Not So Mosaic After All? The Core Genomes of IncP-1 Plasmids Evolve Predominantly Through Vertical Transmission

David Sneddon *[1,2,3,4], Clinton A. Elg*[1,2,3], Jack Sullivan[1,2,3], Eva M. Top[1,2,3]

[1] Bioinformatics and Computational Biology Program, University of Idaho, Moscow, Idaho, USA.

[2] Institute for Interdisciplinary Data Sciences, University of Idaho, Moscow, Idaho, USA.

[3] Department of Biological Sciences, University of Idaho, Moscow, Idaho, USA.

[4] Department of Integrative Biology, Oregon State University, Corvallis, Oregon, USA.

* These two authors contributed equally to this work

Corresponding Authors: David Sneddon

sneddon.david.j@gmail.com

Abstract

Broad-host-range plasmids of the incompatibility group IncP-1 play an important role in bacterial evolution by disseminating genes between diverse bacteria. This includes the notable spread of antibiotic resistance genes to bacterial pathogens that increasingly threaten human health. A better understanding of the evolution of genetic elements like IncP-1 plasmids that underwrite bacterial adaptation is required to counter this threat. Here we examined the evolutionary history of IncP-1 plasmids by utilizing the large increase in published genomes and advancing techniques for lineage-based detection of microbial recombination. We specifically tested whether IncP-1 backbone genes evolve primarily *(1)* through vertical transmission, or *(2)* horizontally through homologous recombination between distinct plasmids. Our key finding is the IncP-1 plasmid backbone evolves predominantly through vertical transmission, which was supported by a nucleotide-level mapping of backbone recombination. We also showed via a phylogeny of 187 IncP-1 plasmids that established subgroup categories are well-supported. Finally, we examined factors that are potentially associated with IncP-1 subgroup diversification, including a gene related to plasmid-host adaptation and accessory genes like antibiotic and resistance metals. While some IncP-1 subgroups do correlate with certain categories of accessory genes, it remains unknown if these are causes or consequences of subgroup diversification.

49 # Introduction

50 Self-transmissible broad-host range (BHR) plasmids are known for their ability to transfer
51 genetic material horizontally between diverse species of bacteria via cell-to-cell transfer called
52 conjugation (Datta and Hedges 1972; Frost et al. 2005; Norman et al. 2009). This horizontal
53 transfer of genes is a key driver of bacterial evolution by allowing bacteria to rapidly acquire
54 new traits, such as virulence factors, catabolic functions, or heavy metal and antibiotic resistance
55 (Goldenfeld and Woese 2007; Soucy et al. 2015; Hall et al. 2017). For example, the rapid spread
56 of antibiotic resistance genes carried on plasmids can lead to the emergence of multidrug-
57 resistant (MDR) pathogens, posing significant challenges for human health (Partridge et al.
58 2018). Despite the threat of plasmids disseminating MDR and virulence factors (Pilla and Tang
59 2018), relatively little is known about the evolution of plasmids and its effect on the threats
60 posed by pathogenic bacteria.

61 The IncP-1 group of plasmids stands out for a broad host range reflected in their ability to
62 transfer between and replicate within diverse species of bacteria (Thomas and Smith 1987). IncP-
63 1 plasmids tested to date can transfer into, establish and maintain themselves within multiple
64 classes of Proteobacteria (Olsen and Shipley 1973; Schmidhauser and Helinski 1985; De Gelder
65 et al. 2007; Shintani et al. 2010; Yano et al. 2012; 2013; Klümper et al. 2015; Kottara et al.
66 2016). They have even been shown to mobilize genes into Gram-positive bacteria (Musovic et al.
67 2006) and eukaryotic microbes (Heinemann and Sprague 1989). The transfer, replication, and
68 maintenance functions needed to benefit these plasmids are encoded by a region of 'backbone'
69 genes that share homology across IncP-1 plasmids. These backbone genes represent the core
70 genomes of this plasmid group. In addition, IncP-1 plasmids often contain genes that encode
71 traits beneficial to their host bacteria including MDR, catabolic pathways, and heavy metal

resistance (Schlüter et al. 2007; Sen et al. 2011; Popowska and Krawczyk-Balska 2013). Such ‘accessory’ cargo genes are highly variable across IncP-1 plasmids and are almost always encoded in two specific accessory gene regions (Sota et al. 2007). Thus, the combined capabilities of IncP-1 plasmids to rapidly acquire and disseminate resistance and other beneficial genes across diverse taxa of bacteria contribute to their role as potent vehicles for bacterial adaptation (Popowska and Krawczyk-Balska 2013).

While the contributions of plasmids like the IncP-1 group to bacterial evolution is increasingly appreciated, our understanding of the evolutionary history of the plasmids themselves remains problematic. There are two possible models for evolution of the IncP-1 plasmid backbone, which we describe as horizontal or vertical. Under the horizontal model, plasmid backbone genes are thought to have different evolutionary histories due to frequent homologous recombination between similar plasmids. This contrasts with the vertical model, in which genes within a plasmid backbone are expected to share the same evolutionary history due to vertical inheritance. These different models have important implications for how IncP-1 plasmids evolve.

IncP-1 plasmids are historically classified into subgroups α, β, γ, δ, ε, ζ, η, θ, κ, ι, ο, μ, λ, ρ (Chikami et al. 1985; Pansegrau et al. 1994; Thorsted et al. 1998; Vedler et al. 2004; Norberg et al. 2011; Haines et al. 2006; Brown et al. 2013; Hayakawa et al. 2022a; Bahl et al. 2007). These classifications are based on overall sequence similarity of sister plasmids, as well as monophyly within phylogenies of individual or few backbone genes. However, these phylogenetic inferences prove troublesome as individual plasmid backbone genes that are putatively inherited via homologous recombination from different ancestral plasmids can significantly bias our estimation of evolutionary history when sampling of loci and plasmids are limited. (Schlüter 2003). The literature presents different findings about the degree of a horizontal versus vertical

95 evolutionary history of the IncP-1 plasmids backbone. While one study concluded horizontal
96 recombination is a prominent feature of IncP-1 backbone evolution (Norberg et al. 2011),
97 another suggested a vertical evolutionary history based on successful reconstruction of resolved
98 phylogeny (Sen et al. 2013). There remains a need to understand whether the IncP-1 plasmid
99 backbone predominately evolves through mutation during vertical inheritance or horizontally
100 through recombination. This will help elucidate the host range of these plasmids, and thereby the
101 host trajectories through which their genetic traits like antibiotic resistance spread within
102 microbial communities.

