Conjugative Plasmids Rarely Confer Antibiotic Resistance or Virulence

Célia P.F. Domingues[1,2], João S. Rebelo[2], Francisco Dionisio[2,**], Teresa Nogueira[1,2,*]

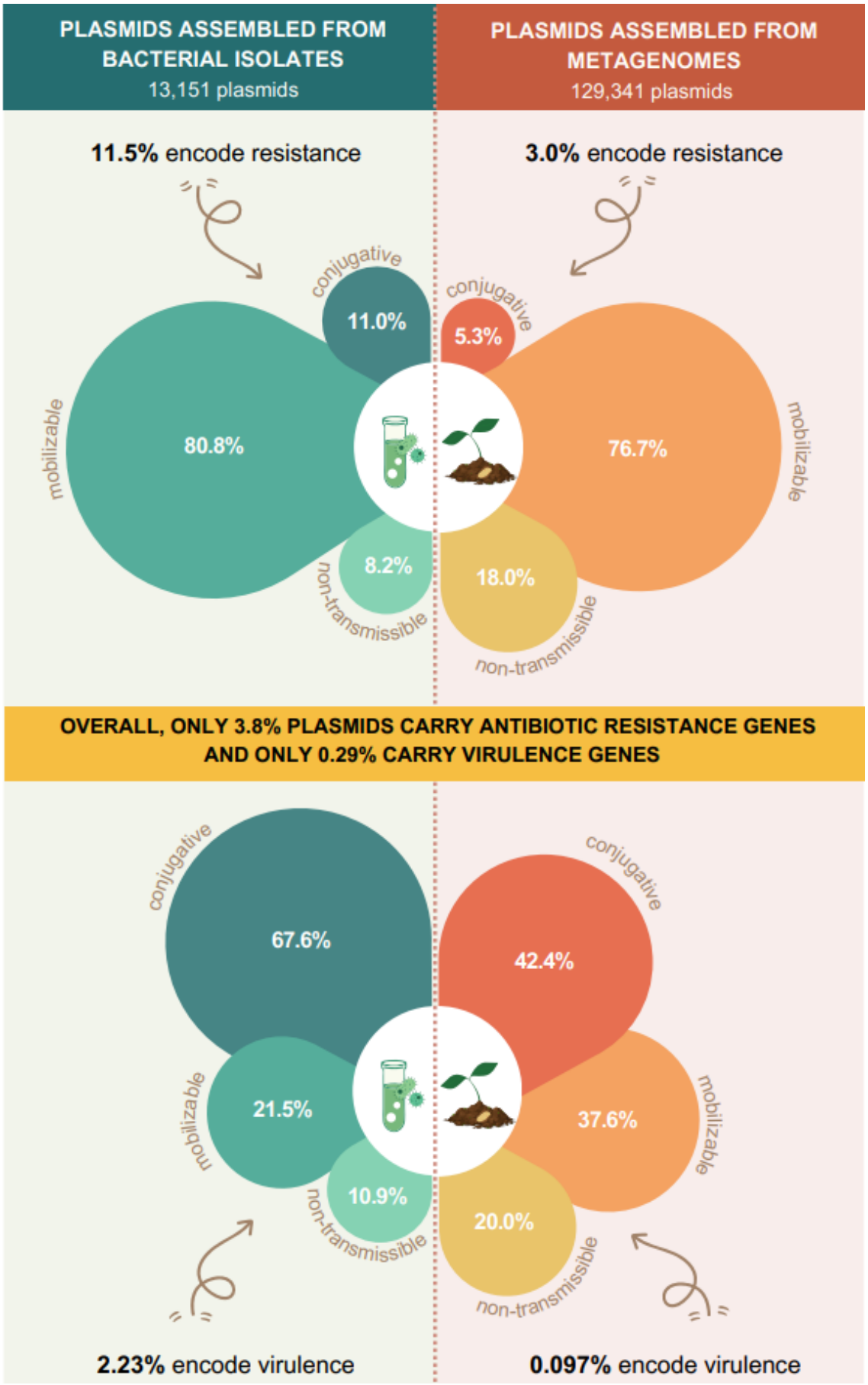
[1] INIAV—Instituto Nacional de Investigação Agrária e Veterinária, 4485-655 Vairão, Portugal

[2] cE3c—Centre for Ecology, Evolution and Environmental Changes & CHANGE, Global Change and Sustainability Institute, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal

Correspondence: * teresa.nogueira@iniav.pt ; ** francisco.dionisio@gmail.com

Summary

Antimicrobial-resistant bacteria cause over a million deaths annually. This study investigates the role of plasmid mobility types in disseminating antimicrobial resistance genes (ARGs) and virulence genes (VGs). By analyzing 142,492 fully sequenced plasmids from bacterial isolates and metagenomes, we offer a comprehensive representation of plasmid distribution in microbial communities. Our findings reveal that conjugative plasmids (pCONJ) are less common than previously estimated, comprising only 4.14% of the dataset. Mobilizable plasmids (pMOB) are the most prevalent (66.82%), followed by non-transmissible plasmids (pNT) at 29.04%. Although plasmids are recognized as key drivers of ARG dissemination, only 3.8% of plasmids encode ARGs. Among pCONJ plasmids, 6.9% harbor ARGs, and just 0.18% carry VGs. These findings contrast with earlier estimates, which overrepresented clinically associated plasmids. Interestingly, ARGs are more frequently associated with pMOB plasmids, which exhibit higher ARG density. In contrast, VGs are primarily found in pCONJ plasmids. This study highlights the significance of pMOB plasmids in ARG dissemination, emphasizing their role in resistance spread. Our results challenge prevailing assumptions about plasmid-mediated gene transfer and underscore the need to prioritize pMOB plasmids in future epidemiological studies. These insights contribute to a deeper understanding of plasmid dynamics and their implications for public health.


PLASMIDS ASSEMBLED FROM BACTERIAL ISOLATES
13,151 plasmids
PLASMIDS ASSEMBLED FROM METAGENOMES
129,341 plasmids
11.5% encode resistance
3.0% encode resistance
conjugative
11.0%
mobilizable
80.8%
non-transmissible
8.2%
conjugative
5.3%
mobilizable
76.7%
non-transmissible
18.0%
OVERALL, ONLY 3.8% PLASMIDS CARRY ANTIBIOTIC RESISTANCE GENES AND ONLY 0.29% CARRY VIRULENCE GENES
conjugative
67.6%
mobilizable
21.5%
non-transmissible
10.9%
conjugative
42.4%
mobilizable
37.6%
non-transmissible
20.0%
2.23% encode virulence
0.097% encode virulence

Keywords

Plasmids, antibiotic resistance, virulence, mobilization mode, conjugation, pCONJ, pMOB, pNT, plasmids assembled from isolates, plasmids assembled from metagenomes

Introduction

Plasmids are extrachromosomal self-replicating genetic elements ubiquitous in bacteria. For example, a recent study has shown that a specific plasmid, namely the pBI143 plasmid, is one of the most prevalent and abundant genetic elements found in the gut microbiomes of industrialized human populations worldwide [1]. Moreover, studies involving a few thousand strains have shown that there are, on average, two to three different plasmids per strain of *Escherichia coli*, *Klebsiella* and *Salmonella*, and almost one plasmid per strain of *Staphylococcus aureus* [2,3].

