RESEARCH ARTICLE

Phylogenetic Distribution of CRISPR-Cas Systems in Antibiotic-Resistant *Pseudomonas aeruginosa*

Alex van Belkum,[a] Leah B. Soriaga,[b] Matthew C. LaFave,[b] Srividya Akella,[b] Jean-Baptiste Veyrieras,[a] E. Magda Barbu,[b] Dee Shortridge,[a] Bernadette Blanc,[a] Gregory Hannum,[b] Gilles Zambardi,[a] Kristofer Miller,[b] Mark C. Enright,[c] Nathalie Mugnier,[a] Daniel Brami,[b] Stéphane Schicklin,[a] Martina Felderman,[b] Ariel S. Schwartz,[b] Toby H. Richardson,[b] Todd C. Peterson,[b] Bolyn Hubby,[b] Kyle C. Cady[b]

bioMérieux, Marcy l'Etoile, France[a]; Synthetic Genomics, Inc., La Jolla, California, USA[b]; Manchester Metropolitan University, Manchester, United Kingdom[c]

A.V.B., L.B.S., and K.C.C. contributed equally to this work.

ABSTRACT **Pseudomonas aeruginosa is an antibiotic-refractory pathogen with a large genome and extensive genotypic diversity. Historically, *P. aeruginosa* has been a major model system for understanding the molecular mechanisms underlying type I clustered regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated protein (CRISPR-Cas)-based bacterial immune system function. However, little information on the phylogenetic distribution and potential role of these CRISPR-Cas systems in molding the *P. aeruginosa* accessory genome and antibiotic resistance elements is known. Computational approaches were used to identify and characterize CRISPR-Cas systems within 672 genomes, and in the process, we identified a previously unreported and putatively mobile type I-C *P. aeruginosa* CRISPR-Cas system. Furthermore, genomes harboring non-inhibited type I-F and I-E CRISPR-Cas systems were on average ~300 kb smaller than those without a CRISPR-Cas system. *In silico* analysis demonstrated that the accessory genome ($n = 22{,}036$ genes) harbored the majority of identified CRISPR-Cas targets. We also assembled a global spacer library that aided the identification of difficult-to-characterize mobile genetic elements within next-generation sequencing (NGS) data and allowed CRISPR typing of a majority of *P. aeruginosa* strains. In summary, our analysis demonstrated that CRISPR-Cas systems play an important role in shaping the accessory genomes of globally distributed *P. aeruginosa* isolates.**

IMPORTANCE ***P. aeruginosa* is both an antibiotic-refractory pathogen and an important model system for type I CRISPR-Cas bacterial immune systems. By combining the genome sequences of 672 newly and previously sequenced genomes, we were able to provide a global view of the phylogenetic distribution, conservation, and potential targets of these systems. This analysis identified a new and putatively mobile *P. aeruginosa* CRISPR-Cas subtype, characterized the diverse distribution of known CRISPR-inhibiting genes, and provided a potential new use for CRISPR spacer libraries in accessory genome analysis. Our data demonstrated the importance of CRISPR-Cas systems in modulating the accessory genomes of globally distributed strains while also providing substantial data for subsequent genomic and experimental studies in multiple fields. Understanding why certain genotypes of *P. aeruginosa* are clinically prevalent and adept at horizontally acquiring virulence and antibiotic resistance elements is of major clinical and economic importance.**

Received 16 October 2015 **Accepted** 26 October 2015 **Published** 24 November 2015

Citation van Belkum A, Soriaga LB, LaFave MC, Akella S, Veyrieras J-B, Barbu EM, Shortridge D, Blanc B, Hannum G, Zambardi G, Miller K, Enright MC, Mugnier N, Brami D, Schicklin S, Felderman M, Schwartz AS, Richardson TH, Peterson TC, Hubby B, Cady KC. 2015. Phylogenetic distribution of CRISPR-Cas systems in antibiotic-resistant *Pseudomonas aeruginosa*. mBio 6(6):e01796-15. doi:10.1128/mBio.01796-15.

Editor Julian Parkhill, Sanger Institute

Address correspondence to Kyle C. Cady, kcady@syntheticgenomics.com.

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

Pseudomonas aeruginosa is a ubiquitous organism that colonizes a wide variety of surfaces and aquatic environments (1). In health care settings, it is a frequent end-stage colonizer in the lungs of cystic fibrosis patients and a common opportunistic pathogen of nosocomial wound, pneumonia, and catheter infections (2). *P. aeruginosa* is capable of acquiring resistance to all antibiotics commonly used against it through horizontal gene transfer and/or by obtaining site-specific mutations within chromosomal elements (3). Determining why some *P. aeruginosa* lineages are clinically widespread and proficient at horizontally acquiring virulence and antibiotic resistance elements is of high scientific importance.

Multilocus sequence typing (MLST) studies have shown that *P. aeruginosa* has a nonclonal population, with globally disseminated epidemic clones of highly successful multidrug-resistant (MDR) *P. aeruginosa* strains that readily accumulate mobile genetic elements (4, 5). As observed with other pathogens (6), certain phylogenetically related *P. aeruginosa* MLST sequence type (ST) groups are more common in nosocomial and environmental settings (7–9). These global clones are multidrug-resistant and

extensively drug resistant, with the most notable having STs 111 and 235 (4, 5, 10–12). The genomic sequence of the PAO1 strain was published in the year 2000 and was soon followed by those of several additional reference strains and clinical isolates (1, 5, 7, 8, 13–16). These sequencing efforts revealed genome sizes ranging from 5.5 to 7 megabase pairs (Mb), GC contents of 65 to 76%, and 79 regions of extensive genome plasticity (16, 17). Over the last year, Kos et al. added an additional 390 genomes, and for 388 strains, susceptibility data for four antibiotics to the public domain, greatly increasing, yet still not saturating, the genomic and phenotypic information available for *P. aeruginosa* (5). Compared to other bacterial pathogens, *P. aeruginosa* has a large genome, encoding many diverse horizontally transferable elements, including plasmids, conjugative elements, prophages, genomic islands, pathogenicity islands, integrons, and transposons (16, 17). The major differences in gene contents between isolates necessitate additional genome sequencing and annotation studies to accurately quantify the pangenome and to associate phenotypic traits with more conserved or less conserved genetic loci (16, 17). Clustered regularly interspaced short palindromic repeat (CRISPR) arrays and CRISPR-associated genes (*cas*) comprise bacterial adaptive immune systems and function as a variable genetic element. Multiple reports have suggested that CRISPR-Cas systems may play a major role in controlling horizontal gene transfer and, consequently, the dynamics of antibiotic resistance gene acquisition in this important human pathogen (5, 12).

CRISPR-Cas bacterial immune modules protect bacteria and archaea from mobile genetic elements, such as plasmids and bacteriophages (phages) (18, 19). Type I CRISPR arrays are transcribed as one large transcript that is processed within an identical repeat sequence into mature, small CRISPR RNAs (crRNAs) by a subtype-specific Cas6 protein (20, 21). Mature crRNAs are then incorporated into subtype-specific, large riboprotein complexes and guided to double-stranded DNA (dsDNA) targets in a sequence-dependent manner (21–24). Once targeted, the riboprotein complex interacts with a subtype-specific Cas3 protein, which enzymatically mediates single-strand nicking and digestion of the target DNA (25–27). Currently, six known type I subtypes have been identified and are abbreviated I-A, I-B, I-C, I-D, I-E, and I-F (28). Importantly, the type I CRISPR-Cas subtypes are evolutionarily distinct; even within a single bacterial species, they have a conserved function, and yet, they have very limited sequence conservation at both the nucleotide and protein level (28). The protein complexes, enzymatic activities, and nucleic acid target preferences of type I, II, and III CRISPR-Cas systems are also highly divergent and are not necessarily analogous between systems (28). The ability of the three phylogenetically diverse and functionally distinct CRISPR-Cas systems to act as prokaryotic adaptive immunity modules is now well established; however, their role in shaping bacterial genomes on a species scale is only now being elucidated. Studies in *Enterococcus* spp. have suggested that type II CRISPR-Cas systems reduce the acquisition of antibiotic resistance and are negatively associated with MDR strains (29). However, recent studies in *Francisella* spp. have indicated that similar type II CRISPR-Cas systems are critical for increased antibiotic tolerance through enhancing envelope integrity (30). *P. aeruginosa* genomes have not been shown to contain type II CRISPR-Cas systems; however, they contain prophages encoding CRISPR-inhibiting genes that are capable of specifically deactivating the type I-F and I-E CRISPR-Cas subtypes (31, 32). Currently, it is still unclear what role, if any, CRISPR-Cas systems play in antibiotic resistance or in shaping the *P. aeruginosa* accessory genome.

Here, we examine the phylogenetic distribution and conservation of CRISPR-Cas systems among 672 *P. aeruginosa* isolates. Our analysis identified multiple distinct lineages harboring an integrative and conjugative element (ICE)-encoded type I-C CRISPR-Cas system, which had previously never been identified in *P. aeruginosa*. Our correlation of phylogeny and spacer content provide a strong foundation for potentially utilizing CRISPR typing to rapidly determine the lineages (ST groups) of a majority of *P. aeruginosa* isolates. Furthermore, the frequency of spacer integration and deletion observed between related strains indicates that CRISPR typing may be useful in tracking strains within outbreaks, even if they are of the same ST group. The global spacer library that was assembled also aided in the identification of difficult-to-characterize mobile genetic elements within next-generation sequencing (NGS) data and demonstrated that the majority of spacer targets resided within the *P. aeruginosa* accessory genome. Interestingly, after correcting for the presence of anti-CRISPR genes, it was found that type I-F and I-E CRISPR-Cas systems were correlated with decreased genome size. Altogether, our analysis reveals that CRISPR-Cas systems play an important role in shaping the accessory genomes of globally distributed *P. aeruginosa* strains and provides a substantial data set for future CRISPR-Cas and genomics studies.

RESULTS

Clinical *P. aeruginosa* strain collection and population structure. To understand the genomic diversity of *P. aeruginosa* clinical isolates, we compared the sequences of 672 strains belonging to three diverse collections. We sequenced 282 clinical isolates obtained from either the private bioMérieux clinical strain collection ($n = 219$) or the previously phenotypically characterized Pirnay collection ($n = 63$) (33). The assemblies of these isolates were submitted to GenBank with BioProject record number PRJNA297679. The vast majority of bioMérieux strains were obtained from the European Union and the United States over a period of 25 years and represent primarily human-derived isolates (214/219). The strains from the globally distributed and temporally wide-ranging Pirnay collection are predominantly clinical in nature; however, five strains from nonhospital settings were included as an environmental outgroup (33). The genomic and/or phenotypic information of strains recently sequenced ($n = 390$) and phenotypically analyzed ($n = 388$) by Kos et al. were included in our analysis to obtain a unified and truly large-scale view of *P. aeruginosa* clinical isolate diversity. The Kos collection clinical isolates were also from diverse geographic locations but temporally range between 2003 and 2012 (5). Importantly, this library of strains focuses largely on isolates obtained from non-cystic fibrosis infections. This choice was made because *P. aeruginosa* isolated from the cystic fibrosis lung are often distinct from the majority of environmental or acute-infection-causing strains. Our analysis showed that the genome sizes range between 5.5 and 7.6 Mb, which is similar to the size range previously reported by smaller-scale studies (16, 17).