103 Sen et al. (2013) recovered four major topologies across 21 backbone genes, one of which was
104 supported by 46% of gene trees. This topology was deemed to be the "true tree," although a
105 single concatenated phylogenetic analysis produced a different topology. This discordance was
106 due to the weighting of an alternative topology represented by three genes, including one
107 disproportionately large gene, *traE*. Here, we expand upon these previous studies by including
108 many more individual plasmids, as well as employing robust methods that are not as susceptible
109 to biases due to discordant evolutionary histories represented in large loci (Zhang et al. 2018).
110 We further explore the presence of and quantify homologous recombination between recognized
111 subgroups of the IncP-1 group of plasmids using techniques for lineage-based detection of
112 microbial recombination (Mostowy et al. 2017). Our key finding is that vertical inheritance is the
113 predominant factor explaining the evolutionary history of IncP-1 plasmid backbones. This
114 conclusion is based on the well-resolved phylogeny of 187 IncP-1 plasmids,as well as a detailed
115 nucleotide-level mapping of recombination among 20 concatenated backbone genes. Our robust
116 phylogeny showed that established IncP-1 subgroups correspond well to monophyletic groups
117 (*i.e.,* they descended from a common ancestor) with detectable vertical history. While our

118 analysis of homologous recombination did identify examples of recombined backbone genes,
119 including several that were previously discovered (Schlüter et al. 2007; Norberg et al. 2011), we
120 found such recombination is comparatively rare (~1.4% of DNA) within the large set of IncP-1
121 plasmids we examined. Thus, both the phylogenetic and recombination analyses provide clear
122 evidence for the predominance of vertical inheritance in the evolution of IncP-1 plasmid
123 backbones.

124 Given that the existing IncP-1 subgroups are genetically distinct through vertical inheritance, we
125 asked what may contribute to the diversification of these subgroups. For example, the bacterial
126 host or the host's environment might associate with subgroups. We took advantage of our
127 curated IncP-1 dataset to analyze how genetic factors correlate with subgroups, including a
128 known host-range factor and classes of genes including antibiotic resistance, metal resistance,
129 and efflux pumps.

130 Results

131 Established IncP-1 subgroups are phylogenetically well-supported

132 We tested if an expanded set of IncP-1 plasmids would produce a well-supported phylogeny
133 indicative of a vertical evolutionary history. Twenty backbone genes were identified to be
134 present in all the plasmids within our dataset and were thereby selected for the analysis. This set
135 included representative genes from each of the four regions of the IncP-1 backbone that
136 collectively encode the core functions of IncP-1 plasmids (red, yellow, green, and blue in Figure
137 1).

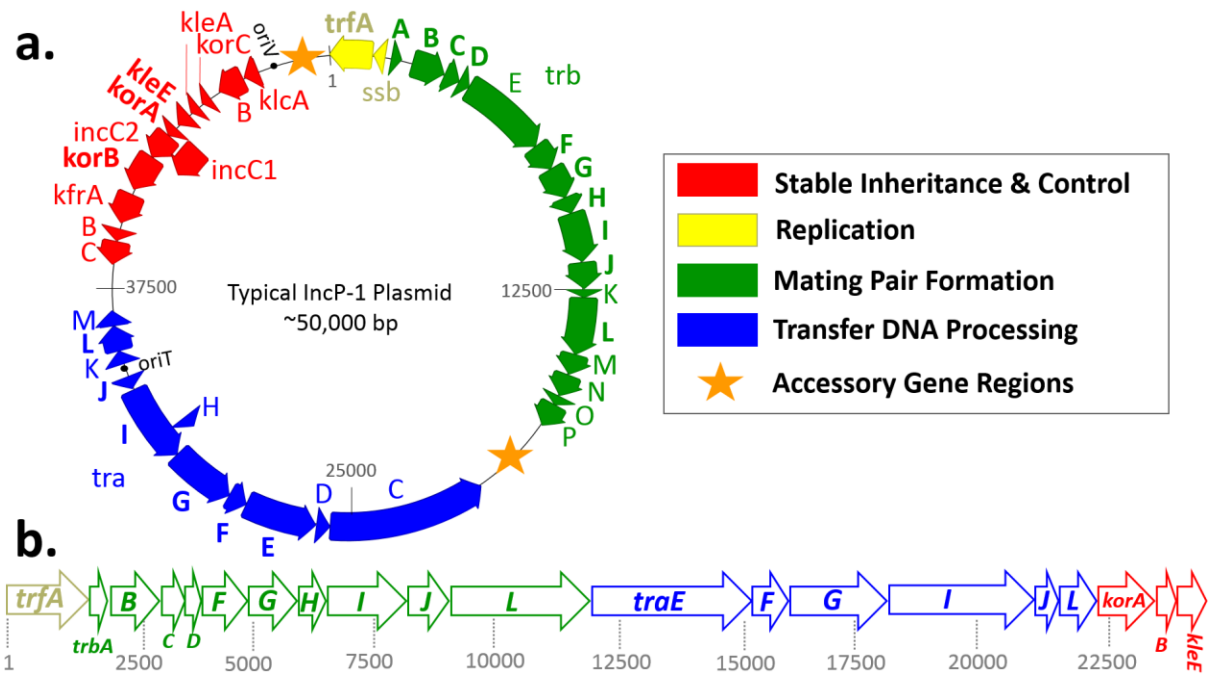

Figure 1. Genomic map of a typical IncP-1 plasmid and the backbone genes used in this study. (a) The four regions of the IncP-1 backbone are shown in red, yellow, green and blue (see legend). An orange star marks each of the two typical accessory gene regions. Bold text signifies backbone genes used in this study. **(b)** A map of the twenty concatenated backbone genes utilized in our lineage analysis.

Among our 20 gene trees we observed expected discordance, with all four gene tree topologies from Sen et al. (2013) represented within our set of loci, as well as a few additional topologies (Supplemental Data S1). The topology selected as the "true" tree in Sen et al. (2013) was represented in 11 of our 20 genes (Supplemental Data S2). This topology was also consistent with our summary lineage tree (Figure 2). Our lineage tree displayed monophyletic resolutions of all historically recognized subgroups with a few exceptions, supporting the hypothesis that all members within most subgroups descended from a common ancestor. All subgroup clades were well supported, with nodal support of ≥ 0.99 posterior probability. High nodal support was present across all deep nodes that define relationships among subgroups except for the nodes that unite IncP-1δ, IncP-1α, and IncP1-ε, in which moderate nodal support was observed (X=0.7,

0.59, 0.71, respectively). This well-supported lineage tree provides evidence of a strong signal of vertical evolutionary history among the 20 backbone genes we retained.

