suggest modest resilience of the human gut microbiota in response to cholera-associated diarrhea.

We created a multivariate logistic regression model, using the known clinical and epidemiological risk factors for cholera in our cohort, to serve as a comparator for the microbial predictive models (Table 1). This logistic regression model did not yield statistically significant results, likely owing to the small sample size; however, the trends that we identified are consistent with findings from several larger cohorts [3, 6], which have identified younger age, lower vibriocidal titer, and blood group O status as predictors of cholera among household contacts.

Development of Machine Learning Model

We first tested whether microbiota-based models established solely on gut community composition or specific OTUs could predict cholera susceptibility. Univariate statistical testing showed that none of 4181 unique OTUs in the cohort was individually associated with susceptibility (based on findings of a false-discovery rate [FDR]–corrected 2-sided Mann-Whitney *U* test). *V. cholerae* infection status was not associated with differences in alpha diversity, as measured by species richness, Shannon index, or evenness across time or infection status ($P > .05$, by a 2-sided Mann-Whitney *U* test), at baseline and over the 30-day follow up period. Alpha-diversity metrics also did not differ between contacts with symptomatic infection and those with asymptomatic infection. Principal coordinate analysis and permutational multivariate analysis of variance (ANOVA) of gut microbiota,

Table 1. Multivariate Logistic Regression Model of Clinical and Epidemiologic Risk Factors for *Vibrio cholerae* Infection

Characteristic	Infected (n = 22)	Uninfected (n = 54)	Univariate OR (95% CI)	*P*	Multivariate Adjusted OR (95% CI)	*P*
Age ≤10 y	7 (32)	6 (11)	3.7 (1.1–13)	.04	3.4 (.88–13)	.08
Blood group O	8 (36)	9 (17)	2.9 (.93–8.8)	.07	3.2 (.96–11)	.06
Baseline vibriocidal titer, geometric mean,[a] (95% CI)	2.38 (1.95–2.9)	3.04 (2.68–3.44)	0.68 (.46–1.0)	.06	0.7 (.47–1.1)	.09
Malnourishment[b]	4 (18)	15 (28)	0.58 (.17–2.0)	.37	…	

Data are no. (%) of participants, unless otherwise indicated. The multivariate logistic regression model was created using generalized estimating equations, with *P* values adjusted for clustering based on household. The final model was based on forward selection with a predetermined cutoff *P* value of < .2 for inclusion. Malnourishment did not meet the predetermined criteria for inclusion in the model.

[a]Vibriocidal titers are log transformed and matched to the serotype of the household cholera case. The OR represents the risk of *V. cholerae* infection per doubling of vibriocidal titer.

[b]Malnourishment is defined per World Health Organization anthropometric thresholds (see Supplementary Methods).

performed using all 4181 OTUs, showed no qualitative distinction in community structure between infected contacts and uninfected contacts. These findings indicate that simple microbiota-based models were unable to predict cholera susceptibility.

In contrast to our findings with simple microbiota-based models, we found that a machine learning model successfully predicted cholera susceptibility when applied to the microbiota data (Figure 2A). We used a hold-out approach that took advantage of the recruitment of study households during 2 distinct periods. The model was trained using contacts recruited during the first period (36 uninfected contacts and 12 infected contacts, recruited during February 2012–December 2012). We then tested the model on the second set of contacts recruited later, using the same protocol (18 uninfected contacts and 10 infected contacts, recruited during February 2013–May 2014). In this approach, we considered the second temporal cohort as an independent validation of the first. Receiver operating characteristic curves showed that the SVM-RFE model built with microbiota data accurately predicted the clinical outcome among contacts (Figure 2A) and outperformed SVM-RFE models built with clinical and epidemiologic risk factors alone (area under the curve [AUC] = 0.80; $P < .01$). The optimal microbiota model used only 143 OTUs of the full set of 4181 input OTUs. A combined model using both microbiota and clinical data did not lead to improved performance ($P > .05$, by a 2-sided Mann-Whitney U test on the distribution of AUCs).

