Article

Systematic In Silico Assessment of Antimicrobial Resistance Dissemination across the Global Plasmidome

Miquel Sánchez-Osuna [1,†,‡] ⓘ**, Jordi Barbé** [1,*] **and Ivan Erill** [2,*] ⓘ

1 Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
2 Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, MD 21250, USA
* Correspondence: jordi.barbe@uab.cat (J.B.); erill@umbc.edu (I.E.); Tel.: +1-410-455-2470 (I.E.); Fax: +1-410-455-3875 (I.E.)
† Current address: Laboratori de Recerca en Microbiologia i Malalties Infeccioses, Institut d'Investigació i Innovació Parc Taulí (I3PT), Hospital Universitari Parc Taulí, Universitat Autònoma de Barcelona, 08208 Sabadell, Spain.
‡ Current address: Institut de Biotecnologia i Biomedicina (IBB), Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.

Abstract: The emergence of pathogenic strains resistant to multiple antimicrobials is a pressing problem in modern healthcare. Antimicrobial resistance is mediated primarily by dissemination of resistance determinants via horizontal gene transfer. The dissemination of some resistance genes has been well documented, but few studies have analyzed the patterns underpinning the dissemination of antimicrobial resistance genes. Analyzing the %GC content of plasmid-borne antimicrobial resistance genes relative to their host genome %GC content provides a means to efficiently detect and quantify dissemination of antimicrobial resistance genes. In this work we automate %GC content analysis to perform a comprehensive analysis of known antimicrobial resistance genes in publicly available plasmid sequences. We find that the degree to which antimicrobial resistance genes are disseminated depends primarily on the resistance mechanism. Our analysis identifies conjugative plasmids as primary dissemination vectors and indicates that most broadly disseminated genes have spread from single genomic backgrounds. We show that resistance dissemination profiles vary greatly among antimicrobials, oftentimes reflecting stewardship measures. Our findings establish %GC content analysis as a powerful, intuitive and scalable method to monitor the dissemination of resistance determinants using publicly available sequence data.

Keywords: antimicrobial resistance; dissemination; evolution; plasmid; mobile genetic element; GC content; horizontal gene transfer

1. Introduction

The discovery of antibacterial compounds is considered one of the great advances in modern medicine, since it has paved the way for the effective treatment of infections caused by pathogenic bacteria. However, these agents have been losing their efficacy due to the emergence and dissemination of resistance among bacterial pathogens [1]. In this sense, the number of infections caused by multi-resistant bacteria is increasing globally, making some of them untreatable. This has prompted the World Health Organization (WHO) to classify antibacterial resistance as one of the three most important public health threats of the 21st century [2].

The development of resistance to antibacterial compounds is associated with their misuse and abuse [3]. The high genetic plasticity of bacteria enables them to rapidly adapt to a wide range of environmental threats, including the one posed by antibacterial agents. In addition to the clinical environment, animal production, agriculture, aquaculture and effluents from pharmaceutical industries and municipal wastewater systems have all been postulated as hotspots for the emergence of antibacterial resistance [4]. These environments

are characterized by high bacterial loads, where constant exposure to sub-inhibitory concentrations of different antimicrobial agents can promote the selection, mobilization and dissemination of genes conferring resistance to them [5]. In these environments, bacteria can leverage two alternative processes to mitigate the effects of antimicrobial agents. On the one hand, spontaneous mutations conferring resistance may occur and be selected for. These mutations may target genes directly associated with the mechanism of action of the drug, impact genes indirectly associated, such as mutations on regulatory processes or interacting partners, or confer cryptic resistance by modifying different aspects of bacterial physiology [6]. On the other hand, bacteria may acquire foreign genetic material encoding resistance to antibacterial compounds through horizontal gene transfer (HGT) [7]. Consequently, the extensive use of antimicrobial agents has led to the co-selection of resistance genes, their mobilization, dissemination and fixation, and to the increased presence of multiple resistance genes in bacterial populations in many ecosystems [8].

It is now widely accepted that the rapid proliferation of multi-resistant pathogens has not been driven by the independent evolution of mutations conferring resistance to different antimicrobials, but through the spreading of mobile genetic elements (MGE) carrying resistance genes via HGT [9,10]. Bacteria classically take up foreign DNA through transformation (direct DNA acquisition from the environment), transduction (acquisition of DNA mediated by bacteriophages), conjugation (acquisition of DNA mediated by cell-cell contact) and vesiduction (cell fusion with membrane vesicles) [11,12]. Among these mechanisms, conjugative plasmids are presumed to be the main drivers of the dissemination of antimicrobial resistance determinants, and they appear to have played an outsized role in the current global antimicrobial resistance crisis [13]. Plasmid-borne antimicrobial resistance factors can modify or protect the antimicrobial target, inactivate the antimicrobial molecule, decrease the intracellular availability of the antibacterial compound or act as alternative enzymes [9].

Several studies have shown that some plasmid-borne genes conferring resistance to antibiotics in clinical isolates have originated by successive horizontal gene transfer of homologous genes available since ancient times [8]. Recent work indicates that this is also true for genes conferring resistance to manmade chemotherapeutic agents, with resistance predating the invention and use of these antimicrobial compounds [14,15]. Therefore, the mobilization of resistance genes from different source microorganisms and their subsequent dissemination by mobile genetic elements has given rise to the multiple variants of antimicrobial resistance determinants present today in pathogenic bacteria [10].

Many studies have elucidated the factors and mechanisms that govern the dispersal of MGEs [16–18], and the dissemination of specific resistance genes has been thoroughly documented [8]. However, relatively few studies have addressed the principles governing the mobilization and subsequent dissemination of antimicrobial resistance genes (ARGs) [19,20]. We have previously shown that analyzing the %GC distribution of resistance genes in plasmids can highlight independent mobilization events [14,15]. Bacterial %GC content is predominantly driven by mutational biases in the replication machinery, and therefore stable among closely related species [21]. The %GC content and codon usage pattern of a horizontally transferred gene will gradually change to match its host genome. This amelioration process is driven by the mutational processes of the host. Over time, introgressed genes accumulate mutations following the host mutational biases, and may also undergo selection to match the host codon usage bias [22]. Phylogenetic evidence and mathematical modeling show that the amelioration process takes millions of years, with estimated rates of divergence in %GC content for coding sequences of 0.045% and 0.91% per million years at nonsynonymous and synonymous codons [15,22]. Plotting host chromosome versus antimicrobial resistance gene %GC content can thus generate two markedly different scenarios, as illustrated in Figure 1. If a resistance gene has been mobilized from the chromosome in different lineages, but has been transferred only among closely related species, the ARG %GC content is expected to align with the host chromosome %GC content, resulting in a diagonal plot (Figure 1A). In contrast, if a resistance gene has been rapidly

and widely disseminated following one or more original mobilization events, the ARG %GC content will be essentially invariant (since amelioration takes millions of years) and hence independent from the host %GC content, presenting as one or more horizontal dissemination bands in the %GC plot (Figure 1B), as reported previously [14].
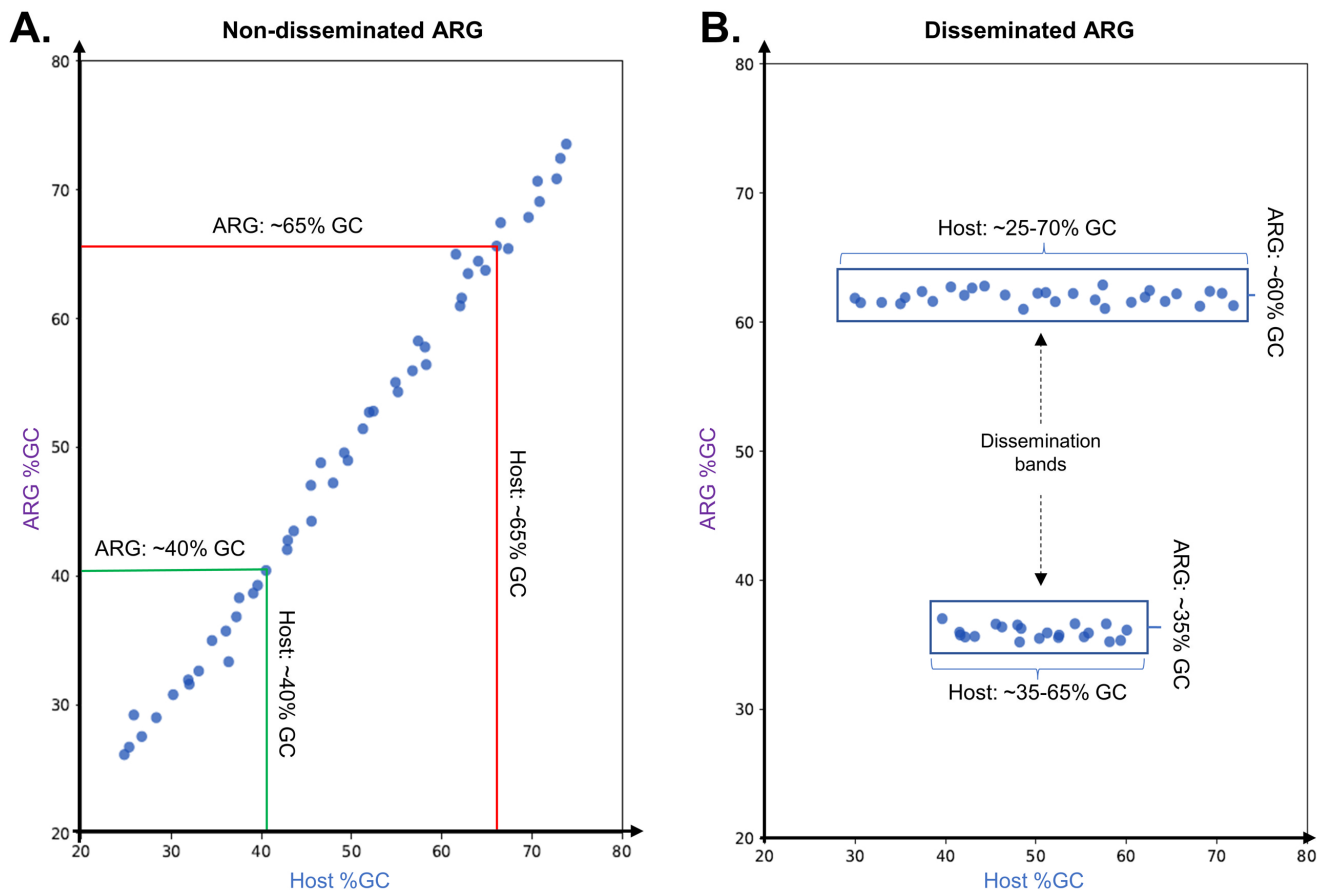
Figure 1. Graphical representation of hypothetical scenarios for the correlation between the %GC content of plasmid-borne ARG and that of their host genome. (**A**) The ARG %GC content is expected to align with the %GC content of the host chromosome when a resistance gene has been mobilized in multiple lineages but has been transferred only among closely related species. (**B**) The ARG %GC content is essentially independent of the host %GC content when a resistance gene has been rapidly disseminated from one or more mobilized instances to a heterogenous group of hosts, generating visible "dissemination bands".