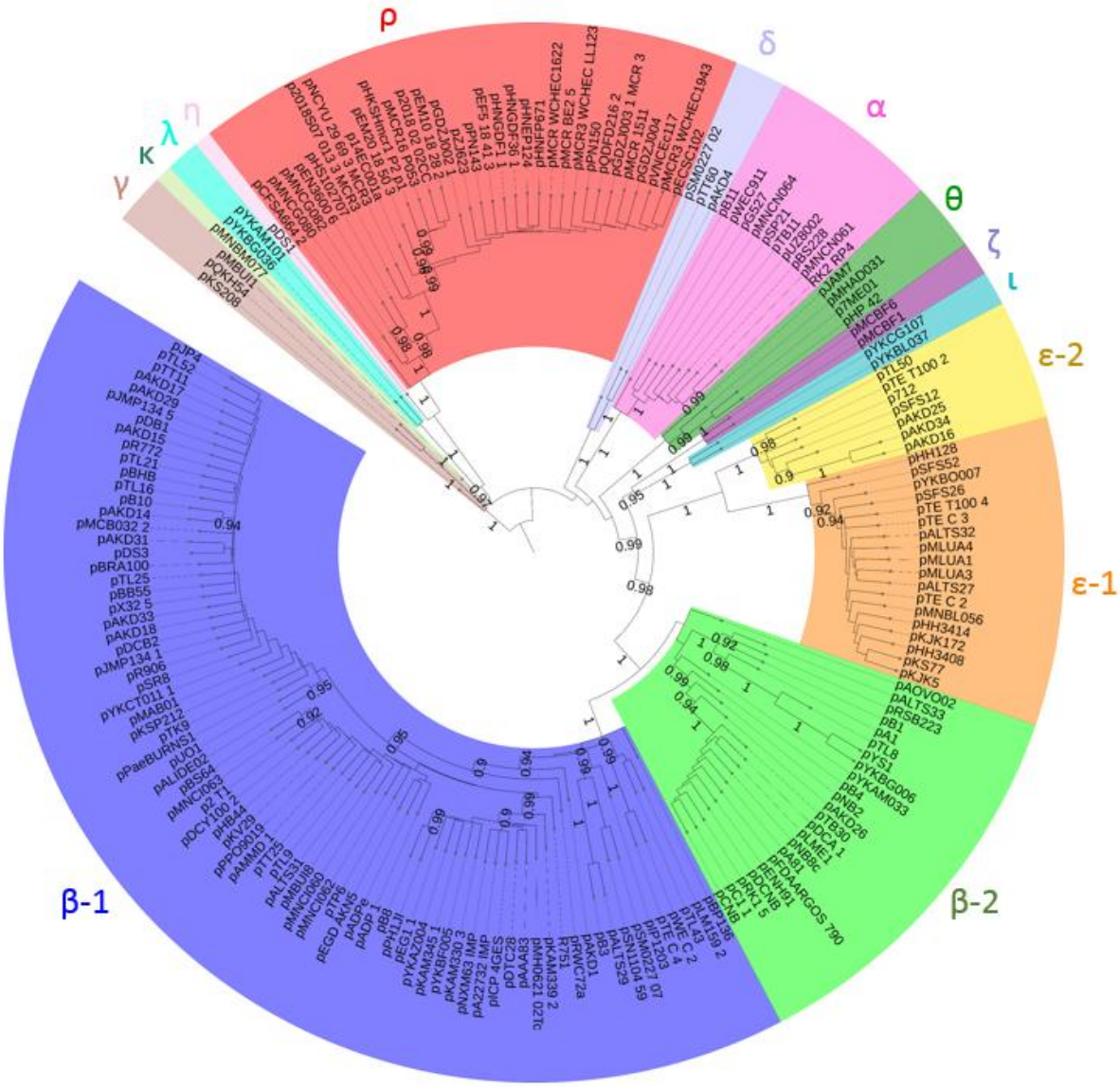

Figure 2. A summary lineage tree of 187 IncP-1 plasmids based on 20 backbone genes. The lineage tree shows monophyletic groups largely in agreement with established IncP-1 subgroups (organized by color). All posterior probabilities > 0.90 are annotated, which demonstrates strong nodal support for the initial branching/diversification, as well as monophyly of distinct subgroups.

152 The individual plasmids present in the monophyletic groups of our tree closely agreed with the
153 14 currently described IncP-1 subgroups: α (alpha; Chikami et al. 1985; Pansegrau et al. 1994),
154 β (beta; Chikami et al. 1985; Thorsted et al. 1998), γ (gamma; Haines et al. 2006; Hill et al.
155 1992), δ (delta; Vedler et al. 2004), ε (epsilon; Bahl et al. 2007), ζ (zeta; Norberg et al. 2011), η
156 (eta; Brown et al. 2013), θ (theta; Yakimov et al. 2016), and a 2022 study (Hayakawa et al.
157 2022b) added ι (iota), κ (kappa), λ (lambda), μ (mu), o (omicron), and ρ (rho). The β and ε
158 subgroups were further divided into β-1 and β-2 (Sen et al. 2010), and ε-1 and ε-2 (Bahl et al.
159 2009), respectively, and are each monophyletic within their respective subgroup. These subgroup
160 divisions are displayed in Figure 2 as independent subgroups to display their monophyly.

161 Our expanded phylogeny suggests a reclassification of two recently proposed subgroups
162 (Hayakawa et al. 2022b), which would reduce the number of subgroups from fourteen to twelve.
163 The proposed o subgroup contains a single plasmid, pYKBG036, which we find is sister with the
164 plasmid from proposed subgroup λ, pYKAM101. We thus propose combining these two
165 plasmids into λ, as the splitting into distinct subgroups does not improve monophyly. Next, the μ
166 group plasmid pMNCG082 renders the ρ subgroup paraphyletic. Thus, we recommend adding
167 pMNCG082 to the ρ subgroup. The net effect of these suggested changes is the merging of
168 subgroup o into λ, and subgroup μ into ρ, leaving twelve subgroups: α, β-1/2, γ, δ, ε-1/2, ζ, η, θ,
169 ι, κ, λ, and ρ.

170 **Recombination of IncP-1 plasmid backbones genes is rare**

171 We next sought to resolve the discrepancy between our lineage tree suggesting a vertical history
172 and previous findings of horizontal recombination in regions of the IncP-1 backbone. In order to
173 assess to what extent horizontal recombination is detectable in the evolutionary history of these

174 IncP-1plasmid segments, we performed a search for recombined regions using a hidden markov
175 model implemented in the software fastGEAR (Mostowy et al. 2017); see Methods). In this
176 analysis, the nucleotide frequency within a subgroup represents a unique genetic signal. Every
177 site of a considered plasmids backbone region was scanned and assigned to the subgroup whose
178 genetic signal it best matches. Under a vertical evolutionary history, the plasmid backbone
179 region will best match the genetic signal of its respective subgroup. Alternatively, under
180 horizontal evolution the recombined regions of a plasmid backbone will match the signal of the
181 subgroup with which recombination occurred. This approach allowed for an increased level of
182 precision in detecting recombination within and between 20 backbone genes across 187 IncP-1
183 plasmids. As in the lineage tree analysis, the β-1/2 and ε-1/2 divisions were treated as
184 independent subgroups.