We performed additional testing to evaluate whether predictive performance could be generalized to other groupings of the contacts or model formulations. Performance of the microbiota model was not reliant on inclusion of young children, who are most susceptible to cholera, as it remained predictive even after exclusion of contacts aged <10 years (AUC = 0.87; $P < .01$). We next tried constructing the SVM-RFE model by using random separations of household contacts into training (90% of samples) and testing (10% of samples) sets. Using this cross-validation scheme, the SVM-RFE model again accurately classified contacts by clinical outcome (Figure 2B). A limited set of 88 OTUs was required in the optimal cross-validation model. The performance of the SVM-RFE model deteriorated below that of the clinical and epidemiological model with the inclusion of fewer (≤26) or more (≥500) OTUs. This likely occurred because the inclusion of OTUs unrelated to cholera susceptibility degrades model performance [20], while the inclusion of too few predictive OTUs cannot account for variation between contacts [21]. An SVM-RFE model constructed using only the presence or absence of OTUs predicted cholera susceptibility (AUC = 0.71; $P < .01$), as did models built using variations of the classification algorithm (Supplementary Figure 2).

Characteristics of Predictive Bacterial Taxa

To learn more about specific bacterial taxa that influenced susceptibility to *V. cholerae* infection, we investigated the characteristics of the subset of gut microbes selected by our SVM-RFE algorithm. The optimal number of OTUs selected by our cross-validation model was 88; to account for possible inaccuracies when identifying taxa, we selected a slightly larger set—the top 100 OTUs—and built an SVM-RFE model using the full cohort of 76 contacts. We term these 100 OTUs the "predictive

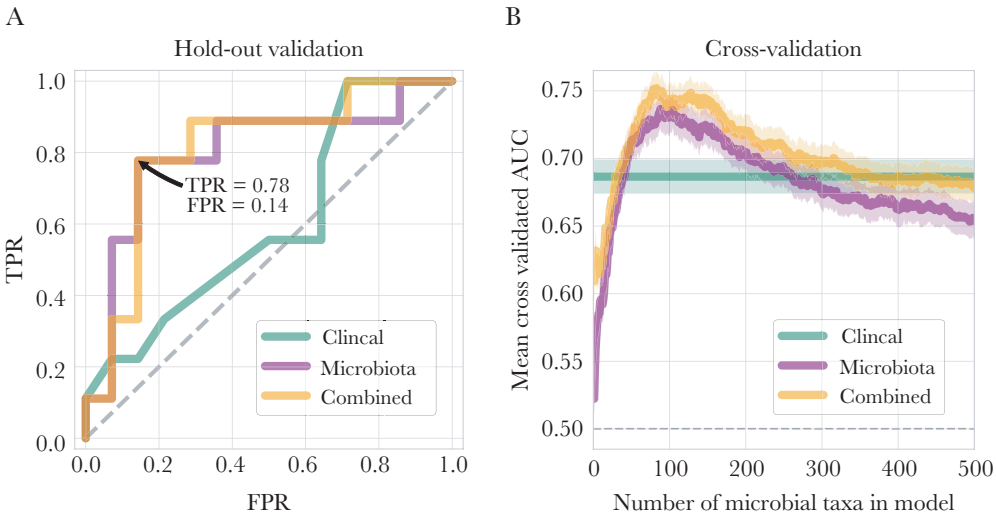

Figure 2. Predictive model performance. *A*, Hold-out models were trained on an initial set of 48 household contacts and tested on a set of 28 household contacts recruited at a later date. Model goodness was assessed using area under the curve (AUC) scores. We evaluated models built using clinical and epidemiological factors (clinical, AUC = 0.60; P = not significant), relative abundance of operational taxonomic units (OTUs; Microbiota, AUC = 0.80; $P < .01$), or a combination of these data types (combined, AUC = 0.81; $P < .01$). P values were computed with permutation tests. *B*, The same set of models was tested using a cross-validation scheme, in which contacts were repeatedly split into training and testing sets. AUCs are shown as a function of OTUs included in the model. Bold lines represent the mean cross-validated AUC, and shaded bands represent the standard error of the mean. Microbiota-based model reached maximum performance at 88 OTUs (AUC = 0.74; $P < .01$). The dashed gray lines indicate the theoretical performance of a model randomly guessing contact susceptibility. FPR, false-positive rate; TPR, true-positive rate.

gut microbiota" (Supplementary Table 2). Principal coordinate analysis based on the predictive gut microbiota demonstrated distinct community structures in the infected and uninfected contacts (Supplementary Figure 3).

