Genomic sequencing-based mutational enrichment analysis identifies motility genes in a genetically intractable gut microbe

Sena Bae[a,b,c], Olaf Mueller[a,b], Sandi Wong[a,b], John F. Rawls[a,b], and Raphael H. Valdivia[a,b,1]

[a]Center for the Genomics of Microbial Systems, Duke University School of Medicine, Durham, NC 27710; [b]Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, NC 27710; and [c]Department of Biomedical Engineering, Duke University, Durham, NC 27708

Edited by Lalita Ramakrishnan, University of Cambridge, Cambridge, United Kingdom, and approved October 19, 2016 (received for review August 1, 2016)

A major roadblock to understanding how microbes in the gastrointestinal tract colonize and influence the physiology of their hosts is our inability to genetically manipulate new bacterial species and experimentally assess the function of their genes. We describe the application of population-based genomic sequencing after chemical mutagenesis to map bacterial genes responsible for motility in *Exiguobacterium acetylicum*, a representative intestinal Firmicutes bacterium that is intractable to molecular genetic manipulation. We derived strong associations between mutations in 57 *E. acetylicum* genes and impaired motility. Surprisingly, less than half of these genes were annotated as motility-related based on sequence homologies. We confirmed the genetic link between individual mutations and loss of motility for several of these genes by performing a large-scale analysis of spontaneous suppressor mutations. In the process, we reannotated genes belonging to a broad family of diguanylate cyclases and phosphodiesterases to highlight their specific role in motility and assigned functions to uncharacterized genes. Furthermore, we generated isogenic strains that allowed us to establish that *Exiguobacterium* motility is important for the colonization of its vertebrate host. These results indicate that genetic dissection of a complex trait, functional annotation of new genes, and the generation of mutant strains to define the role of genes in complex environments can be accomplished in bacteria without the development of species-specific molecular genetic tools.

motility | mutagenesis | comparative genomics | gene annotation

The advent of DNA-sequence-based approaches to survey microbial environments has led to a deepened appreciation for the diversity, ubiquity, and functions of microbial life. For instance, the gastrointestinal tract of humans and other vertebrates is colonized by complex microbial communities that promote gut development, nutrient metabolism, and immune homeostasis (1). Of particular importance to human health, gut microbes have emerged as major risk determinants for obesity and metabolic disorders, in part because of their role in modulating accessibility and absorption of energy-rich dietary nutrients in vertebrates (2). For example, colonization of germ-free zebrafish with *Exiguobacterium* sp. ZWU0009, a Firmicutes bacterium originally isolated from the zebrafish intestine, enhanced the ability of intestinal enterocytes to absorb dietary fat (3). Unfortunately, the molecular bases for how bacteria like *Exiguobacterium* sp. ZWU0009 colonize the intestine and influence host physiology are poorly understood. Indeed, most microbes are not amenable to genetic manipulation because methods for robust DNA transformation, insertional mutagenesis, and *trans*-expression of genes are largely lacking. For a select group of microbial species, including members of the *Bacteroides* genus, some strains are amenable to transposon mutagenesis and have been invaluable in helping decipher the requirement of individual genes in gut colonization and nutrient homeostasis (4, 5). However, genetic tools do not exist for the vast majority of intestinal microbes. As a result, the function of individual genes and their contribution to host–microbe and microbe–microbe interactions within the gut often relies on information inferred from homology to genes characterized in phylogenetically unrelated, but genetically tractable, bacterial systems. This reliance on previously characterized genes has emerged as a major block in the functional annotation of novel genes emerging from metagenomic studies.

Results and Discussion

We sought to develop broad methods for genetic analysis of "genetically intractable" microbes, by using *Exiguobacterium* sp. as a representative gut microbe. We first generated a draft genome sequence of strain ZWU0009 (Taxonomy ID: 1224749) (3, 6). The ZWU0009 genome is ~3.2 Mb and includes 3,289 coding sequences, 30 rRNA operons, and 76 tRNAs. A comparative genome analysis indicated a close relationship with other previously characterized *Exiguobacterium* and that this isolate is a variant of *Exiguobacterium acetylicum* (7). We reasoned that one could apply whole-genome sequencing to monitor experimentally induced genetic variations in *E. acetylicum* ZWU0009 and then derive associations between gene variants and phenotypically selected traits, a process we term "mutational enrichment analysis after phenotypic selection" (MEAPS) (Fig. 1*A*). We chose to perform a MEAPS dissection of *Exiguobacterium* motility, a multigenic complex trait that is important for some bacterial pathogens to colonize the vertebrate gut (8, 9). In contrast, the role of motility in colonization by commensal bacteria is less clear. Metatranscriptomic analysis of healthy gut microbiotas

Significance

There is broad interest in the role microbial communities play in human health. Although DNA-sequencing technologies enabled a broad assessment of microbial diversity and genomic content, our understanding of the molecular mechanisms underlying microbe–microbe and microbe–host interactions has proceeded much more slowly because only a small fraction of microbes are amenable to molecular genetic manipulation. We describe a method, independent of recombinant DNA tools, to perform genetic analysis in any cultivatable microbial species. We identified determinants of motility in a member of the vertebrate microbiome, the Firmicutes *Exiguobacterium acetylicum*, and experimentally determined a role for motility in animal colonization by this previously uncharacterized commensal bacteria that is important for host nutrient homeostasis.

Author contributions: S.B., S.W., J.F.R., and R.H.V. designed research; S.B. and S.W. performed research; S.B., O.M., J.F.R., and R.H.V. analyzed data; and S.B. and R.H.V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the BioProject database, https://www.ncbi.nlm.nih.gov/bioproject (accession no. PRJNA318259). A complete, annotated genome and its analysis have been deposited in the BioProject database, https://www.ncbi.nlm.nih.gov/bioproject (accession no. PRJNA205584) and the GenBank database, https://www.ncbi.nlm.nih.gov/genbank/ (accession no. CP018057).