MLST was employed to compare the phylogenetic diversity of these 672 isolates with the phylogenies found by prior, non-NGS studies (see the supplemental material). MLST relies on comparing the sequences of evolutionarily conserved but polymorphism-

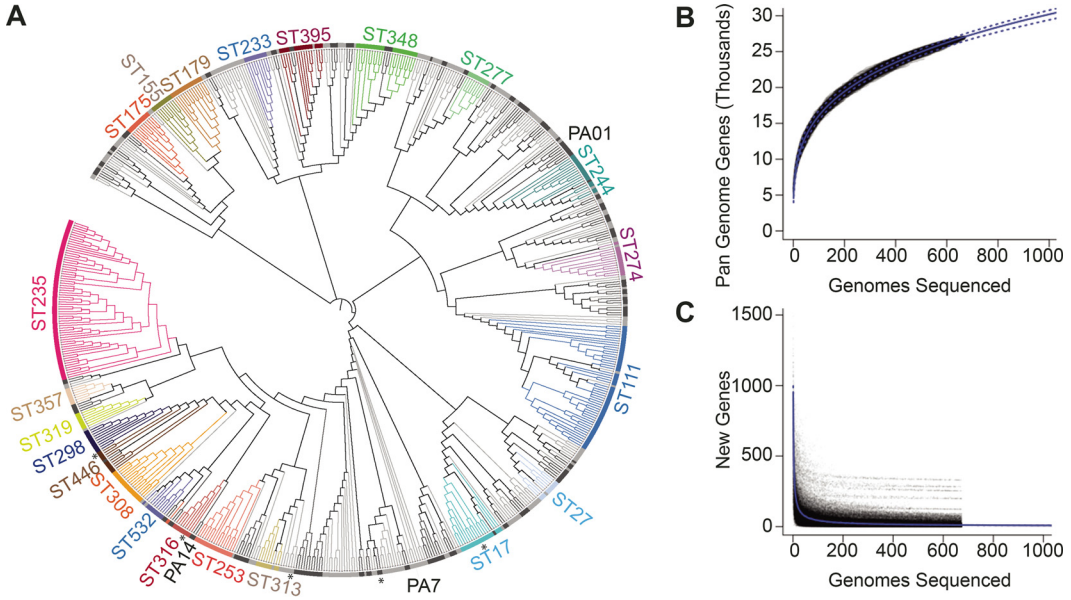


FIG 1 Genomic diversity of *P. aeruginosa* clinical isolates. (A) Phylogenetic tree of *P. aeruginosa* clinical isolates. Phylogenetically related ST groups with ≥7 members are shown in different colors. The prevalent lineages constitute 52% of all strains analyzed. ST groups that have only 1 member are delineated in grey, while black denotes phylogenetic lineages with 2 to 6 members. The phylogenetic locations of the common laboratory strains PAO1, PA14, and PA7 are marked in black for reference, while the 4 nonclinical isolates are labeled with small black asterisks. (B) Rarefaction curve analysis of the total number of genes within the *P. aeruginosa* pangenome as a function of the number of genomes sequenced. Extrapolation is shown with dotted lines. (C) Rarefaction curve analysis of the number of new genes identified as a function of the number of *P. aeruginosa* genomes sequenced. A distribution of values was obtained for each strain count by permuting the order of the strains 500 times. The median trend line and 95% confidence intervals are shown in blue.

harboring genes (*acs*, *aro*, *gua*, *mut*, *nuo*, *pps*, and *trp*) in order to cluster phylogenetically related strains into ST lineages within a single bacterial species (4). MLST coverage of the 672-strain library spanned 216 STs, with 52% of strains falling into 22 STs, clearly demonstrating the diversity of the collection (Fig. 1A). As noted above, the predominant clonal ST lineages were ST235 and ST111 (5, 7, 8). Among the 282 newly sequenced strains, MLST analysis assigned 41 STs that were previously uncatalogued (see Table S1 in the supplemental material). The identification of novel STs is interesting, yet not unexpected for a phylogenetically diverse and genetically promiscuous pathogen like *P. aeruginosa*. Together, these phylogenetic analyses clearly demonstrate both the diversity of *P. aeruginosa* strains capable of causing human infections and the increasing nosocomial fitness of select *P. aeruginosa* lineages as they predominate in the clinic.

Identification and annotation of the *P. aeruginosa* common pangenome. *P. aeruginosa* is well known for its relatively large genome, which contains both a substantial invariable core genome and strain-specific accessory genes that, combined, constitute the pangenome (16). All 672 genomes were uniformly annotated using the commercially available Archetype genomics pipeline. CD-HIT clustering, an algorithm that collapses similar protein sequences into cluters, was used to create a nonredundant catalog of protein-coding sequences from the *P. aeruginosa* genomes (see the supplemental material). This analysis, using a 70% amino acid identity cutoff, reduced the protein-coding pangenome of our *P. aeruginosa* combined library from over 4 million protein sequences to roughly 27,000 (Fig. 1B; see also Table S2 and Text S1 in the supplemental material). Interestingly, a quarter of all pangenome gene clusters are singletons, representing a single unique gene found only in a single genome. Although it is possible that some of these proteins were not annotated in other genomes, these results suggest a substantial reservoir of unique singleton genes within the *P. aeruginosa* population.

Rarefaction curves showed that the number of CD-HIT clusters continues to increase with the number of *P. aeruginosa* genomes sequenced, although at a decreased rate (Fig. 1B). Extrapolation from the rarefaction curve suggests that we have largely identified all accessory genomic elements found in multiple strains (common accessory genome) and that sequencing additional similar *P. aeruginosa* clinical isolate genomes will yield roughly 10 or fewer new singletons per genome sequenced (Fig. 1C). CD-HIT analysis of all 672 genomes identified 5,081 clusters found in at least 90% of genomes (see Table S2 in the supplemental material), closely matching the previous core-genome value of 5,021 clusters identified by comparing five genomes (16). A cutoff of 90% rather than 100% was utilized to prevent exclusion of core-genomic components as a result of the sequencing coverage gaps that are inherent to large-scale NGS studies. These data revealed that the sizes of the core and common accessory genomes will likely vary minimally with additional genome sequences (16). Together, these analyses provide a global view of *P. aeruginosa* clinical isolate genome diversity, including a comprehensive list of the CD-HIT clusters constituting the current pan- and core genomes (see Table S2).

Antibiotic susceptibility testing and resistome identification. *P. aeruginosa* infections refractory to antibiotics are a well-established clinical problem; however, the emergence of multidrug-resistant *P. aeruginosa* (MDRPA), extremely drug-resistant *P. aeruginosa* (XDRPA), and pan-drug-resistant *P. aeruginosa* (PDRPA) isolates has highlighted the need for novel

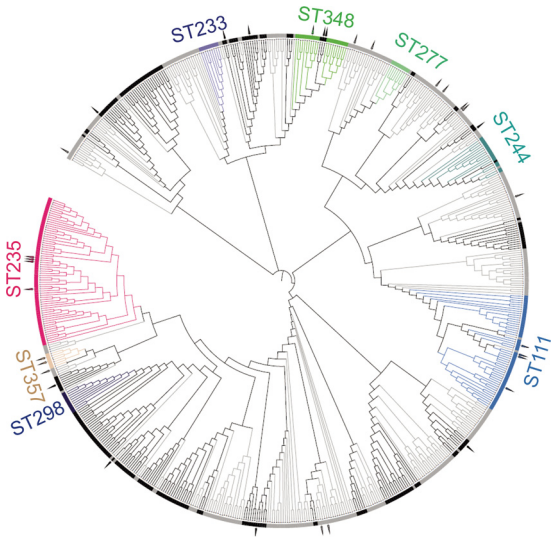


FIG 2 Known and potential new MDRPA lineages. Phylogenetic tree illustrating the locations of ST groups harboring ≥5 MDR strains (colored branches), 2 to 4 MDR strains (black branches), or 0 or 1 MDR strain (grey branches). Black spikes protruding from the circle denote the locations of XDRPA strains.

antipseudomonal drug development (34, 35). Currently, eight antibiotic classes are used to treat *P. aeruginosa* infections (aminoglycosides, carbapenems, cephalosporins, fluoroquinolones, penicillins in combination with β-lactamase inhibitors, monobactams, phosphonic acids, and polymyxins) (see Fig. S1 in the supplemental material) (36). MDRPA isolates are defined as being nonsusceptible to ≥1 agent in ≥3 antimicrobial categories, while XDRPA isolates are nonsusceptible to ≥1 agent in all but ≤2 antimicrobial categories (36). PDRPA isolates, so-called super bugs, are nonsusceptible to all major antipseudomonal antibiotics in clinical use (36). Recent studies have identified ST groups 111 and 235 as major MDR lineages (5, 7, 8).

To better understand the antibiotic resistance profiles of all newly sequenced strains and identify MDRPA and XDRPA isolates, we performed broth dilution assays complemented with Vitek testing for all 282 newly sequenced isolates using four clinically relevant antibiotics from distinct antimicrobial categories (aminoglycosides, penicillins plus beta-lactamase inhibitors, carbapenems, and cephalosporins). Additionally, most strains were tested for multiple other clinically relevant antibiotics from the remaining antipseudomonal antibiotic families (fluoroquinolones, phosphonic acids, monobactams, and polymyxins). The known resistance genes of all newly sequenced strains were computationally identified and combined with phenotypic and genetic data previously obtained for the Kos collection (see Table S1 in the supplemental material) (5). Unsurprisingly, the link between known resistance genes and phenotypic antibiotic resistance was stronger for some antibiotics than for others, an observation previously made by Kos et al. (5).

