RESEARCH ARTICLE

Eco-Evolutionary Dynamics of Episomes among Ecologically Cohesive Bacterial Populations

Hong Xue,[a] Otto X. Cordero,[b] Francisco M. Camas,[c] William Trimble,[d] Folker Meyer,[d] Julien Guglielmini,[e,f] Eduardo P. C. Rocha,[e,f] Martin F. Polz[a]

Parsons Laboratory for Environmental Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA[a]; Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland[b]; Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA[c]; Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois, USA[d]; Microbial Evolutionary Genomics, Institut Pasteur, Paris, France[e]; CNRS, UMR3525, Paris, France[f]

H.X. and O.X.C. contributed equally to this work.

ABSTRACT **Although plasmids and other episomes are recognized as key players in horizontal gene transfer among microbes, their diversity and dynamics among ecologically structured host populations in the wild remain poorly understood. Here, we show that natural populations of marine *Vibrionaceae* bacteria host large numbers of families of episomes, consisting of plasmids and a surprisingly high fraction of plasmid-like temperate phages. Episomes are unevenly distributed among host populations, and contrary to the notion that high-density communities in biofilms act as hot spots of gene transfer, we identified a strong bias for episomes to occur in free-living as opposed to particle-attached cells. Mapping of episomal families onto host phylogeny shows that, with the exception of all phage and a few plasmid families, most are of recent evolutionary origin and appear to have spread rapidly by horizontal transfer. Such high eco-evolutionary turnover is particularly surprising for plasmids that are, based on previously suggested categorization, putatively nontransmissible, indicating that this type of plasmid is indeed frequently transferred by currently unknown mechanisms. Finally, analysis of recent gene transfer among plasmids reveals a network of extensive exchange connecting nearly all episomes. Genes functioning in plasmid transfer and maintenance are frequently exchanged, suggesting that plasmids can be rapidly transformed from one category to another. The broad distribution of episomes among distantly related hosts and the observed promiscuous recombination patterns show how episomes can offer their hosts rapid assembly and dissemination of novel functions.**

IMPORTANCE **Plasmids and other episomes are an integral part of bacterial biology in all environments, yet their study is heavily biased toward their role as vectors for antibiotic resistance genes. This study presents a comprehensive analysis of all episomes within several coexisting bacterial populations of *Vibrionaceae* from the coastal ocean and represents the largest-yet genomic survey of episomes from a single bacterial family. The host population framework allows analysis of the eco-evolutionary dynamics at unprecedented resolution, yielding several unexpected results. These include (i) discovery of novel, nonintegrative temperate phages, (ii) revision of a class of episomes, previously termed "nontransmissible," as highly transmissible, and (iii) surprisingly high evolutionary turnover of episomes, manifest as frequent birth, spread, and loss.**

Received 4 April 2015 **Accepted** 7 April 2015 **Published** 5 May 2015

Citation Xue H, Cordero OX, Camas FM, Trimble W, Meyer F, Guglielmini J, Rocha EPC, Polz MF. 2015. Eco-evolutionary dynamics of episomes among ecologically cohesive bacterial populations. mBio 6(3):e00552-15. doi:10.1128/mBio.00552-15.

Editor Edward G. Ruby, University of Wisconsin-Madison

Address correspondence to Martin F. Polz, mpolz@mit.edu.

This article is a direct contribution from a Fellow of the American Academy of Microbiology.

Most studies on the diversity of plasmids and other episomes have focused on their role as major conduits for the spread of resistance and virulence genes in pathogenic bacteria (1, 2). Only recently have whole genome sequencing of microbial hosts and direct extraction of episomes from environmental samples provided a more unbiased glimpse at their large diversity (3–8). This has shown that, although plasmids are best known for their ability to self-transfer among hosts (9), such conjugative plasmids have recently been suggested to be relatively rare in *Proteobacteria* (3). Most plasmids have been categorized as mobilizable or nontransmissible since they contain only genes that enable them to hitchhike with a conjugative plasmid or have no recognizable transfer function (3). Much remains, however, to be learned about the ecological and evolutionary dynamics of these different types of plasmids in the wild, such as their host range, nucleotide and gene content variation, and frequency and persistence within host populations. Even less well studied are extrachromosomal temperate phages that replicate as plasmid-like structures during the lysogenic phase of their life cycle. Examples of such phage episomes are some *Tectiviridae*, which have been found as linear plasmids in *Bacillus* species (10), and phage N15, which is a relative of phage lambda (11). Interestingly, the role of these phages in the

spread of antibiotic resistance in *Escherichia coli* is currently being reevaluated, since a P1-like phage was found to carry an extended-spectrum β-lactamase (12) and a large-scale analysis of the episome content of antibiotic-resistant strains revealed several extrachromosomal prophages (13). Because plasmid and phage episomes play roles as molecular symbionts or parasites (14) and can mediate horizontal gene exchange (15), their biology must ultimately be studied in the context of the host populations they invade; however, this has remained difficult due to the dearth of suitable model systems of ecologically and genotypically well-constrained bacterial populations.

Here, we take a population-genomic approach to determine carriage of different types of episomes in a recently established model for ecologically and genetically cohesive bacterial populations, asking whether different episomal types (i) are associated primarily to host phylogeny or ecology, (ii) show evidence for distinct transfer (and loss) patterns, and (iii) display different microevolutionary patterns. We use marine bacteria of the family *Vibrionaceae* as our model for environmentally differentiated host populations. These have previously been identified as genotypic clusters with characteristic distribution among environmental samples from the same geographic location, suggesting that they partition resources in the coastal ocean by differential occurrence among the free-living and associated (with suspended organic particles and zooplankton) fractions of bacterioplankton (16–18). Many of these populations do, however, also cooccur on the surfaces and in the guts of filter-feeding and other marine animals (19), providing opportunity for transfer of episomes via occasional contact. Finally, recent analysis of recombination has indicated that these ecological populations display cohesive behavior in terms of gene flow, making it possible for adaptive genes to spread in a population-specific manner (20, 21). Because of these properties, these clusters are hypothesized to represent natural populations and provide a platform to study the diversity and dynamics of episomes.

To explore the diversity of episomes within host populations, we screened a large collection of ecologically characterized *Vibrionaceae* isolates obtained from the coastal ocean in the spring and fall of 2006 (16). We aimed at comprehensively sampling and sequencing all detectable episomes of different sizes to obtain a picture of their diversity as unbiased as possible. Episomes were analyzed in a comparative genomic framework, integrating this analysis with both phylogenetic and habitat information of the bacterial populations in order to identify differential associations and dynamics.

RESULTS AND DISCUSSION

Detection and classification of episomes. We screened 660 *Vibrionaceae* isolates for the presence of episomes using multiple gel electrophoretic assays to resolve DNA of different sizes (see Materials and Methods). This identified 140 DNA bands distributed across 101 of the isolates and varying in size between 1 and 200 kbp. To further investigate these putative episomes, we excised all bands from gels and determined their sequence by the Illumina and 454 technologies. Although in many cases assembly produced single and frequently circular episomes (see Table S1 in the supplemental material), there were instances where several contigs resulted from a single band on electrophoresis gels. Because this may indicate incomplete assembly of a single episome or comigration of multiple, similarly sized episomes, we used additional information contained in the data to differentiate these possibilities. To test for multiple episomes per band, we developed a bioinformatics pipeline that considered whether (i) the combined length of the contigs considerably exceeded band size, (ii) some of the contigs had high similarity to episomes in other *Vibrionaceae* isolates, and (iii) coexisting contigs displayed substantially different coverage (see Materials and Methods). This method identified 187 putative episomes (here called "episomes" for simplicity). To classify these episomes into families, we (i) calculated pairwise similarity values by comparison of the nucleotide similarity among shared proteins (orthologs) normalized by the number of proteins of the larger of the two episomes and (ii) established clusters of episomes of high similarity using the orthoMCL (22) graph clustering algorithm (see Materials and Methods). This analysis enabled exploration of the episome eco-evolutionary dynamics among host populations.