Plasmids have been extensively studied because they may harbor genes that facilitate the degradation of environmental pollutants [4], biofilm formation, and confer bacterial virulence, drug and heavy metal resistance [5–7].

Conjugative plasmids (pCONJ) can transfer themselves by replication between bacterial cells, often between different strains or species. Therefore, several works have emphasized their pivotal role in bacterial evolution by acquiring clinically significant traits [8–12]. pCONJ are self-transmissible because they encode a relaxase (MOB), an endonuclease, that cuts one of the DNA strands at the nic site of the respective *oriT*, and the complete machinery for mating pair formation (MPF), which connects donor and recipient cells while serving as a channel for the plasmid transfer [13,14]. Mobilizable plasmids (pMOB) lack part or all MPF genes, so, for transfer, they must hijack these genes' products from other coresident conjugative plasmids. Plasmids lacking relaxases but containing the *oriT* are also mobilizable and belong to the pMOB group [3]. Finally, non-transmissible plasmids (pNT) have no identifiable *oriT* locus or any of the genes mentioned above, suggesting they cannot transfer horizontally. However, these plasmids may yet contain any other unidentified *oriT* [3,15].

Previous studies estimated that about one-quarter of the plasmids are pCONJ: 23% in a database comprising 11,386 plasmids from *Escherichia coli* and *Staphylococcus aureus*; and 25.2% in a database comprising 14,029 plasmids [9,16]. However, most previously sequenced plasmids are of clinical interest because they were isolated from pathogens or because they confer drug resistance. Therefore, there is a gap in knowledge concerning plasmid biology. In recent years, new and cheaper sequencing techniques have allowed researchers to sequence and study plasmids from different contexts. We thus have an opportunity to fill the gap.

Results

This study aims to elucidate how different plasmid types contribute to disseminating antimicrobial resistance genes (ARGs) and virulence genes (VGs) in natural environments and how that relates to plasmid diversity. We have thus analyzed a large representative sample of plasmids from diverse sources, including plasmids assembled from metagenomes and a vast collection of bacterial isolates. The plasmid sequences were retrieved from the IMG/PR database [17], which contains nearly seven hundred thousand plasmid sequences. We selected the plasmids classified as putatively completed and with direct terminal repeats topology. We retained 142,492 plasmid sequences.

In this large database, we found that only 4.14% of the plasmids are pCONJ, hence a much lower proportion than observed before [9,16]. Most plasmids, 66.82%, are pMOB, and 29.04% are pNT. We divided the database into two datasets, one containing 13,151 plasmids assembled from bacterial isolates and another comprising 129,341 plasmids assembled from metagenomes. This distinction is biologically significant, as different biases may influence each dataset. Assembling plasmid sequences from bacterial isolates (first dataset) is generally less complex and more reliable than metagenomes (second dataset), where sequences' higher diversity and fragmented nature pose additional challenges.

In the isolates' dataset, 12.4% of the plasmids are conjugative, compared to only 3.3% in the metagenome dataset. These differences align with our expectations because they underscore the potential biases associated with each dataset, namely that plasmids from isolates likely overrepresent plasmids of clinical interest. In contrast, plasmids assembled from metagenomes likely contain a greater diversity and represent a random sample of natural plasmids, including many unknown plasmids.

In the isolates' dataset, 60.7% of the plasmids are pMOB, compared to 67.4% in the metagenome dataset. Then, 26.9% of the isolates' dataset, are pNT, compared to 29.3% in the metagenome dataset.

These two datasets have different characteristics. The proportions of pCONJ, pMOB and pNT are statistically different ($\chi^2(2, N = 147{,}492) = 2475.2$, $p < 2.2\text{x}10\text{-}16$, with a small to moderate effect size, $V = 0.13$). Their median sizes are also statistically different (plasmids from metagenomes (Median = 5165 bp, IQR = 3886 bp) is significantly lower than that of isolates (Median = 7938 bp, IQR = 24121 bp) (Mann-Whitney U = 1093778256, $p < 2.2e\text{-}16$, 99.9% CI = [2106; 2483], rg = 0.29, 99.9% CI [0.27; 0.30], hence with a medium effect)). Consequently, we will analyze the two plasmid groups separately.

Plasmid sizes differ between the three plasmid types in each dataset. In both, pCONJ are significantly larger than pMOB and pNT, which was expected because pCONJ carry all the genes required for conjugation (**Figure 1**).