185 We found that homologous recombination among IncP-1 plasmid backbone genes is rare. The
186 most notable result was the degree to which vertical history is conserved in these 20 IncP-1
187 plasmid backbone genes. This was evident in the recombination visualization of Figure 3, in
188 which small patches of between-subgroup recombination contrast with the overall pattern of
189 vertical evolution of IncP-1 plasmid backbone genes. A total of 35 recombined regions were
190 detected, 28 of which are unique to a single plasmid. The seven remaining recombined regions
191 were conserved within lineages, each suggesting an ancestral recombination event that became
192 vertically inherited. Most notable were 22 β-1 plasmids that have identical portions of *korB* and
193 *kleE* from β-2 plasmids (Figure 3 bottom right-hand corner, in light green). To quantify the
194 prevalence of recombination, we counted 65,369 base pairs as originating from external
195 subgroups. The total number of base pairs considered was 4,575,516 (concatenated genes with
196 base pair length of 24,468 x 187 plasmids analyzed). This suggests that only 1.4% of the

analyzed IncP-1 plasmid backbone DNA had a detectable horizontal signal. Thus, 98.6% of the analyzed backbone genes had a vertical evolutionary history while less than 2% was attributable to horizontal between-plasmid recombination (Figure 3; for full resolution image, see Supplemental Data S3).

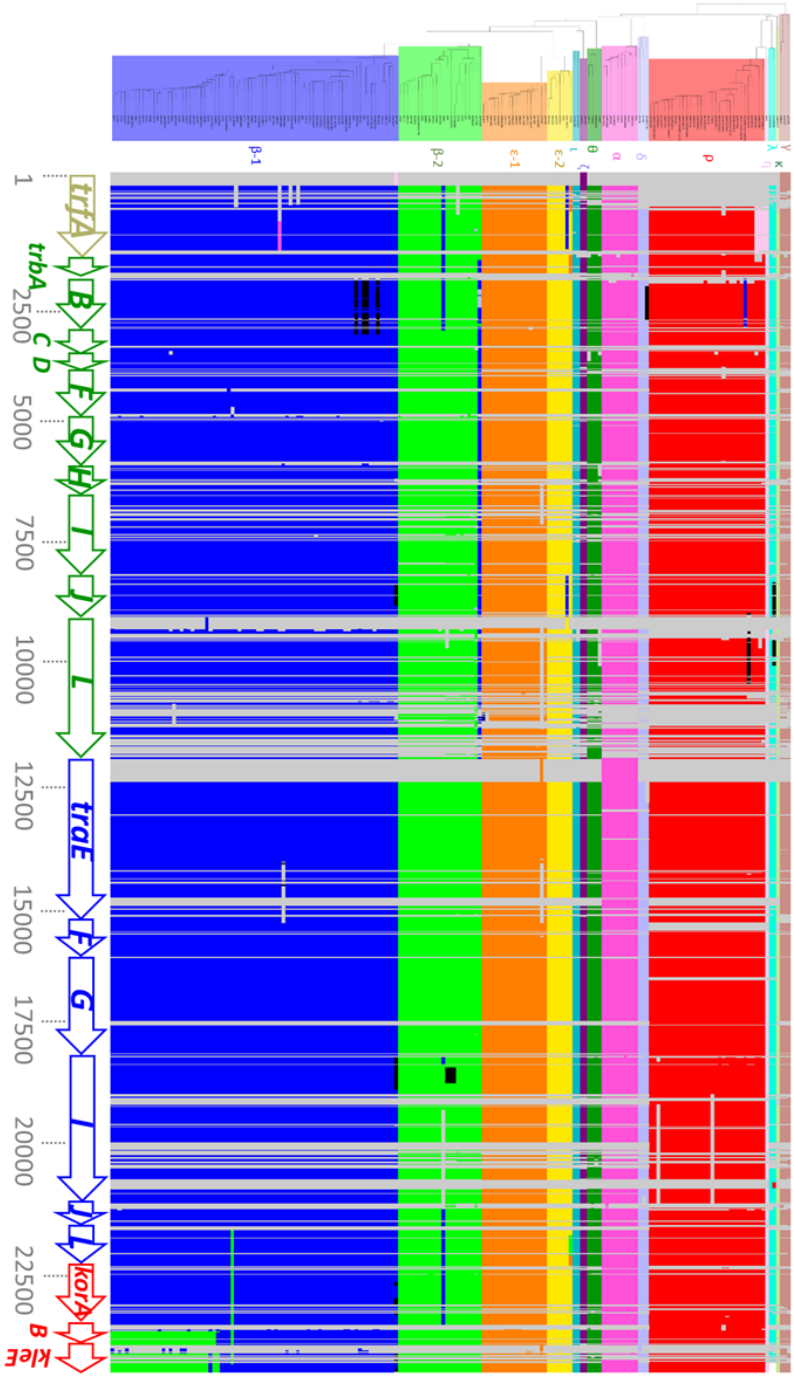

Figure 3. A map of backbone gene recombination between IncP-1 subgroups. Concatenated backbone genes (at bottom) give a nucleotide-level resolution of recombination detected in individual plasmid backbones (phylogenetic tips on left). Vertical grey bars represent gaps in alignment sequences. All other colors represent the genetic signal found within a respective subgroup. Under vertical transmission, the genetic signal of a plasmid's genes will best match that of its subgroup (represented by matching colors). When a region of plasmid backbone is more similar to the genetic signal of a different subgroup, recombination is detected and colored with the plasmid subgroup with which it shares genetic similarity. Black represents regions with sections matching more than one subgroup.

202 To establish the accuracy of our recombination predictions we compared the exact genomic
203 locations of detected recombination to previously reported IncP-1 recombination events. We
204 found four specific examples of IncP-1 backbone region recombination in the literature: plasmids
205 pB10 (Schluter 2003), pIJB1 ((Sen et al. 2010), pB3, pBP136, and pAOVO02 (all from (Norberg
206 et al. 2011). Plasmid pIJB1 is an unusual case of a repeated backbone region and was not
207 included in our study. Plasmid pB10 was the first plasmid in which IncP-1 backbone gene
208 recombination was detected, and our results likewise detected recombination in *kleE* (Schlüter
209 2003). The recombinant *trbB* previously identified in pB3 was present in our recombination data
210 (Schlüter 2003). Norberg et al. (2011) previously predicted an undefined recombination in
211 pAOVO02, and our analysis now provides precision by identifying *trbB* as the site of
212 recombination. We also detected a recombined *trbJ* and *traJ* in pBP13632. In summary, the
213 regions of orthology between subgroups as detected by fastGEAR support and inform the
214 examples of IncP-1 plasmid backbone recombination identified in previous studies. The
215 previously identified recombination was identified and represented in individual gene
216 phylogenies by fastGEAR. The greatly expanded number of plasmids considered here and the
217 increased resolution of our recombination analysis allowed us to demonstrate that recombination
218 between IncP-1 plasmid backbones is rare (<2% of analyzed base pairs).

219 **Factors associated with IncP-1 subgroup diversification**

220 The IncP-1 subgroups have distinct vertically inherited backbones, but what genetic factors
221 might influence diversification into these subgroups? One possible factor could be plasmid
222 adaptation to bacterial hosts, which has been observed in experimental evolution studies.
223 Additional factors that may contribute to IncP-1 diversification are the accessory genes. Plasmids
224 containing accessory genes that are only advantageous to their host in specific environments may

225 themselves become partitioned within that environment and thus become isolated from other
226 IncP-1 plasmids that do not encode those host beneficial genes.