Association with host populations and lifestyle. Comparison of episome incidence across the phylogeny of the *Vibrionaceae* hosts shows that (i) they are abundant in a few populations but relatively sparse in most (Fig. 1A) and (ii) their presence is correlated to host lifestyle across populations (Fig. 1B). For example, episomes were not detected in population no. 5 (*Vibrio* sp. F5), 7 (*Vibrio logei*), 9 (*Vibrio breoganii*), and 10 (*Vibrio* sp. F10), whereas episomes were present in all isolates of population no. 6 (*Vibrio* sp. F6) and 60% of the isolates in population no. 14 (*Vibrio kanaloae*). Whether these distribution differences are due to various degrees of selection for or against episomes within different populations, or due to greater transmission efficiency among some populations, is difficult to determine; however, a strong association of episomes with host lifestyle provides additional information. The *Vibrionaceae* populations display various degrees of free-living or associated existence (e.g., with organic particles, zooplankton). This is expressed in our data as the presence in one of four sequential size fractions, where the large-size fractions contain microbes attached to particles or organisms while the smallest-size fraction contains only unattached, free-living cells (16) (Fig. 1A). We tested whether episomes were enriched in one or more size fractions by calculating the phylogenetic correlation between episome carriage and the association to each of the size fractions (see Materials and Methods). Our approach controls for spurious correlations caused by phylogenetic clustering and calculates confidence intervals based on 100 bootstrap trees of the *hsp60* gene marker used to demarcate *Vibrio* populations. Surprisingly, this analysis showed episome-positive strains to be significantly and strongly biased for the <1-μm size fraction, corresponding to occurrence as free-living cells (Fig. 1B). This association counters the previous suggestion that particle-associated bacteria, which live in diverse and dense communities, are more prone to acquire mobile elements (23, 24). Acquisition of episomes could happen in animal guts within which most of the populations have the potential to encounter each other (19); however, it seems likely that the high incidence of episomes in free-living cells reflects stability within the host and/or environmental selection rather than high transmission.

Episome categories. Sequence annotation identified putative phages and plasmids as the two main episomal categories within the *Vibrionaceae* populations (see Table S1 in the supplemental material). The first was present at a surprisingly high level (22 of 187 episomes) and consisted primarily of two families when a cutoff of 70% sequence similarity was used to define episomal

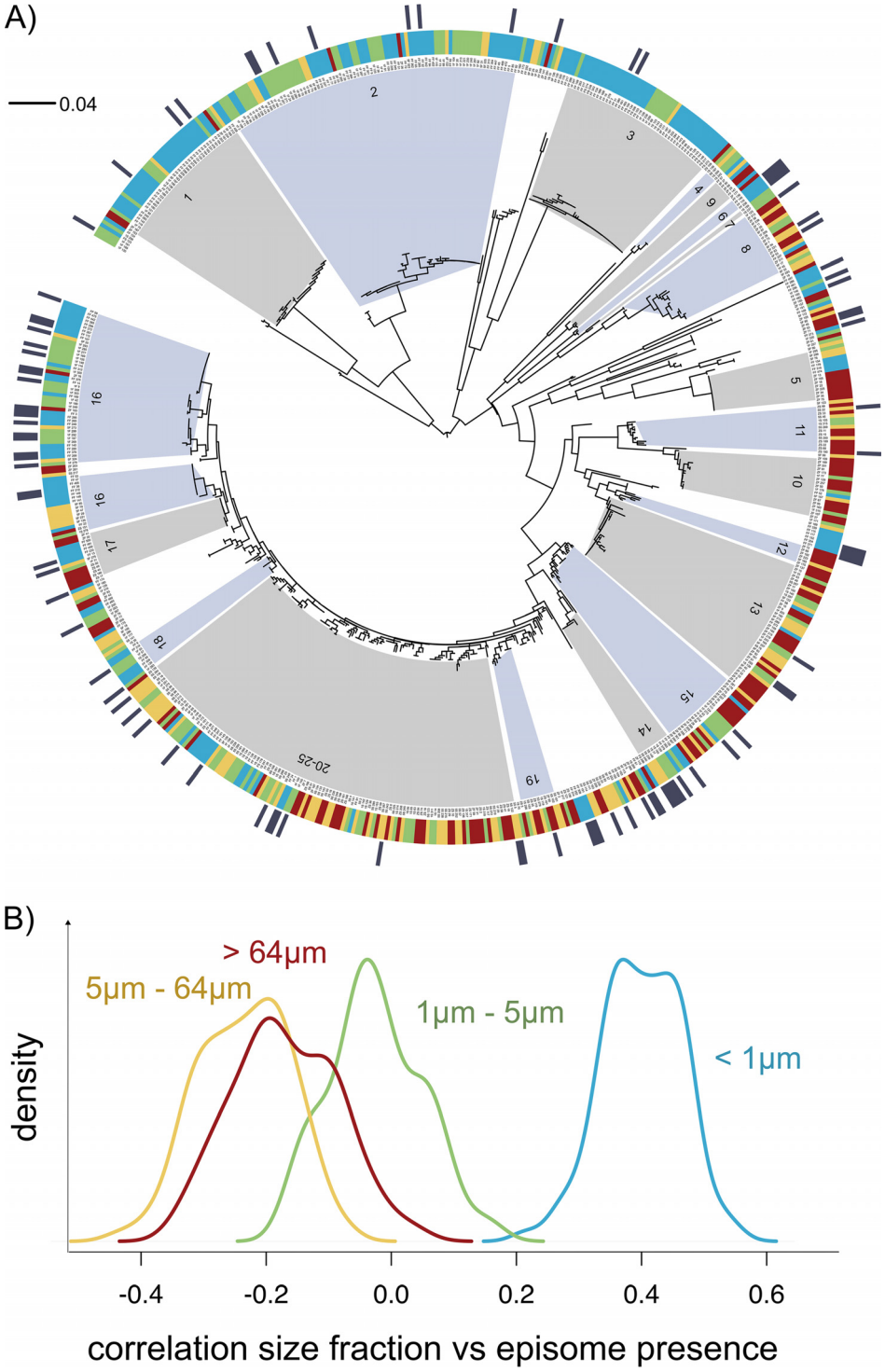

FIG 1 Distribution of episomes on the *Vibrionaceae* phylogeny and relation to environmental metadata. (A) Phylogeny based on the *hsp60* protein-coding gene. *Vibrio* genotypes were isolated from size-fractionated seawater, and colored rings indicate the corresponding size fraction for each isolate (fraction labels in panel B). Dark bars indicate the presence of at least 1 episome. Populations boundaries are indicated by shaded areas, and the closest named species for each population are as follows: 1, *Enterovibrio calviensis*; 2, *Enterovibrio norvegicus*; 3, *Vibrio ordalii*; 4, *Vibrio rumoiensis*-like; 5, *Vibrio* sp. F5; 6, *Vibrio* sp. F6; 7, *Vibrio logei*; 8, *Vibrio fischeri*; 9, *Vibrio breoganii*; 10, *Vibrio* sp. F10; 11, *Vibrio splendidus* cluster 1; 12, *Vibrio* sp. F13; 13, *Vibrio* sp. nov.; 14, *Vibrio kanaloae*; 15, *Vibrio cyclitrophicus*; 16 and 17, *Vibrio tasmaniensis*; and 18 to 25, *Vibrio splendidus*. Taxonomic assignments are as in reference 67 with the exception of population no. 12 and 13, which have been reassigned based on recent genomic comparisons. (B) Phylogenetic correlation between size fractions and presence of episomes. We calculate correlations on the phylogeny using a modified version of the phylogenetic contrast method (54), which allows us to estimate evolutionary linkage between traits (e.g., having an episome and association to one of the size fractions). The correlations are shown as frequency distributions because of the uncertainty in phylogenetic structure. Looking at the position of the distributions on the horizontal axis, we observe that episomes are strongly biased to occur in the free-living lifestyle (occurrence in the smallest-size fraction) and less in the large-size fractions.