Taxa from the predictive gut microbiota could be split into 2 groups based on whether their model coefficients were positive (ie, associated with being infected) or negative (ie, associated with being uninfected). At the individual taxonomic level, the 2 groups showed differences in mean relative abundances and prevalence between the infected and uninfected contacts (Figure 3, Supplementary Figure 4, and Supplementary Table 2), and univariate logistic regression also showed statistically significant effect sizes for some of the top predictive OTUs (Supplementary Table 2). Members of the Bacteroidetes phylum were particularly enriched among taxa associated with being uninfected ($P < .05$, by a FDR-corrected binomial test) and consisted primarily of OTUs from the genus *Prevotella*, which are commonly observed in the human gut microbiota in healthy Bangladeshis [18]. Several taxa, including those from the genera *Blautia* and *Ruminococcus*, were present among OTUs associated with being infected or with being uninfected (Figure 3). Last, 3 infection-associated taxa belonged to the genus *Streptococcus*, which we previously found to be enriched in the gut microbiota of symptomatic patients with cholera [18].

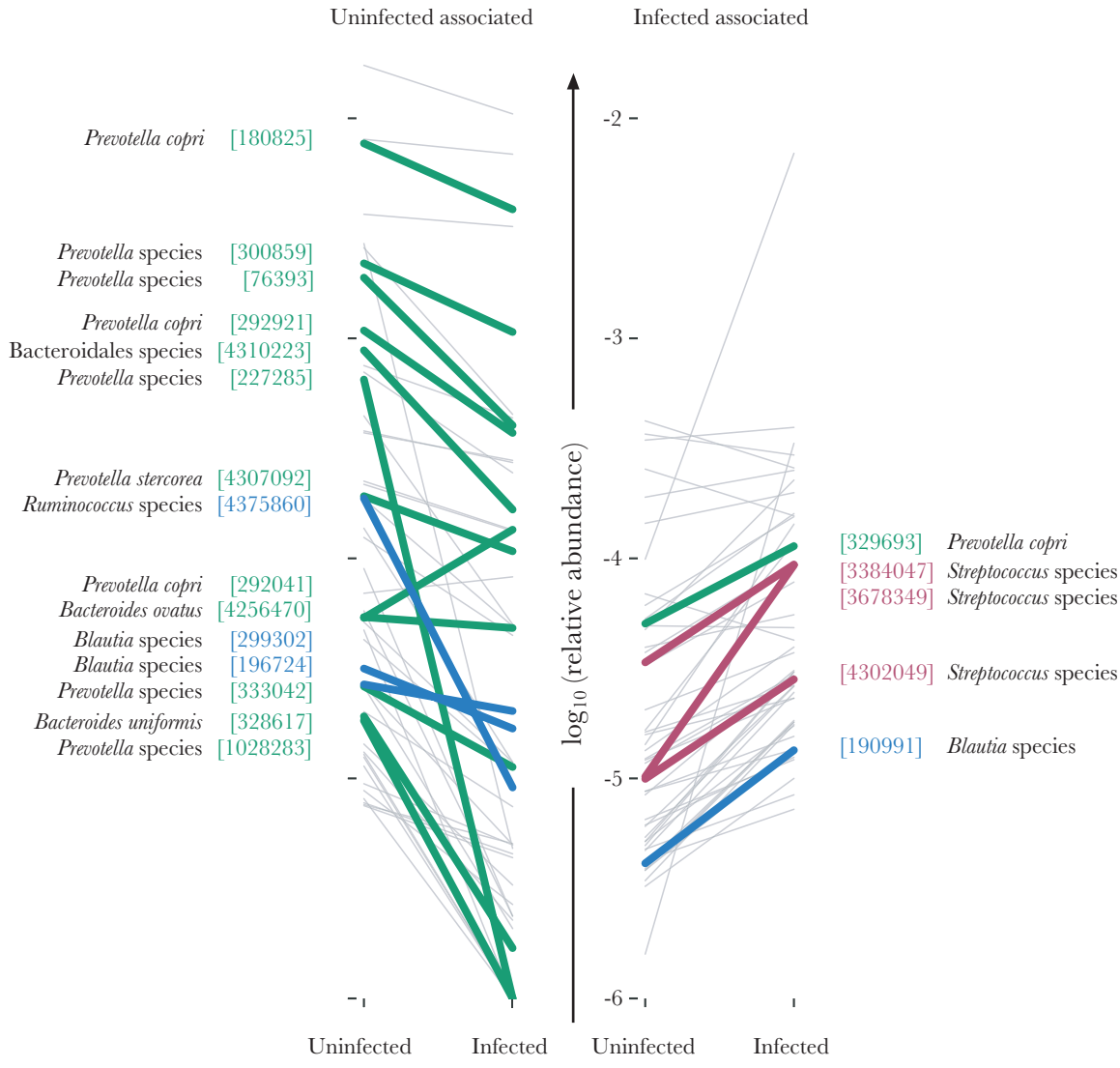

Figure 3. Abundances of predictive gut microbiota in uninfected and infected contacts. Slope graphs are separated into operational taxonomic units (OTUs) with negative model coefficients (uninfected associated; left) and positive model coefficients (infected associated; right). Numbers in brackets indicate the GreenGenes identifiers of representative 16S ribosomal RNA gene sequences. Each gray line depicts an OTU's mean relative abundance in uninfected or infected contacts. Members of the Bacteroidetes phylum (green) were overrepresented among predictive gut microbiota ($P < .05$, by a false-discovery rate–corrected binomial test), members of the genus *Streptococcus* (red) have previously been associated with early stages of cholera [18], and members of the genera *Blautia*/*Ruminococcus* (highlighted in blue) have been previously associated with protection from *Vibrio cholerae* in mice [16].

Relationship Between Gut Microbiota Structure and Age