[1]To whom correspondence should be addressed. Email: raphael.valdivia@duke.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1612753113/-/DCSupplemental.

indicates a broad dampening of the expression of motility genes in Firmicutes and Proteobacteria as a result of innate and adaptive immune responses (10, 11), suggesting that the expression of flagella may confer a disadvantage to commensal bacteria. *E. acetylicum* expresses peritrichous flagella, as assessed by transmission electron microscopy (TEM), when grown in rich medium (Fig. 1*B*).

Based on keyword searches and reciprocal BLAST queries we predicted that 126 *E. acetylicum* genes are potentially involved in motility (Dataset S1). To determine the number of nonmotile *E. acetylicum* mutants we needed to sequence to identify mutations overrepresented in motility loci, we first modeled MEAPS experiments with the assumption that 100 genes are required for motility and that each nonmotile mutant had one motility-disabling mutation. These simulations indicated that, by sequencing the genomes of 400–500 nonmotile mutants with an average of 10 mutations per genome, we can capture ∼70% of all motility genes (Fig. S1). Further changes in the number of mutants sequenced or mutagenesis rates only led to marginal increases to the number of new motility genes that could be identified.

We generated *E. acetylicum* mutants by treating bacteria with ethyl methyl sulfonate (EMS) and used the ability of *E. acetylicum* to swarm in soft agar plates to enrich for nonmotile mutants. *E. acetylicum* inoculated in the center of a 0.3% agar plate will spread as large halos of turbidity emerging from the inoculation site. After serial passages of pools of mutagenized bacteria in soft agar with repeated collection of bacteria from the site of inoculation, we tested bacterial clones for defects in swarming motility (Fig. S2) and stored them as individual clones.

We next applied a pooling strategy to sequence 440 nonmotile *E. acetylicum* mutants and identified all of the mutagen-induced single nucleotide variants (SNVs) by mapping unique sequence reads to the reference genome. To compensate for mutational biases because of gene length and relative %GC content, we also sequenced 700 EMS-generated mutants that had not been selected for the loss of motility. Overall, we identified 4,009 and 5,013 SNVs among the selected (nonmotile) and unselected strains, respectively. Next, we compared mutation frequencies between these sets of mutant strains and identified two regions in the *E. acetylicum* genome that displayed a marked accumulation of nonsynonymous, but not of synonymous, mutations (Fig. 1*C* and Fig. S3). Region I encompasses most predicted flagellar structural genes and chemotaxis genes (Fig. 1*D*). Region II encoded additional predicted flagellar structural components, including two flagellin H-antigen (*Hag*) genes, and the two-component regulatory system, DegS/DegU (12). Using more stringent criteria (*Methods*), we defined 57 genes as most likely required for motility, including 27 genes homologous to genes not commonly associated with motility in Gram-positive bacteria and 7 genes encoding proteins of unknown function. In general, the highest-confidence 21 genes identified by MEAPS (Fig. 1*E* and Dataset S2) were homologous to genes previously associated with the assembly of flagella or the regulation of motility in bacteria (12). Overall, 102 genes were overrepresented among nonmotile strains, including 19% of genes identified as potential motility genes by keyword searches (Dataset S2). Strikingly, only 23% ($n = 24$) and 55% ($n = 56$) of these 102 genes were identified as potential motility genes by reciprocal BLAST homology and Pfam terms, respectively (Fig. 2*A*). These results imply that a significant fraction of genes predicted by bioinformatics to participate in *E. acetylicum* motility do not play a role in motility under the conditions tested. We also monitored intergenic regions (Fig. S4)