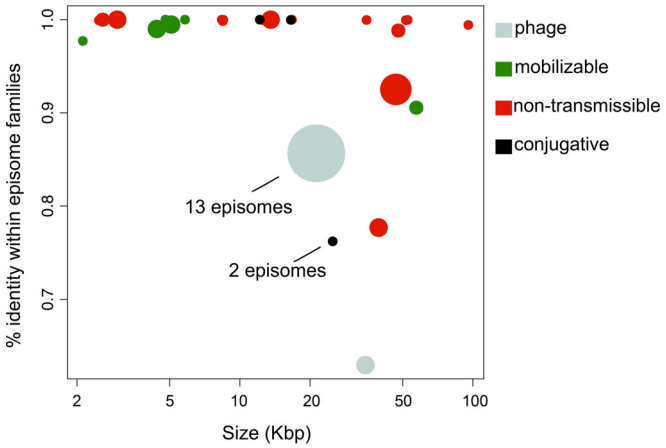

FIG 2 Episomal family age versus size. Average percentage identities are calculated as a proxy for episome family age and plotted against the size of the element in base pairs. The size of the points indicates the number of members in each episome family, which ranges from 2 to 13. Colors indicate episome classification. The analysis shows that most episome families, irrespective of size, are evolutionarily young (little or no DNA divergence).

families. These two have genome sizes ~20 and 40 kbp (see Table S1), with the smaller representing the most numerous episomal family (Fig. 2). Other phage-like episomes were found only once, including one ~80 kbp and four <14 kbp in size. All of these putative phages appear novel, since none were closely related to known phages. Moreover, because of their propagation within the cells during culturing, they presumably represent temperate phages that replicate in a plasmid-like fashion and may not integrate into the host genome, since we found evidence of neither integrase genes nor related integrated phages in ~80 *Vibrionaceae* genomes available from the same collection (25). High proportions of phage-like elements have also been reported from recent plasmid metagenomic studies (8). Such high prevalence suggests an alternative interpretation to the observation of extrachromosomal phage sequences in metagenomic analysis of microbial communities in the ocean. These have been suggested to be lytic phages caught in the act of infecting cells (26) but might, in at least some cases, be plasmid-like lysogens.

Plasmids, the other major category, comprised the majority of episomes and could be divided into conjugative, mobilizable, and nontransmissible according to a previously proposed scheme based on the presence/absence of signature genes (3) (see Fig. S1A and Table S1 in the supplemental material). Twenty-four episomes were judged to be conjugative plasmids since they contained at least 5 key genes of a type IV secretion system (T4SS), which, in combination with a relaxase, is necessary for self-transmission (27). These plasmids are the largest (average of ~60 kbp; see Fig. S1) and carry a high density of genes. Mobilizable plasmids, on the other hand, encode only a relaxase and hence presumably require a T4SS to act in *trans* for mobilization, most likely from a cooccurring conjugative plasmid (3). The 38 plasmids categorized as mobilizable encode, with few exceptions, only relatively few open reading frames (ORFs) and had similar, small average sizes (~11 kbp) (see Fig. S1A). We also detected 103 plasmids lacking relaxases and T4SS and thus classified as putatively nontransmissible. The means of transmission of these elements are usually unclear even though they constitute the majority of the known plasmids in *Proteobacteria* (3). Nontransmissible plasmids displayed large size variation, from a few to over 100 kbp (with an average of ~20 kbp; see Fig. S1A), and were, with 62%, the dominant plasmid category, while conjugative and mobilizable plasmids occurred at 15% and 23%, respectively. This frequency distribution is fairly similar to the proteobacterial average, which is ~20%, 30%, and 50% for conjugative, mobilizable, and nontransmissible plasmids, respectively (3).

Genetic diversity. The episomes detected in this study carry a diversity of genes, albeit with 43% (2,043 ORFs), the largest portion are hypotheticals when annotated using both RAST (28) and the ACLAME database (29) (see Materials and Methods). This is consistent with the notion that mobile elements are enriched in genes with poorly understood function (30, 31). The most important category of known functions is membrane transport, with 296, 70, and 10 genes annotated as members of T4SS, type 6 secretion systems (T6SS), and ABC transporters, respectively. As mentioned above, T4SS is most likely involved in conjugative transfer, and of the 27 T4SS detected, 19 were type F, 6 type T, and 2 type G. Conjugation systems of the F-type have thin, flexible pili that allow high frequency of conjugation in liquid media (32), while type T pili are rigid and are thought to perform better on surfaces (33), providing some indirect support for the biased occurrence (and potential transfer) of episomes among free-living hosts (Fig. 1B). On the other hand, T6SS can inject protein effectors into bacterial and eukaryotic cells and hence likely play a role in predation, pathogenesis, or predation defense (34). Finally, ABC transporters can catalyze translocation of a variety of molecules, including proteins, metabolites, and metals (35, 36). A further 178 proteins are involved in functions ascribed to plasmid maintenance, including 50 resolvases, 29 replicases, 91 partitioning systems, 41 toxin-antitoxin systems, and 17 restriction-modification systems. The large number of partitioning systems detected may indicate that more than half of the plasmids might be low copy number, since partitioning mechanisms are often absent from high-copy-number plasmids (37).

As in previous studies (15, 38, 39), general annotations indicate functions with possible host benefit and highlight the potential role of plasmids in horizontal transfer of a wide variety of genes. Among these are genes involved in the metabolism of amino acids (14 proteins) and carbohydrates (21 proteins) and stress response (21 genes). An example of a full pathway with potential host benefit is the detection of a siderophore gene cluster in a family of large, nontransmissible plasmids. Interestingly, we identified three 5S-rRNA and two tRNA genes. These are embedded in contigs that do not contain any additional ribosomal components and are only ~90% similar in sequence to the equivalents in their host strains, so that the detection of these informational genes is unlikely to be due to host chromosome contamination. Their presence therefore confirms previous findings that plasmids can occasionally contain rRNA genes (40, 41) and can act as transfer vehicles for genes that are thought to be only infrequently involved in exchange among distantly related organisms (42, 43).