Between the Kos collection antibiotic screening (four antimicrobial categories tested) and our analysis, we identified 246 MDRPA isolates and 31 XDRPA isolates (Fig. 2; see also Table S1 in the supplemental material). These data demonstrate that at least 41% of the clinical strains analyzed are at minimum MDRPA according to the breakpoints established by the Clinical and Laboratory Standards Institute (CLSI) (36). Alarmingly, we identified eight isolates that were nonsusceptible to at least one antibiotic from all eight antibiotic categories (see Table S1). After assigning the ST type of each MDRPA and XDRPA isolate, we identified eight STs that appeared enriched in MDRPA/XDRPA isolates (Fig. 2). Importantly, the resistance-enriched STs represented clades containing five or more MDRPA/XDRPA isolates, regardless of the number of strains within the ST. This cutoff by design excludes STs containing less than four isolates and is more likely to include larger STs prevalently found within the clinic and, consequently, our library. As expected, the previously characterized and clinically problematic STs 111 and 235 were both designated as enriched in drug resistance. ST348, which is oddly bisected by the small but highly drug-refractory ST1320, was also shown to harbor multiple drug-resistant strains (Fig. 2). Interestingly, the placement of ST1320 within the ST348 clade likely indicates that these strains are members of a larger clonal complex. Of further interest, the remaining five drug resistance-enriched STs harbor numbers of strains similar to the numbers in many other medium-sized STs; however, they still contain at least five MDRPA or XDRPA strains. Only 4 of the 8 drug-resistant STs (STs 235, 111, 348, and 244) were in the top 10 most prevalent STs identified in the entire collection (Fig. 1A and 2). ST357, which is related to the well-characterized ST235 group, is especially threatening. While our library contains only seven ST357 isolates, three are MDRPA and two are XDRPA. Furthermore, one of the XDRPA strains is nonsusceptible to at least one antibiotic from all eight clinically used antimicrobial categories. The nosocomial origin and antibiotic-refractory nature of these eight STs suggests that these lineages may represent current or emerging high-risk MDRPA clades. Furthermore, the combined antibiotic susceptibility and genomic data set allows a better understanding of the phylogenetic distribution of known resistance elements; however, these data are also invaluable in understanding how CRISPR-Cas systems may modulate antibiotic resistance in *P. aeruginosa*.

Phylogenetic distribution of CRISPR-Cas subtypes and anti-CRISPR within clinical *P. aeruginosa* isolates. *P. aeruginosa* has emerged as a major CRISPR-Cas model system (20, 22, 24, 31, 32, 37); however, only two studies have ever attempted to understand the diversity and conservation of their CRISPR-Cas systems (38, 39). The more comprehensive of these two studies was performed before the widespread use of NGS, and while informative, it lacked the sequencing depth and breadth required to illustrate the scope of CRISPR-Cas-mediated interactions in *P. aeruginosa* (38). Additionally, in the past 5 years, important advancements have been made in understanding the molecular mechanism underlying CRISPR-Cas function, as well as the identification of genes encoding CRISPR-Cas-inactivating proteins (31, 32), enabling a more thorough computational analysis. Utilizing our substantial genomic library, we strove to perform a new large-scale analysis of CRISPR-Cas phylogenetic distribution, conservation, and targeting.

Computational analysis identified intact type I-F and type I-E CRISPR-Cas subtypes in 202 and 81 genomes, respectively (Fig. 3A; see also Table S3 and Fig. S2 in the supplemental material). Both subtypes uniformly localized to previously identified genomic loci (Fig. 3B) (38). We also identified 20 *P. aeruginosa* isolates harboring the type I-C CRISPR-Cas system, which had previously never been identified in *P. aeruginosa* (Fig. 3A and B; see also Table S3 and Fig. S2). The type I-C systems were distrib-

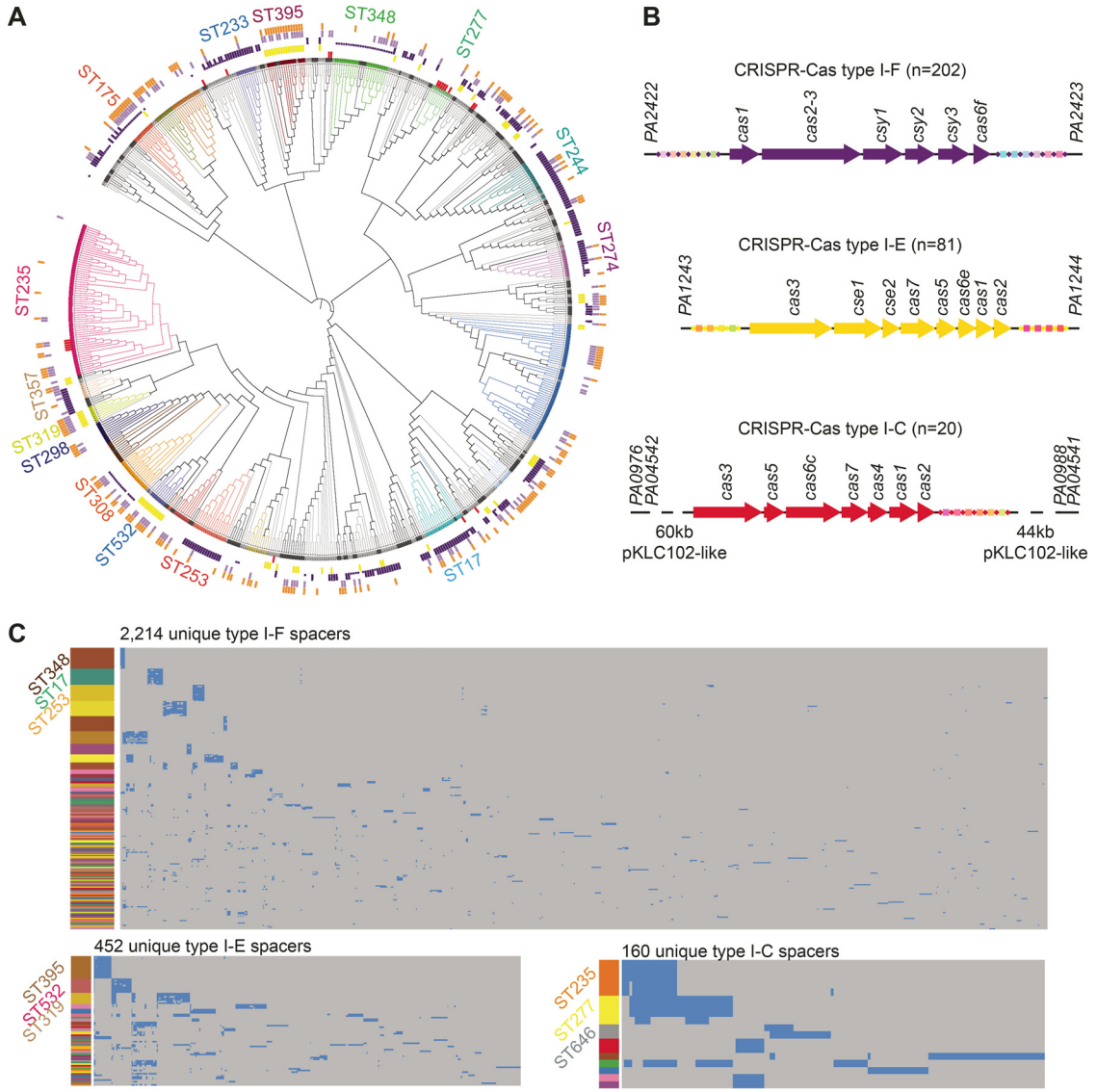


FIG 3 Distribution of CRISPR-Cas systems in *P. aeruginosa*. (A) Locations of CRISPR-Cas subtypes and anti-CRISPR-Cas genes among the 672 genomes analyzed. Purple, yellow, and red markers denote the locations of I-F, I-E, and I-C subtypes, respectively. Strains encoding CRISPR-Cas-inhibitory proteins are denoted in violet for type I-F and orange for type I-E. The locations and colors of major STs are highlighted identically to their depiction in Fig. 1A to aid in orientation. Degenerate type I-F systems lacking *cas* genes are denoted by short purple bars. (B) Diagram of the conserved gene content and location of each CRISPR-Cas subtype. Colored diamonds denote subtype-specific CRISPR repeats, while multicolored squares denote unique spacer content. (C) Heatmap illustrating that CRISPR array spacer content associates with phylogeny. The presence of every unique spacer sequence in each CRISPR-positive strain is denoted by blue squares. All strains are grouped by ST along the *y* axis, as denoted by various colors. The 3 most prevalent STs harboring each CRISPR subtype are shown.

uted among eight ST groups. However, most type I-C-positive strains were clustered into MDRPA groups ST277 and ST235 (Fig. 3A). Unlike the type I-F and I-E subtypes, the type I-C CRISPR-Cas systems were all carried on pKLC102-like ICEs. The pKLC102-like ICEs are >100 kb in size and encode both conjugation and chromosomal integration machinery (40). ICEs are widespread in *P. aeruginosa* and other Gram-negative bacteria and are usually found integrated into the host chromosome at tRNA genes which encode their specific bacterial attachment (*attB*) sites (40). Only 9 of our 20 type I-C-harboring pKLC102-like elements had contig sizes large enough to determine the integration site; however, in all 9 cases, the integration site was in a $tRNA^{Lys}$ gene (PA0976.1 or PA4541.1). As these are the only two $tRNA^{Lys}$ loci in the *P. aeruginosa* genome and perfectly match the established *attB* site for pKLC102-like elements, we can infer that most if not all type I-C-harboring ICEs are integrated at one of these two sites (Fig. 3B). This analysis represents both the first report of a type I-C CRISPR-Cas system in *P. aeruginosa* and the first putatively mobile genetic element with a CRISPR-Cas system in a *Pseudomonas* species.

Specific *P. aeruginosa* lineages lack both CRISPR arrays and *cas* genes (STs 155, 179, 111, 27, 313, 316, 446, and 298), while others harbor degenerate systems containing CRISPR arrays but lacking cognate *cas* genes (STs 308, 175, and 348). The 84 *P. aeruginosa* strains containing CRISPR arrays but lacking cognate *cas* genes do not possess *cas* gene-targeting spacers and no longer encode the

Cas enzymes required to process or utilize the CRISPR array transcripts (see Fig. S2 in the supplemental material). Nearly all strains in the major STs 348 (21/21) and 308 (15/16) have a detectable CRISPR array without cognate *cas* genes (Fig. 3A). Additionally, previous studies in *P. aeruginosa* have identified small proteins which can functionally deactivate CRISPR-Cas type I-E and I-F systems (31, 32, 41). Currently, eight known groups of anti-CRISPR genes (four anti-type I-F and four anti-type I-E genes) have been characterized; however, there are likely to be additional anti-CRISPR genes as yet unidentified, including those targeting the newly identified I-C subtype. Anti-CRISPR genes were widely identified in many diverse strains and groups, with potential CRISPR-Cas type I-E inhibition being particularly apparent in groups ST395 (13/14) and ST319 (7/7) (Fig. 3A). Interestingly, this information indicates that many strains with intact CRISPR-Cas systems are likely to be phenotypically CRISPR-Cas incompetent because they encode a cognate anti-CRISPR protein that deactivates the corresponding CRISPR-Cas system. Subtype-specific anti-CRISPR genes were identified in 40 (19.8%) of 202 strains with I-F CRISPR-Cas subtypes and 43 (53.0%) of 81 strains harboring I-E subtypes (see Fig. S2). While previous studies have demonstrated the function of type I-F and I-E anti-CRISPR genes (31, 32), this is the first study to identify their wide-spread phylogenetic distribution in *P. aeruginosa* clinical isolates.