In this work, we leverage this %GC content-based analysis to systematically assess the dissemination of known resistance determinants across the plasmidome, which represents a major concern for public health [2]. Our results establish the analysis of %GC distribution as a direct and powerful method to monitor the dissemination of resistance determinants. In accordance with recent results, we find that the prevalence of ARG dissemination is determined primarily by the mechanism of action of the resistance gene, that resistance to a significant number of antimicrobial compounds has been disseminated multiple times from independent sources, and that conjugative plasmids play an outsized role in ARG dissemination. We report that different antimicrobial classes present remarkably diverse ARG dissemination profiles, and that plasmid-based ARG dissemination is not detected for many restricted-use antimicrobials.

2. Results and Discussion

2.1. Independent Mobilization and Dissemination of Antimicrobial Resistance Genes

To explore ARG dissemination across the global plasmidome, we used a protein sequence panel of representative ARG from the Comprehensive Antibiotic Resistance Database (CARD) to detect putative ARG homologs in plasmid sequences from the PLSDB database. We then compiled the nucleotide sequences of the genes encoding these proteins and computed their %GC content, as well as the chromosomal %GC content of the bacterial strain, species or genus harboring the plasmid (Supplementary Materials Table S1). Previous studies have used Antibiotic Resistance Ontology (ARO) identifiers for individual ARG to analyze their dissemination [20], but this level of granularity arbitrarily separates protein homologs that should be considered jointly for %GC analysis. Here, we mapped ARG homologs to Clusters of Orthologous Groups (COG) identifiers and we grouped ARG homologs based on their CARD-assigned antimicrobial resistance (AMR) Gene Family and the COG they mapped to (ARO-COG groups). Dissemination bands were then automatically detected in chromosome vs. ARG %GC plots for all ARO-COG groups (Figure S1, Table S1). Our analysis detected 69,679 ARG homologs in 21,789 complete plasmid sequences, which were assigned to 389 distinct ARO-COG groups.

In recent work, we have shown that genes encoding resistance to the chemotherapeutic agents sulfonamide (*sul*) and trimethoprim (*dfrA*) have been mobilized multiple times from different genomic backgrounds and subsequently disseminated across a wide range of bacterial clades, leading to the observation of multiple dissemination bands in %GC content plots [14]. To ascertain the generality of this phenomenon, we analyzed the number of dissemination bands detected across ARO-COG groups. Our results reveal that dissemination bands are only detected in a relatively small fraction (18.7%) of the 389 ARO-COG groups analyzed (Figure 2, Table S2), indicating that most mobilized ARG homologs do not disseminate broadly. Among those ARO-COG groups showing evidence of dissemination, most (63.8%) display only a single dissemination band (Figure 3A), suggesting that unique mobilization events are responsible for a majority of broadly disseminated antibiotic resistance genes. Widespread dissemination has been experimentally reported for 29 of the 69 ARO-COG groups with detected dissemination bands, validating the ability of this method to detect ARG dissemination (Table S3).
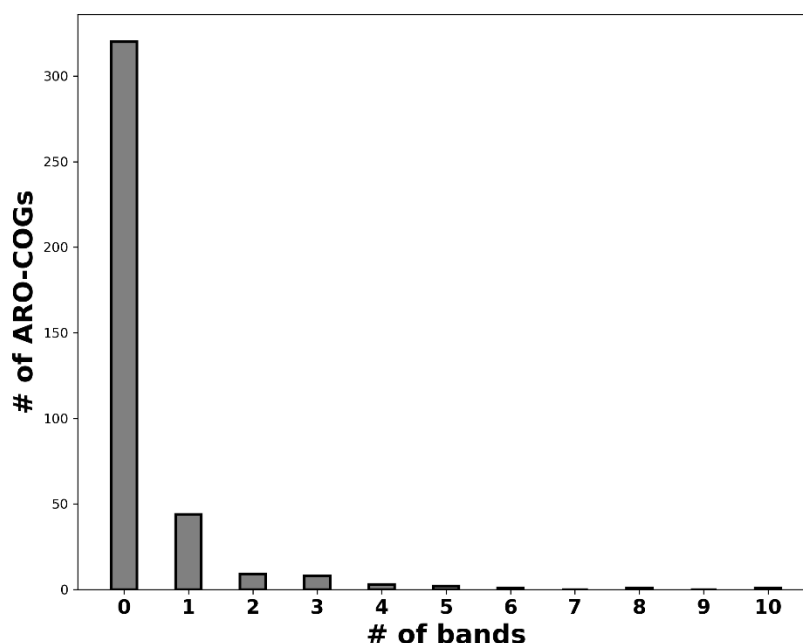
Figure 2. Histogram depicting the distribution of the number of predicted dissemination bands per ARO-COG group.
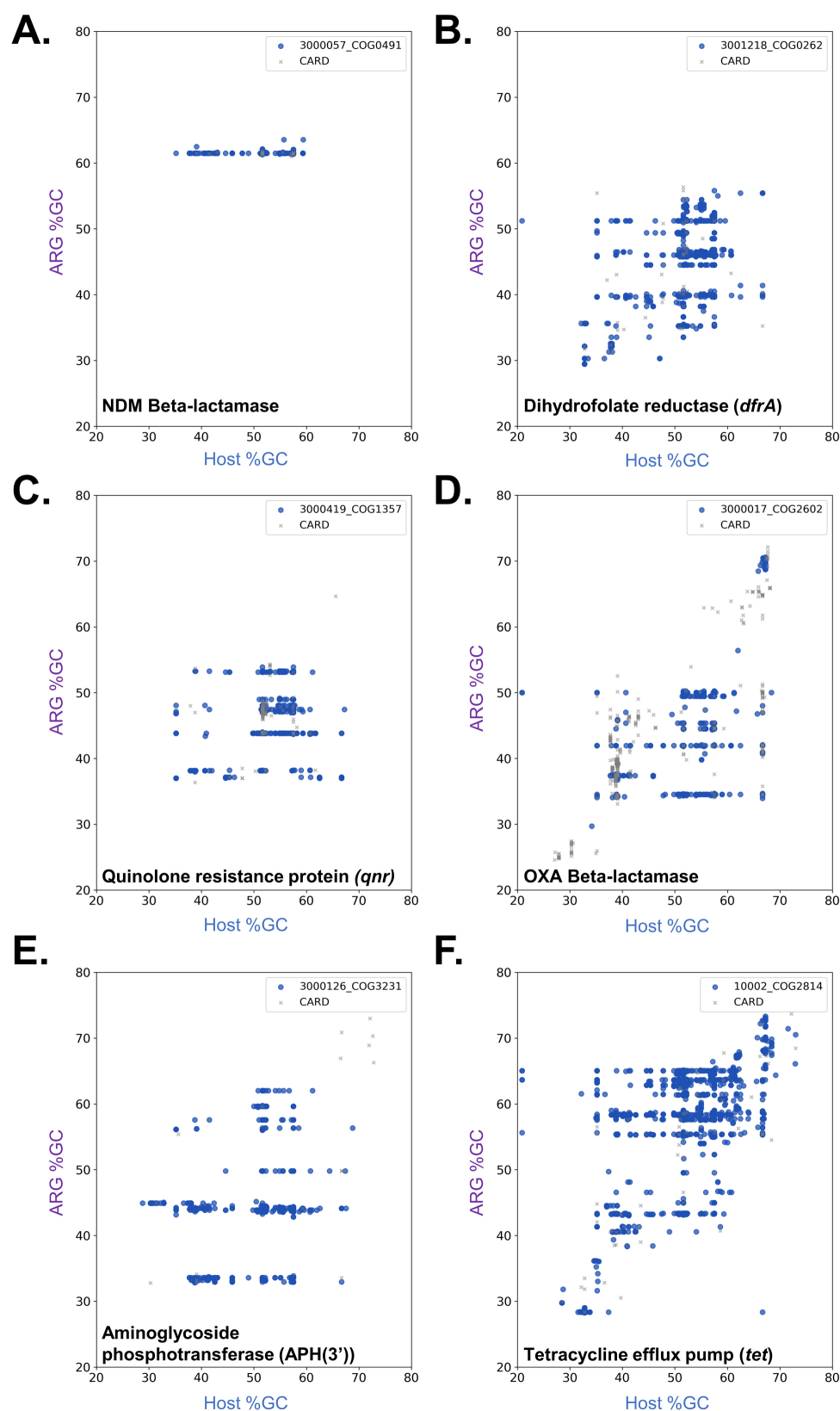
Figure 3. Scatter plots showing the correlation between the %GC content of mobile NDM beta-lactamase (**A**), *dfrA* (**B**), *qnr* (**C**), OXA beta-lactamase (**D**), APH(3′) (**E**) and *tet* (**F**) genes and that of their host chromosome. Only genes encoding redundant protein homologs (>90% sequence similarity) are displayed. Experimentally validated resistance genes annotated in CARD that map to each ARO-COG group are shown as grey crosses.