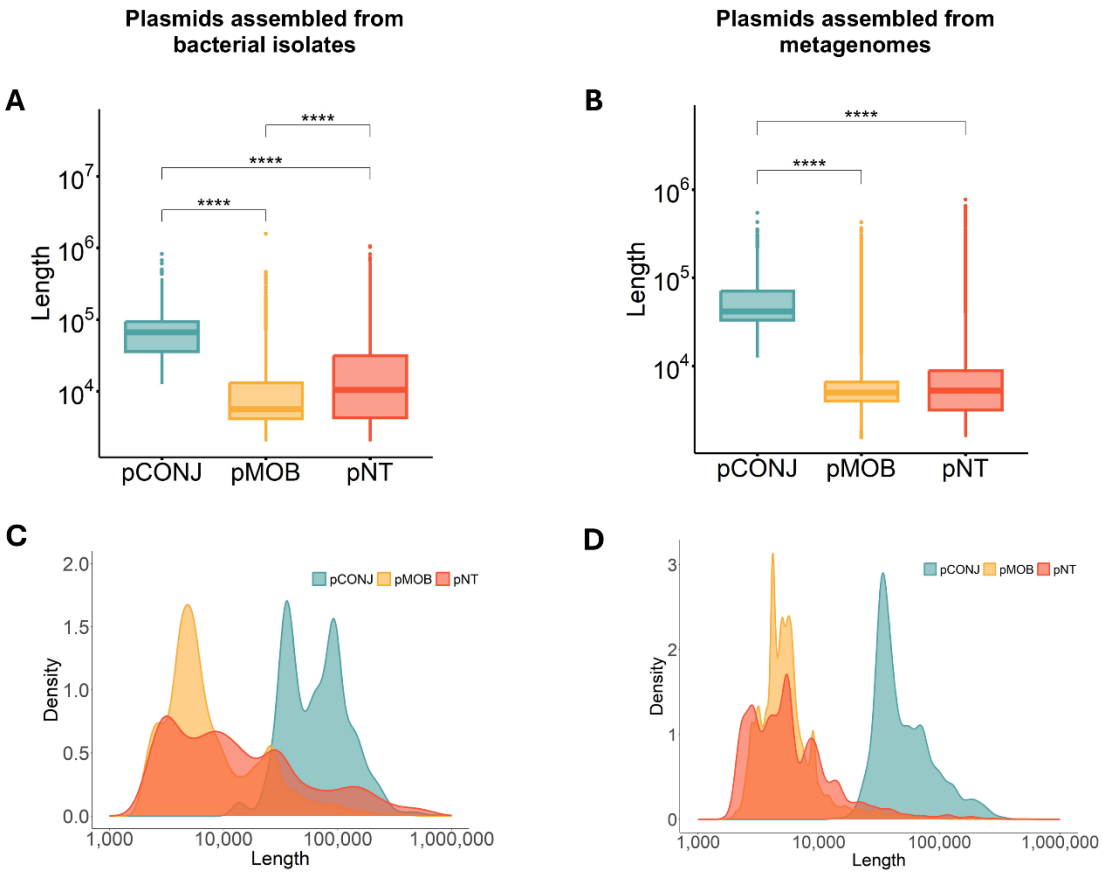


Figure 1. Plasmid size of the three plasmid types: pCONJ, pMOB, and pNT

(A) and (C) Plasmids assembled from bacterial isolates. Figures showing densities normalize the area under the curve to one for each plasmid type. Sizes of the three plasmid groups are different (Kruskal-Wallis analyses: $\chi^2(2, N = 13151) = 3371$, $p < 2.2\text{x}10^{-16}$, $\eta^2 = 0.256$, with large effect size). A Dunn's test revealed that pCONJ are larger than pMOB ($Z = 57.61$, p.adj < $2.2\text{x}10^{-16}$) and pNT ($Z = 38.72$, p.adj < $2.2\text{x}10^{-16}$), and that pMOB are larger than pNT ($Z = -20.15$, p.adj < $2.2\text{x}10^{-16}$).

(B) and (D) Plasmids assembled from metagenomes. Sizes of the three plasmid groups are different ($\chi^2(2, N=129{,}341) = 11424$, $p < 2.2\text{x}10^{-16}$). A Dunn's test revealed that pCONJ are also larger than pMOB ($Z = 106.11$, p.adj < $2.2\text{x}10^{-16}$) and pNT ($Z = 103.07$, p.adj < $2.2\text{x}10^{-16}$), but that the sizes of pMOB and pNT do not differ significantly ($Z = 0.15$, adjusted $p \approx 1$).

*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Distribution of ARGs Across pCONJ, pMOB, and pNT

We then performed a comprehensive screening of plasmids to detect ARGs. To do that, we used the ResFinder database, which contains 3,146 distinct sequences of ARGs derived from whole genome datasets.

Strikingly, only 3.8% (i.e., 5,389 plasmids) of all the plasmids carry ARGs. Moreover, among plasmids carrying ARGs, only 6.9% are conjugative. Most plasmids with ARGs are pMOB (77.9%), followed by pNT (15.2%) (see abstract figure). The isolates' dataset contains 1,517 plasmids (out of 13,151, hence 11.54% ($CI_{95\%}$ = [10.99%, 12.09%]) of the plasmids) with ARGs, and the dataset of plasmids isolated from metagenomes, with 129,341 plasmids contains 3,872 plasmids with ARGs (2.99%, $CI_{95\%}$ = [2.90%, 3.09%]). The proportion of plasmids with ARGs in the metagenomes' dataset is lower than in the isolates' dataset ($\chi^2(1) = 2391$, $p < 2.2\text{x}10^{-16}$, with a small effect size, Cramér's V = 0.130, $CI_{95\%}$ = [0.122, 0.138]). Thus, ARGs are more often linked to plasmids derived from the isolates dataset than to those assembled from metagenomes.

Among plasmids of the isolates' dataset carrying ARGs, 11.0% are pCONJ, 80.8% are pMOB, and 8.2% are pNT. In this dataset, the association between plasmid type and the presence of ARGs is statistically significant ($\chi^2(2) = 336.96$, $p < 2.2\text{x}10^{-16}$, with a weak to moderate association, Cramér's V = 0.160, $CI_{95\%}$ = [0.148, 0.173]). There are more pMOB with ARGs than expected (adj. res. = 17.0), but fewer pNT with ARGs than expected (adj. res. = -17.4). Although not statistically significant, the number of pCONJ with ARGs is lower than expected (adj. res. = -1.7). Therefore, contrary to our initial expectations, there seems to be no selective pressure for pCONJ to carry ARGs.

In the metagenomes' dataset, 5.3% of plasmids encoding ARGs are pCONJ, 76.7% are pMOB, and 18.0% are pNT. This dataset also shows an association between plasmid type and ARGs' presence ($\chi^2(2) = 274.79$, $p < 2.2\text{x}10^{-16}$). There are more pMOB with ARGs than expected (adj res. = 12.5), but fewer pNT with ARGs than expected (adj. res. = -15.7). The number of pCONJ with ARGs is higher than expected (adj. res. = 7.1). However, the statistical significance is driven by a large sample size because the strength of the association is irrelevant (Cramér's V = 0.046, $CI_{95\%}$ = [0.041, 0.051]). Thus, the frequency of ARGs among the three plasmid types appears to be primarily driven by the relative abundance of each plasmid type in the datasets rather than by the preferential selection of these genes for a specific plasmid type. Again, we conclude that the selective pressure on pCONJ to acquire or retain ARGs does not appear stronger than on pMOB.

Plasmids encoding ARGs: Analysis of Size and ARG Density

Among plasmids harboring ARGs, pCONJ are longer (number of bp) than plasmids from the other types and, in the metagenomes' dataset, pNT are longer than pMOB (**Figure 2A** and **2B**). ARGs' density (defined as the number of ARGs/ plasmid size (bp)) of pCONJ is lower than in the other two plasmid groups (**Figure 2C-F**). In the dataset of plasmids from metagenomes, the density of ARGs in pMOB is higher than in pNT (**Figure 2F**).