We next tested for links between the predictive gut microbiota and previously identified risk factors for cholera. At the community level, older individuals had greater gut microbiota richness (Spearman rho = 0.26; $P \leq .05$), and the predictive gut microbiota was associated with age ($R^2 = 0.024$; $P < .05$, by permutational multivariate ANOVA). At the individual taxonomic level, 2 predictive OTUs (members of the genus *Clostridium* and the family Lachnospiraceae) were enriched in younger contacts ($P < .05$, by a FDR–corrected 2-sided Mann-Whitney *U* test; Supplementary Table 3), and 4 clusters of predictive OTUs were significantly associated with age ($P < .05$, by a FDR-corrected 2-sided Mann-Whitney *U* test; Supplementary Table 3). Additionally, an SVM-RFE model that included both microbiota and age (AUC = 0.75; $P < .01$, by a classification permutation test) ultimately discarded age as a model feature, suggesting that gut microbial communities and age encode redundant information.

Association Between Microbes of Interest and *V. cholerae* In Vitro

We performed in vitro experiments, using several members of the predictive taxa (Figure 3) and a second group of taxa identified in a preliminary model of subjects recruited during the first enrollment period (Supplementary Table 4). This preliminary model included subjects with evidence of *V. cholerae* infection at enrollment, distinct from the models identifying the predictive taxa. We focused on organisms for which acquisition of a representative isolate was possible. SCS from predictive taxa, including *B. obeum, B. ovatus, B. uniformis*, and *P. buccalis* (closely related to the predictive taxa *Prevotella copri*), was not found to influence *V. cholerae* growth or biofilm production. *V. cholerae* growth was significantly increased when inoculated into *P. aminovorans* SCS in nutrient-rich medium but not in minimal medium (Figure 4A and B; $P < .001$, by a 2-sided Mann-Whitney *U* test). Because we noted bacterial agglutination during the *V. cholerae* and *P. aminovorans* coculture, we hypothesized that *V. cholerae* biofilm production may be stimulated by *P. aminovorans* SCS. To evaluate this, we measured biofilm production in the setting of *P. aminovorans* SCS by wild-type *V. cholerae* and a *V. cholerae* strain with an in-frame deletion in *vpsA* (*Vibrio* polysaccharide A) [22, 23]. There was no difference between biofilm production of *V. cholerae* in the presence of *P. aminovorans* SCS.