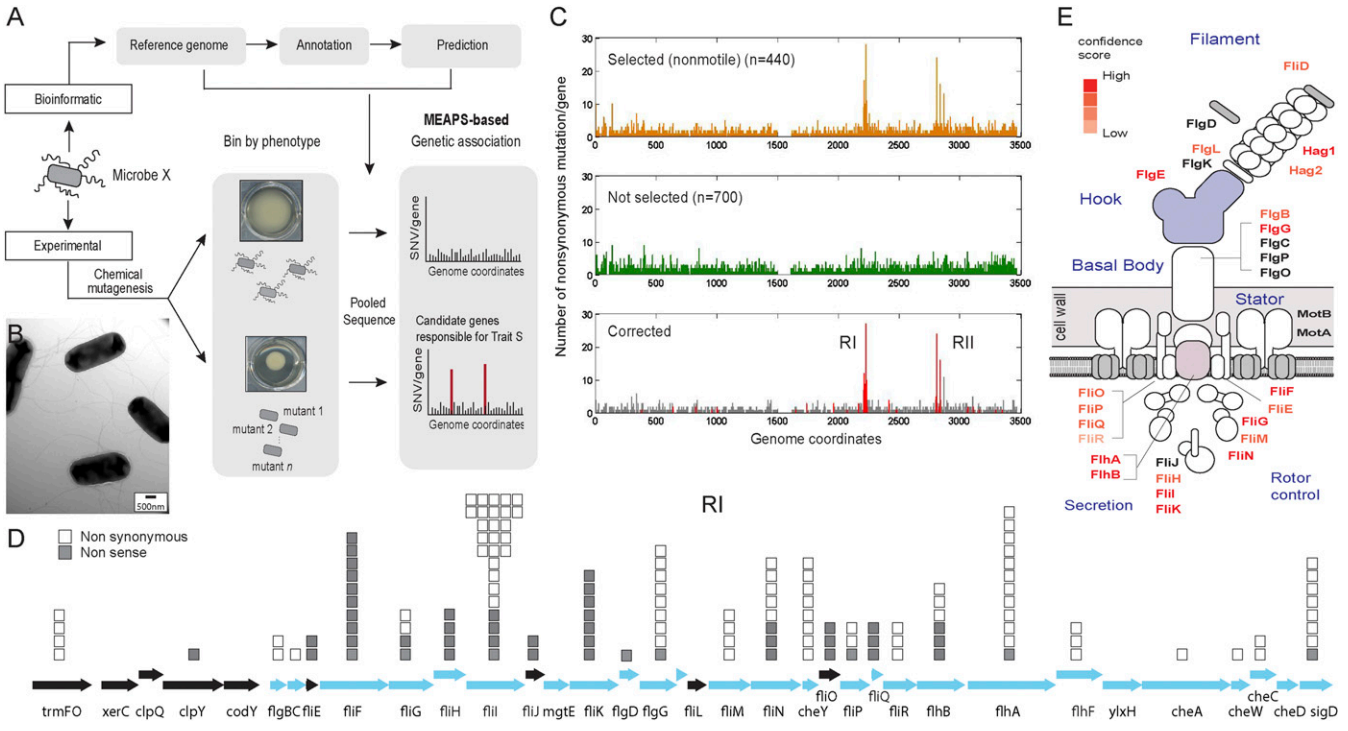

Fig. 1. Identification of genes required for motility in *E. acetylicum*. (*A*) Schematic of the approach coupling chemical mutagenesis and DNA sequencing to perform a MEAPS and identify genes required for swarming motility. (*B*) Transmission electron micrograph of *E. acetylicum* flagella. (Scale bar: 0.5 μm.) (*C*) MEAPS of *E. acetylicum* strains defective for motility. *E. acetylicum* mutants were selected based on the loss of swarming ($n = 440$) or not selected ($n = 700$), and their genomes were sequenced. The normalized frequency of nonsynonymous mutations reveals two chromosomal regions (RI and RII) that preferentially accumulate mutations in nonmotile *E. acetylicum* strains. Red bars highlight predicted motility genes based on their similarities to motility genes in other bacteria. (*D*) Genetic map of ORFs in motility region I. Genes with homology to predicted motility genes (Dataset S1) are shown as blue arrows, and the number of nonsense and nonsynonymous mutations identified, after correction, are represented by gray and white squares, respectively. (*E*) Cartoon schematic of the Gram-positive flagellar apparatus displaying all components conserved in *E. acetylicum*. Components identified by MEAPS are shown in different font colors to reflect confidence of their relative association with motility (Dataset S2).

and identified mutations in three loci that were overrepresented in nonmotile mutants, including SNVs mapping to the predicted ribosome-binding site of *cheY* and *flhA*, encoding putative chemotaxis and flagellar structural proteins, respectively.

Although MEAPS led to strong associations between mutations in specific genes and the loss of motility, it is difficult to unequivocally assign causality to any one mutation, especially because the average number of SNVs per mutant strains is 9.3. We applied another basic tool in microbial genetics, the isolation of spontaneous genetic suppressors mutations, to determine whether the genes identified were responsible for the loss of motility. For this analysis we isolated strains from among our mutant collection with nonsense mutations in *flhA*, required for flagellar biosynthesis (13); *fliE*, *fliF*, *fliK*, and *flgG1*, encoding core components of the basal structure (14); *fliM*, encoding the flagellar M-ring and switch component (15); *flgN*, encoding a secretion chaperone (16); and *hag1*, encoding one of the two predicted flagellin subunits. Mutants were passaged on soft agar plates, and bacteria were collected from the leading edge of the inoculation spot. After three to four passages, clonal populations of strains that had regained motility were isolated, and the DNA region spanning the predicted motility gene was sequenced. In all instances, the suppressor mutations either reversed the original nonsense mutation or changed adjacent nucleotides to generate a reading codon (Fig. 2*B*), indicating that for loss-of-function mutations in predicted structural flagellar genes and accessory factor(s), the only path to restore motility was to repair the nonsense mutation. In a representative *hag1* nonsense mutant, flagella were observed by TEM only after intragenic repair of the nonsense lesion to generate a Q-to-W codon switch (Fig. 2*C*). These findings provide compelling genetic evidence that the mutations identified by MEAPS were indeed responsible for the loss of motility.

The availability of genetically defined nonmotile and coisogenic suppressor strains allowed us to test whether flagellar motility is required for *E. acetylicum* colonization of zebrafish hosts. We isolated rifampin-resistant variants of a strain bearing a *hag1* (*ea2793*) nonsense mutation ($Hag1^{Q222*}$) and its suppressed sister strain ($Hag1^{Q222W}$) and performed competitive colonization experiments in germ-free zebrafish larvae. The relative enrichment of one strain over the other at 0 and 3 d postinoculation was then