Phylogenetic distribution of CRISPR-targeting spacer content. CRISPR arrays contain a series of sequence-specific conserved repeats flanking unique inserts known as spacers. New spacers are derived from invading mobile genetic elements and integrated into CRISPR arrays through the activity of Cas proteins (19). Each spacer encodes the targeting sequence of a given CRISPR array-derived small crRNA (18, 24). The spacer sequences of a given bacterial strain therefore constitute a history of past CRISPR-Cas-mediated interactions. Studies in *Salmonella*, *Campylobacter*, *Mycobacterium*, and *Corynebacterium* isolates have shown that comparing the spacer sequences of a given strain can provide a phylogenetic bar code to rapidly track pathogenic strains during outbreaks in a process termed CRISPR typing (42–47).

CRISPR typing is of particular interest to clinicians attempting to track outbreaks as it only requires the sequencing of a few loci rather than the ~7 loci used for MLST typing. However, a strong link between phylogeny and CRISPR array sequence must first be experimentally established for a given pathogen, as spacer content may change over time (19, 22). To determine if there is a strong association between phylogeny and subtype-specific spacer content in this diverse *P. aeruginosa* library, we directly compared the spacer content of all strains within our collection. A full spacer library was created by orienting and extracting all spacer sequences based on the conserved direct repeat motif (Fig. 4; see also Table S4 in the supplemental material). The full spacer library of 11,041 sequences was then collapsed by CD-HIT clustering to remove all redundant spacers with a 90% identity cutoff, thereby creating a nonredundant spacer library (Fig. 4; see also Table S4). By mapping the presence or absence of every unique spacer sequence found in the nonredundant spacer library (see Table S4) (2,214 type I-F, 452 type I-E, and 160 type I-C) for each strain harboring a CRISPR array (see also Table S3), we identified strains harboring similar spacer contents (Fig. 3C; see also Fig. S3). The clustering of spacer-positive strains by ST grouping in all three CRISPR-Cas subtypes indicates a strong link between phylogeny and spacer content in *P. aeruginosa* clinical isolates (Fig. 3C; see also Fig. S3), similar to observations made for select *P. aeruginosa* cystic fibrosis isolates (39). Usefully, spacer content can be slightly variable even within an ST group due to random spacer deletion or new spacer integration (Fig. 3C). This inter-ST diversity could provide a powerful tool for tracking sublineages of genetically similar strains within a hospital outbreak. As CRISPR arrays are found in 14 of the 22 most common clinical ST groups analyzed, they may provide considerable value in tracking select *P. aeruginosa* lineages in nosocomial environments. Together, these analyses provide both the MLST-based phylogenetic data and CRISPR array sequences required to aid in future *P. aeruginosa* CRISPR typing.

Novel approach for highlighting previously uncharacterized mobile genetic elements in large NGS genome datasets. While comparison of CRISPR spacer sequences to previously studied elements within the NCBI database identified potential targets for roughly 30% of unique spacers, comparison of the remaining ~70% of spacers against non-CRISPR accessory genome elements within our library identified potential targets for a further ~55% of spacer sequences (Fig. 4). These data indicate that our genome library contains a massive archive of unstudied CRISPR-Cas-targeted genetic elements hidden within the larger accessory genomic sequences. Unfortunately, as the full or partial sequences of these unstudied mobile genetic elements are unknown and their full lengths can be split between multiple contigs, standard homology-based identification is challenging.

As unique CRISPR spacer content primarily represents independently acquired and evolutionarily selected short ribonucleoprotein-targeting sequences with low genomic background complementarity (48), their sequences can be used to highlight previously encountered genetic elements within complex genomic data. Akin to antibodies, spacer sequences are the adaptive targeting element of an immune system and are rarely retained within the host if they are cross-reactive. By identifying NGS contigs harboring large clusters of novel spacer target sequences, we can rapidly highlight *P. aeruginosa* mobile genetic elements not currently annotated within the NCBI database (Fig. 4). As a proof of principle, we utilized our spacer library to identify novel mobile genetic elements commonly contained within the accessory genomes of the clinically problematic MDRPA ST111 lineage of *P. aeruginosa*.

As ST111 forms a globally disseminated, highly antibiotic-resistant, CRISPR-Cas-devoid, and clinically prevalent lineage of *P. aeruginosa*, it represents the perfect test case to identify novel mobile genetic elements. By comparing the unique spacer library against all contigs within ST111 *P. aeruginosa* genomes, we identified a few mobile genetic elements commonly found in ST111 strains (Fig. 4; see also Table S5 in the supplemental material). We identified one pKLC102-like ICE with sequence similarity to a previously sequenced element found in *P. aeruginosa* strain PA7 (15); interestingly, it was present in 67% of ST111 strains. Additionally, three novel prophages were identified, which were found in 27%, 12%, and 48% of ST111 strains and were most similar to LESB58 prophage 5, M18 prophage I, and PA7 prophage RGP78, respectively (Fig. 4; see also Table S5) (14, 15, 49). As both ICEs and bacteriophages have been shown to modulate the virulence of *P. aeruginosa*, the role of these four novel elements in ST111 virulence should be more thoroughly analyzed in future studies. This small-scale examination demonstrates that large spacer libraries can be used to highlight novel NGS elements for further analysis

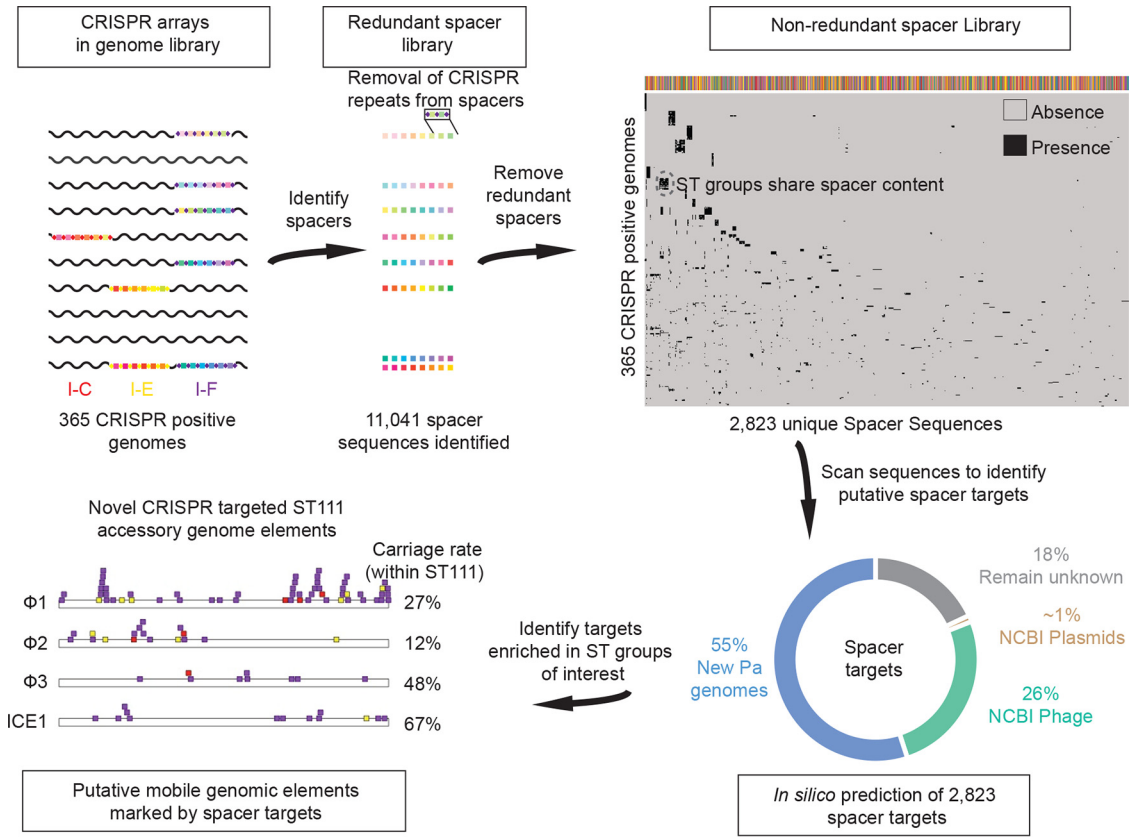


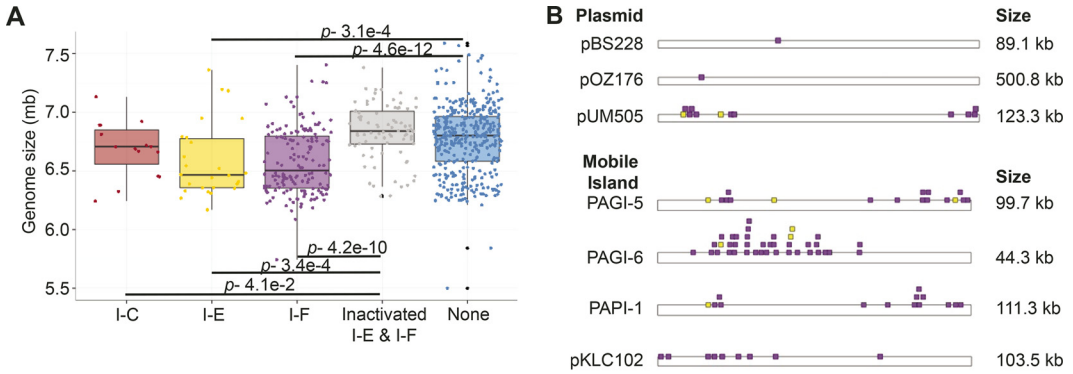
FIG 4 *In silico* CRISPR target prediction. CRISPR arrays were computationally identified using PILER-CR and subsequently oriented uniformly. Wavy lines denote individual *P. aeruginosa* genomes. Sequences containing type I-C CRISPR arrays are shown with red diamond repeats, while those with I-E or I-F harbor yellow and purple diamonds, respectively. Conserved repeats of all three subtypes (I-C, I-E, and I-F) were computationally extracted, leaving only spacer content (rainbow-colored boxes), which provides the sequence specificity required for CRISPR-Cas targeting. Removal of spacer sequences with greater than 90% identity to other spacers provides a nonredundant spacer library. Unique spacer sequences are shown as different colors across the top of the image. Black squares indicate the presence of that spacer in a given strain. Spacer complementarity to known viral and plasmid databases indicates substantial CRISPR targeting; however, the majority of putative spacer targets were found in the accessory genomes of *P. aeruginosa* strains analyzed in this study. CRISPR-Cas target sequences (spacers) reside largely in the nonannotated accessory genomes of *P. aeruginosa* strains analyzed in this study. The pie chart denotes the percentage of unique spacers with complementarity to a given type of target sequence. Spacer sequences can be used to highlight previously uncharacterized accessory elements that are likely to be mobile genetic elements. Purple, red, and yellow squares represent the sites of complementarity for nonredundant type I-F, I-C, and I-F spacer sequences, respectively. φ1, φ2, and φ3 represent prophages predominantly found in the highly MDR and clinically prevalent ST111 *P. aeruginosa* lineage, while ICE1 represents a pKLC102-like element that is highly prevalent in ST111 strains.