227 Searching the MEGAres (Bonin et al. 2023) database against our IncP-1 dataset yielded 401
228 antibiotic drug resistance, 79 metal resistance, and 54 multi-compound resistance genes
229 (Supplemental Data S4). There was an average of 2.1 drug resistance, 0.42 metals resistance, and
230 0.29 multi-compound genes per plasmid. For each gene category, we took the average number of
231 genes found within a subgroup and divided it by the average number of genes for all plasmids.
232 The results are presented in Figure 4, where a value of 1 means that the subgroup has an average
233 number of category genes, whereas a value of 2 shows this subgroup has on average twice as
234 many category genes compared to all plasmids, and so on. A value of zero represents that class
235 of genes was not found in a particular subgroup. Figure 4 shows clear differential patterns of
236 accessory gene categories across subgroups. Of particular interest is the division of the ε
237 subgroup: ε-1 plasmids have on average ~> 3-fold more drug and multi-compound genes than
238 average, while ε-2 had ~0.5 the average number of hits in the same categories. This is consistent
239 with the finding of Li et al. (2016) that in contrast to the IncP-1ε-1 plasmids, the ε-2 plasmids
240 known at the time did not contain antibiotic resistance genes. Another interesting observation is
241 the slightly higher gene content for all three categories in the β-1 group compared to β-2. This is
242 consistent with a large clade of β-2 plasmids being catabolic plasmids that do not encode
243 antibiotic resistance genes (Król et al. 2012).

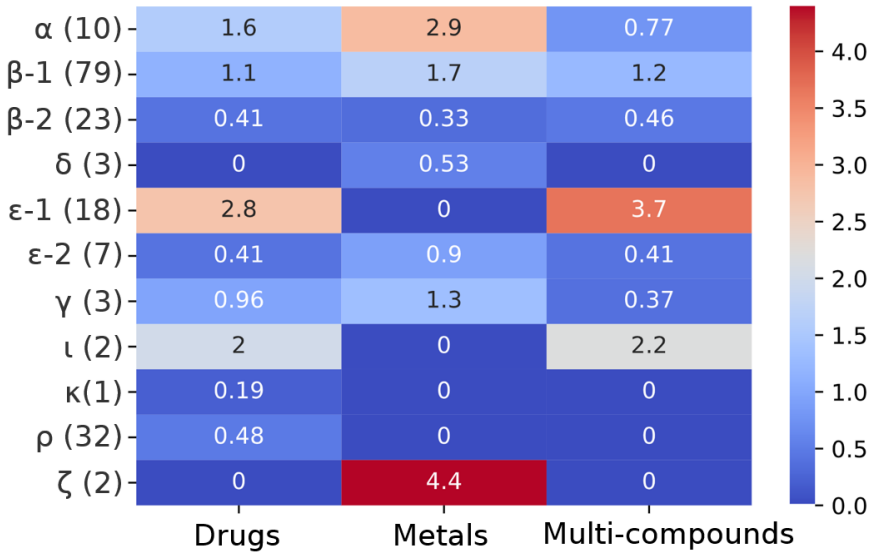

Figure 4. The relative frequency of classes of accessory region genes in 187 IncP-1 plasmids. Accessory regions genes were divided into sets of genes conferring resistance to drugs, or metals, or 'multi-compound' genes like efflux pumps which may confer some resistance to both (horizontal axis). Subgroups are on the vertical axis, with the parenthesis containing the sum of plasmids in that subgroup that contain at least one of the classes of genes examined. The numbers on the right-side scale are calculated from the average number of genes within a subgroup divided by the average number of genes across all plasmids.

Discussion

Plasmids represent a molecular cornerstone, allowing rapid bacterial adaptation to changing environments, and are of particular importance in conferring MDR to pathogens that threaten human health (Castañeda-Barba et al. 2024). A better understanding of the evolution of IncP-1 plasmids, which can broadly disseminate such resistances between diverse species (Popowska and Krawczyk-Balska 2013b), is therefore of importance. We found that IncP-1 plasmid backbones primarily evolved through vertical transmission and were able to quantify the low prevalence and genetic location of horizontal recombination among a subset of backbone genes. Of particular interest are the factors that may lead to subgroup diversification (Li et al. 2016), as understanding the dynamics of IncP-1 plasmid diversification can provide crucial context to their

255 role in bacterial adaptation, including the spread of genetic traits like MDR within bacterial
256 communities.

257 Our lineage tree of 187 IncP-1 plasmids leveraged the increase of available plasmid sequences to
258 provide new insights into their evolutionary history. Overall, we observed a clear vertical history
259 of evolution represented in our individual-level phylogeny summarized across 20 plasmid
260 backbone genes. The topology of our summary tree displays the same topology represented by
261 the majority of our genes (11 genes; Supplemental Data S2); however, we do recover discordant
262 topologies across other genes within our dataset. This signal is represented in the deep nodal
263 support of our tree, in which discordant nodes that unite certain subgroups have significantly
264 lower support than other deep nodes. In all but a select few cases as illustrated by our
265 recombination analysis, IncP-1 subgroups are monophyletic regardless of the deeper resolution
266 between subgroups. This indicates that this discordance among gene trees is likely caused by rare
267 single recombination events deep in the tree that are vertically inherited by all plasmids in the
268 recipient subgroup, or random sorting of ancestral alleles (incomplete lineage sorting).

269 As observed by Sen et al. (2013), the majority of individual gene tree topologies (11 of 20)
270 support a single underlying topology in which IncP-1δ, IncP-1α, and IncP-1ε are all paraphyletic
271 with respect to each other, and IncP-1ε is sister to IncP-1β (Topology 1 in Sen et al. 2013). The
272 second most represented gene tree topology places the α subgroup sister to δ (5 genes). In
273 previous studies similar to ours, large loci (e.g., *traE*) that may be involved in backbone
274 recombination have biased overall phylogenetic signal towards this alternative topology, which
275 prompted the removal of these loci from their analyses. Our summary-based phylogenomic
276 approach is more robust to these alternative topologies, as all genes are weighted equally

277 regardless of sequence length. This allows us to use an expanded set of genes, thus retaining any
278 signal of inter-subgroup recombination deep in the tree.

279 Interestingly, the recent addition of newer subgroups (ρ, θ, ζ, ι, λ, η, κ) does not disrupt the
280 overall topology recovered by Sen et al (2013) as the ρ, η, κ, and λ all form a monophyletic
281 group and are sister to the rest of the IncP-1 ingroup. Subgroups ι and ζ are sister but are not
282 monophyletic with any other individual subgroups. This is the case for θ as well. Across
283 individual gene trees, resolution of κ, which is represented by a single plasmid (pMNBM077) is
284 highly discordant. This could be due to inter-subgroup recombination that is not represented in
285 our fastGEAR output. As this plasmid is the only representative of this subgroup, and is highly
286 dissimilar to any other subgroup, our analytical framework lacks the statistical power to detect
287 any recombination. To explore the recombination history of this plasmid under the framework
288 presented here, more individuals belonging to the κ subgroup must be sampled.