DISCUSSION

More than 1 billion people are at risk of cholera, and cholera has become more geographically widespread in recent years. An improved understanding of susceptibility will support the design of better prevention strategies for cholera. Here, we used a prospective human study to show that the gut microbiota is associated with susceptibility to *V. cholerae* infection. This is consistent with the emerging understanding that the gut microbiome is an active determinant of outcomes after pathogen exposure [11–14].

Our models found cholera susceptibility to be characterized by depletion of bacteria normally found in healthy individuals. Members of the phylum Bacteroidetes, including 10 OTUs from the genus *Prevotella*, were among the taxa that our machine learning models identified in uninfected household contacts. *Prevotella* has been shown to dominate the gut microbiota of healthy individuals in developing countries [18, 24]. These

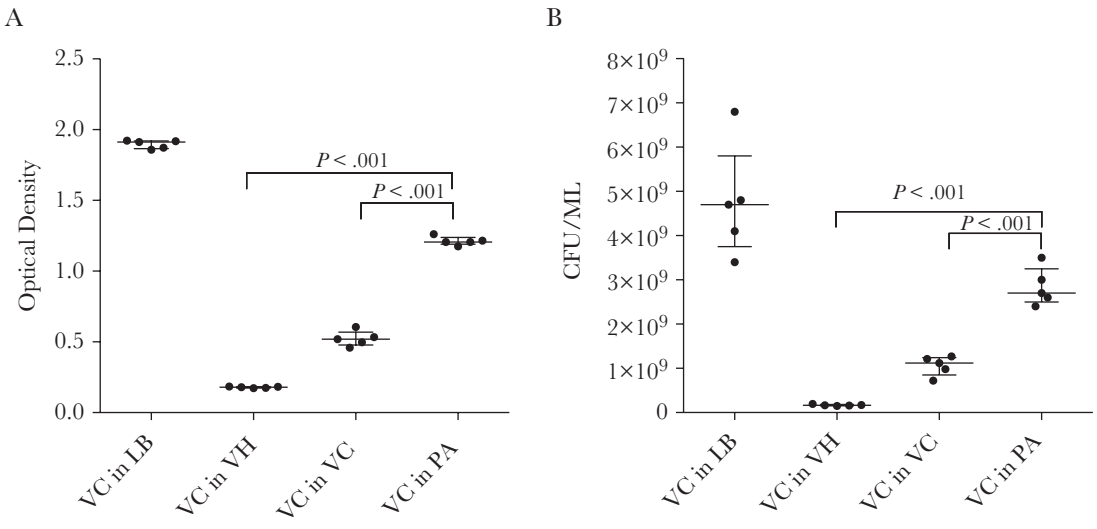

Figure 4. *Vibrio cholerae* (VC) growth in the presence of *Paracoccus aminovorans* (PA) cell-free spent culture supernatant (SCS). *A*, VC demonstrated increased growth in SCS of PA. The first initials on the *x*-axis represent the cultured organism grown in the SCS from the organism represented by the second set of initials, or in fresh medium. *B*, Colony-forming units (CFU) of *V. cholerae* grown in SCS and Luria-Bertani (LB) broth. Median values with interquartile ranges are shown. *P* values are based on a 2-sided Mann-Whitney *U* test. VH, *Vibrio harveyi* (control).

findings support the hypothesis that resident microbes may resist *V. cholerae* infection [16].

Some bacterial taxa identified within our predictive microbiota were surprising. For instance, we observed that members of the Enterobacteriaceae taxa were encountered in both uninfected and infected contacts. Other reports have determined that this family is enriched during enteric infection and have suggested that gut bacteria are more likely to enable colonization of phylogenetically related organisms; we therefore anticipated that this family might be more common in infected household contacts [18, 25]. Additionally, several taxa of the genus *Blautia* and *Ruminococcus* were present among OTUs associated with both the infected and uninfected contacts. A recent in vitro study found that *B. obeum* can inhibit *V. cholerae* virulence [16]. These discordant observations may be due to functional heterogeneity between bacterial OTUs within the same genus. Indeed, model performance degraded when we classified cholera outcomes using data from the species or genus levels; this has been noted previously in microbiota studies that used machine learning tools [26].