Episome cooccurrences within hosts. Most host cells containing episomes harbored either a single (~60%) or two (~30%) episomes; however, several isolates contained a large number of episomes, and some of these represented unusual combinations (see Fig. S1B in the supplemental material). For example, annotation suggested that in strain FF472, a conjugative, two mobilizable, and four nontransmissible plasmids were present along with

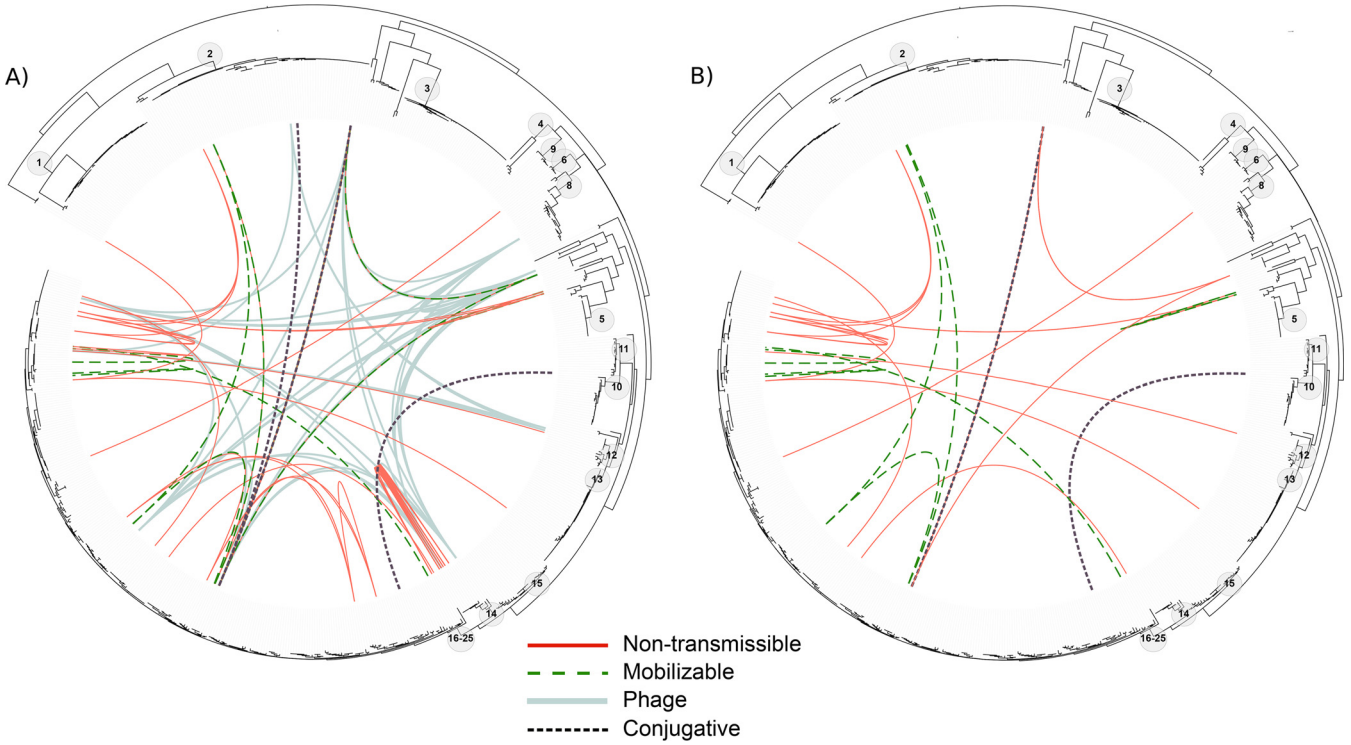

FIG 3 Episome family network across the *Vibrionaceae* phylogeny calculated for different nucleotide similarities as cutoffs for family membership. The phylogeny is annotated (bubbles with population identifiers matching those in Fig. 1) to indicate the origin of the known ecological populations. Links connect strains that share episomes in the same family. Colors of links indicate whether an episome family is putatively classified as phage or conjugative, mobilizable, or nontransmissible plasmid. Episome families were defined with 70% (A) and 97% (B) nucleotide similarity cutoff (as a reference, the average gene content overlap between unrelated strains is only 40% [25]). The analysis shows that episomes are distributed among distantly related hosts, indicating spread by horizontal gene transfer. Restriction to families with only closely related members (97% sequence identity) preserves this pattern for most episomes except phage and some conjugative and nontransmissible plasmids.

a phage (see Table S2 in the supplemental material). Another strain (FF112) contained a phage and two different types of conjugative plasmids. Finally, systematic exploration of cooccurrence patterns across host isolates did not suggest any obvious codependencies of episomes on each other, since many were also detected as single episomes and never in the same combinations in multiple hosts.

Inferred inheritance dynamics. Our data also allow estimation of the inheritance dynamics of episomes within and between populations. Considering the dominance of nontransmissible episomes among the *Vibrionaceae* populations and *Proteobacteria* in general (3), we were particularly interested in whether these plasmids are primarily vertically inherited and hence present in closely related isolates or whether there is evidence for their transfer among distantly related host populations. To differentiate these possibilities, we first classified episomes into families based on sequence similarity (Fig. 2; see Table S3 in the supplemental material) and then constructed a network visualizing the occurrence of these families on the phylogeny of their hosts (see Materials and Methods) (Fig. 3).

This analysis suggests that episomes have spread primarily by horizontal transmission rather than vertical inheritance (Fig. 3A). We identified 31 multimember families (with family size ranging from 2 to 13 members), while 94 episomes remained singletons, suggesting that there is a large pool of rare episomal types within these populations. In fact, low frequencies of genetic variants in a lineage are usually the result of recent introductions, suggesting that many of these elements have been recently transferred. Such transfers could originate from sporadic contact with other microbes on particles or guts of animals. Surprisingly, only a few families, most notably the two containing phage, have accumulated high nucleotide diversity, while most consist of highly similar elements with, on average, ≥98% nucleotide identity (Fig. 2). Mapping the occurrence of these families onto the host phylogeny shows that the majority are distributed across distantly related hosts (overall gene content overlap of <40% [25]), implying that episomes have spread horizontally. This is the case for all phages, which may therefore possess broad host range, and, surprisingly, also for a large number of nontransmissible plasmids (Fig. 3A), whose horizontal transfer was previously suggested to depend on chance events and hence to be rare.

Further consideration of inheritance patterns of episomes suggests that many, but especially nontransmissible, plasmids are subject to rapid evolutionary turnover, i.e., they arise, spread, and are lost frequently. This conclusion is based on restricting the network analysis to families with high sequence similarity (>97%) and reanalyzing their distribution. The resultant network shows that a very high percentage of episomes that have been transferred among *Vibrionaceae* populations are closely related (Fig. 3B). In fact, many of these have identical nucleotide sequences, suggesting that they have spread in a time frame that has not permitted the accumulation of nucleotide changes (Fig. 2).