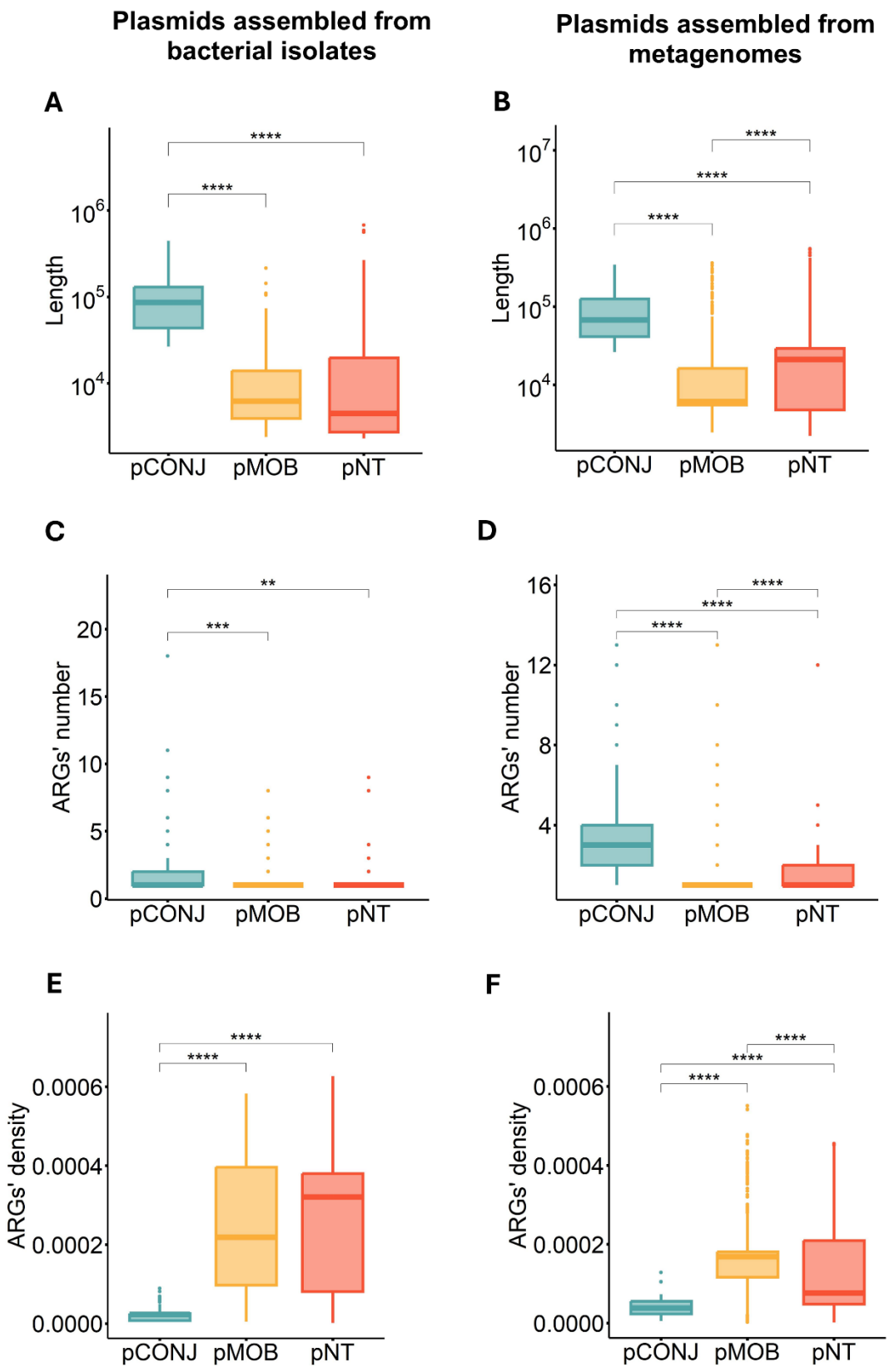


Figure 2. Size of plasmids with ARGs, number of ARGs in plasmids, and density of ARGs

(A) Plasmids assembled from bacterial isolates: Sizes (bp) of the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 1517) = 418.1$, $p < 2.2x10^{-16}$, $\eta^2 = 0.275$ hence of large magnitude). The Dunn test revealed that pCONJ are significantly larger than pMOB ($Z = 20.24$, $p_{adj} < 2.2x10^{-16}$) and pNT ($Z = 14.74$, $p_{adj} < 2.2x10^{-16}$) and that the length of pNT is not significantly different from that of pMOB ($Z = 0.78$, $p_{adj} \approx 1$).
(B) Plasmids assembled from metagenomes: Sizes (bp) of the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 3872) = 567.89$, $p < 2.2x10^{-16}$, $\eta^2 = 0.146$ hence of large magnitude). The Dunn test revealed that pCONJ are significantly larger than pMOB ($Z = 22.82$, $p_{adj} < 2.2x10^{-16}$) and pNT ($Z = 15.76$, $p_{adj} < 2.2x10^{-16}$) and that pNT are larger than pMOBs ($Z = -9.35$, $p_{adj} < 2.2x10^{-16}$).
(C) Plasmids assembled from bacterial isolates: The number of ARGs in the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 1517) = 16.27$, $p = 0.0003$, $\eta^2 = 0.0094$ hence of small magnitude). The Dunn test revealed that pCONJ harbor significantly more ARGs than pMOB ($Z = 3.98$, $p_{adj} = 0.00021$) and pNT ($Z = 2.97$, $p_{adj} = 0.0090$) and that the number of ARGs in pMOB is not significantly different from that in pNT ($Z = 0.24$, $p_{adj} \approx 1$).
(D) Plasmids assembled from metagenomes: The number of ARGs in the three plasmid groups is different (Kruskal-Wallis test, $\chi^2(2, N = 3872) = 411.13$, $p_{adj} < 2.2x10^{-16}$, $\eta^2 = 0.106$, hence of moderate magnitude). The Dunn test revealed that pCONJ harbor significantly more ARGs than pMOB ($Z = 19.63$, $p_{adj} < 2.2x10^{-16}$) and pNT ($Z = 13.98$, $p_{adj} < 2.2x10^{-16}$) and that the number of ARGs in pNT is significantly higher than in pMOB ($Z = -7.24$, $p_{adj} = 1.8x10^{-12}$).
(E) Plasmids assembled from bacterial isolates: Densities (number of ARGs/ plasmid size) of the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 1517) = 376.11$, $p < 2.2x10^{-16}$, $\eta^2 = 0.247$ hence of large magnitude). Dunn tests revealed that the density of ARGs is significantly lower in pCONJ than in pMOB ($Z = -19.32$, $p_{adj} < 2.2x10^{-16}$) or pNT ($Z = -13.18$, $p_{adj} < 2.2x10^{-16}$). The densities of ARGs in pMOB and pNT are not significantly different ($Z = 0.37$; $p_{adj} \approx 1$).
(F) Plasmids assembled from metagenomes: Densities (number of ARGs/ plasmid size) of the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 3872) = 494.22$, $p < 2.2x10^{-16}$, $\eta^2 = 0.127$, hence of moderate effect). Dunn tests revealed that the density of ARGs is significantly lower in pCONJ than in pMOB ($Z = -20.44$, $p_{adj} < 2.2x10^{-16}$) or pNT ($Z = -12.74$, $p_{adj} < 2.2x10^{-16}$). The density of ARGs in pMOB is higher than in pNT ($Z = 10.96$; $p_{adj} < 2.2x10^{-16}$).
*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Plasmid sizes and ARGs' DNA content are positively correlated in both datasets **(Figure 3)**. However, the rate of increase is modest (slopes much lower than one), as most ARGs are distributed across plasmids regardless of size. Therefore, the density of ARGs in larger plasmids is lower than in shorter plasmids.