Several ARO-COG groups display multiple dissemination bands. We have previously shown that *dfrA* genes (ARO:3001218 | COG0262) were disseminated following independent mobilization events in different donor strains [14]. This is reflected in the detection of six different mobilization bands for this ARO-COG group, with putative donor genera %GC ranging from 30% to 53% (Figure 3B). Similarly, four independent dissemination bands are detected in the ARO-COG group for *qnr* genes providing resistance to quinolones (ARO:3000419 | COG1357) (Figure 3C), in agreement with previous reports for multiple chromosomal origins for these genes [23,24].

Independent mobilization events have also been postulated as the source of the heavily disseminated CTX-M β-lactamase genes [25]. Even though these mobilization events have been traced to chromosomal genes in different *Kluyvera* species, our analysis detects three distinct dissemination bands corresponding to previously established CTX-M clusters (Figure 3D). In agreement with previous phylogenetic analyses, this suggests that *Kluyvera* species acquired these genes from different donors prior to the mobilization events that fostered their widespread dissemination [25]. Similar results are obtained for OXA β-lactamase genes (ARO:3000017 | COG2602), where four dissemination bands are detected. This is consistent with prior reports on their likely divergent chromosomal origins [26].

Among the ARO-COG groups presenting multiple dissemination bands, those conferring resistance to aminoglycosides are significantly overrepresented. Independent dissemination bands are detected for aminoglycoside modifying enzyme classes APH(3′) (ARO:3000126 | COG3231) (Figure 3E), ANT(3″) (ARO:3004275 | COG1708), AAC(6′) (ARO:3000345 | COG1670), APH(6) (ARO:3000151 | COG3570) and AAC(3) (ARO:3000322 | COG2746), in agreement with previous reports on their wide phylogenetic distribution and postulated multiple independent origins [27,28]. The ARO-COG groups with the largest number of detected dissemination bands correspond to efflux pumps conferring resistance to tetracycline and phenicol (ARO:10002 | COG2814; 10 bands) and quaternary ammonium compounds (ARO:10002 | COG0477; 8 bands). The %GC plot profiles for these ARO-COG groups reveal extensive mobilization of homologs from these transporter families (Figure 3F). This indicates that efflux pumps conferring resistance to these compounds are present in most chromosomal backgrounds, from which they have been repeatedly disseminated [20,29].

Overall, our analysis indicates that widespread dissemination of ARGs from multiple independent mobilization events is relatively infrequent and restricted to a fairly specific set of ARG families but can be readily detected with %GC plot profiles.

2.2. Resistance Mechanisms Determine the Prevalence of ARG Dissemination

The ability to consistently detect dissemination bands in %GC plots for ARO-COG groups led us to develop a Dissemination Index (DI) as a means to systematically quantify and analyze the dissemination of different ARG families. This index measures the proportion of ARG homologs mapping to dissemination bands, with respect to the total number of ARG homologs detected for a given ARO-COG group. As noted earlier, dissemination bands are not detected for most ARO-COG groups, leading to zero values for the dissemination index in 81.3% of ARO-COG groups (Table S2). For ARO-COG groups with detected dissemination bands, there is no significant correlation between DI values and the number of dissemination bands detected (Pearson $r = 0.068$) (Figure S2). This indicates that the selective pressure underpinning the widespread dissemination of a particular ARG does not necessarily elicit its independent mobilization from different chromosomal backgrounds. This apparently counterintuitive result can be explained by multiple overlapping rationales. For instance, efflux pumps conferring resistance to tetracycline and phenicol are highly mobilized (Figure 3F), leading to repeated dissemination events but also to a low DI due to the large number of non-disseminated ARG homologs detected in plasmids for this ARO-COG group. The high degree of mobilization of efflux pumps may be explained by their cryptic interaction with chromosomal elements coordinating transport processes, which may induce changes in bacterial physiology that promote mobilization [29,30]. On

the other hand, genes displaying a single dissemination band, like the NDM β-lactamases (Figure 3A), might have fewer diverse chromosomal sources (not differentiable in a %GC analysis) and may have associated early with highly transmissible MGEs, relieving the pressure for dissemination from independent sources [31].

To identify trends in dissemination index values among the analyzed ARO-COG groups, we analyzed the distribution of DI values when mapping ARO-COG groups to their CARD Resistance Mechanism. The results shown in Figure 4 reveal dramatic differences in DI values for different resistance mechanisms. For obvious reasons, no dissemination is observed for ARG mediating resistance through the absence of the antimicrobial target. Similarly, no dissemination is observed for resistance mediated by reduced permeability to antimicrobials, reflecting both its low prevalence and the fact that in most cases it is achieved through the loss or knockdown mutation of a porin [32]. In contrast, three resistance mechanisms (antibiotic inactivation, target replacement and target protection) display remarkably high DI values (Figure 4). Among these three primary resistance mechanisms, antibiotic inactivation genes are the most consistently disseminated, with an average DI of 81.2%. Several ARO-COG groups encompassing different β-lactamase classes (OKP, TEM, KPC and NDM) and aminoglycoside modifying enzymes (APH(3″) and AAC(6′)-Ib-cr) display DI values in the 99.0–100.0% range, indicating that ARG homologs belonging to these ARO-COG groups are almost exclusively detected in dissemination bands (Table S2). This remarkable degree of dissemination may be linked to the early adoption of both β-lactams and aminoglycosides in the clinical practice, which led to heightened pressure for the selection of resistant strains [33]. Other highly disseminated ARO-COG groups mediating antibiotic inactivation are the macrolide phosphotransferases (ARO:3000333 | COG3173; DI: 97.7%), the fosfomycin thiol-transferases (ARO:3000133 | COG2514; 87.5%) and, to a lesser extent, the chloramphenicol acetyltransferases (ARO:3000122 | COG0110; DI: 70.3%).
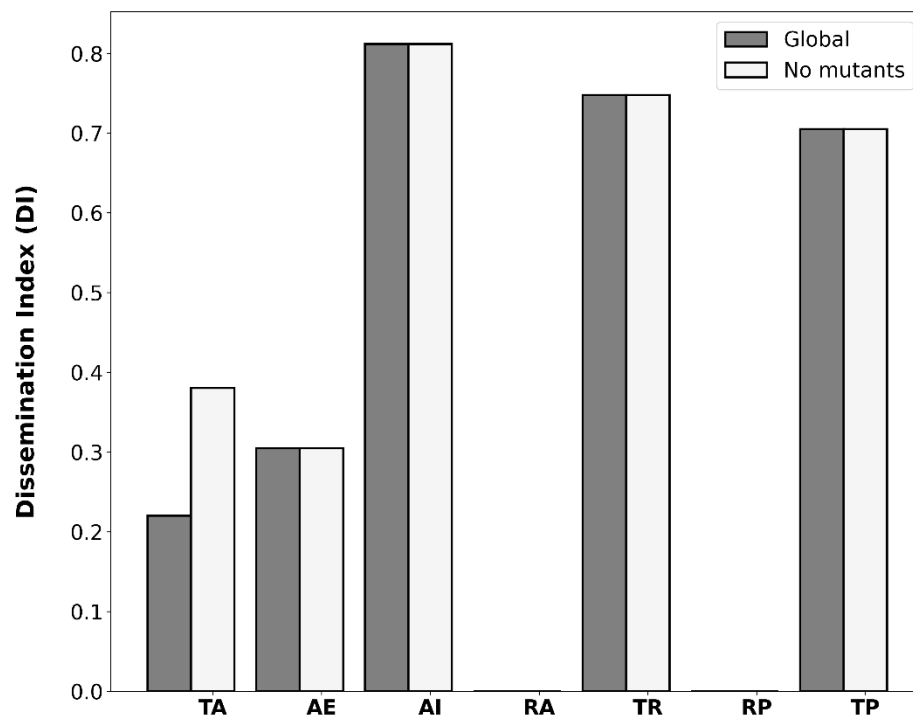
Figure 4. Bar plot of DI values aggregated on CARD Resistance Mechanism: target alteration (TA), antibiotic efflux (AE), antibiotic inactivation (AI), resistance by absence (RA), target replacement (TR), reduced permeability (RP) or target protection (TP). The dissemination status (in band, not in band) of each ARG homolog mapping to a given resistance mechanism is used for the computation of the aggregate DI value. The DI values of each resistance mechanism when removing any ARG homologs annotated as conferring resistance via mutation are also reported.