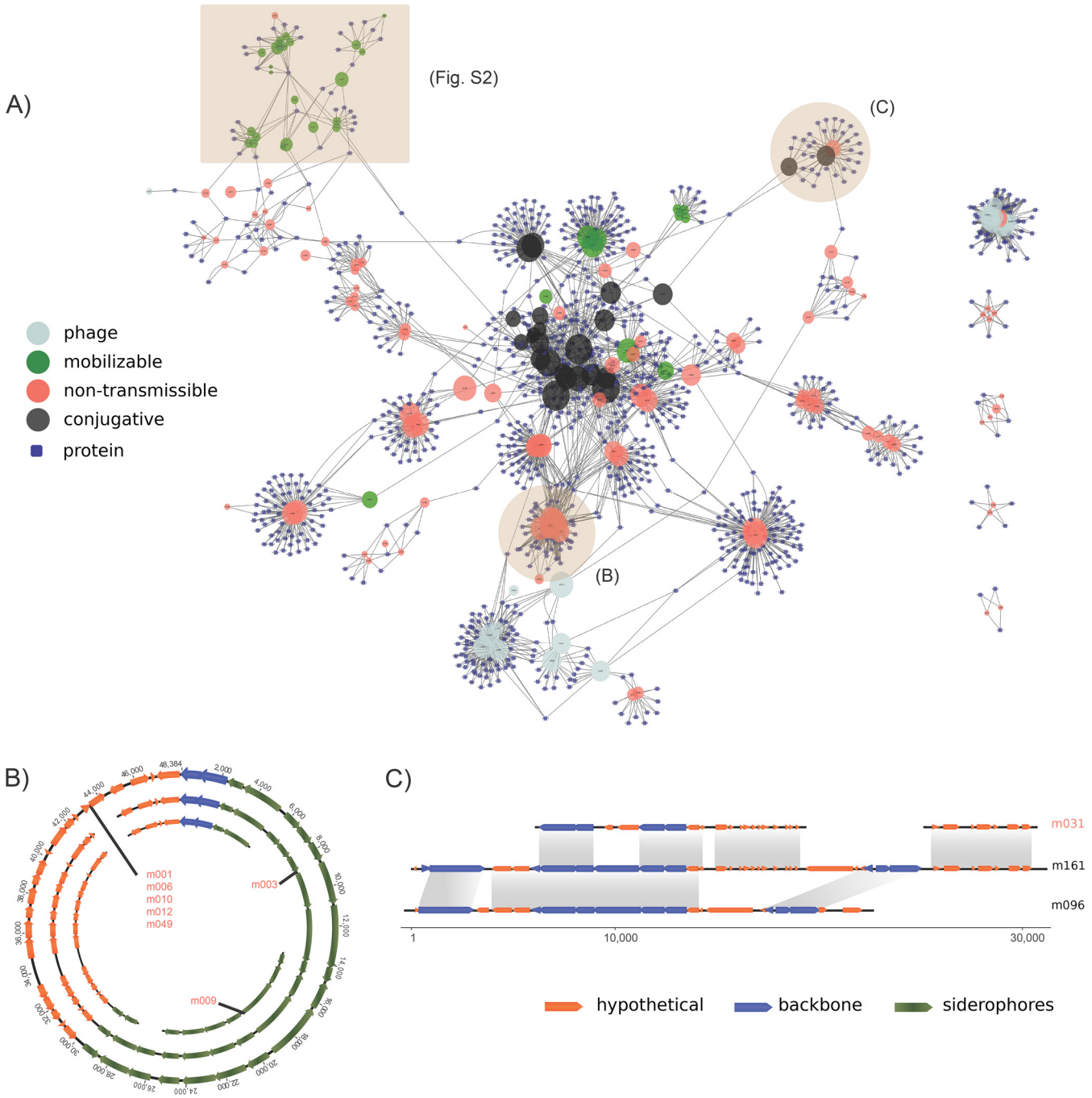

FIG 4 (A) Network of recent horizontal gene transfer among episomes. Episomes are connected by proteins (blue dots) shared by at least two episomes at ≥97% sequence similarity. The diameter of episome symbols indicates the size of the genome. The analysis shows that nearly all episomes have exchanged genes with a cluster of conjugative plasmids forming a hub at the center. (B) Family of nontransmissible plasmids containing siderophore biosynthesis genes. These elements are characterized by the absence of genes involved in self-transmission and have partitioning systems only in their backbones. (C) Gene content comparison of a mixed episome family (containing both conjugative and nontransmissible) reveals that the two episome categories can evolve from each other by either gain or loss of the conjugation machinery.

This pattern is especially puzzling for "nontransmissible" plasmids and suggests that a currently unrecognized, direct transfer mechanism enables their rapid dissemination among bacterial populations. Several such mechanisms have been proposed. They include DNA vesicles (44, 45), nanotubes (46), natural transformation (47), and transduction (48).

Gene exchange among episomes. Because of the apparently rapid turnover of episomes, we investigated to what extent episomes themselves are evolutionarily stable entities by constructing a network of recently exchanged genes (Fig. 4A). To restrict the analysis to events of fairly recent transfer, we first clustered genes into closely related families (>97% in sequence identity) and then

determined how many episomes share these families. This shows a network that connects most episomal families by recent gene exchange with a hub of strongly connected conjugative plasmids at the center (Fig. 4A). Although these share many types of genes, T4SS appear most frequently shared. Strongly connected to this hub are many nontransmissible plasmids, while the remainder of episomes are relatively sparsely connected. Closer inspection, however, reveals that, with the exception of phage, the limited number of connections has to be seen in the context of the small size of many mobilizable and nontransmissible plasmids. The fact that many of these small plasmids share only backbone genes at high nucleotide similarity, while having completely different functional gene content (Fig. S2), provides evidence for the rapid evolution of these elements consistent with the known modular rapid evolution of mobile genetic elements (49, 50).

Although episomes appear evolutionarily unstable and are subject to frequent reassortment of genes by recombination, the two phage, one conjugative, and two nontransmissible plasmid families are exceptions. We highlight the example of a family of large nontransmissible plasmids mentioned above (Fig. 4B) consisting of three modules. The first encodes a plasmid partition system (blue ORFs) that ensures reliable distribution into daughter cells. The function of the second module (orange) remains unidentified. The third module (green) encodes a protein complex that shares high sequence identity with siderophore biosynthesis genes from a previously characterized plasmid family from *Vibrio vulnificus* and *Vibrio parahaemolyticus* (51), suggesting that these genes are mobile among episomes. However, the nucleotide diversity of 92.5% across the plasmid family identified here suggests that the acquisition of the siderophore operon was a more ancient event and that these nontransmissible plasmids have persisted over longer evolutionary times than many of the other nontransmissible plasmids, which consist of clusters of highly identical genomes.

Gene gain or loss may also change one plasmid category into another and may, in part, explain the rapid evolutionary turnover of most plasmids. For example, episome m161 and m096 are conjugative plasmids that share almost all of the backbone genes (blue ORFs), which are responsible for self-transmission (Fig. 4C). They differ, however, in two relatively large regions, which are present in m161 but absent in m096. These regions are also shared by the nontransmissible plasmid m031, which is overall more similar to m161 except for the lack of genes responsible for conjugative transfer. This confirms the view that plasmid gene repertoires change rapidly (52, 53). It further suggests that nontransmissible plasmids may originate from loss of T4SS and relaxase genes.

Conclusions. Overall, our data reveal surprising results about the diversity and distribution of episomes among *Vibrionaceae* populations. First, contrary to the expectation that, due to the requirement of cell-to-cell contact for transmission, plasmids should be preferentially associated with isolates recovered from surfaces, microcolonies, or biofilms (23, 24), we show that they are significantly enriched in free-living, planktonic cells. Second, our data suggest that plasmids previously categorized as nontransmissible are subject to high evolutionary turnover and transfer frequently among populations. It is therefore likely that a currently unrecognized transfer mechanism is at work. Candidates are transformation, conjugation by cointegration into a conjugative element, and packaging into phage or membrane vesicles. Regardless, we propose that the name "nontransmissible" should be abandoned. Third, the high incidence of putative temperate phages that appear to propagate as plasmids is unexpected. Although phages have previously been described in plasmid metagenomes (8), there was no indication that these might be temperate phages, as suggested here by the stable propagation in our isolates. Such plasmid-like temperate phages have been previously described in only a very small number of studies. Accordingly, the phages detected here appear novel. Their prevalence suggests a previously unanticipated important role in the marine environment. Finally, analysis of gene transfer among plasmids indicates that genes involved in plasmid maintenance and transfer are a frequently exchanged, rapidly changing categorization of plasmids. This exchange also offers host strains a rich supply of external genetic materials that may allow the assembly of different functions on a backbone of plasmid functions perhaps adapted to specific host populations.

MATERIALS AND METHODS

Isolation, sequencing, and assembly of episomes are described in the supplemental material.