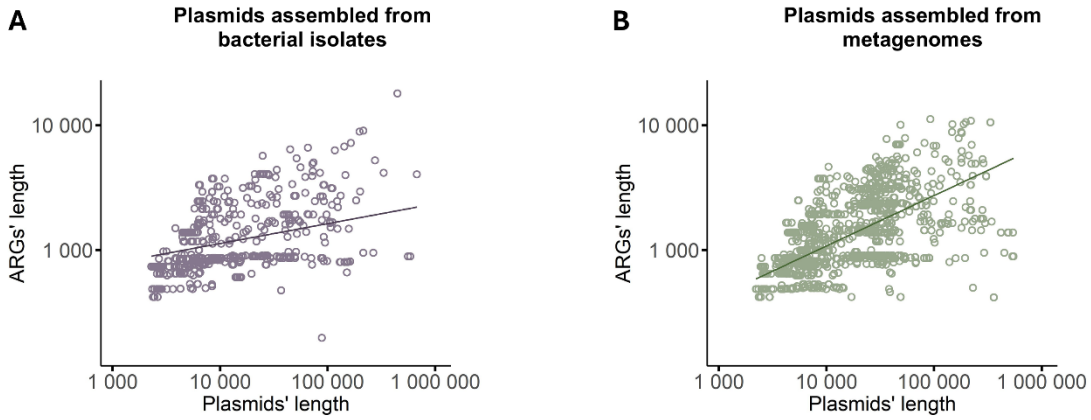


Figure 3. ARGs' DNA content increases with plasmid size

(A) In the dataset of plasmids assembled from bacterial isolates, there is a moderate positive correlation between plasmid size and ARGs' DNA content (Spearman's ρ = 0.53), with a slightly positive trend (slope = 0.16).

(B) In the dataset of plasmids assembled from metagenomes, the correlation is moderate to strong (Spearman's ρ = 0.67), with a slope of 0.40.

Distribution of VF Across pCONJ, pMOB, and pNT

Subsequently, we conducted a comprehensive screening of plasmids to identify VGs. We used the Virulence Factors Database, which contains 4,360 distinct VG sequences [18]. VGs were present in only 418 plasmids among the 142,492 of the entire datasets (0.29%). Among plasmids carrying VGs, 60.1% are pCONJ, 26.3% are pMOB, and 13.6% are pNT (see abstract figure). The isolates' dataset contains 293 plasmids with VGs (out of 13,151, hence 2.23%, $CI_{95\%}$ = [1.98%, 2.49%], of the plasmids), whereas the dataset of plasmids isolated from metagenomes contains 125 plasmids with VGs (out of 129,341, hence 0.097%, $CI_{95\%}$ = [0.080%, 0.115%], of the plasmids). These percentages of plasmids with VGs among plasmids assembled from isolates and metagenomes are statistically different ($\chi^2(1) = 1847$, $p < 2.2\times10^{-16}$, with a small effect size, Cramér's V = 0.114, $CI_{95\%}$ = [0.104, 0.123]). Therefore, VGs are more associated with plasmids from isolates than the ones assembled from metagenomes.

Among plasmids of the isolates' dataset carrying VGs, 67.6% are pCONJ, 21.5% are pMOB, and 10.9% are pNT. In this dataset, the association between plasmid type and the presence of VGs is statistically significant ($\chi^2(2) = 842.59$, $p < 2.2\times10^{-16}$, with a moderate association between plasmid type and VG presence, Cramér's V = 0.253, $CI_{95\%}$ = [0.225, 0.281]). pCONJ with VGs were more represented than expected (adj res. = 29.0), but pMOB (adj. res. = -13.9) and pNT (adj. res. = -6.2) were significantly underrepresented. Therefore, pCONJ plasmids dominate the subset of plasmids with VGs of the isolates' dataset, with pMOB and pNT plasmids occurring less frequently than expected.

In the metagenomes' dataset, 42.4% of the plasmids encoding VGs are pCONJ, 37.6% are pMOB, and 20.0% are pNT. The distribution of plasmids is significantly different from an equal distribution ($\chi^2(2)= 599.4$, $p < 2.2x10^{-16}$, Cramér's V = 0.068, $CI_{95\%}$ = [0.053, 0.083], indicating a very small but statistically significant association between plasmid type and VG presence). pCONJ with VGs are overrepresented (adj. res. = 24.5), but pMOB (adj. res. = -7.1) and pNT with VGs are significantly underrepresented (adj. res: −2.3). Like in the plasmids from the isolates' dataset, pCONJ with VGs are overrepresented among plasmids assembled from metagenomes.

Plasmids encoding VGs: Analysis of Size and VG Density

Among plasmids harboring VGs in the metagenomes' dataset, pCONJ are the shorter (number of bp), and pMOB are shorter than pNT, (**Figure 4**). Regarding the isolates' dataset, there are no significant differences. In the dataset of plasmids from metagenomes, VGs' density (defined as the number of VGs/ plasmid size (bp)) of pCONJ is higher than in the other two plasmid groups (**Figure 4**). Regarding the isolates' dataset, there are no significant differences.

We found that in the large dataset used in this study, less than 0.3% of plasmids are pCONJ and carry ARGs, only 3% are pMOB and carry ARGs and only 0.6% are pNT and carry ARGs. Furthermore, only 0.18% of the plasmids in this dataset are pCONJ and carry VGs, only 0.08% are pMOB and carry VGs, and only 0.04% are pNT and carry VGs.