289 We are indebted to the early work of previous studies that explored the evolution of IncP-1
290 plasmid backbone genes, often finding cases of recombination by painstaking manual analysis
291 (Schluter 2003; Norberg et al. 2011; Sen et al. 2013). It was striking to find four cases where
292 previous predictions matched our HMM predicted recombined regions (in pB3, pB10,
293 pAOVO02, pBP136). These studies were limited at the time by the number of IncP-1 genomes
294 available (ranging between 25 and 65), and the less precise methods of detecting recombination.
295 For example, utilizing the pairwise homoplasy index (PHI; Bruen et al. 2006) allowed for
296 statistical testing of backbone regions for recombination (Norberg et al. 2011), but this could not
297 provide any genetic context on where or to what degree regions of recombination exist.
298 Additionally, only four subgroups had enough representative plasmids to be examined for
299 recombination, compared to the twelve subgroups we analyzed here. Remarkably, six of these

300 subgroups have been defined in just the last couple of years (Hayakawa et al. 2022). Our present
301 work considering 187 plasmids and the nucleotide by nucleotide fine-resolution of
302 recombination regions also agrees with the previous finding of a detectable vertical history in
303 IncP-1 backbone genes (Sen et al. 2013). The low frequency and limited cases of backbone
304 recombination found in our expanded study suggest that IncP-1 plasmid backbones primarily
305 evolve by mutations that are vertically inherited, rather than by homologous recombination
306 between plasmids.

307 Thus, it becomes apparent that the distinct IncP-1 subgroups generally display "recombination
308 isolation," which raises important questions about the host range, ecology, and evolution of IncP-
309 1 plasmids. Are IncP-1 plasmid subgroups physically separated by environment? Our analysis of
310 genes encoding antibiotic and metal resistance and efflux pumps shows clear differences
311 between subgroups, suggesting the accessory genes of a subgroup may collectively reflect the
312 environment of the bacterial plasmid hosts (Popowska and Krawczyk-Balska 2013; Król et al.
313 2012; Li et al. 2016). However, it is unclear if this represents a cause or consequence of
314 subgroup diversification. Others have hypothesized that subgroup diversification is driven by
315 plasmid-host specialization (Yano et al. 2013). Unfortunately, only a limited number of
316 examples of plasmid adaption to bacterial hosts currently exist to test this (Bouma and Lenski
317 1988; Dahlberg and Chao 2003; Stalder et al. 2017; Bottery et al. 2017; Sota et al. 2010; Hughes
318 et al. 2012; Yano et al. 2016; De Gelder et al. 2008). We recently showed that a IncP-1 plasmid
319 gene that imposed a cost to *E. coli* strains and was rapidly lost in these hosts in the laboratory,
320 was, only found in the IncP-1β subgroup, with one exception (Elg et al. 2025). This suggests at
321 least some of the IncP-1β plasmids, namely those with an intact *upf31* (an accessory gene), have
322 not spent considerable time previously in *E. coli* or closely related taxa, unless they accumulated

323 mutations that inactivated the gene. Such differences in host range may well lead to plasmid
324 backbone diversification. Since *upf31* is only found within the IncP-1β subgroup, this might
325 suggest that the gene factored into a more ancient diversification event relative to the
326 diversification of existing subgroups. However, the presence of an intact *upf31* within the
327 phylogeny of the IncP-1 β was paraphyletic (Elg et al. 2025), indicating it may not be actively
328 contributing to diversification within or between established subgroups. Since *upf31* is known to
329 have a very harmful effect on laboratory *E. coli* strains, and many captured plasmids are mated
330 into *E. coli*, it is possible more plasmids contained full-length *upf31* existed but lost the gene in
331 *E. coli* laboratory passage before sequencing. Further discovery of plasmid adaptation to
332 bacterial hosts, and examination of how they correspond to known subgroup lineages, is needed
333 to assess the role of plasmid-bacteria adaptation in IncP-1 plasmid diversification.

334 Alternatively, it is possible that IncP-1 plasmids from different subgroups regularly co-exist in
335 bacterial communities but are not jointly present within cells long enough for homologous
336 recombination to frequently occur. IncP-1 plasmids belong to the same incompatibility group,
337 which indicates their general inability to co-exist within a cell lineage for multiple generations
338 due to shared replication or partitioning machinery (Novick 1987). Furthermore, all IncP-1
339 plasmids in our study contain at least two entry exclusion genes, *trbJ* and *trbL*, that lower the
340 frequency of conjugation of another plasmid into the same cell (Garcillán-Barcia and de la Cruz
341 2008; Cabezón et al. 2015). Thus, the incompatibility between plasmids combined with entry
342 exclusion provides a rational explanation for the low frequency of detectable recombination
343 observed between IncP-1 plasmids backbones. While one study incidentally found some IncP-1
344 plasmids could co-exist for up to fifty generations (Chikami et al. 1985), we are otherwise
345 unaware of any systematic examination of the time various IncP-1 plasmids can co-exist intra-

346 cellularly. Such a study is warranted, especially if it explored *(1)* the degree of plasmid backbone
347 similarity, and *(2)* the period of intra-cellular co-existence required for recombination to
348 influence plasmid evolution.

349 Despite a primarily vertical history with strong recombination isolation, we cannot discount that
350 rare recombination events in plasmid backbone genes could meaningfully impact IncP-1 plasmid
351 evolution. Based on data from IncF and IncW plasmids, it has been hypothesized (Fer Andez-
352 Opez et al. 2006; Boyd et al. 1996) that rare recombination events set conjugative plasmids on
353 different evolutionary trajectories leading to diversification. Thus, the rare recombination we see
354 in IncP-1 plasmids might reflect a more generalizable mechanism of plasmid diversification. As
355 shown by Luria and Delbrück in their classic study (Luria and Delbrück 1943), a rare and early
356 evolutionary contingency can have outsized influence on later population genetics (Blount et al.
357 2018). It is left to be discovered if the recombined *korB* and *kleE* regions found in 22 β-1
358 plasmids, which appear to be vertically inherited after an initial recombination event, may
359 represent an early subgroup diversification event. More broadly, our predicted regions of
360 recombination and the lineage tree of backbone genes should aid future research into the factors
361 that lead to subgroup diversification of this prevalent and broad-host-range group of plasmids.

362 Methods

363 Curation of IncP-1 Dataset and Concatenation of Backbone Genes

364 The PLSDB (Galata et al. 2019) plasmid database was used along with a literature review to
365 identify IncP-1 plasmid sequences. Our initial set of 206 plasmids was reduced to 187 when
366 partial sequences and those with transposon insertions in backbone genes were removed
367 (Supplemental Data S5).