Phylogenetic correlation between size fractions and episome incidence. To calculate the evolutionary linkage between traits (such as episome carriage), it is necessary to correct for the fact that traits are linked through a phylogeny and therefore not independently distributed. To obtain independently distributed variables, we calculated the frequency of association with episomes and with size fractions across all clades in the phylogeny and applied the phylogenetic contrast method (54) to remove phylogenetic autocorrelations. The resulting contrast vectors contain independently distributed variables (evolutionary transitions) associated to each branch of the phylogeny. We then used these vectors to calculate Spearman correlations between evolutionary transitions in frequency of episomes and in each of the size classes. Figure 1B shows the distribution of correlation values obtained from repeating this process for 100 bootstrap trees of the *hsp60* gene used for phylogenetic analysis of populations.

Annotation of proteins. We annotated the ORFs and the corresponding function of the encoded proteins (in terns of FIGFAMS and Subsystems) using the RAST tools (28). Ten short bands for which no ORFs could be annotated were removed from any further analysis. From the total set of 5,598 proteins annotated in the remaining bands, we built families of protein orthologs using orthoMCL (22). Using Blast (55) (*BLASTp* with an E value of $\leq$1e$-$10), we compared the protein sequences in our set with those in the database of mobile elements ACLAME (v. 0.4) (29), which includes records for a total of 122,154 proteins from phage (23.1%), prophage (21.2%), and plasmid (55.6%) origin. This way, we labeled each protein as virus or plasmid associated.

Identification of episomes from contigs and delineation of episomal families. Because each band can, in principle, contain more than one independent episome, we developed a bioinformatics pipeline to differentiate episomes that were broken into multiple contigs from multiple episomes comigrating in a gel band (see Text S1 in the supplemental material). This pipeline also enabled identification of episome families distributed across different hosts. Briefly, we constructed a network of sequence similarity based on the gene content overlap across all contigs from all bands. To calculate the similarity measure between pairs of contigs, we normalized the DNA similarity to the size of the largest contig in order to give a maximum score only to perfect matches and not to nested pairs. We then applied the well-known MCL algorithm (56) to cluster contigs based on the patterns of connectivity in the network using a cutoff of 70% identity and an inflation parameter of 1.5. These clusters were taken as our initial episomal families.

We then focused on those pairs of contigs with nested similarity to improve the assembly of multicontig bands where the challenge was to differentiate single episomes broken into multiple contigs from multiple episomes comigrating in the same gel band. When two contigs from the

same band matched two different segments of one larger contig, we took this as evidence to support the consolidation of the two contigs into a single episome. The new putative multicontig episome was validated by checking that joined contigs had similar read coverages using SSAHA2 (57) for mapping the reads and SAMtools (58) to compute the mapping coverage. Multicontig episomes and their matching references were subsequently used to define families by applying variable sequence cutoffs. Episomes without matches in the similarity network were considered singletons.

Classification of episomes. Based on previous identification of broad categories of episomes, we searched our data for evidence of conjugative, mobilizable, and nontransmissible plasmids and of phage. To identify conjugative plasmids, we searched for cooccurrence of genes encoding the type IV secretion system (T4SS) and the relaxosome. Plasmids encoding both components were classified as conjugative, while plasmids encoding a relaxase only were classified as mobilizable. The episomes that did not encode either of these elements and could not be identified as phage (see below) were classified as putatively nontransmissible plasmids following the logic of Smillie et al. (3).

For plasmid identification, we used an update of the protein profiles used in reference 59. For conjugative plasmids, we first searched for matches to TraU/VirB4 from each of the mating pair formation (MPF) system families defined previously (27), since this is the only protein that is associated with all known T4SS (or at least the only sufficiently conserved in sequence). We then gathered all proteins found within a frame of −20/+20 ORFs around TraU/VirB to determine whether a functional T4SS was present. For each MPF type, we carried out similarity searches between all proteins and clustered them into families. These families were aligned, analyzed, and curated. We iterated based on criteria such as sensitivity and specificity and then made multiple alignments that were used to build protein profiles with HMMER (60, 61). This led to a database of protein profiles associated with conjugation that correspond in general to the known essential proteins in each system (albeit a few evolve too fast and give poor sequence similarity hits). The profiles were then searched using HMMER in the proteomes of all episomes. We filtered the results by using an E value of ≤0.01 and coverage (ratio between the target and query lengths) of more than 0.5. Clusters are based on the findings in the entire replicon.

For phage identification, we blasted the episome contigs against the ACLAME Database of Mobile Elements v. 0.4 (29) using BLASTp with an E value of ≤1e−10. We unified the phage and prophage categories into a single "phage" category so that we broadly classified proteins into phage- and plasmid-like. Phage episomes were identified based on the density of proteins matching phage proteins in the ACLAME database. Because some proteins can have homologs in both plasmids and phage (e.g., regulators), we selected those episomes with a clear bias to phage-only proteins, with at least 50% more hits to phage only. Also, episomes were required to have at least 20% of all coding regions annotated as phage. Because most proteins do not have a clear hit to known phage or plasmid genes, this *ad hoc* rule represents a conservative measure. Inspection of the elements classified as phage confirmed that this method allowed us to systematically classify episomes that had clear phage characteristics based on annotation.

Network of gene sharing among families. To determine the relationship among the episome genomes in terms of shared genes, episomes were assigned to groups based on clustering of shared proteins. In brief, ORFs were identified using Glimmer 3.0 (62), followed by the clustering of ORFs using OrthoMCL (63, 64), in which a minimum 97% coverage of the longer sequence was required as well as an E value of ≤1e−5. Whole genomes of episomes were then clustered based on these shared proteins, using the FT ClustNSee clustering algorithm in cytoscape (65, 66).

Nucleotide sequence accession numbers. Sequence data can be retrieved from GenBank under accession numbers KP795445 to KP795714.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00552-15/-/DCSupplemental.

Text S1, DOCX file, 0.2 MB.
Figure S1, DOCX file, 0.1 MB.
Figure S2, DOCX file, 0.1 MB.
Table S1, DOCX file, 0.1 MB.
Table S2, DOCX file, 0.1 MB.
Table S3, DOCX file, 0.1 MB.

ACKNOWLEDGMENT

This work was supported by a grant from the National Science Foundation Biological Oceanography Program.