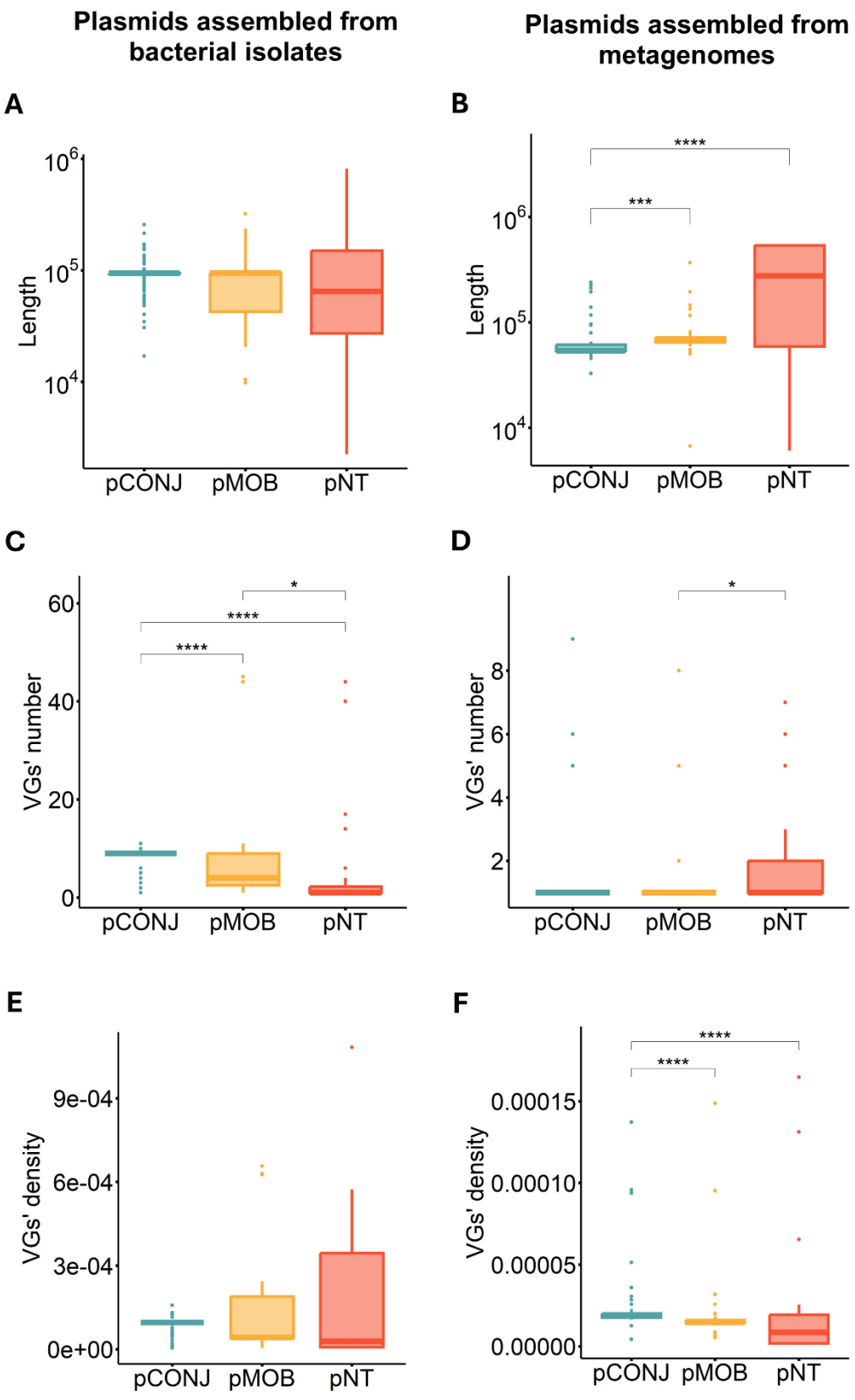


Figure 4. Size of plasmids with VGs, number of VGs in plasmids, and density of VGs

(A) Plasmids assembled from bacterial isolates: Sizes (bp) of the three plasmid groups are not different (Kruskal-Wallis test, $\chi^2(2, N = 293) = 0.359$, $p = 0.8358$, $\eta^2 = -0.006$ hence of small magnitude).
(B) Plasmids assembled from metagenomes: Sizes (bp) of the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 125) = 30.184$, $p = 2.8\times10^{-7}$, $\eta^2 = 0.231$ hence of large magnitude). The Dunn test revealed that pCONJ are significantly smaller than pMOB ($Z= -3.97$, $p_{adj} = 2.1\times10^{-4}$) and pNT ($Z= -5.04$, $p_{adj} = 1.4\times10^{-6}$).
(C) Plasmids assembled from bacterial isolates: the number of VGs in the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 293) = 56$, $p = 6.9\times10^{-13}$, $\eta^2 = 0.186$ hence of large magnitude). The Dunn test revealed that pCONJ harbors significantly more VGs than pMOB ($Z= 4.85$, $p_{adj} = 3.64\times10^{-6}$) and pNT ($Z= 6.49$, $p_{adj} = 2.6\times10^{-10}$) and that the number of VGs in pMOB is higher than in pNT ($Z=2.46$, $p_{adj} = 4.2\times10^{-2}$).
(D) Plasmids assembled from metagenomes: The number of ARGs in the three plasmid groups is different (Kruskal-Wallis test, $\chi^2(2, N = 125) = 7.9$, $p = 0.02$, $\eta^2 = 0.05$, hence of small magnitude). The Dunn test revealed that pNT harbor significantly more VGs than pMOB ($Z= -2.81$, $p_{adj} = 0.015$).
(E) Plasmids assembled from bacterial isolates: Densities (number of VGs/ plasmid size) of the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 293) =7.77$, $p = 0.02$, $\eta^2 = 0.02$, hence of small magnitude). The Dunn test revealed that the density of VGs is marginally higher in pCONJ than in pMOB ($Z= 2.30$, $p_{adj} = 0.064$).
(F) Plasmids assembled from metagenomes: Densities (number of VGs/ plasmid size) of the three plasmid groups are different (Kruskal-Wallis test, $\chi^2(2, N = 125) = 31.99$, $p = 1.13\times10^{-7}$, $\eta^2 = 0.246$, hence of large effect). The Dunn test revealed that the density of VGs is significantly higher in pCONJ than in pMOB ($Z= 5.00$, $p_{adj} = 1.74\times10^{-6}$) or pNT ($Z= 4.38$, $p_{adj} = 3.55\times10^{-5}$).
$^*p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

Discussion

The aim of this work was to evaluate the role of different plasmid mobility types in the spread of antimicrobial resistance and virulence traits. To achieve this, we analyzed a comprehensive collection of fully sequenced plasmids.