(see the supplemental material). Lastly, these data support the idea that *P. aeruginosa* CRISPR-Cas systems interact both with known accessory genomic elements deposited in the NCBI database and with many previously uncharacterized elements contained within *P. aeruginosa* clinical isolates.

CRISPR-Cas systems affect *P. aeruginosa* genome size and target many previously studied mobile genetic elements. We observed that *P. aeruginosa* genomes harboring intact CRISPR-Cas type I-F or I-E systems and lacking known cognate anti-CRISPR genes were on average significantly smaller in size (~300 kb) than those without a CRISPR-Cas system (Fig. 5A). Few type I-C-harboring genomes were available; however, a similar trend in genome size was also observed for this less prevalent CRISPR-Cas subtype. These data are consistent with the proposed role of CRISPR-Cas systems acting to filter horizontal gene transfer into CRISPR-Cas-encoding strains. To better understand the types of genetic elements being blocked by the CRISPR-Cas systems, we compared the unique spacer library obtained from all *P. aeruginosa* strains analyzed (see Table S4 in the supplemental material) against the NCBI mobile genetic element libraries.

Computationally calculating the complementarity between nonredundant spacers and the NCBI plasmid (8,915 plasmids) and phage (9,448 phages) databases assigned targets for 1% and 26% of the unique spacers, respectively. Interestingly, no spacers were identified to have complementarity to known integrons, which are mobile genetic elements encoding a site-specific recombinase and, often, antibiotic resistance elements. This was perhaps expected as integrons are small, often integrated into larger mobile genetic elements, such as plasmids and transposons, and endow selective advantages upon host cells. However, integrons are widespread in *P. aeruginosa* and are a major mechanism by which antibiotic resistance is spread within MDRPA. Conversely, many spacers were identified against known virulence islands and other ICE elements, likely as a result of their substantial size and metabolic burden on host cells. As it is impossible to talk about all of the interesting CRISPR-Cas targets identified within the NCBI databases, we will focus our analysis on a few interesting representatives.

Similar to the results of previous smaller-scale studies, more spacers are complementary to known temperate bacteriophages than to other mobile genetic elements, such as plasmids, genomic



C

CRISPR-Cas subtype	Correlation	Antibiotic Susceptibility	+ST111/ST235 p-value	+ST111/ST235 FDR	-ST111/ST235 p-value	-ST111/ST235 FDR
I-F	+	Amikacin	1.6e-3	0.016	0.19	0.82
	+	Tobramycin	2.3e-3	0.016	0.4	0.82
I-E	+	Piperacillin/Tazobactam	5.9e-5	5.2e-4	7.5e-4	6.9e-3
	+	Aztreonam	7.4e-5	5.2e-4	9.9e-4	6.9e-3
	+	Ciprofloxacin	4.8e-3	0.022	5.2e-2	0.18
	+	Amikacin	6.2e-3	0.022	0.12	0.24
	+	Levofloxacin	8.9e-3	0.025	0.11	0.24
I-C	-	Amikacin	4.6e-3	0.065	1.6e-3	0.022

D

CRISPR-Cas subtype	Correlation	Resistance gene	+ST111/ST235 p-value	+ST111/ST235 FDR	-ST111/ST235 p-value	-ST111/ST235 FDR
I-F	-	*sul1*	6.5e-5	0.008	1	1
	-	*CARB*	3.1e-4	0.02	0.45	1
I-E	-	*sul1*	1.3e-5	0.002	0.026	1
I-C	+	*rmtD*	5.5e-8	3.6e-6	3.7e-8	2.4e-6
	+	*aadA7*	5.5e-8	3.6e-6	3.7e-8	2.4e-6
	+	*OXA*	2.8e-5	0.004	4.9e-4	0.016

FIG 5 Correlations between CRISPR-Cas subtypes and antibiotic resistance. (A) Clinical isolates containing CRISPR-Cas systems have, on average, smaller genomes than those lacking the same systems. Red, yellow, and purple dots represent the genome sizes of individual strains harboring the respective CRISPR-Cas system. Blue and grey dots denote the genome sizes of strains either lacking CRISPR-Cas systems or predicted to have CRISPR-Cas systems that are inactivated by anti-CRISPR-Cas genes, respectively. Box plots illustrate the distribution of genome sizes in each category. The upper and lower "hinges" of each box correspond to the first and third quartiles of the genome sizes for the clinical isolates in each category. Upper and lower whiskers extend from the given hinge to the highest or lowest value that is within 1.5* IQR, where IQR is the inter-quartile range, or distance between the first and third quartiles. The Wilcoxon rank-sum test and *t* test were used to compare the distributions of sizes between pairs, and multiple-hypothesis-adjusted *P* values are provided. (B) Locations of spacer target sequences in previously sequenced plasmids and pathogenicity islands. The colors correspond to the CRISPR-Cas subtypes depicted in panel A. (C) Correlation of antibiotic susceptibility with the presence or absence of a given CRISPR-Cas subtype. (D) Correlation of antibiotic resistance genes with the presence or absence of a given CRISPR-Cas subtype. (C and D) Data are corrected for the presence of anti-CRISPR-Cas genes. The *P* values are from the Wilcoxon rank-sum test. The false discovery rate (FDR) adjustment of the *P* values was performed using the Benjamini-Hochberg method with the function p.adjust in R.

islands, or transposons (see Table S4 in the supplemental material) (38). The mosaic nature of bacteriophages coupled with the rise of bacteriophage sequencing has resulted in an abundance of spacers that potentially target many different bacteriophages. Interestingly, the ability of a single spacer to target many distinct bacteriophages is likely an evolutionary advantage and possibly a reason some spacers are retained longer than others. To make these data understandable and relatively easy to interpret, we have chosen to list only a single target for each individual spacer, even though they may provide resistance to other, similar elements. As previously observed, spacers were identified targeting nearly the full diversity of known *P. aeruginosa* temperate bacteriophages (see Table S4) (38). However, unlike previous studies, we identified multiple spacers against previously identified and sequenced lytic bacteriophages (KPP23, KPP25, EL, and P3_CHA) (50–53).

While bacteriophages seem to be the major target of *P. aeruginosa* CRISPR-Cas systems, many spacer targets were still identified in previously characterized *P. aeruginosa* plasmids (Fig. 4 and 5B). *Pseudomonas* plasmids play a key role in the horizontal transfer of antibiotic resistance genes between strains (54, 55). In total, 29 previously sequenced plasmids were identified from the NCBI database with some level of spacer target sequences (see Table S4 in the supplemental material). Two of these plasmids (pBS228 and pOZ176) have been demonstrated to endow antibiotic resistance (56, 57). Plasmid pBS228 is virtually identical to the conjugative Birmingham IncP-1α resistance plasmid PK2/RP4/RP1; however, pBS228 contains an extra transposon conferring resistance to streptothricin, streptomycin, spectinomycin, and trimethoprim (56). Similarly, pOZ176 is a conjugative IncP-2 resistance plasmid harboring two integrons encoding resistance to aminoglycosides, carbapenems, chloramphenicol, and carbenicillin (57). Interestingly, no pBS228-like plasmids were identified in our genome library, and while two MDRPA CRISPR-Cas-deficient strains harbored pOZ176-like plasmids, these plasmids did not contain antibiotic resistance genes. In aggregate, these data indicate that while *P. aeruginosa* CRISPR-Cas systems have historically been

found to target elements within known antibiotic resistance plasmids (16), we could not detect these same resistance plasmids in a library containing a substantial number of MDRPA strains.

As another focused example of the plethora of interesting CRISPR-Cas-genomic island interactions observed in our data set (see Table S4 in the supplemental material), we looked closely at a widely distributed spacer targeting the *pilV2* gene of the conjugative type IVb pilus of pKLC102, PAPI-1, and PAGI-5. While we identified multiple spacers in many strains targeting these three widely distributed conjugative elements (Fig. 5B), we chose to focus on a single spacer targeting *pilV2*. PilV2 is the minor subunit of the pilus filament and is required for functional conjugation-mediated horizontal gene transfer of these and other important accessory genomic elements (58). As *pilV2* is required for function and is highly conserved between similar conjugative elements, it provides both a perfect target for the CRISPR-Cas bacterial immune system and an excellent marker for determining the carriage rate of similar elements in our clinical library. Roughly 40% of clinical isolates within our library contained conjugative elements harboring a 100% conserved target and protospacer-adjacent motif (PAM) sequence within the *pilV2* gene, clearly indicating the ubiquity of these elements in the *P. aeruginosa* global population.

The diversity of genetic elements potentially targeted by *P. aeruginosa* CRISPR-Cas systems is striking. However, well characterized mobile genetic elements account for only ~30% of unique spacer matches, while 55% are complementary to largely uncharacterized portions of the *P. aeruginosa* accessory genome. These data clearly illustrate the need for more accurate curation of *P. aeruginosa* accessory genome elements, a challenging process due to the chimeric nature of many mobile elements and the enormous number of *P. aeruginosa* strains being sequenced. However, as this study establishes a massive library of spacer sequences to highlight these interesting elements within NGS data, this challenge should be more manageable.

Correlations between antibiotic resistance and the presence or absence of CRISPR-Cas systems in *P. aeruginosa*. The hypothesis that CRISPR-Cas bacterial adaptive immune systems prevent clinical isolates from acquiring antibiotic resistance elements has been widely discussed; however, it has not been documented outside the enterococci (29). To assess whether CRISPR-Cas systems play a role in *P. aeruginosa* antibiotic resistance, we correlated their presence or absence with phenotypic susceptibility to clinically relevant antibiotics (31, 32) (Fig. 5C). Importantly, all of these correlations were done after correcting for the presence of subtype-specific anti-CRISPR genes (Fig. 3A and 5A).