REFERENCES

1. **Carattoli A.** 2009. Resistance plasmid families in *Enterobacteriaceae*. Antimicrob Agents Chemother **53:**2227–2238. http://dx.doi.org/10.1128/AAC.01707-08.
2. **Johnson TJ, Nolan LK.** 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. Microbiol Mol Biol Rev **73:**750–774. http://dx.doi.org/10.1128/MMBR.00015-09.
3. **Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EP, de la Cruz F.** 2010. Mobility of plasmids. Microbiol Mol Biol Rev **74:**434–452. http://dx.doi.org/10.1128/MMBR.00020-10.
4. **Jones BV, Sun F, Marchesi JR.** 2010. Comparative metagenomic analysis of plasmid encoded functions in the human gut microbiome. BMC Genomics **11:**46. http://dx.doi.org/10.1186/1471-2164-11-46.
5. **Zhang T, Zhang XX, Ye L.** 2011. Plasmid metagenome reveals high levels of antibiotic resistance genes and mobile genetic elements in activated sludge. PLoS One **6:**e26041. http://dx.doi.org/10.1371/journal.pone.0026041.
6. **Brown Kav A, Sasson G, Jami E, Doron-Faigenboim A, Benhar I, Mizrahi I.** 2012. Insights into the bovine rumen plasmidome. Proc Natl Acad Sci U S A **109:**5452–5457. http://dx.doi.org/10.1073/pnas.1116410109.
7. **Sentchilo V, Mayer AP, Guy L, Miyazaki R, Green Tringe S, Barry K, Malfatti S, Goessmann A, Robinson-Rechavi M, van der Meer JR.** 2013. Community-wide plasmid gene mobilization and selection. ISME J **7:**1173–1186. http://dx.doi.org/10.1038/ismej.2013.13.
8. **Sentchilo V, Mayer AP, Guy L, Miyazaki R, Green Tringe S, Barry K, Malfatti S, Goessmann A, Robinson-Rechavi M, van der Meer JR.** 2014. Community-wide plasmid gene mobilization and selection. ISME J **7:**1173–1186. http://dx.doi.org/10.1038/ismej.2013.13.
9. **Aachmann FL, Eijsink VG, Vaaje-Kolstad G.** 2011. 1H, 13C, 15N resonance assignment of the chitin-binding protein CBP21 from *Serratia marcescens*. Biomol NMR Assign **5:**117–119. http://dx.doi.org/10.1007/s12104-010-9281-2.
10. **Verheust C, Jensen G, Mahillon J.** 2003. pGIL01, a linear tectiviral plasmid prophage originating from *Bacillus thuringiensis* serovar Israelensis. Microbiology **149:**2083–2092. http://dx.doi.org/10.1099/mic.0.26307-0.
11. **Ravin NV.** 2011. N15: the linear phage-plasmid. Plasmid **65:**102–109. http://dx.doi.org/10.1016/j.plasmid.2010.12.004.
12. **Billard-Pomares T, Fouteau S, Jacquet ME, Roche D, Barbe V, Castellanos M, Bouet JY, Cruveiller S, Médigue C, Blanco J, Clermont O, Denamur E, Branger C.** 2014. Characterization of a P1-like bacteriophage carrying an SHV-2 extended-spectrum beta-lactamase from an *Escherichia coli* strain. Antimicrob Agents Chemother **58:**6550–6557. http://dx.doi.org/10.1128/AAC.03183-14.
13. **Brolund A, Franzén O, Melefors O, Tegmark-Wisell K, Sandegren L.** 2013. Plasmidome-analysis of ESBL-producing *Escherichia coli* using conventional typing and high-throughput sequencing. PLoS One **8:**e65793. http://dx.doi.org/10.1371/journal.pone.0065793.
14. **Rankin DJ, Rocha EP, Brown SP.** 2011. What traits are carried on mobile genetic elements, and why? Heredity **106:**1–10. http://dx.doi.org/10.1038/hdy.2010.24.
15. **Frost LS, Leplae R, Summers AO, Toussaint A.** 2005. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol **3:**722–732. http://dx.doi.org/10.1038/nrmicro1235.
16. **Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polz MF.** 2008. Resource partitioning and sympatric differentiation among closely related