Our findings underscore the significance of using metagenome-derived plasmids to achieve a more accurate representation of plasmid distribution in natural microbial communities, rather than relying on inherently biased collections of isolated strains. Although assembling plasmids from metagenomes avoids the selective pressures associated with strain isolation and provides a more representative distribution, it may also lead to an increased risk of false negatives [19] and still fail to capture the full diversity of the plasmidome. The observed differences in the distribution of plasmid types highlight the impact of these challenges on the resulting data. On the other hand, the IMG/PR database [17], used in this study provided an unparalleled diversity of plasmid sequences from a wide range of sources, thereby strengthening the robustness of our conclusions.

After categorizing all plasmids according to their mobility type, pCONJ, pMOB or pNT, we found that only 4.14% of plasmids are classified as pCONJ, in contrast to previous estimates of 23-25% based on smaller datasets [2,3] However, this high proportion of pCONJ reported in the literature may seem inflated. While these plasmids can diversify more frequently than less mobile ones through the acquisition of DNA from other mobile genetic elements and host chromosomes [20–22], three key factors suggest that their actual proportion should be lower. First, with a few exceptions [23–25], several studies have shown that the transfer rate of pCONJ is too low to compensate for their fitness cost [26]. Secondly, if plasmid transfer imposes a cost on bacteria, pCONJ are likely more costly than pMOB, which, in turn, should be more costly than pNT [27]. Third, pCONJ may suffer deactivating mutations (i.e., any DNA changes, such as SNPs, or DNA rearrangements) in transfer genes, making them pMOB or pNT [3,8–11,14,28]. Recent studies corroborate this view. Phylogenetic analyses of relaxases indicate more transitions from pCONJ to pMOB than the reverse [16]. Pointing in the same direction, a recent analysis of pseudogenes in enterobacterial plasmids shows that the non-functionalization of the conjugation machinery has led to the emergence of pMOB [29]. According to these mutational events, the frequency of pCONJ is expected to decrease in favor of pMOB or pNT. Likewise, through mutational events, the frequency of pMOB may decrease in favor of pNT [29].

Furthermore, we observed that antibiotic ARGs are more commonly associated with plasmids in the isolate dataset than in those from metagenomes. This was expected since bacterial genomes in databases mostly come from clinically relevant strains subjected to greater selective pressure to maintain antimicrobial resistance. Our results confirm that plasmids exhibit a similar trend, emphasizing their pivotal role in driving genome evolution under antibiotic selection pressure.

Although conjugative plasmids are more easily mobilized, contrary to our initial expectations, there appears to be no selective pressure for pCONJ plasmids to carry more ARGs than other plasmid mobilization types. Indeed, with the large dataset used in this study, we observe that only 0.26% of the plasmids are pCONJ and carry ARGs, only 2.94% are pMOB and carry ARGs, and only 0.58% are pNT and carry ARGs.

pMOB plasmids exhibit a high density of ARGs and are the most abundant plasmid type in both datasets. Their prevalence makes them the primary carriers of ARGs, a trend not observed for virulence genes (VGs), for example, which are predominantly associated with pCONJ plasmids. Only 0.18% of the plasmids in this dataset are pCONJ and carry VGs, only 0.08% are pMOB and carry VG, and only 0.04% are pNT and carry VG. These percentages are even lower in the dataset of plasmids assembled from metagenomes, which is likely a more representative set of natural plasmids. These findings clearly show that very few plasmids are conjugative, and very few encode drug resistance or virulence factors.

Antibiotic-resistant bacteria cause over one million deaths annually and represent a critical global health challenge. pMOBs impose lower fitness costs on their hosts than pCONJs, allowing them to spread and mobilize ARGs more effectively. This is particularly important when pMOB plasmids coexist with pCONJ plasmids with fewer genes.

This study highlights the important role of pMOB plasmids in the spread of ARGs and the urgent need to prioritize these genetic elements in future research to combat the spread of antibiotic resistance.

In conclusion, this study has provided a more accurate assessment of the distribution of mobile plasmids, both conjugative and mobilizable, in natural environments, revealing a lower frequency of conjugative plasmids than previously expected. It also highlights the important role of pMOB plasmids in the large-scale mobilization of antibiotic resistance. This finding underscores the importance of focusing also on pMOB plasmids in future epidemiologic studies to better understand their contribution to the spread of resistance.

Methods

Plasmid sequences were obtained from the IMG/PR database on December 19, 2023 (https://genome.jgi.doe.gov/portal/IMG_PR/IMG_PR.home.html), which includes almost seven hundred thousand sequences [17]. We focused on the putatively completed plasmids with a topology characterized by direct terminal repeats, assembled from bacterial isolates (13,151 sequences) and metagenomic samples (129,341 sequences). Our dataset comprises a total of 142,492 plasmids.

Antibiotic resistance genes (ARGs) were identified using ABRicate version 1.0.1, with the ResFinder database containing 3,146 nucleotide sequences [31], Virulence genes (VGs) were identified using ABRicate with the Virulence Factors Database (VFDB), which includes 4,360 sequences [32]. Both databases were last updated on 1 February 2024.

Plasmids were classified as conjugative (pCONJ), mobilizable (pMOB), or non-transmissible (pNT) based on the machinery required for DNA transfer. Plasmids encoding mobility genes, type IV coupling protein, type IV secretion systems, and a mating pair formation (MPF) system were classified as pCONJ. If plasmids do not fall within the pCONJ category, they are classified as pMOB if they contain mobility genes or an origin of transfer (oriT). The remaining plasmids are pNT. Nevertheless, these non-transmissible plasmids may have an unidentified oriT, owing to the limited understanding of these regions *aureus* [3].

Statistics

We performed the statistical analyses in R, version 4.4.2 [33]. We used Pearson's chi-squared test to analyze the dependence between two categories and the respective residuals, measuring effect size using Cramér's V from rcompanion package, version 2.4.36. We performed a z-test to compare the plasmid proportions and Cohen’s h to measure the effect size. We used the nonparametric Mann-Whitney U test, also known as the Wilcoxon rank-sum test, to compare the means of two samples, and calculate the effect size using the Glass rank biserial correlation coefficient (r) from effectsize package version 1.0.0. We used Dunn’s post-hoc test from FSA

package version 0.9.5 to identify which groups differ. Visualizations were performed with the R package ggplot2 version v3.5.1.

Acknowledgments

Célia P. F. Domingues and João S. Rebelo acknowledge FCT - Fundação para a Ciência e a Tecnologia, IP for their fellowships (PhD grants UI/BD/153078/2022, and SFRH/BD/04631/2021, respectively). FCT also supports cE3c through the strategic funding (DOI:10.54499/UIDB/00329/2020) and the associate lab CHANGE (LA/P/0121/2020).