The presence of type I-E and I-F CRISPR-Cas subtypes strongly correlated with the susceptibility of the strain to multiple clinically relevant antibiotics; however, the presence of these correlations was dependent on the two prevalent MDRPA STs 111 and 235 (Fig. 5C). Puzzlingly, the less prevalent and ICE-localized type I-C subtype showed a strong correlation with resistance to the antibiotic amikacin, a clinically used aminoglycoside (Fig. 5C) (36). To determine whether known antibiotic resistance genes also correlated with the presence or absence of each CRISPR-Cas subtype, statistical analysis was performed for each known resistance gene and allele. We observed that the sulfonamide resistance gene *sul1* was negatively associated with both the type I-E and I-F system; however, the significance of this correlation was again dependent on MDRPA STs 111 and 235. The *sul1* gene is commonly associated with the 3′ region of integrons (59), which, again, often carry a diverse array of other clinically important antibiotic resistance elements (60, 61). A significant negative correlation was also observed between the type I-F subtype and the beta-lactamase CARB. Interestingly, we observed positive correlations between the type I-C CRISPR-Cas subtype and the *rmtD*, *aadA7*, and OXA-encoding resistance genes. These data indicate that correlations between antibiotic resistance and the presence or absence of CRISPR-Cas systems in *P. aeruginosa* are being driven largely by select phylogenetic lineages.

The combined data and associations within this study are consistent with a role for CRISPR-Cas systems in modulating the movement of mobile genetic elements through the global *P. aeruginosa* population. However, we could only identify two resistance plasmids with complementarity to spacer sequences, and we found no spacer complementary to integrons or antibiotic resistance genes. Additionally, both the phenotypic and genotypic antibiotic resistance/susceptibility correlations with CRISPR-Cas systems were heavily dependent on select MDRPA lineages. Taken together, these data are most consistent with *P. aeruginosa* CRISPR-Cas systems acting to shield bacteria from temperate bacteriophages or genomic islands rather than blocking the transfer of resistance genes directly.

DISCUSSION

CRISPR-Cas bacterial immune systems are emerging as major players in bacterial physiology and are increasingly being co-opted for use in molecular biology, genome engineering, and translational applications (62). *P. aeruginosa* is a major CRISPR-Cas model system (20, 22, 24, 31, 32, 37), an important opportunistic human pathogen (3), and a common target for phage therapy applications (63). Together, these attributes highlight the necessity of understanding the ubiquity, conservation, and potential roles of *P. aeruginosa* CRISPR-Cas systems. In this study, we assembled and characterized a 672-strain genomic and phenotypic data set to better understand how CRISPR-Cas systems influence the composition of the *P. aeruginosa* accessory genome on a global scale.

The *P. aeruginosa* pangenome contains numerous accessory elements, among which prophages, genomic islands, plasmids, and CRISPR-Cas systems are some of the most interesting (16, 17). The rarefaction analysis performed in this study demonstrated that the global pangenome is now better defined and that we have likely obtained the majority of common elements (nonsingletons). Our combined analysis provided useful phylogenetic information and the experimental sequences for 27,117 unique pangenome genes. Interestingly, this analysis also found that the majority of unique spacer sequences (~55%) were complementary to elements within the pangenome and not the NCBI plasmid/phage libraries. These spacer-target data illustrated that a large reservoir of understudied genetic elements resides within these genomes. Furthermore, spacer sequences were shown to have substantial value in rapidly identifying novel mobile genetic elements in NGS data. As a proof of principle, three novel prophages and one ICE that are enriched in ST111 were identified. Notably, these 4 mobile elements constitute roughly 215 kb of accessory DNA per strain, nearly matching the average discrepancy in overall genome size between CRISPR-Cas-intact and -deficient strains. While determining the role of these elements is outside the scope of this work, it is tempting to speculate that they

may be virulence-enhancing elements, as has been observed for other plasmids and temperate bacteriophages (14, 64, 65).

As CRISPR spacer sequences have also been shown to be valuable in characterizing bacterial phylogeny and tracking pathogenic bacterial outbreaks in a process termed CRISPR typing (42–45, 66), we aimed to establish a global link between MLST and CRISPR array sequences in *P. aeruginosa*. By creating CRISPR spacer libraries and analyzing the spacer content of each strain against all others, we were able to establish a strong relationship between spacer content and phylogeny in CRISPR array-positive *P. aeruginosa* isolates. The combination of these data and the sequence library potentially supplies the analytical data needed for clinicians and scientists to utilize CRISPR typing to rapidly identify and track roughly 60% (152/253) of *P. aeruginosa* ST lineages. CRISPR typing, while not universal in *P. aeruginosa*, could substantially increase the speed and accuracy of phylogenetic characterization of *P. aeruginosa* clinical or environmental isolates.

The extensive antibiotic susceptibility testing performed on our panel of 282 *P. aeruginosa* isolates coupled with the antibiotic screening performed by Kos and colleagues has allowed us to better characterize MDRPA lineages (5). We found that 41% of strains analyzed were phenotypically MDRPA or XDRPA, substantially higher than documented in previous reports (5). Confirming previous studies, we identified STs 111 and 235 as the predominant clinical and MDRPA clonal groups (5, 7, 8). However, we also identified STs 233, 348, 244, 277, 298, and 357 as being enriched in drug resistance and poised to emerge as high-risk clones. Furthermore, this data set allowed for a robust analysis of how CRISPR-Cas systems impact *P. aeruginosa* antibiotic resistance, a topic of major interest in multiple fields.

Confirming previous small-scale studies (38, 39), the most common intact CRISPR-Cas system found in *P. aeruginosa* is type I-F (30%), followed distantly by type I-E (12%). We also identified a previously unobserved type I-C CRISPR-Cas system encoded on ICEs within the genomes of 20 strains. Prior to this work, no study had identified type I-C or putatively mobile CRISPR-Cas systems in *P. aeruginosa*. Phylogenetic analysis demonstrated that certain common *P. aeruginosa* lineages completely lacked any detectable CRISPR arrays, while others had degenerate systems harboring CRISPR arrays but lacking cognate *cas* genes. Interestingly, some lineages are nearly uniformly endowed with CRISPR-Cas I-F, I-E, or I-C subtypes. By further overlaying the presence of established anti-CRISPR genes (31, 32), we also identified strains and lineages that are CRISPR-Cas intact, yet likely functionally suppressed. The wide distribution of known anti-CRISPR genes within the *P. aeruginosa* population provides important insights for scientists attempting to make CRISPR-Cas-based sequence-specific antimicrobials to combat drug-resistant pathogens (67–71). As these antimicrobials are likely to be ineffective against bacteria that express cognate anti-CRISPR proteins, additional studies must be performed to determine the diversity and distribution of new anti-CRISPR proteins in other antibiotic-resistant pathogens. Importantly, Cas9-based antimicrobials are likely to be functional in *P. aeruginosa* because it harbors only type I CRISPR-Cas systems and, presumably, only type I anti-CRISPR genes.

Previous highly influential studies have implicated type II CRISPR-Cas systems both in hindering the acquisition of antibiotic resistance in enterococci and in enabling antibiotic resistance in *Francisella* (29, 30). However, no link has been identified for type I CRISPR-Cas systems in antibiotic resistance (72). Utilizing our diverse data set, we identified strong correlations between resistance to multiple antibiotics and the absence of type I-F and I-E CRISPR-Cas systems; however, these correlations largely disappeared when the prevalent and largely CRISPR-Cas-deficient MDRPA STs 111 and 235 were removed. Conversely, the strong correlations between antibiotic resistance and the rare type I-C CRISPR-Cas system likely result from these systems being present primarily in MDRPA STs 235 and 277. Consistent with the antibiotic resistance correlations, associations were also observed for specific antibiotic resistance genes and alleles. Together, these data demonstrate a strong correlation between specific CRISPR-Cas systems and antibiotic resistance/susceptibility; however, they do not illustrate that CRISPR-Cas systems are directly causing these lineage-biased correlations. No CRISPR-Cas-targeting spacer sequences complementary to any known antibiotic resistance genes were identified, indicating that *P. aeruginosa* CRISPR-Cas systems are not directly targeting resistance elements. While a small number of spacer target sequences did match known and previously studied antibiotic resistance-encoding plasmids, no spacer target sequences were identified against the small but highly prevalent antibiotic resistance-encoding elements known as integrons. CRISPR-Cas systems were inversely correlated with genome size in *P. aeruginosa*, supporting their now established role in shielding CRISPR-Cas-intact bacteria from horizontal gene transfer (22, 32). Collectively, these data are most consistent with *P. aeruginosa* CRISPR-Cas systems shielding strains from temperate phage infection rather than blocking the flow of antibiotic resistance elements.

The outlined analysis of 672 predominantly clinical *P. aeruginosa* isolates has yielded multiple massive, uniformly annotated datasets for future investigation. These include an abundance of experimental antibiotic resistance data for a sequenced strain collection, a considerable accessory genome library, and a wealth of CRISPR-Cas data. Our phylogenetic analysis of CRISPR-Cas systems and their cognate anti-CRISPR genes provides a bird's eye view of their distribution within the *P. aeruginosa* population and will be highly useful to researchers attempting to further our understanding of the diversity, function, and environmental role of CRISPR-Cas systems.

MATERIALS AND METHODS

Additional details are provided in Text S1 in the supplemental material.

Clinical isolate strain panels. The bacterial strains analyzed in this study were derived from either the bioMérieux private clinical strain collection (219 isolates), the Pirnay collection (62 isolates) (33), or the Kos collection (388 isolates) (5). Information pertaining to each *P. aeruginosa* isolate is available in Table S1 in the supplemental material.

Antibiotic susceptibility testing. Broth dilution assays complemented with Vitek testing (bioMérieux, Marcy-l'Étoile, France) were used to obtain antibiotic resistance data for all isolates contained within the bioMérieux and Pirnay strain collections. Resistance data from previously characterized Kos isolates were imported from publically available data (5). Clinical and Laboratory Standards Institute (CLSI) guidelines were applied to determine susceptibility, intermediate resistance, or resistance to individual antibiotics (see Table S1 in the supplemental material).

DNA isolation and genome sequencing. DNA was extracted from *P. aeruginosa* cells cultured overnight in LB broth at 37°C under 220 rpm agitation. DNA samples were prepared using the UltraClean microbial DNA isolation kit (Mo Bio Laboratories, Carlsbad, CA) essentially according to the manufacturer's instructions. Sequencing was performed using the Illumina paired-end method with read lengths of 150 bp (Illumina HiSeq 2500; Ambry Genetics, Aliso Viejo, CA). Paired-end libraries

were prepared using KAPA kits according to the manufacturer's instructions (KAPA Biosystems, Wilmington, MA).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.01796-15/-/DCSupplemental.