bacterioplankton. Science **320:**1081–1085. http://dx.doi.org/10.1126/science.1157890.
17. **Preheim SP, Timberlake S, Polz MF.** 2011. Merging taxonomy with ecological population prediction: a case study of *Vibrionaceae*. Appl Environ Microbiol **77:**7195–7206. http://dx.doi.org/10.1128/AEM.00665-11.
18. **Szabo G, Preheim SP, Kauffman KM, David LA, Shapiro J, Alm EJ, Polz MF.** 2013. Reproducibility of *Vibrionaceae* population structure in coastal bacterioplankton. ISME J **7:**509–519. http://dx.doi.org/10.1038/ismej.2012.134.
19. **Preheim SP, Boucher Y, Wildschutte H, David LA, Veneziano D, Alm EJ, Polz MF.** 2011. Metapopulation structure of *Vibrionaceae* among coastal marine invertebrates. Environ Microbiol **13:**265–275. http://dx.doi.org/10.1111/j.1462-2920.2010.02328.x.
20. **Shapiro BJ, Friedman J, Cordero OX, Preheim SP, Timberlake SC, Szabó G, Polz MF, Alm EJ.** 2012. Population genomics of early events in the ecological differentiation of bacteria. Science **336:**48–51. http://dx.doi.org/10.1126/science.1218198.
21. **Polz MF, Alm EJ, Hanage WP.** 2013. Horizontal gene transfer and the evolution of bacterial and archaeal population structure. Trends Genet **29:**170–175. http://dx.doi.org/10.1016/j.tig.2012.12.006.
22. **Chen F, Mackey AJ, Stoeckert CJ, Jr, Roos DS.** 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res **34:**D363–D368. http://dx.doi.org/10.1093/nar/gkj123.
23. **Król JE, Nguyen HD, Rogers LM, Beyenal H, Krone SM, Top EM.** 2011. Increased transfer of a multidrug resistance plasmid in *Escherichia coli* biofilms at the air-liquid interface. Appl Environ Microbiol **77:**5079–5088. http://dx.doi.org/10.1128/AEM.00090-11.
24. **Molin S, Tolker-Nielsen T.** 2003. Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. Curr Opin Biotechnol **14:**255–261. http://dx.doi.org/10.1016/S0958-1669(03)00036-3.
25. **Cordero OX, Wildschutte H, Kirkup B, Proehl S, Ngo L, Hussain F, Le Roux F, Mincer T, Polz MF.** 2012. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. Science **337:**1228–1231. http://dx.doi.org/10.1126/science.1219385.
26. **DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard NU, Martinez A, Sullivan MB, Edwards R, Brito BR, Chisholm SW, Karl DM.** 2006. Community genomics among stratified microbial assemblages in the ocean's interior. Science **311:**496–503. http://dx.doi.org/10.1126/science.1120250.
27. **Guglielmini J, de la Cruz F, Rocha EP.** 2013. Evolution of conjugation and type IV secretion systems. Mol Biol Evol **30:**315–331. http://dx.doi.org/10.1093/molbev/mss221.
28. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9:**75. http://dx.doi.org/10.1186/1471-2164-9-75.
29. **Leplae R, Lima-Mendez G, Toussaint A.** 2004. ACLAME: a classification of mobile genetic elements, update 2010. Nucleic Acids Res **38:**D57–D61. http://dx.doi.org/10.1093/nar/gkh084.
30. **Leplae R, Lima-Mendez G, Toussaint A.** 2006. A first global analysis of plasmid encoded proteins in the ACLAME database. FEMS Microbiol Rev **30:**980–994. http://dx.doi.org/10.1111/j.1574-6976.2006.00044.x.
31. **Lima-Mendez G, Toussaint A, Leplae R.** 2007. Analysis of the phage sequence space: the benefit of structured information. Virology **365:**241–249. http://dx.doi.org/10.1016/j.virol.2007.03.047.
32. **Clarke M, Maddera L, Harris RL, Silverman PM.** 2008. F-pili dynamics by live-cell imaging. Proc Natl Acad Sci U S A **105:**17978–17981. http://dx.doi.org/10.1073/pnas.0806786105.
33. **Bradley DE, Taylor DE, Cohen DR.** 1980. Specification of surface mating systems among conjugative drug resistance plasmids in *Escherichia coli* K-12. J Bacteriol **143:**1466–1470.
34. **Ho BT, Dong TG, Mekalanos JJ.** 2014. A view to a kill: the bacterial type VI secretion system. Cell Host Microbe **15:**9–21. http://dx.doi.org/10.1016/j.chom.2013.11.008.
35. **Klein JS, Lewinson O.** 2011. Bacterial ATP-driven transporters of transition metals: physiological roles, mechanisms of action, and roles in bacterial virulence. Metallomics **3:**1098–1108. http://dx.doi.org/10.1039/c1mt00073j.
36. **Cui J, Davidson AL.** 2011. ABC solute importers in bacteria. Essays Biochem **50:**85–99. http://dx.doi.org/10.1042/bse0500085.
37. **Schumacher MA.** 2012. Bacterial plasmid partition machinery: a minimalist approach to survival. Curr Opin Struct Biol **22:**72–79. http://dx.doi.org/10.1016/j.sbi.2011.11.001.
38. **Top EM, Springael D.** 2003. The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds. Curr Opin Biotechnol **14:**262–269. http://dx.doi.org/10.1016/S0958-1669(03)00066-1.
39. **Brown Kav A, Sasson G, Jami E, Doron-Faigenboim A, Benhar I, Mizrahi I.** 2012. Insights into the bovine rumen plasmidome. Proc Natl Acad Sci U S A **109:**5452–5457. http://dx.doi.org/10.1073/pnas.1116410109.
40. **Battermann A, Disse-Krömker C, Dreiseikelmann B.** 2003. A functional plasmid-borne *rrn* operon in soil isolates belonging to the genus *Paracoccus*. Microbiology **149:**3587–3593. http://dx.doi.org/10.1099/mic.0.26608-0.
41. **Kunnimalaiyaan M, Stevenson DM, Zhou Y, Vary PS.** 2001. Analysis of the replicon region and identification of an rRNA operon on pBM400 of Bacillus megaterium QM B1551. Mol Microbiol **39:**1010–1021. http://dx.doi.org/10.1046/j.1365-2958.2001.02292.x.
42. **Cohen O, Gophna U, Pupko T.** 2011. The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer. Mol Biol Evol **28:**1481–1489. http://dx.doi.org/10.1093/molbev/msq333.
43. **Jain R, Rivera MC, Lake JA.** 1999. Horizontal gene transfer among genomes: the complexity hypothesis. Proc Natl Acad Sci U S A **96:**3801–3806. http://dx.doi.org/10.1073/pnas.96.7.3801.
44. **Rumbo C, Fernández-Moreira E, Merino M, Poza M, Mendez JA, Soares NC, Mosquera A, Chaves F, Bou G.** 2011. Horizontal transfer of the OXA-24 carbapenemase gene via outer membrane vesicles: a new mechanism of dissemination of carbapenem resistance genes in *Acinetobacter baumannii*. Antimicrob Agents Chemother **55:**3084–3090. http://dx.doi.org/10.1128/AAC.00929-10.
45. **Yaron S, Kolling GL, Simon L, Matthews KR.** 2000. Vesicle-mediated transfer of virulence genes from *Escherichia coli* O157:H7 to other enteric bacteria. Appl Environ Microbiol **66:**4414–4420. http://dx.doi.org/10.1128/AEM.66.10.4414-4420.2000.
46. **Dubey GP, Ben-Yehuda S.** 2011. Intercellular nanotubes mediate bacterial communication. Cell **144:**590–600. http://dx.doi.org/10.1016/j.cell.2011.01.015.
47. **Frischer ME, Thurmond JM, Paul JH.** 1990. Natural plasmid transformation in a high-frequency-of-transformation marine *Vibrio* strain. Appl Environ Microbiol **56:**3439–3444.
48. **Ruhfel RE, Robillard NJ, Thorne CB.** 1984. Interspecies transduction of plasmids among *Bacillus anthracis*, *B. cereus*, and *B. thuringiensis*. J Bacteriol **157:**708–711.
49. **Botstein D.** 1980. A theory of modular evolution for bacteriophages. Ann N Y Acad Sci **354:**484–490. http://dx.doi.org/10.1111/j.1749-6632.1980.tb27987.x.
50. **Fernández-López R, Garcillán-Barcia MP, Revilla C, Lázaro M, Vielva L, de la Cruz F.** 2006. Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. FEMS Microbiol Rev **30:**942–966. http://dx.doi.org/10.1111/j.1574-6976.2006.00042.x.
51. **Naka H, Liu M, Actis LA, Crosa JH.** 2013. Plasmid- and chromosome-encoded siderophore anguibactin systems found in marine vibrios: biosynthesis, transport and evolution. Biometals Int J Role Met Ions Biol Biochem Med **26:**537–547. http://dx.doi.org/10.1007/s10534-013-9629-z.
52. **Osborn AM, Böltner D.** 2002. When phage, plasmids, and transposons collide:genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. Plasmid **48:**202–212. http://dx.doi.org/10.1016/S0147-619X(02)00117-8.
53. **Tamminen M, Virta M, Fani R, Fondi M.** 2012. Large-scale analysis of plasmid relationships through gene-sharing networks. Mol Biol Evol **29:**1225–1240. http://dx.doi.org/10.1093/molbev/msr292.
54. **Felsenstein J.** 1985. Phylogenies and the comparative method. Am Nat **125:**1–15. http://dx.doi.org/10.1086/284325.
55. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J Mol Biol **215:**403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
56. **Enright AJ, Van Dongen S, Ouzounis CA.** 2002. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res **30:**1575–1584. http://dx.doi.org/10.1093/nar/30.7.1575.
57. **Ning Z, Cox AJ, Mullikin JC.** 2001. SSAHA: a fast search method for large

DNA databases. Genome Res **11:**1725–1729. http://dx.doi.org/10.1101/gr.194201.
58. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing Subgroup**. 2009. The sequence alignment/map format and SAMtools. Bioinformatics **25:**2078–2079. http://dx.doi.org/10.1093/bioinformatics/btp352.
59. **Guglielmini J, Quintais L, Garcillán-Barcia MP, de la Cruz F, Rocha EP**. 2011. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. PLoS Genet **7:**e1002222. http://dx.doi.org/10.1371/journal.pgen.1002222.
60. **Finn RD, Clements J, Eddy SR**. 2011. HMMER Web server: interactive sequence similarity searching. Nucleic Acids Res **39:**W29–W37. http://dx.doi.org/10.1093/nar/gkr367.
61. **Eddy SR**. 2011. Accelerated profile HMM Searches. PLOS Comput Biol **7:**e1002195. http://dx.doi.org/10.1371/journal.pcbi.1002195.
62. **Delcher AL, Bratke KA, Powers EC, Salzberg SL**. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. Bioinformatics **23:**673–679. http://dx.doi.org/10.1093/bioinformatics/btm009.
63. **Fischer S, Brunk BP, Chen F, Gao X, Harb OS, Iodice JB, Shanmugam D, Roos DS, Stoeckert CJ, Jr**. 2011. Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. Curr Protoc Bioinformatics Chapter **6:**Unit 6.12.1–Unit 6.12.19. http://dx.doi.org/10.1002/0471250953.bi0612s35.
64. **Li L, Stoeckert CJ, Jr, Roos DS**. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res **13:**2178–2189. http://dx.doi.org/10.1101/gr.1224503.
65. **Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T**. 2011. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics **27:**431–432. http://dx.doi.org/10.1093/bioinformatics/btq675.
66. **Spinelli L, Gambette P, Chapple CE, Robisson B, Baudot A, Garreta H, Tichit L, Guénoche A, Brun C**. 2013. Clust&See: a cytoscape plugin for the identification, visualization and manipulation of network clusters. Biosystems **113:**91–95. http://dx.doi.org/10.1016/j.biosystems.2013.05.010.
67. **Preheim SP, Perrotta AR, Martin-Platero AM, Gupta A, Alm EJ**. 2013. Distribution-based clustering: using ecology to refine the operational taxonomic unit. Appl Environ Microbiol **79:**6593–6603. http://dx.doi.org/10.1128/AEM.00342-13.