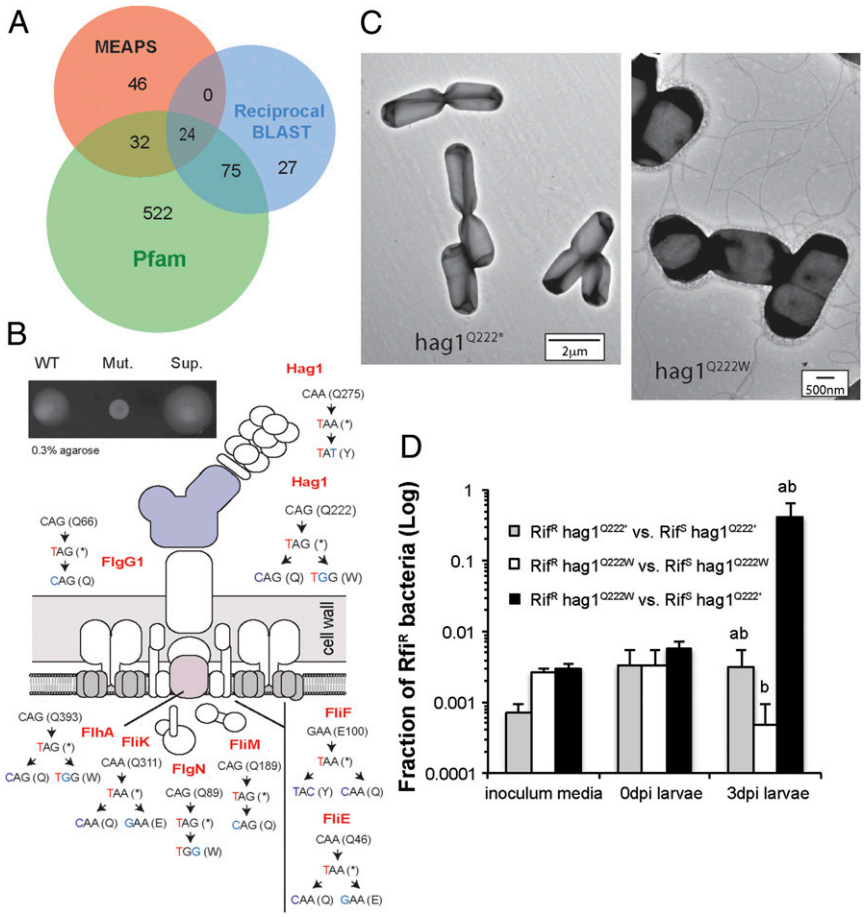

Fig. 2. Functional characterization of motility genes in *E. acetylicum*. (*A*) Overlapping set of putative *E. acetylicum* motility genes identified by reciprocal BLAST homology searches, Pfam associated with motility, or MEAPS. (*B*) Suppressors of nonsense mutations in putative structural flagellar genes confirm their role in motility. Strains with nonsense mutations in flagellar components were passaged in soft agar to enrich for spontaneous motile variants. (*B*, *Inset*) Hag^{Q222*} and its Hag^{Q222W} suppressor. Sequence analysis indicated the presence of reversions and intragenic suppressor mutations that restored the reading frame. The relevant sequence of mutated and suppressed codons is shown. (*C*) Intragenic suppressor of a *hag1* nonsense allele restores the formation of wild-type flagellar structures in *E. acetylicum*. (*D*) Motility enhances *E. acetylicum* colonization of germ-free zebrafish. Rifampin (Rif)-resistant and -sensitive versions of an *E. acetylicum* strain with a nonsense mutation in Hag1 ($Hag1^{Q222*}$) and its motile suppressor isogenic derivative ($Hag1^{Q222W}$) were placed in direct competition for colonization of 6 d postfertilization germ-free zebrafish embryos. Inoculum medium consisted of Rif-resistant strains mixed with sensitive strains at a 1:500 ratio. The relative frequency of each strain in the inoculum media at 6 d postfertilization, in association with animals immediately after colonization at 0 d postinoculation (0 dpi) and after 3 d of association (3 dpi) were determined by assessing the percentage of Rif-resistant bacteria. Error bars represent SD. Letters indicate $P < 0.05$ compared with respective inoculum medium (a) or 0-dpi larvae (b) using the Kruskal–Wallis test.

assessed by plating larvae-associated bacteria and enumerating rifampin-resistant colonies. The nonmotile *hag1* mutants were rapidly outcompeted by their motile suppressors, even when the majority of the starting inoculum consisted of nonmotile strains (Fig. 2*D*). Similar results were obtained with a competition between a spontaneous *flhA* mutant and a wild-type strain (Fig. S5), emphasizing the role for *E. acetylicum* motility in stable colonization of its vertebrate host.

We extended genetic-suppression analysis to define the role of genes annotated with ambiguous or unknown functions that we identified as motility genes by MEAPS. We isolated suppressors of nonsense alleles of uncharacterized genes in RII: *ea2862*, encoding a protein with domains with predicted diguanylate cyclase (GGDEF) and phosphodiesterase activities (EAL), respectively (17); *ea2619*, encoding a hypothetical protein; and *ftsX*, encoding a protein associated with septation and sporulation (18) (Fig. 3*A*). GGDEF/EAL domain proteins regulate the formation of cyclic dimeric GMP (c-di-GMP), a signaling molecule that controls multiple cellular behaviors, including motility and biofilm formation (17). The *E. acetylicum* genome encodes 10 proteins with tandem GGDEF/EAL domains, but only two genes encoding GGDEF/EAL proteins displayed a significantly higher mutational load among nonmotile mutants (Fig. 3*B*). For the *ftsX* nonsense mutant, as with structural flagellar genes, the suppressor mutations isolated were intragenic and consisted of reversions of the original mutation or nucleotide changes that restored translation (Fig. 3*C*). In *Bacillus subtilis*, the FtsXE complex is required for the secretion of peptidoglycan hydrolases CwlO and LytE and proper septum assembly during sporulation and elongation (19), and peptidoglycan remodeling is important for flagellar biosynthesis and function in Gram-positive bacteria (20). Interestingly, although FtsX in *B. subtilis* and *E. acetylicum* are 49% identical and share similar chromosomal location adjacent to motility genes, FtsX has not been associated with motility in *B. subtilis*. Flagella were readily apparent in *ftsX* mutants, indicating that CwlO, LytE, and related peptidoglycan-remodeling enzymes are not required for flagellar assembly (Fig. 3*D* and Fig. S6). We propose that the negative impact of *ftsX* mutations on motility may be the result of improper assembly of additional factors required for flagellar function, as has been suggested for gliding motility in Flavobacteria (20, 21). Consistent with this premise, mutations in the putative ATPase FtsE and multiple peptidoglycan lyticases (Lyt) were also overrepresented among nonmotile *E. acetylicum* mutants (Dataset S2).