Consistent with recent reports, genes conferring resistance to sulfonamides (ARO: 3004238 | COG0294) and trimethoprim (ARO:3001218 | COG0262) via target replacement also showed high DI values (89.0% and 84.3%, respectively) [14]. Sulfonamides have been used in combination with trimethoprim and other diaminopyrimidines since the 1960's [33], and genes conferring resistance to both antimicrobial agents are frequently found together on highly transmissible MGEs, such as class 1 integrons and conjugative plasmids [34]. In contrast, alternate variants of the penicillin-binding proteins that confer resistance to methicillin (ARO:3001208 | COG0768) showed no evidence of widespread dissemination (Table S2), as reported previously [35]. Genes *mecA* and *mecC* are typically found as members of the staphylococcal chromosomal cassettes (SCCmec) and known to be mobilizable [36], but the almost exclusive clinical use of methicillin against *Staphylococcus aureus* has likely limited the dissemination of these penicillin-binding proteins to this genus.

Some target protection genes also display relatively large DI values, leading to an average DI of 70.5% for this resistance mechanism. Representative ARO-COG groups for this mechanism include homologs of the *qnr* determinants conferring resistance to fluoroquinolones (ARO:3000419 | COG1357; DI: 89.9%), the *tetM* and *tetO* genes providing resistance to tetracyclines (ARO:0000002 | COG0480; DI: 78.5%) and the ribosomal protection genes *vgaA* and *vlmR* providing resistance to lincosamide and streptogramin (ARO:3004469 | COG0488; DI: 46.5%) (Table S2). The high dissemination indices of antibiotic inactivation, target replacement and target protection mechanisms are likely associated with their mode of action. These mechanisms confer resistance via highly specific and autonomous processes, providing a rapid and efficient adaptation to the antimicrobial challenge that can be subsequently discarded to maximize reproductive yield [37].

ARGs conferring resistance via target alteration present an average DI of 22.0%, a significantly lower value than those discussed above (Figure 4). Interestingly, however, several target alteration ARGs present large DI values. The colistin-resistance *mcr* genes (ARO:3004465 | COG2194) have one of the highest DI values (99.5%), and the *erm* (ARO:3000560 | COG0030) and *cfr* genes (ARO:3000202 | COG0820), conferring cross-resistance to macrolides–lincosamides–class B streptogramins (MLS$_B$) using 23S ribosomal RNA methylation, also display relatively large DI values (62.1% and 71.4%, respectively) (Table S2). In contrast, many ARG homologs mapping to the target alteration resistance mechanism present very low or zero DI values. Such is the case of homologs of the *ndh* gene conferring resistance to isoniazid (ARO:3003460 | COG1252; DI: 19.0%) or the *gyrB* gene conferring resistance to fluoroquinolones (ARO:3000864 | COG0187; 0.0%).

The apparently contradictory results for ARO-COG groups mapping to the target alteration mechanism are partly explained by the fact that this CARD Resistance Mechanism category includes both spontaneous mutations to targeted enzymes as well as enzymes that induce chemical modifications in the antimicrobial target protein. The average DI for ARO-COG groups mapping to the target alteration resistance mechanism rises from 22.0% to 38.0% when spontaneous mutants are excluded (Figure 4). This indicates that mutated antimicrobial target variants are rarely disseminated, in contrast to enzymes that mediate the chemical modification of antimicrobial targets. The low dissemination of many ARGs conferring resistance via antimicrobial target alteration, and in particular of mutated variants, may be due to the necessary integration of these altered targets into the host biochemical pathways and physiological processes, which can restrict their ability to confer resistance in unrelated chromosomal backgrounds [38].

Efflux pumps providing resistance to antimicrobial compounds also present a low average dissemination index (30.5%), indicating that they are not generally disseminated (Figure 4). This was also suggested in a recent study, which showed that the chromosomal-to-plasmid ratio for this ARGs is low [20]. Despite this low ratio, our results reveal that the number of plasmid-borne antibiotic efflux ARG homologs vastly outnumbers the number of ARG homologs observed for any other resistance mechanism (Table S4). Only a few ARO-COG groups mapping to antibiotic efflux, like the one conferring resistance to tetracycline and phenicol, show evidence of widespread dissemination (Figure 3F). A few efflux pumps

display very high DI values, indicating that they are predominantly disseminated and not frequently mobilized from taxonomically close chromosomal backgrounds. Some variants of the *qac* genes (ARO:10002 | COG2076; DI: 98.8%) are among the most notable of these instances (Table S2). These efflux pumps are integron-borne and confer resistance to quaternary ammonium agents. Their high dissemination index value is driven by the most prevalent variant, *qacEΔ1*, which is typically found in the widespread class 1 integrons with *sul* and *dfr* genes [39].

Overall, our results support a general framework wherein the degree of dissemination of ARGs is proportional to their mutational accessibility and their availability in the bacterial pangenome. Intuitively, specialized enzymes capable of selectively inactivating antimicrobials or of protecting or replacing antimicrobial targets without impacting on their function cannot be readily evolved and must be acquired via HGT. This leads to high dissemination indices that reflect their widespread dissemination across multiple bacterial clades. In contrast, the ability to repurpose existing efflux pumps or to generate viable resistant target variants through mutation implies that chromosomal sources for resistance genes are likely available within any given bacterial clade. This leads to high rates of mobilization, but simultaneously lessens the selective pressure for the widespread dissemination of these resistance genes.

2.3. Antimicrobial Agent Classes Present Different Dissemination Patterns

The analysis of %GC content to monitor ARG dissemination also reveals different patterns of dissemination as a function of the antimicrobial agent they confer resistance to (Figure 5; Table S5). Among antimicrobial classes, the different patterns observed reflect the availability and degree of dissemination of different resistance mechanisms for any given drug class. Sulfonamides (Figure 6A), for instance, present an aggregate DI of 66.1%. ARG homologs conferring resistance to these chemotherapeutic agents are dominated by highly disseminated *sul* genes' mediating target replacement, which have been heavily disseminated through their association with class 1 integrons [15]. The highly disseminated *sul* genes are sometimes annotated in CARD as specific for sulfonamides, but more frequently as having multiple targets (sulfonamides and sulfones), yielding DI values of 94.7% and 88.8%, respectively. Sulfonamide resistance can also be conferred by nonspecific efflux pumps, but these ARG homologs present very low dissemination (3.8%), driving down the aggregate DI value for this drug class.

Other drug classes, such as the β-lactams (Figure 6B; aggregate DI: 43.9%) present profiles similar to the one observed for sulfonamides. For β-lactams, heavily disseminated β-lactamases mediating specific antibiotic inactivation present high DI values (75.6%), but the overall dissemination index for this antibiotic class is impacted by the abundance of nonspecific, infrequently disseminated efflux mechanisms also targeting β-lactams. The contribution of non-disseminated efflux pumps to the overall plasmid resistome dissemination profile is even larger for other antimicrobial classes, such as the fluoroquinolones (Figure 6C; aggregate DI: 29.7%) or the macrolides (Figure 6D; aggregate DI: 33.5%), where multiple resistance mechanisms are detected. In these cases, ARGs linked to specific resistance mechanisms (antibiotic inactivation and target protection, respectively) display very high dissemination indices when compared to nonspecific mechanisms. This is particularly true for nonspecific antibiotic efflux mechanisms conferring cross-resistance to these antimicrobials. Homologs for these efflux pumps are detected abundantly in the plasmidome, but present low levels of dissemination in agreement with recent reports [20].

Several antimicrobial classes display a remarkable absence of homologs mediating resistance by either efflux pumps or specific mechanisms. The streptogramins, for instance, show substantial dissemination of nonspecific ARG homologs mapping to target alteration (DI: 63.4%) and protection mechanisms (DI: 45.8%) (Figure 6E). These ARG homologs mediate, respectively, targeted methylation of ribosomal subunits [40,41] and ribosomal protection [42], conferring resistance to multiple antimicrobial classes including lincosamides, oxazolidinones, phenicol and streptogramins [43]. In contrast, a substantial