Author contributions

Conceptualization, F.D and T.N.; Software, C.P.F.D. and J.S.R.; Methodology, C.P.F.D., J.S.R., F.D. and T.N.; Investigation, C.P.F.D. and J.S.R.; Formal Analysis, C.P.F.D., J.S.R., F.D. and T.N.; Writing – Original Draft, F.D., C.P.F.D. and J.S.R. and T.N.; Resources, T.N.; Visualization, C.P.F.D. and J.S.R.; Supervision, F.D. and T.N.

Declaration of interests

The authors declare no competing interests.

References

1. Fogarty, E. C. *et al.* A cryptic plasmid is among the most numerous genetic elements in the human gut. *Cell* **187**, 1206-1222.e16 (2024).
2. Wang, Y. & Dagan, T. The evolution of antibiotic resistance islands occurs within the framework of plasmid lineages. *Nat Commun* **15**, 4555 (2024).
3. Ares-Arroyo, M., Coluzzi, C. & Rocha, E. P. C. Origins of transfer establish networks of functional dependencies for plasmid transfer by conjugation. *Nucleic Acids Research* **51**, 3001–3016 (2023).
4. Heuer, H. & Smalla, K. Plasmids foster diversification and adaptation of bacterial populations in soil. *Fems Microbiol Rev* **36**, 1083–1104 (2012).
5. Donlan, R. M. Biofilms: Microbial Life on Surfaces. *Emerg. Infect. Dis.* **8**, 881–890 (2002).
6. Johnson, T. J. & Nolan, L. K. Pathogenomics of the Virulence Plasmids of *Escherichia coli*. *Microbiol Mol Biol Rev* **73**, 750–774 (2009).
7. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin Microbiol Rev* **31**, e00088-17 (2018).
8. Bottery, M. J. Ecological dynamics of plasmid transfer and persistence in microbial communities. *Current Opinion in Microbiology* **68**, 102152 (2022).

9. Che, Y. *et al.* Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2008731118 (2021).
10. Dimitriu, T. Evolution of horizontal transmission in antimicrobial resistance plasmids. *Microbiology* **168**, (2022).
11. Fraikin, N., Couturier, A. & Lesterlin, C. The winding journey of conjugative plasmids toward a novel host cell. *Current Opinion in Microbiology* **78**, 102449 (2024).
12. San Millan, A. Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context. *Trends in Microbiology* **26**, 978–985 (2018).
13. Schröder, G. & Lanka, E. The mating pair formation system of conjugative plasmids—A versatile secretion machinery for transfer of proteins and DNA. *Plasmid* **54**, 1–25 (2005).
14. Smillie, C., Garcillán-Barcia, M. P., Francia, M. V., Rocha, E. P. C. & De La Cruz, F. Mobility of Plasmids. *Microbiol Mol Biol Rev* **74**, 434–452 (2010).
15. Ares-Arroyo, M., Nucci, A. & Rocha, E. P. C. Expanding the diversity of origin of transfer-containing sequences in mobilizable plasmids. *Nat Microbiol* (2024) doi:10.1038/s41564-024-01844-1.
16. Coluzzi, C., Garcillán-Barcia, M. P., De La Cruz, F. & Rocha, E. P. C. Evolution of Plasmid Mobility: Origin and Fate of Conjugative and Nonconjugative Plasmids. *Molecular Biology and Evolution* **39**, msac115 (2022).
17. Camargo, A. P. *et al.* IMG/PR: a database of plasmids from genomes and metagenomes with rich annotations and metadata. *Nucleic Acids Research* **52**, D164–D173 (2024).
18. Chen, L. *et al.* VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res* **33**, D325-328 (2005).
19. Antipov, D., Raiko, M., Lapidus, A. & Pevzner, P. A. Plasmid detection and assembly in genomic and metagenomic data sets. *Genome Res.* **29**, 961–968 (2019).
20. Halary, S., Leigh, J. W., Cheaib, B., Lopez, P. & Bapteste, E. Network analyses structure genetic diversity in independent genetic worlds. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 127–132 (2010).

21. Tamminen, M., Virta, M., Fani, R. & Fondi, M. Large-Scale Analysis of Plasmid Relationships through Gene-Sharing Networks. *Molecular Biology and Evolution* **29**, 1225–1240 (2012).
22. Wang, X. *et al.* Global Increase of Antibiotic Resistance Genes in Conjugative Plasmids. *Microbiol Spectr* **11**, e04478-22 (2023).
23. Alonso-del Valle, A. *et al.* Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities. *Nat Commun* **12**, 2653 (2021).
24. Dionisio, F., Matic, I., Radman, M., Rodrigues, O. R. & Taddei, F. Plasmids Spread Very Fast in Heterogeneous Bacterial Communities. *Genetics* **162**, 1525–1532 (2002).
25. Hall, J. P. J., Wood, A. J., Harrison, E. & Brockhurst, M. A. Source–sink plasmid transfer dynamics maintain gene mobility in soil bacterial communities. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8260–8265 (2016).
26. Carroll, A. C. & Wong, A. Plasmid persistence: costs, benefits, and the plasmid paradox. *Can. J. Microbiol.* **64**, 293–304 (2018).
27. San Millan, A. & MacLean, R. C. Fitness Costs of Plasmids: a Limit to Plasmid Transmission. *Microbiol Spectr* **5**, 5.5.02 (2017).
28. Porse, A., Schønning, K., Munck, C. & Sommer, M. O. A. Survival and Evolution of a Large Multidrug Resistance Plasmid in New Clinical Bacterial Hosts. *Mol Biol Evol* **33**, 2860–2873 (2016).
29. Hanke, D. M., Wang, Y. & Dagan, T. Pseudogenes in plasmid genomes reveal past transitions in plasmid mobility. *Nucleic Acids Research* **52**, 7049–7062 (2024).
30. Coluzzi, C., Garcillán-Barcia, M. P., de la Cruz, F. & Rocha, E. P. C. Evolution of Plasmid Mobility: Origin and Fate of Conjugative and Nonconjugative Plasmids. *Mol Biol Evol* **39**, msac115 (2022).
31. Florensa, A. F., Kaas, R. S., Clausen, P. T. L. C., Aytan-Aktug, D. & Aarestrup, F. M. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microbial Genomics* **8**, 000748 (2022).
32. Liu, B., Zheng, D., Jin, Q., Chen, L. & Yang, J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Research* **47**, D687–D692 (2019).

33. Team, R. C. R: A language and environment for statistical computing. *MSOR connections* **1**, (2014).