For nonsense mutations in *ea2862* and *ea2619*, we obtained multiple extragenic suppressor mutations that restored motility (Dataset S3). Common suppressor mutations of at least three independent nonsense alleles of *ea2862* and of *ea2619* included point mutations in the flagellar switch proteins FliM and FliN, regulators of chemotaxis (CheY and PtsI) (22, 23), and the kinase DegS (24) (Fig. 4*A*). Mutations in these genes were also common in spontaneous *E. acetylicum* mutant strains that had been selected for hypermotility on soft agar (Dataset S4). Overall, these results

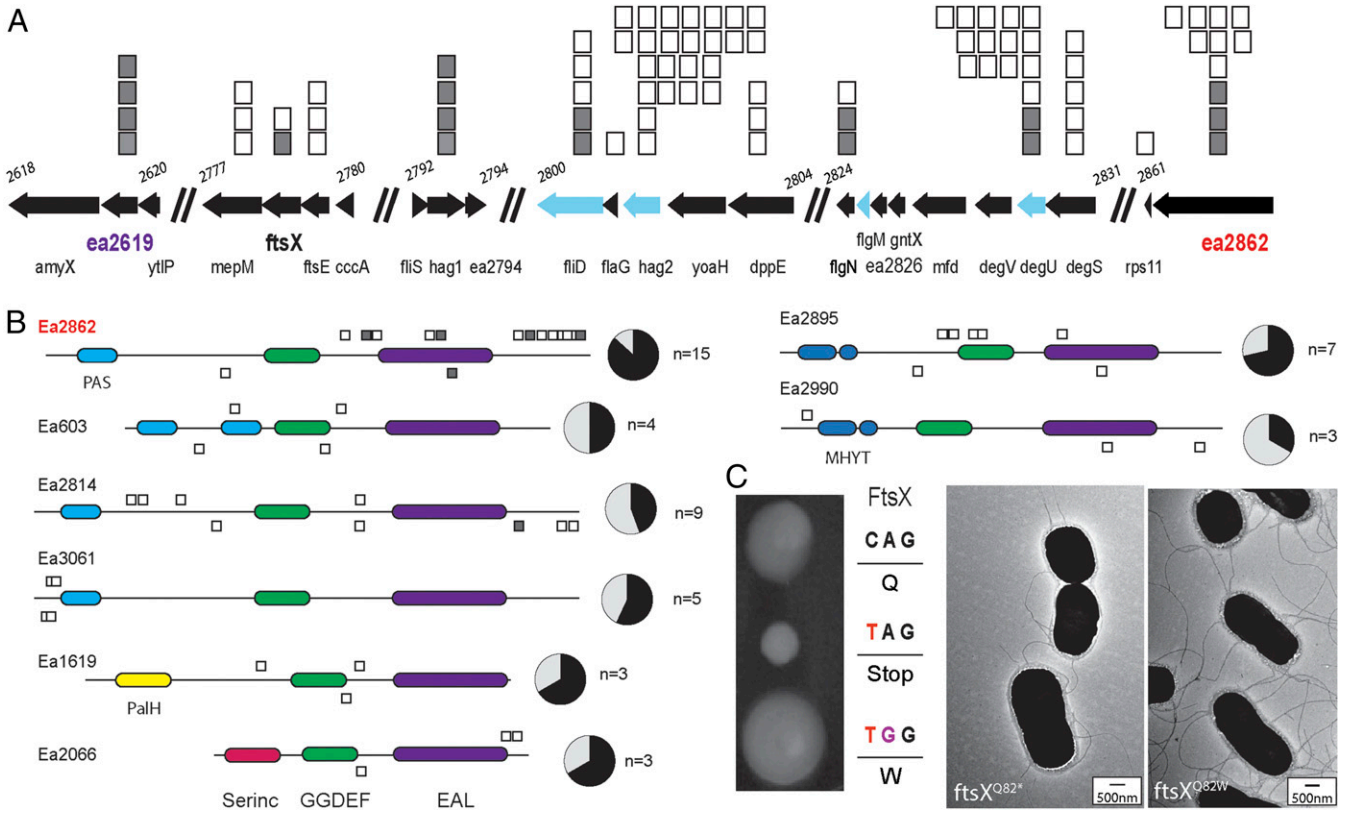

Fig. 3. Motility region II of *E. acetylicum* encodes for new motility genes. (*A*) Nonsynonymous mutations in uncharacterized genes and close homologs of predicted motility genes (blue arrows) in the RII motility region are overrepresented among nonmotile *E. acetylicum* mutants. The number of independent nonsense and nonsynonymous mutations identified are represented by gray and white squares, respectively. (*B*) GGDEF/EAL domain proteins in *E. acetylicum*. Only genes with more than three total nonsynonymous mutations are shown. Pie charts indicate the proportion of total mutations identified among nonmotile (black) and unselected (gray) mutant pools. Squares on top and bottom of each gene represent the location of mutations identified in the selected (nonmotile) and unselected group, respectively. EAL, phosphodiesterase domain; GGDEF, diaguanylate cyclase domain; MYHT, bacterial signaling (Pfam03707)); PalC, proteolytic processing (pfam08733); PAS, sensor for signal transduction; Serinc, serine incorporator (pfam03348. (*C*) Motility behavior and flagellar assembly of a FtsX^{Q82*} nonsense *E. acetylicum* mutant and a spontaneous variant that regained motility (FtsX^{Q82W}). TEM analysis of mutants and their suppressor indicated that FtsX is not required for flagellar assembly.

indicate that mutations in *ea2862* and *ea2619* likely regulate the frequency and direction of motility as opposed to flagellar assembly or function. Consistent with this observation, the suppressor mutations we identified in the FliM cluster in regions predicted to regulate interactions with FliG and other motility regulators (25, 26) and potentially the ability for FliM to self-assemble (27). Similarly, mutations in the homolog of the *B. subtilis* Ser/Thr kinase PrkC, which is linked to cell-wall metabolism (28), also emerged as common suppressors. Additional suppressors of *ea2619* included mutations in a lipoteichoic acid biosynthetic protein (LtaS2) and MprF, a phosphatidyl lysil transferase, which are involved in maintenance of the cell envelope (29, 30). We speculate that cell-wall modifications in these strains alter the rotation capacities of the flagellum (Fig. 4*B*).