368 The number of backbone genes vary by subgroup with an average of 33 genes (Sen et al. 2013).
369 Among our 187 functional plasmid backbone genomes, we identified a set of 20 reliable,
370 conserved backbone genes based on presence and annotation quality that were included in our
371 phylogenetic analyses (Supplemental Data S6). These 20 candidate loci include representatives
372 from the four regions of the plasmids backbone that collectively encode the core functions of
373 IncP-1 plasmids; replication, stable inheritance, control, and the conjugative transfer regions *tra*,
374 and *trb* (Thorsted et al. 1996; Macartney et al. 1997; Figure 1). The replication initiation gene
375 *trfA1* was utilized, as the *ssb* gene was interestingly missing from some plasmids in our dataset.
376 From the central control region, which includes genes for the stable inheritance and regulation of
377 replication and transfer function, we utilized *korB*, *korA*, and *kleE*. For the *tra* region that
378 encodes genes for processing plasmid transfer DNA we included *traE, F, G, I, J, L*. In the *trb*
379 region that encodes the mating pore and conjugative pilus for plasmid transmission we used
380 *trbA, B, C, D, F, G, H, I, J, L*. These individual gene sequences were aligned using mafft v7.429
381 (Katoh and Standley 2013). The alignment file for concatenated backbone genes of all 187
382 plasmids is available in Supplemental Data S7.

383 **Inference of the IncP-1 phylogeny**

384 We reconstructed individual gene trees for the 187 IncP-1 plasmids for all candidate loci. To
385 distinguish between the random sorting of ancestral alleles (incomplete lineage sorting; ILS) and
386 homologous recombination, we utilized summary-based methods of phylogenetic reconstruction
387 under the multispecies coalescent model (MSC; Pamilo and Nei 1988; Rannala and Yang 2003;
388 Knowles and Carstens 2007). First, we generated maximum-likelihood phylogenies from each of
389 the 20 aligned backbone genes using IQ-TREE v2.2.5 (Minh et al. 2020). Models of nucleotide
390 evolution were selected under Bayesian information criteria (BIC) using ModelFinder

391 (Kalyaanamoorthy et al. 2017) as implemented within IQ-TREE, and maximum-likelihood
392 phylogenies were estimated for each gene independently using 1000 ultra-fast bootstrap
393 replicates (Hoang et al. 2018). From our 20 independent gene tree topologies, we estimated a
394 single lineage tree using ASTRAL-III v5.7.367 (Zhang et al. 2018). The final ASTRAL-III
395 lineage tree was rooted on the IncP-1γ clade as outgroup as done in Sen et al (2013). Our final
396 summary lineage tree was visualized and annotated using the Interactive Tree of Life (Letunic
397 and Bork 2024) and are publicly available to view at
398 https://itol.embl.de/shared/1SsM2kpxu4zAB.

399 **Detection of Backbone Gene Recombination**

400 We identified regions of the plasmid backbone that are involved in recombination at a nucleotide
401 level using fastGEAR (Mostowy et al. 2017). This program uses a lineage-informed hidden
402 Markov model (HMM) to detect microbial recombination by using polymorphic sites between an
403 individual sequence and other sequences in same lineage to determines whether it is more likely
404 that it originated from an external lineage. The HMM will then determine the alternative lineage
405 from which it was most likely received. We first used fastGEAR's built-in Bayesian analysis of
406 population structure (BAPS; Corander et al. 2003) to perform lineage clustering, which produced
407 clusters consistent with observed clades in our IncP-1 lineage tree. We then took our IncP-1
408 clades as illustrated in Figure 2 as inputs for fastGEAR using the 'current' setting, resulting in
409 the output for Figure 3 and Supplemental Data S8.

410 **Comparison of accessory gene patterns between subgroups**

411 Different patterns of resistance gene acquisition between plasmids can contribute to environment
412 specificity potentially leading to the isolation and diversification of plasmid subgroups hosts

413 (Popowska and Krawczyk-Balska 2013; Król et al. 2012; Li et al. 2016). To examine this in our
414 set of 187 IncP-1 plasmid genomes we searched our genomes for genes categorized as *(i)* drugs
415 resistance, *(ii)* metal resistance, and *(iii)* 'multi-compound' resistance as classified within the
416 MEGAres database (Bonin et al. 2023). The latter refers to genes like those encoding efflux
417 pumps, which confer at least some resistance to multiple factors (drugs, metals, biocides, *etc.).*
418 Genes within our plasmid genomes that had greater than 90% coverage to gene accessions within
419 the database were identified and returned. The gene accession with the highest coverage was
420 chosen in cases where a plasmid gene had greater than 90% coverage to multiple resistant gene
421 accessions. For the full list of unique gene accessions returned by our search, see Supplemental
422 Data S4. To show differential patterns of resistance gene categories between subgroups, we took
423 the average number of accessory genes for each category found in each subgroup and scaled it to
424 the average number of genes for all plasmids.

425 Acknowledgements

426 We are thankful to Olivia Kosterlitz (University of Washington, University of Idaho) and Luke
427 Harmon (University of Idaho) for making important improvements to this manuscript. Thank
428 you to Thibault Stalder (INSERM) for assisting with resistance gene identification. E.M.T. and
429 C.A.E. received support for this work from National Institute of Allergy and Infectious Diseases
430 Extramural Activities grant no. R01AI084918 from the National Institutes of Health. C.A.E. was
431 supported by the NSF Graduate Research Fellowship (GRFP) grant no. DGE-2019265372.
432 C.A.E and D.J.S. were both supported by the Bioinformatics and Computational Biology (BCB)
433 Fellowship and the Paul Joyce Memorial BCB Fellowship Endowment at the University of
434 Idaho.

References

Bahl, M. I., Hansen, L. H., Goesmann, A., & Sørensen, S. J. (2007). The multiple antibiotic resistance IncP-1 plasmid pKJK5 isolated from a soil environment is phylogenetically divergent from members of the previously established α, β and δ sub-groups. *Plasmid, 58*(1), 31-43.

Bahl, M. I., Burmølle, M., Meisner, A., Hansen, L. H., & Sørensen, S. J. (2009). All IncP-1 plasmid subgroups, including the novel ε subgroup, are prevalent in the influent of a Danish wastewater treatment plant. *Plasmid*, *62*(2), 134-139.

Blount, Z. D., Lenski, R. E., & Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life's tape. *Science, 362*(6415), eaam5979.

Bonin, N., Doster, E., Worley, H., Pinnell, L. J., Bravo, J. E., Ferm, P., Marini, S. Prosperi, M., Noyes, N., Morley, P. S., & Boucher, C. (2023). MEGARes and AMR++, v3. 0: an updated comprehensive database of antimicrobial resistance determinants and an improved software pipeline for classification using high-throughput sequencing. *Nucleic acids research, 51*(D1), D744-D752.

Bottery, M. J., Wood, A. J., & Brockhurst, M. A. (2017). Adaptive modulation of antibiotic resistance through intragenomic coevolution. *Nature ecology & evolution, 1*(9), 1364-1369.

Bouma, J. E., & Lenski, R. E. (1988). Evolution of a bacteria/plasmid association. *Nature, 335*(6188), 351-352.