We confirmed the biological validity of our machine learning model of cholera susceptibility through study of *P. aminovorans*, a bacterial species identified in the microbiota of infected household contacts. *P. aminovorans* has previously been reported to be abundant in the skin microbiota of individuals with rash associated with *Haemophilus ducreyi* infection [27, 28]. We observed that exposure to *P. aminovorans* SCS increased growth of *V. cholerae* in nutrient-rich medium but not in nutrient-poor medium. This suggests that metabolites made by *P. aminovorans* in the nutrient-rich environment may provide a substrate for *V. cholerae* growth. Other commensal bacteria have been shown to facilitate pathogen growth in mice through metabolite generation or through reduction of reactive oxygen species [29].

In addition to directly influencing *V. cholerae* pathogenesis, predictive taxa identified by our models may also reflect physiological determinants of cholera susceptibility. Older age and preexisting immunity to *V. cholerae* are well-known protective factors for cholera [2, 6]. Recent studies suggest that gut microbes may serve as a sensitive biomarker of host maturity, including in Bangladeshi infants [30]. Here, we observed that the gut microbiota in our cohort varied with age, demonstrating both alpha- and beta-diversity differences. In addition, 4 OTUs from our predictive gut microbiota (*Dialister succinatiphilus*, *Prevotella copri*, *Ruminococcus gnavus*, and *Weissella cibaria*) were previously shown to change during childhood among Bangladeshi youth [31]. Last, our machine learning results suggest that age and microbiota data encode similar information regarding outcomes of *V. cholerae* exposure. These findings together support the concept that the gut microbiota reflects the developmental status of individuals in our cohort. More broadly, our results are consistent with the model that gastrointestinal maturity influences susceptibility to enteric diseases like cholera.

Our study has limitations. We assessed rectal swab samples that likely represent microbiota from the lumen of the large intestine. The composition of these rectal swab samples may differ from the microbes that live on or near the gut mucosa [32]. Furthermore, after ingestion and passage through the stomach, *V. cholerae* adheres to the small intestine, and this is where *V. cholerae* is immunologically and metabolically active [33, 34]. The small intestine has a lower pH and higher oxygen tension and harbors a gut microbiota structure distinct from that of the colon. Our study also does not establish a causal relationship between the gut microbiota and in vivo *V. cholerae* susceptibility. The *P. aminovorans* used in our experiments was an ATCC strain, and the genomic relatedness between our tested strain and ones from Bangladeshi study participants is not known. Future work could isolate strains from patients with cholera, as well as test the influence of these taxa on *V. cholerae* in multispecies interaction experiments, which likely better resemble conditions involving gut microbial communities. The predictive microbiota we identified could also be a marker for another factor responsible for increased susceptibility to *V. cholerae,* such as a nutritional deficiency, environmental enteropathy, or surreptitious antibiotic use, which is common in Bangladesh [35].

Ultimately, our results suggest that future studies to understand host factors responsible for *V. cholerae* susceptibility should include assessment of the gut microbiota. The most widely used oral cholera vaccine, Shanchol, generates varying levels of protection between infants, children, and adults; the gut microbiota may be a contributing factor to this observed variation [36, 37]. This hypothesis is supported by a recent observational study of live oral rotavirus vaccines in Ghanaian infants, in which gut microbial profiles correlated with differences in vaccine immunogenicity [38]. Our findings therefore suggest that consideration should be given to the inclusion of gut microbiota data in cholera vaccine efficacy trials. New strategies to protect vulnerable populations from enteric diseases are needed, and our findings suggest that a focus on protective features of the host microbiota could be fruitful.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. We thank Dr Paula Watnick and Dr Wai-Leung Ng, for contributing materials to this study; Sayan Mukherjee and Florian Wagner, for helpful discussions; the people of Dhaka, where our study was undertaken; and the

field, laboratory, and data management staff, who provided tremendous effort to make the study successful.