Some suppressor mutations were gene-specific. For instance, two independent amino acid changes in Ea2157, a PilZ domain protein predicted to bind c-di-GMP, specifically suppressed two independent nonsense mutations in *ea2862*. The N terminus of Ea2157 has homology to the YcgR family of proteins, which in enteric bacteria control chemotaxis and swimming velocity by interacting directly with stator proteins (25, 31). These findings confirm the important role of c-di-GMP in the regulation of flagellar motility in Firmicutes (32). In addition to GGDEF and EAL domains, Ea2862 also has a PAS domain that is predicted to sense small molecules such as metabolites and gases (33). Indeed, the automated PROKKA annotation program ascribed Ea2862 as a homolog of the oxygen-sensing protein DosP (34). In *B. subtilis*, loss of the EAL protein PdeH causes elevated levels of c-di-GMP that inhibit motility through the PilZ domain protein YpfA and the flagellar motor protein MotA (31, 35, 36). Aside from the EAL domains in Ea2862, this protein does not share any homology with its *B. subtilis* counterpart. The role of Ea2862, or Ea2157, in motility could have not been predicted solely on bioinformatic analysis and highlights the utility of MEAPS for the functional annotation of a microbial genome (Fig. 4*B*). We hypothesize that the role of Ea2862 in integrating multiple signals, through its PAS domain, to regulate c-di-GMP levels and motility is more complex than that performed by PdeH in *B. subtilis*. The observation that many *E. acetylicum* GGDEF/EAL proteins also have PAS domains is intriguing (Fig. 3*B*), given that this domain can bind diffusible signal factors, such as unsaturated fatty acids (37), to potentially regulate c-di-GMP signaling and intraspecies and interspecies cell communication in the intestine.

The paucity of experimental tools to manipulate bacterial genomes has emerged as a major roadblock to understand how microbial communities assemble and influence human and environmental health. Here, we describe how coupling of chemical mutagenesis, phenotypic selection, suppression analysis, and genomic sequencing-based mutational mapping, can be applied to rapidly derive strong phenotype–genotype correlations in a microbe with no preestablished molecular genetic tools, leading to a functional annotation of previously uncharacterized genes. A similar conceptual framework has been proposed to explore gene function in the obligate intracellular pathogen *Chlamydia trachomatis* (38), for conservation in protein function in bacteriophages (39), and to identify genes required for magnetosome formation (40) and microbial drug resistance (41, 42). By extending the application of genomic sequencing to the large-scale analysis of suppressor mutations, we genetically confirmed the contribution of specific genes to a complex trait and further revealed functions for new proteins whose activity could not be inferred solely from sequence homologies. We anticipate that this approach will significantly improve our ability to define the molecular mechanisms by which any culturable bacterial species interacts with their environments, aid with genome annotations, and provide biological tools to probe the function of specific genes in complex microbial communities.

Methods

Genome Assembly and Prediction of Motility Genes. *Exiguobacterium* sp. strain ZWU0009 (6) was grown in brain heart infusion (BHI) broth (BD Biosciences) overnight at 30 °C, and genomic DNA was isolated with a DNeasy blood and tissue kit (Qiagen). The genome was sequenced with PacBio RS (Pacific Bioscience Inc.). PacBio reads were assembled with PacBio SMRT analysis software, by using the HGAP2 protocol. Subsequent gene annotation was performed with PROKKA (43). Potential motility genes were identified by using the terms "firmicutes" and "motility" in reciprocal BLAST searches and Pfam domains in the European Molecular Biology Laboratory–European Bioinformatics Institute database by using the keywords "motility," "flagellar," and "chemotaxis." Redundancies in the retrieved set of protein sequences were removed by clustering with USEARCH8 (44).

Chemical Mutagenesis and Screening for Nonmotile *Exiguobacterium* Mutants. *E. acetylicum* was mutagenized with 2.5–20 mg/mL EMS or *N*-ethyl-*N*-nitrosourea (ENU) (Sigma-Aldrich) in PBS for 1 h. Mutant bacteria were inoculated in 0.3% BHI agar plates. After overnight incubation, bacteria at the initial site of inoculation were collected and reinoculated in the center of another 0.3% agar plate to enrich for nonswarming mutants (45). After three rounds of enrichment, individual bacterial colonies were tested in a 96-well plating assay (46) to identify mutants that failed to swarm in low-percentage agar.