proportion of detected ARG homologs conferring resistance to tetracyclines map to highly disseminated specific efflux pumps and their transcriptional regulators (DI: 61.9%), as well as nonspecific efflux pumps with low dissemination index (DI: 19.5%). For tetracyclines, specific target alteration (DI: 78.1%) and protection (DI: 50.1%) mechanisms are highly disseminated but appear to play a secondary role in terms of the amount of ARG homologs detected in the plasmidome (Figure 6F).
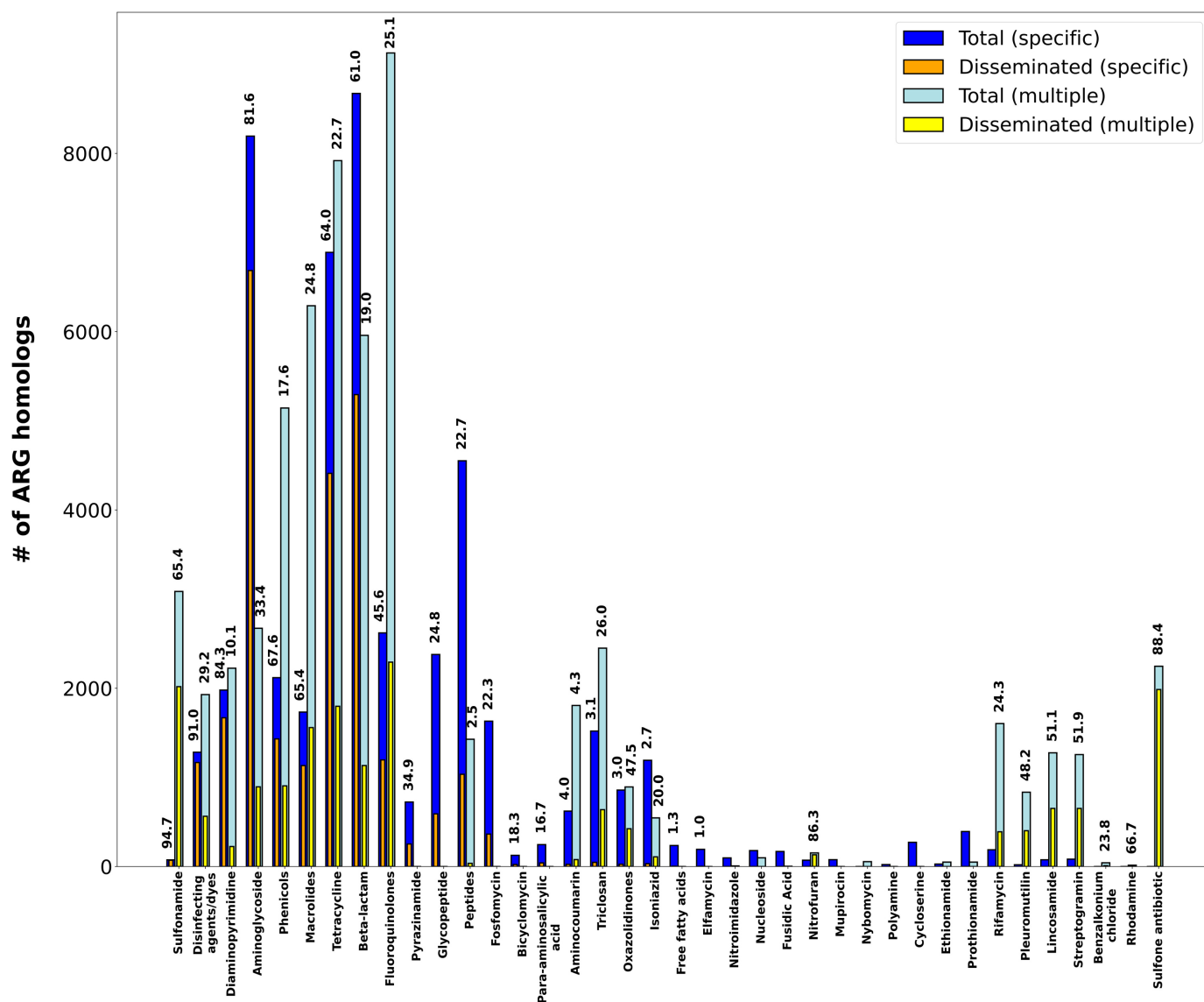
Figure 5. Bar plot showing the total vs. disseminated number of ARG homologs aggregated by CARD Drug Class. "Specific" denotes DI values computed solely on ARG homologs matching a single CARD Drug Class. "Multiple" denotes DI values computed solely on ARG homologs matching several CARD Drug Classes. Dissemination index values for each group are superimposed when larger than 0%.
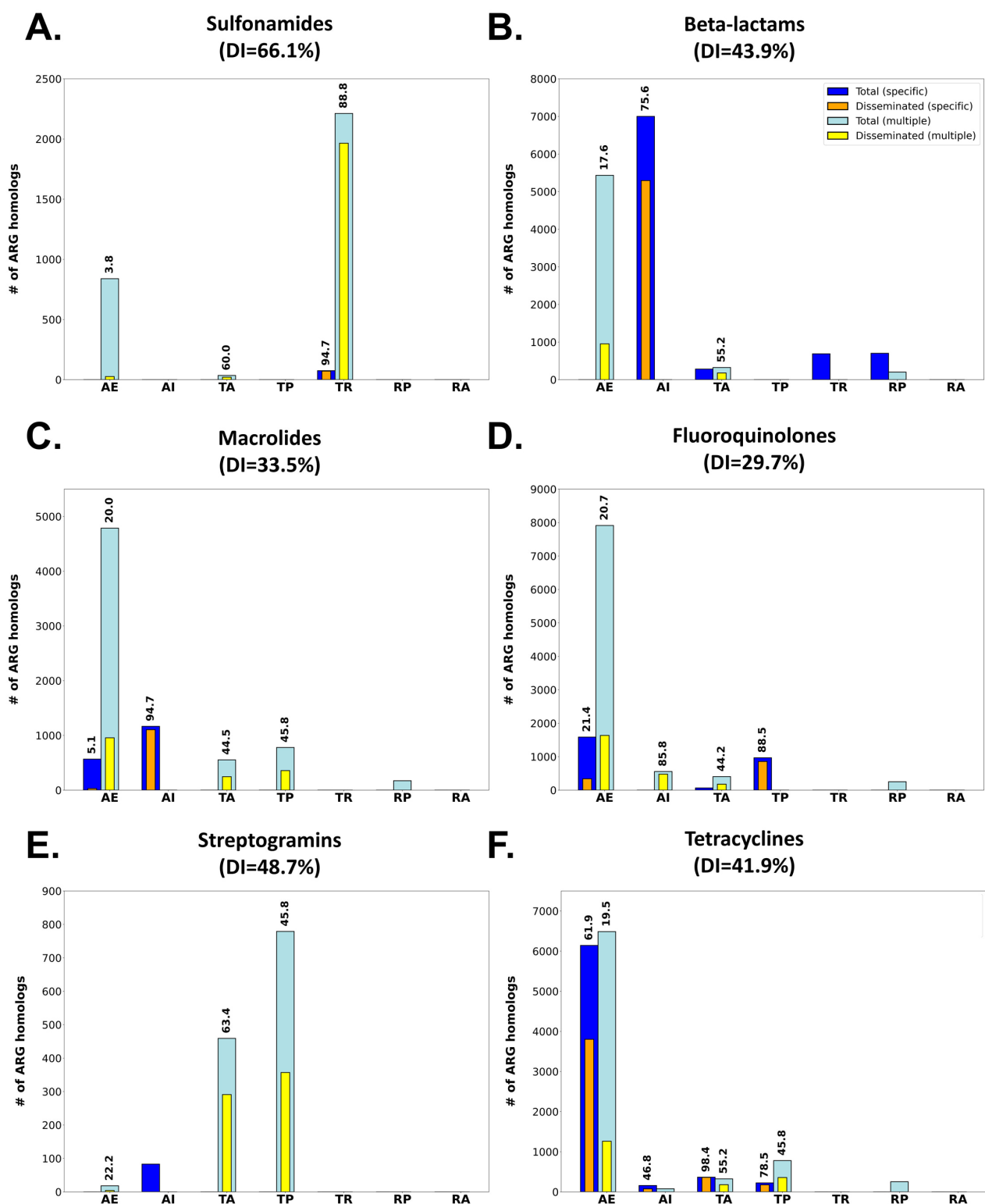
Figure 6. Bar plots showing the total vs. disseminated number of ARG homologs matching single (Specific) or several (Multiple) CARD Drug Classes and conferring resistance to sulfonamides (**A**), beta-lactams (**B**), macrolides (**C**) fluoroquinolones (**D**), streptogramins (**E**) or tetracyclines (**F**). ARG homologs are also split according to their CARD Resistance Mechanism: target alteration (TA), antibiotic efflux (AE), antibiotic inactivation (AI), resistance by absence (RA), target replacement (TR), reduced permeability (RP) or target protection (TP). Dissemination index values for each group are superimposed if larger than 0%.

Interestingly, several antimicrobial classes display low or negligible dissemination profiles (Figure 5). For some of these antimicrobials, the most likely explanation for their low dissemination index is their organism-specific usage, typically restricted to the clinical setting. This is the case of isoniazid, ethionamide and prothionamide, which are almost exclusively used to treat *Mycobacterium tuberculosis* infections [44]. For other antimicrobial classes, the observed low DI values are likely associated with resistance being mediated almost exclusively by nonspecific efflux pumps (e.g., antibacterial fatty acids, aminocoumarin or bicyclomycin) or by the mutational accessibility of variants conferring resistance (e.g., cycloserine or polyamine antibiotics) [45].

The different patterns of dissemination observed for antimicrobial classes may reflect different evolutionary processes. Exposure to high doses of an antimicrobial agent can promote incorporation via HGT of specific resistance genes that provide an immediate evolutionary advantage [46]. These ARGs often act through largely autonomous mechanisms, such as antibiotic inactivation, and target specific drug classes. In contrast, ARGs providing cross-resistance often operate through nonspecific mechanisms such as efflux pumps. These ARGs are frequently encoded by the bacterial chromosome and operate in coordination with other physiological processes. This facilitates their mobilization, even in the absence of selective pressure, but may limit their widespread dissemination to unrelated chromosomal backgrounds [20,29,30].