Text S1, DOCX file, 0.05 MB.
Table S1, XLSX file, 0.4 MB.
Table S2, TXT file, 25 MB.
Table S3, XLSX file, 0.1 MB.
Table S4, XLSX file, 1.1 MB.
Table S5, XLSX file, 0.01 MB.
Figure S1, PDF file, 0.7 MB.
Figure S2, PDF file, 0.1 MB.
Figure S3, PDF file, 1 MB.
Data set S1, PDF file, 0.4 MB.

ACKNOWLEDGMENTS

We thank Chris Huff, Stephanie Lewis and Kathy Wilkey (bioMérieux) for their help with the cultivation, identification, and shipment of the *P. aeruginosa* strains. David Creely (bioMérieux) performed the antibiotic susceptibility testing for the newly sequenced isolates. We also thank Christen DiPetrillo (Synthetic Genomics) for her critical review and editing during the preparation of the manuscript.

All authors are employees of bioMerieux or Synthetic Genomics, Inc., and, hence, have a business implication in all work presented here. However, the study was designed and executed in an open manner and all data generated have been deposited in the public domain.

REFERENCES

1. **Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrock-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock RE, Lory S, Olson MV.** 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature **406:**959–964. http://dx.doi.org/10.1038/35023079.
2. **Bodey GP, Bolivar R, Fainstein V, Jadeja L.** 1983. Infections caused by *Pseudomonas aeruginosa*. Rev Infect Dis **5:**279–313. http://dx.doi.org/10.1093/clinids/5.2.279.
3. **Lister PD, Wolter DJ, Hanson ND.** 2009. Antibacterial-resistant *Pseudomonas aeruginosa*: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. Clin Microbiol Rev **22:**582–610. http://dx.doi.org/10.1128/CMR.00040-09.
4. **Curran B, Jonas D, Grundmann H, Pitt T, Dowson CG.** 2004. Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa*. J Clin Microbiol **42:**5644–5649. http://dx.doi.org/10.1128/JCM.42.12.5644-5649.2004.
5. **Kos VN, Déraspe M, McLaughlin RE, Whiteaker JD, Roy PH, Alm RA, Corbeil J, Gardner H.** 2015. The resistome of *Pseudomonas aeruginosa* in relationship to phenotypic susceptibility. Antimicrob Agents Chemother **59:**427–436. http://dx.doi.org/10.1128/AAC.03954-14.
6. **Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG.** 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J Clin Microbiol **38:**1008–1015.
7. **Witney AA, Gould KA, Pope CF, Bolt F, Stoker NG, Cubbon MD, Bradley CR, Fraise A, Breathnach AS, Butcher PD, Planche TD, Hinds J.** 2014. Genome sequencing and characterization of an extensively drug-resistant sequence type 111 serotype O12 hospital outbreak strain of *Pseudomonas aeruginosa*. Clin Microbiol Infect **20:**O609–O618. http://dx.doi.org/10.1111/1469-0691.12528.
8. **Viedma E, Villa J, Juan C, Oliver A, Chaves F.** 2014. Draft genome sequence of colistin-only-susceptible *Pseudomonas aeruginosa* strain ST235, a hypervirulent high-risk clone in Spain. Genome Announc **2:**e01097-14. http://dx.doi.org/10.1128/genomeA.01097-14.
9. **Wiehlmann L, Wagner G, Cramer N, Siebert B, Gudowius P, Morales G, Kohler T, van Delden C, Weinel C, Slickers P, Tummler B.** 2007. Population structure of *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A **104:**8101–8106. http://dx.doi.org/10.1073/pnas.0609213104.
10. **Woodford N, Turton JF, Livermore DM.** 2011. Multiresistant Gram-negative bacteria: the role of high-risk clones in the dissemination of antibiotic resistance. FEMS Microbiol Rev **35:**736–755. http://dx.doi.org/10.1111/j.1574-6976.2011.00268.x.
11. **Samuelsen O, Toleman MA, Sundsfjord A, Rydberg J, Leegaard TM, Walder M, Lia A, Ranheim TE, Rajendra Y, Hermansen NO, Walsh TR, Giske CG.** 2010. Molecular epidemiology of metallo-beta-lactamase-producing *Pseudomonas aeruginosa* isolates from Norway and Sweden shows import of international clones and local clonal expansion. Antimicrob Agents Chemother **54:**346–352. http://dx.doi.org/10.1128/AAC.00824-09.
12. **Turton JF, Wright L, Underwood A, Witney AA, Chan YT, Al-Shahib A, Arnold C, Doumith M, Patel B, Planche TD, Green J, Holliman R, Woodford N.** 2015. High-resolution analysis by whole-genome sequencing of an international lineage (sequence type 111) of *Pseudomonas aeruginosa* associated with metallo-carbapenemases in the United Kingdom. J Clin Microbiol **53:**2622–2631. http://dx.doi.org/10.1128/JCM.00505-15.
13. **He J, Baldini RL, Deziel E, Saucier M, Zhang Q, Liberati NT, Lee D, Urbach J, Goodman HM, Rahme LG.** 2004. The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. Proc Natl Acad Sci U S A **101:**2530–2535. http://dx.doi.org/10.1073/pnas.0304622101.
14. **Winstanley C, Langille MGI, Fothergill JL, Kukavica-Ibrulj I, Paradis-Bleau C, Sanschagrin F, Thomson NR, Winsor GL, Quail MA, Lennard N, Bignell A, Clarke L, Seeger K, Saunders D, Harris D, Parkhill J, Hancock REW, Brinkman FSL, Levesque RC.** 2009. Newly introduced genomic prophage islands are critical determinants of *in vivo* competitiveness in the Liverpool epidemic strain of *Pseudomonas aeruginosa*. Genome Res **19:**12–23. http://dx.doi.org/10.1101/gr.086082.108.
15. **Roy PH, Tetu SG, Larouche A, Elbourne L, Tremblay S, Ren Q, Dodson R, Harkins D, Shay R, Watkins K, Mahamoud Y, Paulsen IT.** 2010. Complete genome sequence of the multiresistant taxonomic outlier *Pseudomonas aeruginosa* PA7. PLoS One **5:**e8842. http://dx.doi.org/10.1371/journal.pone.0008842.
16. **Mathee K, Narasimhan G, Valdes C, Qiu X, Matewish JM, Koehrsen M, Rokas A, Yandava CN, Engels R, Zeng E, Olavarietta R, Doud M, Smith RS, Montgomery P, White JR, Godfrey PA, Kodira C, Birren B, Galagan JE, Lory S.** 2008. Dynamics of *Pseudomonas aeruginosa* genome evolution. Proc Natl Acad Sci U S A **105:**3100–3105. http://dx.doi.org/10.1073/pnas.0711982105.
17. **Klockgether J, Cramer N, Wiehlmann L, Davenport CF, Tümmler B.** 2011. *Pseudomonas aeruginosa* genomic structure and diversity. Front Microbiol **2:**150. http://dx.doi.org/10.3389/fmicb.2011.00150.
18. **Marraffini LA, Sontheimer EJ.** 2008. CRISPR interference limits horizontal gene transfer in *staphylococci* by targeting DNA. Science **322:**1843–1845. http://dx.doi.org/10.1126/science.1165771.
19. **Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P.** 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science **315:**1709–1712. http://dx.doi.org/10.1126/science.1138140.
20. **Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA.** 2010. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. Science **329:**1355–1358. http://dx.doi.org/10.1126/science.1192272.
21. **Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin EV, van der Oost J.** 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science **321:**960–964. http://dx.doi.org/10.1126/science.1159689.
22. **Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, O'Toole GA.** 2012. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. J Bacteriol **194:**5728–5738. http://dx.doi.org/10.1128/JB.01184-12.
23. **Wiedenheft B, Lander GC, Zhou K, Jore MM, Brouns SJJ, van der Oost J, Doudna JA, Nogales E.** 2011. Structures of the RNA-guided surveillance complex from a bacterial immune system. Nature **477:**486–489. http://dx.doi.org/10.1038/nature10402.
24. **Wiedenheft B, van Duijn E, Bultema JB, Waghmare SP, Zhou K, Barendregt A, Westphal W, Heck AJR, Boekema EJ, Dickman MJ, Doudna JA.** 2011. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc Natl