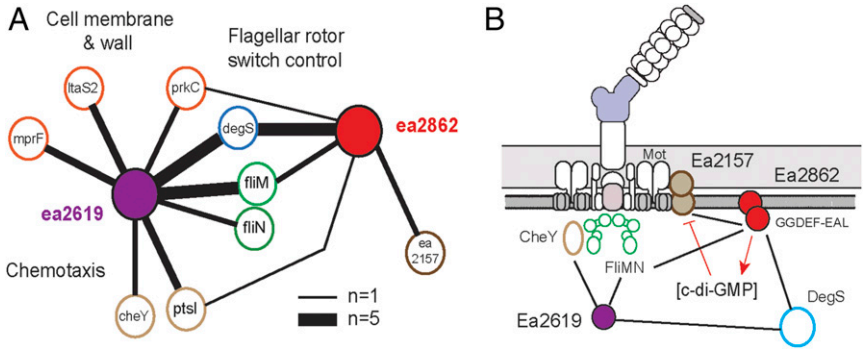

Fig. 4. Extragenic suppression analysis of nonmotile *E. acetylicum* mutants identifies a role for cell-wall modifications and c-di-GMP sensing in commensal Firmicutes motility. (*A*) Genetic suppressors of loss-of-function alleles in two novel motility genes reveal that loss of motility in mutants defective for ea2619 and ea2862 can be bypassed by changes in flagellar rotor switch control and chemotaxis. Motile variants of mutants bearing independent nonsense alleles of *ea2619* or *ea2862* were isolated. Common suppressor mutations (open circles) were mapped to chemotaxis genes (brown), rotor control genes (green), regulators of flagellar gene transcription (blue), and cell membrane homeostasis (orange). Thickness of lines connecting nodes is proportional to the number of independent suppressor mutations identified (Dataset S3). (*B*) Schematic of suppressor mutations linking Ea2619 and Ea2862 to the regulation of swimming speed and direction. Ea2157 indicate a direct link for the levels of c-di-GMP regulated by Ea2862 as central to the control of flagellar motility in *E. acetylicum* (red arrows). *ea2157* was independently identified by MEAPS as a putative motility gene.

Computational Modeling, Genomic Sequencing, and Mutational Enrichment Analysis. These methods and associated code for computational modeling are described in *SI Methods*.

Isolation of Spontaneous Suppressors of Nonmotile *E. acetylicum* Strains. Each nonmotile mutant strain was grown independently in BHI broth, placed in the center of 0.3% agar BHI plates, and incubated overnight. Bacteria at the edge of the growth zone were collected and reinoculated in the center of another 0.3% agar plate. The cycle was repeated until a clear enrichment for motile variants was observed.

Gnotobiotic Zebrafish Colonization. All zebrafish experiments were conducted in conformity with the Public Health Service Policy on Humane Care and Use of Laboratory Animals using protocols approved by the Institutional Animal Care and Use Committee of Duke University. Germ-free zebrafish were reared in sterile tissue culture flasks. At 6 d postfertilization, overnight shaking cultures of *E. acetylicum* motile and nonmotile rifampin-resistant and -sensitive strains were mixed in 1:500 ratios. After 30 min of immersion in the *E. acetylicum* strain mixtures, animals were rinsed in growth medium. Bacterial loads were determined at 0 or 3 d postinoculation (dpi). The number of colony-forming units was assessed after overnight growth at 28 °C.

Other methods and statistical approaches are described in *SI Methods*.

ACKNOWLEDGMENTS. We thank Joshua Granek and Joseph Lucas for assistance in statistical methods; Tim Reddy for helpful discussions; Per Malkus for critical reading of the manuscript; Sara Miller for technical support with electron microscopy; and the Duke GCB Sequencing and Genomic Technologies Core. This work was supported in part by pilot funding from the Duke Center for the Genomics of Microbial Systems and National Institutes of Health Grant R01-DK081426.