Disclaimer. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Financial support. This work was supported by National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH; grants R01AI103055 [to J. B. H. and R. C. L.], U01AI0589354 [to E. T. R. and S. B. C.], and T32A1070611976 and K08AI123494 [to A. A. W.]); the National Institute of General Medical Sciences, NIH (grants T32GM071340 [to F. S. M.] and T32GM007171 [to J. D. S.]); the NSF (grants DGE-1545220 [to F.S.M.] and DGE-1106401 [to A. T. R.]); the Searle Scholars Program (to L. A. D.); the Alfred P. Sloan Foundation (research fellowship to L. A. D.); the North Carolina Biotechnology Center (grant 2016-IDG-1013 [HARDAC+: Reproducible HPC for Next-generation Genomics]); the Government of the People's Republic of Bangladesh (to the International Centre for Diarrhoeal Disease Research [icddr,b]); Global Affairs Canada (to the icddr,b); the Swedish International Development Cooperation Agency (to the icddr,b); and the UK Department for International Development (to the icddr,b).

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Ali M, Lopez AL, You YA, et al. The global burden of cholera. Bull World Health Organ **2012**; 90:209–218A.
2. Weil AA, Khan AI, Chowdhury F, et al. Clinical outcomes in household contacts of patients with cholera in Bangladesh. Clin Infect Dis **2009**; 49:1473–9.
3. Harris JB, LaRocque RC, Chowdhury F, et al. Susceptibility to *Vibrio cholerae* infection in a cohort of household contacts of patients with cholera in Bangladesh. PLoS Negl Trop Dis **2008**; 2:e221.
4. Mosley WH, Ahmad S, Benenson AS, Ahmed A. The relationship of vibriocidal antibody titre to susceptibility to cholera in family contacts of cholera patients. Bull World Health Organ **1968**; 38:777–85.
5. Glass RI, Becker S, Huq MI, et al. Endemic cholera in rural Bangladesh, 1966–1980. Am J Epidemiol **1982**; 116:959–70.
6. Harris JB, Khan AI, LaRocque RC, et al. Blood group, immunity, and risk of infection with *Vibrio cholerae* in an area of endemicity. Infect Immun **2005**; 73:7422–7.
7. Larocque RC, Sabeti P, Duggal P, et al. A variant in long palate, lung and nasal epithelium clone 1 is associated with cholera in a Bangladeshi population. Genes Immun **2009**; 10:267–72.
8. Glass RI, Svennerholm AM, Khan MR, Huda S, Huq MI, Holmgren J. Seroepidemiological studies of El Tor cholera in Bangladesh: association of serum antibody levels with protection. J Infect Dis **1985**; 151:236–42.
9. Lindsay B, Oundo J, Hossain MA, et al. Microbiota that affect risk for shigellosis in children in low-income countries. Emerg Infect Dis **2015**; 21:242–50.
10. Kampmann C, Dicksved J, Engstrand L, Rautelin H. Composition of human faecal microbiota in resistance to Campylobacter infection. Clin Microbiol Infect **2016**; 22:61.e1–8.
11. Moroni O, Kheadr E, Boutin Y, Lacroix C, Fliss I. Inactivation of adhesion and invasion of food-borne Listeria monocytogenes by bacteriocin-producing Bifidobacterium strains of human origin. Appl Environ Microbiol **2006**; 72:6894–901.
12. Kamada N, Kim YG, Sham HP, et al. Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. Science **2012**; 336:1325–9.
13. Raqib R, Sarker P, Bergman P, et al. Improved outcome in shigellosis associated with butyrate induction of an endogenous peptide antibiotic. Proc Natl Acad Sci U S A **2006**; 103:9178–83.
14. Fukuda S, Toh H, Hase K, et al. Bifidobacteria can protect from enteropathogenic infection through production of acetate. Nature **2011**; 469:543–7.
15. Lawson PA, Finegold SM. Reclassification of *Ruminococcus obeum* as *Blautia obeum* comb. nov. Int J Syst Evol Microbiol **2015**; 65:789–93.
16. Hsiao A, Ahmed AM, Subramanian S, et al. Members of the human gut microbiota involved in recovery from *Vibrio cholerae* infection. Nature **2014**; 515:423–6.
17. Weil AA, Begum Y, Chowdhury F, et al. Bacterial shedding in household contacts of cholera patients in Dhaka, Bangladesh. Am J Trop Med Hyg **2014**; 91:738–42.
18. David LA, Weil A, Ryan ET, et al. Gut microbial succession follows acute secretory diarrhea in humans. mBio **2015**; 6:e00381-15.
19. Statnikov A, Henaff M, Narendra V, et al. A comprehensive evaluation of multicategory classification methods for microbiomic data. Microbiome **2013**; 1:11.
20. Weston J, Mukherjee S, Chapelle O, Pontil M, Poggio T, Vapnik V. Feature selection for SVMs. Adv Neur In **2001**; 13:668–74.
21. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. Nature **2012**; 486:207–14.
22. Yildiz FH, Schoolnik GK. *Vibrio cholerae* O1 El Tor: identification of a gene cluster required for the rugose colony type, exopolysaccharide production, chlorine resistance, and biofilm formation. Proc Natl Acad Sci U S A **1999**; 96:4028–33.