Acad Sci U S A **108:**10092–10097. http://dx.doi.org/10.1073/pnas.1102716108.
25. **Westra E, van Erp PB, Künne T, Wong S, Staals RJ, Seegers CC, Bollen S, Jore M, Semenova E, Severinov K, de Vos W, Dame R, de Vries R, Brouns SJ, van der Oost J.** 2012. CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by cascade and Cas3. Mol Cell **46:**595–605. http://dx.doi.org/10.1016/j.molcel.2012.03.018.
26. **Mulepati S, Bailey S.** 2011. Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). J Biol Chem **286:**31896–31903. http://dx.doi.org/10.1074/jbc.M111.270017.
27. **Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V.** 2011. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. EMBO J **30:**1335–1342. http://dx.doi.org/10.1038/emboj.2011.41.
28. **Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, Moineau S, Mojica FJM, Wolf YI, Yakunin AF, van der Oost J, Koonin EV.** 2011. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol **9:**467–477. http://dx.doi.org/10.1038/nrmicro2577.
29. **Palmer KL, Gilmore MS.** 2010. Multidrug-resistant enterococci lack CRISPR-cas. mBio **1:**e00227-10. http://dx.doi.org/10.1128/mBio.00227-10.
30. **Sampson TR, Weiss DS.** 2014. CRISPR-Cas systems: new players in gene regulation and bacterial physiology. Front Cell Infect Microbiol **4:**37. http://dx.doi.org/10.3389/fcimb.2014.00037.
31. **Bondy-Denomy J, Pawluk A, Maxwell KL, Davidson AR.** 2013. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. Nature **493:**429–432. http://dx.doi.org/10.1038/nature11723.
32. **Pawluk A, Bondy-Denomy J, Cheung VHW, Maxwell KL, Davidson AR.** 2014. A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. mBio **5:**e00896-14. http://dx.doi.org/10.1128/mBio.00896-14.
33. **Pirnay J, Bilocq F, Pot B, Cornelis P, Zizi M, Van Eldere J, Deschaght P, Vaneechoutte M, Jennes S, Pitt T, De Vos D.** 2009. *Pseudomonas aeruginosa* population structure revisited. PLoS One **4:**e7740. http://dx.doi.org/10.1371/journal.pone.0007740.
34. **Hsueh PR, Tseng SP, Teng LJ, Ho SW.** 2005. Pan-drug-resistant *Pseudomonas aeruginosa* causing nosocomial infection at a university hospital in Taiwan. Clin Microbiol Infect **11:**670–673. http://dx.doi.org/10.1111/j.1469-0691.2005.01196.x.
35. **Willmann M, Bezdan D, Zapata L, Susak H, Vogel W, Schroppel K, Liese J, Weidenmaier C, Autenrieth IB, Ossowski S, Peter S.** 2015. Analysis of a long-term outbreak of XDR *Pseudomonas aeruginosa*: a molecular epidemiological study. J Antimicrob Chemother **70:**1322–1330. http://dx.doi.org/10.1093/jac/dku546.
36. **Magiorakos A-, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J, Struelens MJ, Vatopoulos A, Weber JT, Monnet DL.** 2012. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect **18:**268–281. http://dx.doi.org/10.1111/j.1469-0691.2011.03570.x.
37. **Rollins MF, Schuman JT, Paulus K, Bukhari HST, Wiedenheft B.** 2015. Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from *Pseudomonas aeruginosa*. Nucleic Acids Res **43:**2216–2222. http://dx.doi.org/10.1093/nar/gkv094.
38. **Cady KC, White AS, Hammond JH, Abendroth MD, Karthikeyan RSG, Lalitha P, Zegans ME, O'Toole GA.** 2011. Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. Microbiology **157:**430–437. http://dx.doi.org/10.1099/mic.0.045732-0.
39. **Essoh C, Blouin Y, Loukou G, Cablanmian A, Lathro S, Kutter E, Thien HV, Vergnaud G, Pourcel C.** 2013. The susceptibility of *Pseudomonas aeruginosa* strains from cystic fibrosis patients to bacteriophages. PLoS One **8:**e60575. http://dx.doi.org/10.1371/journal.pone.0060575.
40. **Klockgether J, Reva O, Larbig K, Tummler B.** 2004. Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. J Bacteriol **186:**518–534. http://dx.doi.org/10.1128/JB.186.2.518-534.2004.
41. **Bondy-Denomy J, Garcia B, Strum S, Du M, Rollins MF, Hidalgo-Reyes Y, Wiedenheft B, Maxwell KL, Davidson AR.** 2015. Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. Nature **526:**136–139. http://dx.doi.org/10.1038/nature15254.
42. **Fabre L, Zhang J, Guigon G, Le Hello S, Guibert V, Accou-Demartin M, de Romans S, Lim C, Roux C, Passet V, Diancourt L, Guibourdenche M, Issenhuth-Jeanjean S, Achtman M, Brisse S, Sola C, Weill F.** 2012. CRISPR typing and subtyping for improved laboratory surveillance of *Salmonella* infections. PLoS One **7:**e36995. http://dx.doi.org/10.1371/journal.pone.0036995.
43. **Mokrousov I, Limeschenko E, Vyazovaya A, Narvskaya O.** 2007. *Corynebacterium diphtheriae* spoligotyping based on combined use of two CRISPR loci. Biotechnol J **2:**901–906. http://dx.doi.org/10.1002/biot.200700035.
44. **Schouls LM, Reulen S, Duim B, Wagenaar JA, Willems RJL, Dingle KE, Colles FM, Van Embden JDA.** 2003. Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. J Clin Microbiol **41:**15–26. http://dx.doi.org/10.1128/JCM.41.1.15-26.2003.
45. **Filliol I, Ferdinand S, Negroni L, Sola C, Rastogi N.** 2000. Molecular typing of *Mycobacterium tuberculosis* based on variable number of tandem DNA repeats used alone and in association with spoligotyping. J Clin Microbiol **38:**2520–2524.
46. **Groenen PMA, Bunschoten AE, van Soolingen D, van Embden JDA.** 1993. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. Mol Microbiol **10:**1057–1065. http://dx.doi.org/10.1111/j.1365-2958.1993.tb00976.x.
47. **Pourcel C, Salvignol G, Vergnaud G.** 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology **151:**653–663. http://dx.doi.org/10.1099/mic.0.27437-0.
48. **Stern A, Keren L, Wurtzel O, Amitai G, Sorek R.** 2010. Self-targeting by CRISPR: gene regulation or autoimmunity? Trends Genet **26:**335–340. http://dx.doi.org/10.1016/j.tig.2010.05.008.
49. **Wu D, Ye J, Ou H, Wei X, Huang X, He Y, Xu Y.** 2011. Genomic analysis and temperature-dependent transcriptome profiles of the rhizosphere originating strain *Pseudomonas aeruginosa* M18. BMC Genomics **12:**438. http://dx.doi.org/10.1186/1471-2164-12-438.
50. **Yamaguchi K, Miyata R, Shigehisa R, Uchiyama J, Takemura-Uchiyama I, Kato S-, Ujihara T, Sakaguchi Y, Daibata M, Matsuzaki S.** 2014. Genome analysis of *Pseudomonas aeruginosa* bacteriophage KPP23, belonging to the family *Siphoviridae*. Genome Announc **2:**e00233-14. http://dx.doi.org/10.1128/genomeA.00233-14.
51. **Miyata R, Yamaguchi K, Uchiyama J, Shigehisa R, Takemura-Uchiyama I, Kato S, Ujihara T, Sakaguchi Y, Daibata M, Matsuzaki S.** 2014. Characterization of a novel *Pseudomonas aeruginosa* bacteriophage, KPP25, of the family *Podoviridae*. Virus Res **189:**43–46. http://dx.doi.org/10.1016/j.virusres.2014.04.019.
52. **Hertveldt K, Lavigne R, Pleteneva E, Sernova N, Kurochkina L, Korchevskii R, Robben J, Mesyanzhinov V, Krylov VN, Volckaert G.** 2005. Genome comparison of *Pseudomonas aeruginosa* large phages. J Mol Biol **354:**536–545. http://dx.doi.org/10.1016/j.jmb.2005.08.075.
53. **Morello E, Saussereau E, Maura D, Huerre M, Touqui L, Debarbieux L.** 2011. Pulmonary bacteriophage therapy on *Pseudomonas aeruginosa* cystic fibrosis strains: first steps towards treatment and prevention. PLoS One **6:**e16963. http://dx.doi.org/10.1371/journal.pone.0016963.
54. **Schlüter A, Szczepanowski R, Pühler A, Top EM.** 2007. Genomics of IncP-1 antibiotic resistance plasmids isolated from wastewater treatment plants provides evidence for a widely accessible drug resistance gene pool. FEMS Microbiol Rev **31:**449–477. http://dx.doi.org/10.1111/j.1574-6976.2007.00074.x.
55. **Kung VL, Ozer EA, Hauser AR.** 2010. The accessory genome of *Pseudomonas aeruginosa*. Microbiol Mol Biol Rev **74:**621–641. http://dx.doi.org/10.1128/MMBR.00027-10.
56. **Haines AS, Jones K, Batt SM, Kosheleva IA, Thomas CM.** 2007. Sequence of plasmid pBS228 and reconstruction of the IncP-1alpha phylogeny. Plasmid **58:**76–83. http://dx.doi.org/10.1016/j.plasmid.2007.01.001.
57. **Xiong J, Alexander DC, Ma JH, Deraspe M, Low DE, Jamieson FB, Roy PH.** 2013. Complete sequence of pOZ176, a 500-kilobase IncP-2 plasmid encoding IMP-9-mediated carbapenem resistance, from outbreak isolate *Pseudomonas aeruginosa* 96. Antimicrob Agents Chemother **57:**3775–3782. http://dx.doi.org/10.1128/AAC.00423-13.
58. **Carter MQ, Chen J, Lory S.** 2010. The *Pseudomonas aeruginosa* pathoge-

nicity island PAPI-1 is transferred via a novel type IV pilus. J Bacteriol **192:**3249–3258. http://dx.doi.org/10.1128/JB.00041-10.

59. **Hall RM, Brown HJ, Brookes DE, Stokes HW**. 1994. Integrons found in different locations have identical 5′ ends but variable 3′ ends. J Bacteriol **176:**6286–6294.
60. **Fonseca ÃL, Vieira VV, Cipriano R, Vicente ACP**. 2005. Class 1 integrons in *Pseudomonas aeruginosa* isolates from clinical settings in Amazon region, Brazil. FEMS Immunol Med Microbiol **44:**303–309. http://dx.doi.org/10.1016/j.femsim.2005.01.004.
61. **Kouda S, Ohara M, Onodera M, Fujiue Y, Sasaki M, Kohara T, Kashiyama S, Hayashida S, Harino T, Tsuji T, Itaha H, Gotoh N, Matsubara A, Usui T, Sugai M**. 2009. Increased prevalence and clonal dissemination of multidrug-resistant *Pseudomonas aeruginosa* with the blaIMP-1 gene cassette in Hiroshima. J Antimicrob Chemother **64:**46–51. http://dx.doi.org/10.1093/jac/dkp142.
62. **Sander JD, Joung JK**. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol **32:**347–355. http://dx.doi.org/10.1038/nbt.2842.
63. **Debarbieux L, Leduc D, Maura D, Morello E, Criscuolo A, Grossi O, Balloy V, Touqui L**. 2010. Bacteriophages can treat and prevent *Pseudomonas aeruginosa* lung infections. J Infect Dis **201:**1096–1104. http://dx.doi.org/10.1086/651135.
64. **Wagner PL, Waldor MK**. 2002. Bacteriophage control of bacterial virulence. Infect Immun **70:**3985–3993. http://dx.doi.org/10.1128/IAI.70.8.3985-3993.2002.
65. **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.
66. **Van Belkum A, Scherer S, van Alphen L, Verbrugh H**. 1998. Short-sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev **62:**275–293.
67. **Yosef I, Manor M, Kiro R, Qimron U**. 2015. Temperate and lytic bacteriophages programmed to sensitize and kill antibiotic-resistant bacteria. Proc Natl Acad Sci U S A **112:**7267–7272. http://dx.doi.org/10.1073/pnas.1500107112.
68. **Citorik RJ, Mimee M, Lu TK**. 2014. Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. Nat Biotechnol **32:**1141–1145. http://dx.doi.org/10.1038/nbt.3011.
69. **Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, Fischetti VA, Marraffini LA**. 2014. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nat Biotechnol **32:**1146–1150. http://dx.doi.org/10.1038/nbt.3043.
70. **Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL**. 2014. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. mBio **5:**e00928-13. http://dx.doi.org/10.1128/mBio.00928-13.
71. **Vercoe RB, Chang JT, Dy RL, Taylor C, Gristwood T, Clulow JS, Richter C, Przybilski R, Pitman AR, Fineran PC**. 2013. Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. PLoS Genet **9:**e1003454. http://dx.doi.org/10.1371/journal.pgen.1003454.
72. **Touchon M, Charpentier S, Pognard D, Picard B, Arlet G, Rocha EPC, Denamur E, Branger C**. 2012. Antibiotic resistance plasmids spread among natural isolates of *Escherichia coli* in spite of CRISPR elements. Microbiology **158:**2997–3004. http://dx.doi.org/10.1099/mic.0.060814-0.