1. McFall-Ngai M, et al. (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci USA* 110(9):3229–3236.
2. Nieuwdorp M, Gilijamse PW, Pai N, Kaplan LM (2014) Role of the microbiome in energy regulation and metabolism. *Gastroenterology* 146(6):1525–1533.
3. Semova I, et al. (2012) Microbiota regulate intestinal absorption and metabolism of fatty acids in the zebrafish. *Cell Host Microbe* 12(3):277–288.
4. Goodman AL, et al. (2009) Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* 6(3):279–289.
5. Wu M, et al. (2015) Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides. *Science* 350(6256):aac5992.
6. Rawls JF, Mahowald MA, Ley RE, Gordon JI (2006) Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell* 127(2):423–433.
7. Vishnivetskaya TA, et al. (2014) Draft genome sequences of 10 strains of the genus exiguobacterium. *Genome Announc* 2(5):e01058-14.
8. Erhardt M (2016) Strategies to block bacterial pathogenesis by interference with motility and chemotaxis. *Curr Top Microbiol Immunol*, 10.1007/82_2016_493.
9. Rawls JF, Mahowald MA, Goodman AL, Trent CM, Gordon JI (2007) In vivo imaging and genetic analysis link bacterial motility and symbiosis in the zebrafish gut. *Proc Natl Acad Sci USA* 104(18):7622–7627.
10. Cullender TC, et al. (2013) Innate and adaptive immunity interact to quench microbiome flagellar motility in the gut. *Cell Host Microbe* 14(5):571–581.
11. Schwab C, et al. (2014) Longitudinal study of murine microbiota activity and interactions with the host during acute inflammation and recovery. *ISME J* 8(5):1101–1114.
12. Mukherjee S, Kearns DB (2014) The structure and regulation of flagella in *Bacillus subtilis*. *Annu Rev Genet* 48:319–340.
13. Carpenter PB, Ordal GW (1993) *Bacillus subtilis* FlhA: A flagellar protein related to a new family of signal-transducing receptors. *Mol Microbiol* 7(5):735–743.
14. Zuberi AR, Ying C, Bischoff DS, Ordal GW (1991) Gene-protein relationships in the flagellar hook-basal body complex of *Bacillus subtilis*: Sequences of the flgB, flgC, flgG, fliE and fliF genes. *Gene* 101(1):23–31.
15. Zuberi AR, Bischoff DS, Ordal GW (1991) Nucleotide sequence and characterization of a *Bacillus subtilis* gene encoding a flagellar switch protein. *J Bacteriol* 173(2):710–719.
16. Cairns LS, et al. (2014) FlgN is required for flagellum-based motility by *Bacillus subtilis*. *J Bacteriol* 196(12):2216–2226.
17. Schirmer T, Jenal U (2009) Structural and mechanistic determinants of c-di-GMP signalling. *Nat Rev Microbiol* 7(10):724–735.
18. Garti-Levi S, Hazan R, Kain J, Fujita M, Ben-Yehuda S (2008) The FtsEX ABC transporter directs cellular differentiation in *Bacillus subtilis*. *Mol Microbiol* 69(4):1018–1028.
19. Meisner J, et al. (2013) FtsEX is required for CwlO peptidoglycan hydrolase activity during cell wall elongation in *Bacillus subtilis*. *Mol Microbiol* 89(6):1069–1083.
20. Chen R, Guttenplan SB, Blair KM, Kearns DB (2009) Role of the sigmaD-dependent autolysins in *Bacillus subtilis* population heterogeneity. *J Bacteriol* 191(18):5775–5784.
21. Kempf MJ, McBride MJ (2000) Transposon insertions in the *Flavobacterium johnsoniae* ftsX gene disrupt gliding motility and cell division. *J Bacteriol* 182(6):1671–1679.
22. Szurmant H, Muff TJ, Ordal GW (2004) *Bacillus subtilis* CheC and FliY are members of a novel class of CheY-P-hydrolyzing proteins in the chemotactic signal transduction cascade. *J Biol Chem* 279(21):21787–21792.
23. Neumann S, Grosse K, Sourjik V (2012) Chemotactic signaling via carbohydrate phosphotransferase systems in *Escherichia coli*. *Proc Natl Acad Sci USA* 109(30):12159–12164.
24. Amati G, Bisicchia P, Galizzi A (2004) DegU-P represses expression of the motility flache operon in *Bacillus subtilis*. *J Bacteriol* 186(18):6003–6014.
25. Paul K, Nieto V, Carlquist WC, Blair DF, Harshey RM (2010) The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism. *Mol Cell* 38(1):128–139.
26. Park SY, Lowder B, Bilwes AM, Blair DF, Crane BR (2006) Structure of FliM provides insight into assembly of the switch complex in the bacterial flagella motor. *Proc Natl Acad Sci USA* 103(32):11886–11891.
27. Pandini A, Kleinjung J, Rasool S, Khan S (2015) Coevolved mutations reveal distinct architectures for two core proteins in the bacterial flagellar motor. *PLoS One* 10(11):e0142407.
28. Libby EA, Goss LA, Dworkin J (2015) The eukaryotic-like Ser/Thr kinase PrkC regulates the essential WalRK two-component system in *Bacillus subtilis*. *PLoS Genet* 11(6):e1005275.
29. Levefaudes M, et al. (2015) Diaminopimelic acid amidation in corynebacteriales: New insights into the role of LtsA in peptidoglycan modification. *J Biol Chem* 290(21):13079–13094.
30. Samant S, Hsu FF, Neyfakh AA, Lee H (2009) The *Bacillus anthracis* protein MprF is required for synthesis of lysylphosphatidylglycerols and for resistance to cationic antimicrobial peptides. *J Bacteriol* 191(4):1311–1319.
31. Boehm A, et al. (2010) Second messenger-mediated adjustment of bacterial swimming velocity. *Cell* 141(1):107–116.
32. Guttenplan SB, Kearns DB (2013) Regulation of flagellar motility during biofilm formation. *FEMS Microbiol Rev* 37(6):849–871.
33. Henry JT, Crosson S (2011) Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annu Rev Microbiol* 65:261–286.
34. Tuckerman JR, et al. (2009) An oxygen-sensing diguanylate cyclase and phosphodiesterase couple for c-di-GMP control. *Biochemistry* 48(41):9764–9774.
35. Gao X, et al. (2013) Functional characterization of core components of the *Bacillus subtilis* cyclic-di-GMP signaling pathway. *J Bacteriol* 195(21):4782–4792.
36. Chen Y, Chai Y, Guo JH, Losick R (2012) Evidence for cyclic Di-GMP-mediated signaling in *Bacillus subtilis*. *J Bacteriol* 194(18):5080–5090.
37. Ryan RP, An SQ, Allan JH, McCarthy Y, Dow JM (2015) The DSF family of cell-cell signals: An expanding class of bacterial virulence regulators. *PLoS Pathog* 11(7):e1004986.
38. Nguyen BD, Valdivia RH (2012) Virulence determinants in the obligate intracellular pathogen *Chlamydia trachomatis* revealed by forward genetic approaches. *Proc Natl Acad Sci USA* 109(4):1263–1268.
39. Robins WP, Faruque SM, Mekalanos JJ (2013) Coupling mutagenesis and parallel deep sequencing to probe essential residues in a genome or gene. *Proc Natl Acad Sci USA* 110(9):E848–E857.
40. Rahn-Lee L, et al. (2015) A genetic strategy for probing the functional diversity of magnetosome formation. *PLoS Genet* 11(1):e1004811.
41. Wride DA, et al. (2014) Confirmation of the cellular targets of benomyl and rapamycin using next-generation sequencing of resistant mutants in *S. cerevisiae*. *Mol Biosyst* 10(12):3179–3187.
42. Harper M, Gronenberg L, Liao J, Lee C (2014) Comprehensive detection of genes causing a phenotype using phenotype sequencing and pathway analysis. *PLoS One* 9(2):e88072.
43. Seemann T (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30(14):2068–2069.
44. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19):2460–2461.
45. Girgis HS, Liu Y, Ryu WS, Tavazoie S (2007) A comprehensive genetic characterization of bacterial motility. *PLoS Genet* 3(9):1644–1660.
46. Malapaka VR, Barrese AA, Tripp BC, Tripp BC (2007) High-throughput screening for antimicrobial compounds using a 96-well format bacterial motility absorbance assay. *J Biomol Screen* 12(6):849–854.
47. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.
48. Wei Z, Wang W, Hu P, Lyon GJ, Hakonarson H (2011) SNVer: A statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. *Nucleic Acids Res* 39(19):e132.
49. Pham LN, Kanther M, Semova I, Rawls JF (2008) Methods for generating and colonizing gnotobiotic zebrafish. *Nat Protoc* 3(12):1862–1875.