23. Wai SN, Mizunoe Y, Takade A, Kawabata SI, Yoshida SI. *Vibrio cholerae* O1 strain TSI-4 produces the exopolysaccharide materials that determine colony morphology, stress resistance, and biofilm formation. Appl Environ Microbiol **1998**; 64:3648–55.
24. Yatsunenko T, Rey FE, Manary MJ, et al. Human gut microbiome viewed across age and geography. Nature **2012**; 486:222–7.
25. Stecher B, Chaffron S, Käppeli R, et al. Like will to like: abundances of closely related species can predict susceptibility to intestinal colonization by pathogenic and commensal bacteria. PLoS Pathog **2010**; 6:e1000711.
26. Pasolli E, Truong DT, Malik F, Waldron L, Segata N. Machine learning meta-analysis of large metagenomic datasets: tools and biological insights. PLoS Comput Biol **2016**; 12:e1004977.
27. van Rensburg JJ, Lin H, Gao X, et al. The human skin microbiome associates with the outcome of and is influenced by bacterial infection. MBio **2015**; 6:e01315–15.
28. Urakami T, Araki H, Oyanagi H, Suzuki K, Komagata K. *Paracoccus aminophilus* sp. nov. and *Paracoccus aminovorans* sp. nov., which utilize N,N-dimethylformamide. Int J Syst Bacteriol **1990**; 40:287–91.
29. Yoon MY, Min KB, Lee KM, et al. A single gene of a commensal microbe affects host susceptibility to enteric infection. Nat Commun **2016**; 7:11606.
30. Subramanian S, Huq S, Yatsunenko T, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature **2014**; 510:417–21.
31. Blanton LV, Charbonneau MR, Salih T, et al. Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. Science **2016**; 351:pii:aad3311. doi: 10.1126/science.aad3311.
32. Araújo-Pérez F, McCoy AN, Okechukwu C, et al. Differences in microbial signatures between rectal mucosal biopsies and rectal swabs. Gut Microbes **2012**; 3:530–5.
33. Holmgren J, Lönnroth I, Månsson J, Svennerholm L. Interaction of cholera toxin and membrane GM1 ganglioside of small intestine. Proc Natl Acad Sci U S A **1975**; 72:2520–4.
34. Yamamoto T, Kamano T, Uchimura M, Iwanaga M, Yokota T. *Vibrio cholerae* O1 adherence to villi and lymphoid follicle epithelium: in vitro model using formalin-treated human small intestine and correlation between adherence and cell-associated hemagglutinin levels. Infect Immun **1988**; 56:3241–50.
35. Larson CP, Saha UR, Islam R, Roy N. Childhood diarrhoea management practices in Bangladesh: private sector dominance and continued inequities in care. Int J Epidemiol **2006**; 35:1430–9.
36. Desai SN, Akalu Z, Teferi M, et al. Comparison of immune responses to a killed bivalent whole cell oral cholera vaccine between endemic and less endemic settings. Trop Med Int Health **2016**; 21:194–201.
37. Qadri F, Wierzba TF, Ali M, et al. Efficacy of a single-dose, inactivated oral cholera vaccine in Bangladesh. N Engl J Med **2016**; 374:1723–32.
38. Harris VC, Armah G, Fuentes S, et al. Significant correlation between the infant gut microbiome and rotavirus vaccine response in Rural Ghana. J Infect Dis **2017**; 215:34–41.