2.4. Conjugative Plasmids Mediate Dissemination of Antimicrobial Resistance Genes

To study the contribution of different plasmid classes on plasmid-borne antimicrobial resistance dissemination, we used the MOB_suite to predict plasmid mobility types. We then constructed a bipartite network using Gephi with the predicted ARGs and PLSDB plasmids as nodes. The presence of ARG homologs in plasmids was depicted as network connections and node coloring accounted for ARO-COG dissemination indices and predicted plasmid mobility class, respectively. In accordance with previous results, our analysis shows that conjugation is the main dissemination driver of most ARGs [1], providing independent validation to the proposed DI as an estimator of ARG dissemination.

The association pattern of ARO-COG groups and plasmids observed in the network (Figure 7, Table S6) reveals that non-disseminated gene families (DI = 0) are more common in non-mobilizable plasmids (44.3%) than in mobilizable (33.7%) and conjugative plasmids (22.0%). In contrast, ARO-COG groups with very large dissemination indices (DI > 75%) are mainly associated with conjugative plasmids (69.4%). The rest of ARO-COG groups showing evidence of dissemination (0% < DI < 75%) are also more likely to be found in non-mobilizable plasmids (38.4%) than in conjugative (34.0%) and mobilizable plasmids (27.6%). The latter case is of interest because ARO-COG groups with evidence of dissemination appear to be more common in non-mobilizable than in mobilizable plasmids. This suggests that conjugation mediated by helper *tra* genes might not play a fundamental role in the widespread dissemination of ARG in natural communities.

Differences in chromosomal %GC content are a well-established proxy for phylogenetic divergence, which is known to limit the efficiency of conjugation [21,47]. Our results illustrate that the %GC content of many ARG homologs is effectively independent of the %GC of the host on which the plasmid harboring them has been isolated, providing the means to identify highly disseminated ARGs (Figure 1). The observed %GC independence profile of highly disseminated ARGs can be explained by two complementary processes. On the one hand, selective pressure driven by exposure to antimicrobials could overcome phylogenetic barriers to plasmid conjugation, facilitating the proliferation of the ARG-harboring plasmid on a nonnative host. On the other hand, following an initial conjugation event from a nonnative plasmid harboring ARGs in integrons or other MGEs, the same selective pressure might promote the transmission of the ARGs to a plasmid compatible with the host, facilitating its subsequent spread. Analysis of plasmid %GC content relative to chromosomal host %GC content reveals that plasmid and chromosome %GC content are strongly correlated (Pearson $r = 0.893$) when all ARG homologs are considered. This

robust correlation is maintained (Pearson *r* = 0.792) when only plasmids harboring ARG homologs in dissemination bands are considered (Table S7). This suggests that widespread dissemination of ARGs is predominantly mediated by transmission of ARG homologs to plasmids compatible with their hosts and highlights the importance of integrons and other plasmid-borne MGEs in the spread of antimicrobial resistance [48,49].
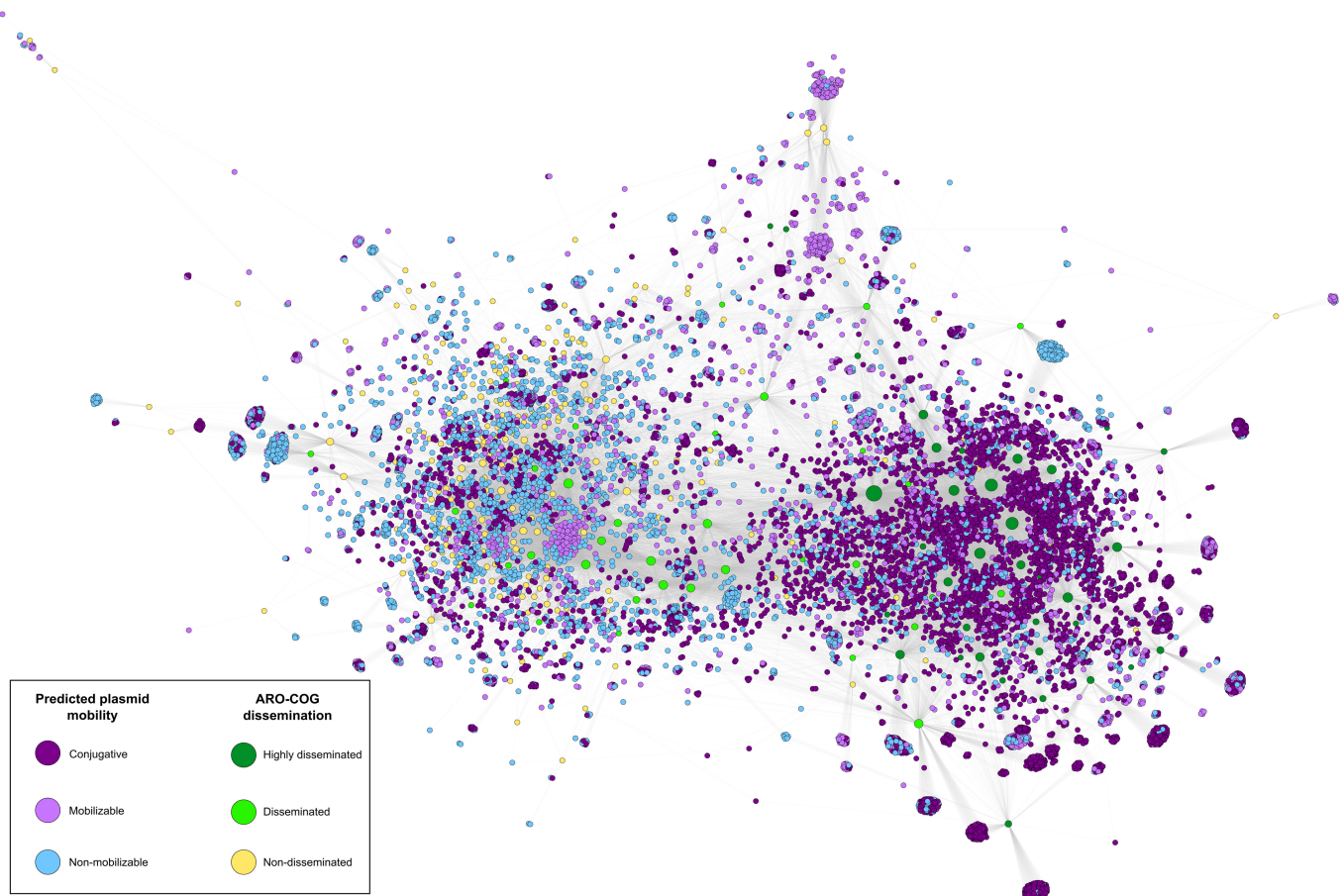
Figure 7. Bipartite network illustrating the correlation between predicted plasmid mobility and ARG dissemination. Plasmid mobility was predicted using the MOB_suite. "Highly disseminated" denotes ARO-COGs with DI > 75%, "Disseminated" stands for ARO-COGs with 0% > DI > 75% and "Non-disseminated" for ARO-COGs with DI = 0%. An interactive version of the network is available at https://miquelsanchezosuna.github.io/arg_gc_network/ (accessed on 14 December 2022).

3. Materials and Methods

3.1. Antimicrobial Resistance Gene Prediction, Classification and Plasmid Analysis

Putative antimicrobial resistance genes (ARGs) were predicted on plasmid sequences from the PLSDB plasmid database [50] using all available protein sequences from the Comprehensive Antibiotic Resistance Database (CARD) collection as queries [51]. The PLSDB complete plasmid assemblies and CARD protein sequences were downloaded from the NCBI Nucleotide and Protein databases in FASTA protein format. Homologs of CARD protein sequences in the PLSDB were detected with BLASTP using an e-value $<1 \times 10^{-20}$ and a query coverage >75% as limiting thresholds [52]. Each putative ARG identified in PLSDB was linked to a CARD antimicrobial resistance (AMR) Gene Family, Drug Class and Resistance Mechanism using the accessory information from the CARD query with the lowest e-value in the BLASTP analysis. CARD annotation on genes conferring resistance through mutation was compiled and assigned to corresponding homologs. Plasmid type and host information for each putative ARG were obtained from the PLSDB. Predicted ARGs were mapped to Clusters of Orthologous Groups (COG) [53] using HMMER (hmm-

scan) [54] with a limiting e-value of 1×10^{-5} and otherwise default parameters. Putative ARGs not mapping to any COG were removed from the analysis. Predicted ARGs were then grouped based on their CARD AMR Gene Family and their mapping COG. These groups were assigned compound identifiers using the Antibiotic Resistance Ontology (ARO) accession for the AMR Gene Family and the COG accession and are henceforth referred to by their ARO-COG accessions.

3.2. Antimicrobial Resistance Gene Dissemination Analysis

ARG dissemination analysis is based on the study of correlation between the %GC of ARG homologs identified in plasmids and the chromosomal %GC content of the plasmid isolation hosts. For widely disseminated genes, it is expected that ARG %GC will be independent of host chromosomal %GC, displaying dissemination bands that can be detected via computational analysis (Figure 1). For each ARO-COG group with more than 10 ARG homologs mapping to it, protein and coding DNA sequences for the ARG homologs identified in PLSDB were downloaded from the NCBI GenBank database for %GC content analysis. %GC content for ARG homologs was computed using BioPython [55]. The host %GC content for each predicted ARG was determined by matching the reported plasmid host strain, species or genus against a RefSeq representative complete genome from the NCBI Assembly database. The correlation between the %GC content of the ARGs and that of their host genome was depicted as a scatter plot using a custom Python script. Redundant and nonredundant sets of antibiotic resistance coding sequences were generated by clustering all ARG homologs identified in PLSDB using Usearch with a 90% similarity threshold and otherwise default parameters [56].

For the systematic analysis of the dissemination of antimicrobial resistance genes, we computed the Dissemination Index (DI) for each ARO-COG group. The DI is defined as the fraction of redundant ARG homologs in the ARO-COG group that map to dissemination bands, relative to the total number of ARG homologs in the ARO-COG group. Dissemination bands were calculated using a custom algorithm inspired by the gapless extension phase of BLAST [52] (Figure S3). Briefly, for a given scatter plot contrasting host (*x*-axis) and redundant ARG (*y*-axis) %GC content, the algorithm identifies a dissemination band as any segment of ARG homologs spanning more than a detection span *DS* (15% by default) in the *x*-axis (i.e., sets of ARG homologs distributed among hosts with a %GC divergence greater than 15%). Scanning proceeds by 1% increments on the *y*-axis until a dissemination band meeting the condition above is detected. Upon detecting a dissemination band, a bidirectional extension phase ensues. The band *x*-axis span is saved as the current span, and subsequent 1% increments on the *y*-axis are incorporated to the band if their span is at the most *d* (5 by default) percentual points smaller than the current span. Otherwise, extension of the dissemination band in that direction is terminated. To limit the occurrence of false positives, the average number of ARG homologs in any given 1% increment containing data is computed, and only dissemination bands exceeding this expected value are considered for analysis. To compute DI values for CARD metadata-based groupings, such as Resistance Mechanism or Drug Class, the CARD metadata and dissemination band status associated with each predicted ARG homolog is tabulated independently to generate a DI value for the entire metadata grouping.

The scripts used for data collection and analysis are available at the GitHub ErillLab repository (https://github.com/ErillLab/, accessed on 14 December 2022).

4. Conclusions

The ability to assess and monitor ARG dissemination is critical to inform antibiotic policy and develop predictive models of resistance dissemination. This work shows that analysis of the correlation between ARG and plasmid-host %GC content provides the means to efficiently leverage publicly available plasmid sequence data to quantitatively assess ARG dissemination with a visually interpretable and easily scalable dissemination index (DI). Furthermore, our analysis pipeline can be generalized to study other genetic elements

relevant to antimicrobial resistance and microbial adaptation, such as siderophores or genes determining plasmid incompatibility. Results from the %GC content analysis show that ARG dissemination is driven predominantly by resistance mechanism. This finding confirms previous results and provides insights into the selective pressures underpinning ARG dissemination. Our analysis also reveals that most broadly disseminated genes originated with a single mobilization event, and that their dissemination has been mediated primarily by conjugative plasmids. Our data on ARG dissemination reveals substantial differences among antimicrobial classes, reflecting in some cases antibiotic policy. Overall, our findings illustrate how a fast and intuitive analysis can be used to reveal fundamental aspects of antimicrobial resistance dissemination and provide the basis for subsequent experimental work to gauge the impact of disseminated ARGs on antibiotic resistance in natural and clinical isolates.

References

1. Khan, A.; Miller, W.R.; Arias, C.A. Mechanisms of Antimicrobial Resistance among Hospital-Associated Pathogens. *Expert Rev. Anti-Infect. Ther.* **2018**, *16*, 269–287. [CrossRef] [PubMed]
2. World Health Organization (Ed.) *Antimicrobial Resistance: Global Report on Surveillance*; World Health Organization: Geneva, Switzerland, 2015.
3. Berglund, B. Environmental Dissemination of Antibiotic Resistance Genes and Correlation to Anthropogenic Contamination with Antibiotics. *Infect. Ecol. Epidemiol.* **2015**, *5*, 28564. [CrossRef] [PubMed]
4. Saima, S.; Fiaz, M.; Zafar, R.; Ahmed, I.; Arshad, M. Chapter 6—Dissemination of Antibiotic Resistance in the Environment. In *Antibiotics and Antimicrobial Resistance Genes in the Environment*; Hashmi, M.Z., Ed.; Advances in Environmental Pollution Research series; Elsevier: Amsterdam, The Netherlands, 2020; Volume 1, pp. 99–116. ISBN 978-0-12-818882-8.
5. Baym, M.; Lieberman, T.D.; Kelsic, E.D.; Chait, R.; Gross, R.; Yelin, I.; Kishony, R. Spatiotemporal Microbial Evolution on Antibiotic Landscapes. *Science* **2016**, *353*, 1147–1151. [CrossRef] [PubMed]
6. Suarez, S.A.; Martiny, A.C. Gene Amplification Uncovers Large Previously Unrecognized Cryptic Antibiotic Resistance Potential in E. Coli. *Microbiol. Spectr.* **2021**, *9*, e0028921. [CrossRef]
7. van Hoek, A.; Mevius, D.; Guerra, B.; Mullany, P.; Roberts, A.; Aarts, H. Acquired Antibiotic Resistance Genes: An Overview. *Front. Microbiol.* **2011**, *2*, 203. [CrossRef] [PubMed]
8. Aminov, R.I.; Mackie, R.I. Evolution and Ecology of Antibiotic Resistance Genes. *FEMS Microbiol. Lett.* **2007**, *271*, 147–161. [CrossRef]
9. Davies, J.; Davies, D. Origins and Evolution of Antibiotic Resistance. *Microbiol. Mol. Biol. Rev.* **2010**, *74*, 417–433. [CrossRef]
10. Peterson, E.; Kaur, P. Antibiotic Resistance Mechanisms in Bacteria: Relationships Between Resistance Determinants of Antibiotic Producers, Environmental Bacteria, and Clinical Pathogens. *Front. Microbiol.* **2018**, *9*, 2928. [CrossRef]
11. Domingues, S.; Nielsen, K.M. Membrane Vesicles and Horizontal Gene Transfer in Prokaryotes. *Curr. Opin. Microbiol.* **2017**, *38*, 16–21. [CrossRef]
12. Soler, N.; Forterre, P. Vesiduction: The Fourth Way of HGT. *Environ. Microbiol.* **2020**, *22*, 2457–2460. [CrossRef]
13. Hernández-Beltrán, J.C.R.; San Millán, A.; Fuentes-Hernández, A.; Peña-Miller, R. Mathematical Models of Plasmid Population Dynamics. *Front. Microbiol.* **2021**, *12*, 606396. [CrossRef]
14. Sánchez-Osuna, M.; Cortés, P.; Llagostera, M.; Barbé, J.; Erill, I. Exploration into the Origins and Mobilization of Di-Hydrofolate Reductase Genes and the Emergence of Clinical Resistance to Trimethoprim. *Microb. Genom.* **2020**, *6*, mgen000440. [CrossRef] [PubMed]
15. Sánchez-Osuna, M.; Cortés, P.; Barbé, J.; Erill, I. Origin of the Mobile Di-Hydro-Pteroate Synthase Gene Determining Sulfonamide Resistance in Clinical Isolates. *Front. Microbiol.* **2019**, *9*, 3332. [CrossRef] [PubMed]
16. Jiang, X.; Ellabaan, M.M.H.; Charusanti, P.; Munck, C.; Blin, K.; Tong, Y.; Weber, T.; Sommer, M.O.A.; Lee, S.Y. Dissemination of Antibiotic Resistance Genes from Antibiotic Producers to Pathogens. *Nat. Commun.* **2017**, *8*, 15784. [CrossRef] [PubMed]
17. Møller, T.S.B.; Liu, G.; Boysen, A.; Thomsen, L.E.; Lüthje, F.L.; Mortensen, S.; Møller-Jensen, J.; Olsen, J.E. Treatment with Cefotaxime Affects Expression of Conjugation Associated Proteins and Conjugation Transfer Frequency of an IncI1 Plasmid in Escherichia Coli. *Front. Microbiol.* **2017**, *8*, 2365. [CrossRef]
18. Zeng, X.; Lin, J. Factors Influencing Horizontal Gene Transfer in the Intestine. *Anim. Health Res. Rev.* **2017**, *18*, 153–159. [CrossRef]
19. Ellabaan, M.M.H.; Munck, C.; Porse, A.; Imamovic, L.; Sommer, M.O.A. Forecasting the Dissemination of Antibiotic Resistance Genes across Bacterial Genomes. *Nat. Commun.* **2021**, *12*, 2435. [CrossRef]
20. Nielsen, T.K.; Browne, P.D.; Hansen, L.H. Antibiotic Resistance Genes Are Differentially Mobilized According to Resistance Mechanism. *GigaScience* **2022**, *11*, giac072. [CrossRef]
21. Musto, H. How Many Factors Influence Genomic GC Content Among Prokaryotes? *J. Mol. Evol.* **2022**, *91*, 6–9. [CrossRef]
22. Lawrence, J.G.; Ochman, H. Amelioration of Bacterial Genomes: Rates of Change and Exchange. *J. Mol. Evol.* **1997**, *44*, 383–397. [CrossRef]
23. Albornoz, E.; Tijet, N.; De Belder, D.; Gomez, S.; Martino, F.; Corso, A.; Melano, R.G.; Petroni, A. QnrE1, a Member of a New Family of Plasmid-Located Quinolone Resistance Genes, Originated from the Chromosome of Enterobacter Species. *Antimicrob. Agents Chemother.* **2017**, *61*, e02555-16. [CrossRef] [PubMed]
24. Cantón, R. Antibiotic Resistance Genes from the Environment: A Perspective through Newly Identified Antibiotic Resistance Mechanisms in the Clinical Setting. *Clin. Microbiol. Infect.* **2009**, *15* (Suppl. S1), 20–25. [CrossRef] [PubMed]
25. Cantón, R.; González-Alba, J.M.; Galán, J.C. CTX-M Enzymes: Origin and Diffusion. *Front. Microbiol.* **2012**, *3*, 110. [CrossRef]
26. Walther-Rasmussen, J.; Høiby, N. OXA-Type Carbapenemases. *J. Antimicrob. Chemother.* **2006**, *57*, 373–383. [CrossRef] [PubMed]
27. Ramirez, M.S.; Tolmasky, M.E. Aminoglycoside Modifying Enzymes. *Drug Resist. Updates* **2010**, *13*, 151–171. [CrossRef] [PubMed]
28. Salipante, S.J.; Hall, B.G. Determining the Limits of the Evolutionary Potential of an Antibiotic Resistance Gene. *Mol. Biol. Evol.* **2003**, *20*, 653–659. [CrossRef]
29. Depardieu, F.; Podglajen, I.; Leclercq, R.; Collatz, E.; Courvalin, P. Modes and Modulations of Antibiotic Resistance Gene Expression. *Clin. Microbiol. Rev.* **2007**, *20*, 79–114. [CrossRef]

30. Martínez, J.L.; Coque, T.M.; Baquero, F. What Is a Resistance Gene? Ranking Risk in Resistomes. *Nat. Rev. Microbiol.* **2015**, *13*, 116–123. [CrossRef]

31. Luo, X.; Yin, Z.; Zeng, L.; Hu, L.; Jiang, X.; Jing, Y.; Chen, F.; Wang, D.; Song, Y.; Yang, H.; et al. Chromosomal Integration of Huge and Complex Bla NDM-Carrying Genetic Elements in Enterobacteriaceae. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 690799. [CrossRef]

32. Shropshire, W.C.; Aitken, S.L.; Pifer, R.; Kim, J.; Bhatti, M.M.; Li, X.; Kalia, A.; Galloway-Peña, J.; Sahasrabhojane, P.; Arias, C.A.; et al. IS26-Mediated Amplification of BlaOXA-1 and BlaCTX-M-15 with Concurrent Outer Membrane Porin Disruption Associated with de Novo Carbapenem Resistance in a Recurrent Bacteraemia Cohort. *J. Antimicrob. Chemother.* **2021**, *76*, 385–395. [CrossRef]

33. Aminov, R.I. A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. *Front. Microbiol.* **2010**, *1*, 134. [CrossRef]

34. Shin, H.W.; Lim, J.; Kim, S.; Kim, J.; Kwon, G.C.; Koo, S.H. Characterization of Trimethoprim-Sulfamethoxazole Resistance Genes and Their Relatedness to Class 1 Integron and Insertion Sequence Common Region in Gram-Negative Bacilli. *J. Microbiol. Biotechnol.* **2015**, *25*, 137–142. [CrossRef] [PubMed]

35. Lee, A.S.; de Lencastre, H.; Garau, J.; Kluytmans, J.; Malhotra-Kumar, S.; Peschel, A.; Harbarth, S. Methicillin-Resistant Staphylococcus Aureus. *Nat. Rev. Dis. Prim.* **2018**, *4*, 18033. [CrossRef] [PubMed]

36. Liu, J.; Chen, D.; Peters, B.M.; Li, L.; Li, B.; Xu, Z.; Shirliff, M.E. Staphylococcal Chromosomal Cassettes Mec (SCCmec): A Mobile Genetic Element in Methicillin-Resistant Staphylococcus Aureus. *Microb. Pathog.* **2016**, *101*, 56–67. [CrossRef] [PubMed]

37. Christaki, E.; Marcou, M.; Tofarides, A. Antimicrobial Resistance in Bacteria: Mechanisms, Evolution, and Persistence. *J. Mol. Evol.* **2020**, *88*, 26–40. [CrossRef] [PubMed]

38. MacLean, R.C.; Hall, A.R.; Perron, G.G.; Buckling, A. The Population Genetics of Antibiotic Resistance: Integrating Molecular Mechanisms and Treatment Contexts. *Nat. Rev. Genet.* **2010**, *11*, 405–414. [CrossRef] [PubMed]

39. Slipski, C.J.; Jamieson-Datzkiw, T.R.; Zhanel, G.G.; Bay, D.C. Characterization of Proteobacterial Plasmid Integron-Encoded Qac Efflux Pump Sequence Diversity and Quaternary Ammonium Compound Antiseptic Selection in Escherichia Coli Grown Planktonically and as Biofilms. *Antimicrob. Agents Chemother.* **2021**, *65*, e0106921. [CrossRef]

40. Nannini, E.; Murray, B.E.; Arias, C.A. Resistance or Decreased Susceptibility to Glycopeptides, Daptomycin, and Linezolid in Methicillin-Resistant Staphylococcus Aureus. *Curr. Opin. Pharm.* **2010**, *10*, 516–521. [CrossRef]

41. Park, A.K.; Kim, H.; Jin, H.J. Phylogenetic Analysis of RRNA Methyltransferases, Erm and KsgA, as Related to Antibiotic Resistance. *FEMS Microbiol. Lett.* **2010**, *309*, 151–162. [CrossRef]

42. Sharkey, L.K.R.; O'Neill, A.J. Antibiotic Resistance ABC-F Proteins: Bringing Target Protection into the Limelight. *ACS Infect. Dis.* **2018**, *4*, 239–246. [CrossRef]

43. Long, K.S.; Poehlsgaard, J.; Kehrenberg, C.; Schwarz, S.; Vester, B. The Cfr RRNA Methyltransferase Confers Resistance to Phenicols, Lincosamides, Oxazolidinones, Pleuromutilins, and Streptogramin A Antibiotics. *Antimicrob. Agents Chemother.* **2006**, *50*, 2500–2505. [CrossRef]

44. Ushtanit, A.; Kulagina, E.; Mikhailova, Y.; Makarova, M.; Safonova, S.; Zimenkov, D. Molecular Determinants of Ethionamide Resistance in Clinical Isolates of Mycobacterium Tuberculosis. *Antibiotics* **2022**, *11*, 133. [CrossRef] [PubMed]

45. Woodford, N.; Ellington, M.J. The Emergence of Antibiotic Resistance by Mutation. *Clin. Microbiol. Infect.* **2007**, *13*, 5–18. [CrossRef]

46. Baquero, F.; Alvarez-Ortega, C.; Martinez, J.L. Ecology and Evolution of Antibiotic Resistance. *Environ. Microbiol. Rep.* **2009**, *1*, 469–476. [CrossRef] [PubMed]

47. Alderliesten, J.B.; Duxbury, S.J.N.; Zwart, M.P.; de Visser, J.A.G.M.; Stegeman, A.; Fischer, E.A.J. Effect of Donor-Recipient Relatedness on the Plasmid Conjugation Frequency: A Meta-Analysis. *BMC Microbiol.* **2020**, *20*, 135. [CrossRef] [PubMed]

48. Stalder, T.; Barraud, O.; Casellas, M.; Dagot, C.; Ploy, M.-C. Integron Involvement in Environmental Spread of Antibiotic Resistance. *Front. Microbiol.* **2012**, *3*, 119. [CrossRef] [PubMed]

49. Bennett, P.M. Plasmid Encoded Antibiotic Resistance: Acquisition and Transfer of Antibiotic Resistance Genes in Bacteria. *Br. J. Pharm.* **2008**, *153* (Suppl. S1), S347–S357. [CrossRef]

50. Galata, V.; Fehlmann, T.; Backes, C.; Keller, A. PLSDB: A Resource of Complete Bacterial Plasmids. *Nucleic Acids Res.* **2019**, *47*, D195–D202. [CrossRef]

51. Alcock, B.P.; Raphenya, A.R.; Lau, T.T.Y.; Tsang, K.K.; Bouchard, M.; Edalatmand, A.; Huynh, W.; Nguyen, A.-L.V.; Cheng, A.A.; Liu, S.; et al. CARD 2020: Antibiotic Resistome Surveillance with the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **2020**, *48*, D517–D525. [CrossRef]

52. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]

53. Galperin, M.Y.; Wolf, Y.I.; Makarova, K.S.; Vera Alvarez, R.; Landsman, D.; Koonin, E.V. COG Database Update: Focus on Microbial Diversity, Model Organisms, and Widespread Pathogens. *Nucleic Acids Res.* **2021**, *49*, D274–D281. [CrossRef] [PubMed]

54. Eddy, S.R. Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **2011**, *7*, e1002195. [CrossRef] [PubMed]

55. Cock, P.J.A.; Antao, T.; Chang, J.T.; Chapman, B.A.; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; et al. Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423. [CrossRef]
56. Edgar, R.C. Search and Clustering Orders of Magnitude Faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [CrossRef] [PubMed]