Boyd, E. F., Hill, C. W., Rich, S. M., & Hartl, D. L. (1996). Mosaic structure of plasmids from natural populations of *Escherichia coli. Genetics, 143*(3), 1091-1100.

Brown, C. J., Sen, D., Yano, H., Bauer, M. L., Rogers, L. M., Van der Auwera, G. A., & Top, E. M. (2013). Diverse broad-host-range plasmids from freshwater carry few accessory genes. *Applied and environmental microbiology, 79*(24), 7684-7695.

Bruen, T. C., Philippe, H., & Bryant, D. (2006). A simple and robust statistical test for detecting the presence of recombination. *Genetics, 172*(4), 2665-2681.

Cabezón, E., Ripoll-Rozada, J., Peña, A., De La Cruz, F., & Arechaga, I. (2015). Towards an integrated model of bacterial conjugation. *FEMS microbiology reviews, 39*(1), 81-95.

Carstens, B. C., & Knowles, L. L. (2007). Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from Melanoplus grasshoppers. *Systematic biology, 56*(3), 400-411.

Castañeda-Barba, S., Top, E. M., & Stalder, T. (2024). Plasmids, a molecular cornerstone of antimicrobial resistance in the One Health era. *Nature Reviews Microbiology, 22*(1), 18-32.

Chikami, G. K., Guiney, D. G., Schmidhauser, T. J., & Helinski, D. R. (1985). Comparison of 10 IncP plasmids: homology in the regions involved in plasmid replication. *Journal of bacteriology, 162*(2), 656-660.

471 Corander, J., Waldmann, P., & Sillanpää, M. J. (2003). Bayesian analysis of genetic
472 differentiation between populations. *Genetics, 163*(1), 367-374.

473 Dahlberg, C., & Chao, L. (2003). Amelioration of the cost of conjugative plasmid carriage in
474 *Escherichia coli* K12. *Genetics, 165*(4), 1641-1649.

475 Datta, N., & Hedges, R. W. (1972). Host ranges of R factors. *Microbiology*, *70*(3), 453-460.

476 De Gelder, L., Ponciano, J. M., Joyce, P., & Top, E. M. (2007). Stability of a promiscuous
477 plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology, 153*(2),
478 452-463.

479 De Gelder, L., Williams, J. J., Ponciano, J. M., Sota, M., & Top, E. M. (2008). Adaptive plasmid
480 evolution results in host-range expansion of a broad-host-range plasmid. *Genetics, 178*(4),
481 2179-2190.

482 Elg, C. A., Mack, E., Rolfsmeier, M., McLean, T. C., Sneddon, D., Kosterlitz, O., Soderling, E.,
483 Narum, S., Rowley, P. A., Sullivan, J., Thomas, C. M., & Top, E. M. (2025) Evolution of a
484 Plasmid Regulatory Circuit Ameliorates Plasmid Fitness Cost. *Molecular biology and*
485 *evolution*, *42*(4) msaf062.

486 Fernández-López, R., Garcillán-Barcia, M. P., Revilla, C., Lázaro, M., Vielva, L., & De La Cruz,
487 F. (2006). Dynamics of the IncW genetic backbone imply general trends in conjugative
488 plasmid evolution. *FEMS microbiology reviews, 30*(6), 942-966.

489 Frost, L. S., Leplae, R., Summers, A. O., & Toussaint, A. (2005). Mobile genetic elements: the
490 agents of open source evolution. *Nature Reviews Microbiology, 3*(9), 722-732.

491 Galata, V., Fehlmann, T., Backes, C., & Keller, A. (2019). PLSDB: a resource of complete
492 bacterial plasmids. *Nucleic acids research, 47*(D1), D195-D202.

493 Garcillán-Barcia, M. P., & de la Cruz, F. (2008). Why is entry exclusion an essential feature of
494 conjugative plasmids? *Plasmid, 60*(1), 1-18.

495 Goldenfeld, N., & Woese, C. (2007). Biology's next revolution. *Nature, 445*(7130), 369.

496 Haines, A. S., Akhtar, P., Stephens, E. R., Jones, K., Thomas, C. M., Perkins, C. D., Williams, J.
497 R., Day, M. J., & Fry, J. C. (2006). Plasmids from freshwater environments capable of IncQ
498 retrotransfer are diverse and include pQKH54, a new IncP-1 subgroup archetype.
499 *Microbiology, 152*(9), 2689-2701.

500 Hall, J. P., Brockhurst, M. A., & Harrison, E. (2017). Sampling the mobile gene pool: innovation
501 via horizontal gene transfer in bacteria. *Philosophical Transactions of the Royal Society B:*
502 *Biological Sciences, 372*(1735), 20160424.

503 Hayakawa, M., Tokuda, M., Kaneko, K., Nakamichi, K., Yamamoto, Y., Kamijo, T., Umeki, H.,
504 Chiba, R., Yamada, R., Mori, M., Yanagiya, K., Moriuchi, R., Yuki, M., Dohra, H.,
505 Futamata, H., Ohkuma, M., Kimbara, K., & Shintani, M. (2022). Hitherto-unnoticed self-
506 transmissible plasmids widely distributed among different environments in Japan. *Applied*
507 *and Environmental Microbiology, 88*(18), e01114-22.

508 Heinemann, J. A., & Sprague Jr, G. F. (1989). Bacterial conjugative plasmids mobilize DNA
509 transfer between bacteria and yeast. *Nature, 340*(6230), 205-209.

510 Hill, K. E., Weightman, A. J., & Fry, J. C. (1992). Isolation and screening of plasmids from the
511 epilithon which mobilize recombinant plasmid pD10. *Applied and Environmental*
512 *Microbiology, 58*(4), 1292-1300.

513 Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2:
514 improving the ultrafast bootstrap approximation. *Molecular biology and evolution, 35*(2),
515 518-522.

516 Hughes, J. M., Lohman, B. K., Deckert, G. E., Nichols, E. P., Settles, M., Abdo, Z., & Top, E.
517 M. (2012). The role of clonal interference in the evolutionary dynamics of plasmid-host
518 adaptation. *MBio, 3*(4), 10-1128.

519 Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., & Jermiin, L. S. (2017).
520 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods,*
521 *14*(6), 587-589.

522 Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7:
523 improvements in performance and usability. *Molecular biology and evolution, 30*(4), 772-
524 780.

525 Klümper, U., Riber, L., Dechesne, A., Sannazzarro, A., Hansen, L. H., Sørensen, S. J., & Smets,
526 B. F. (2015). Broad host range plasmids can invade an unexpectedly diverse fraction of a soil
527 bacterial community. *The ISME journal, 9*(4), 934-945.

528 Kottara, A., Hall, J. P., Harrison, E., & Brockhurst, M. A. (2016). Multi-host environments select
529 for host-generalist conjugative plasmids. *BMC Evolutionary Biology, 16*, 1-4.

530 Król, J. E., Penrod, J. T., McCaslin, H., Rogers, L. M., Yano, H., Stancik, A. D., Dejonghe, W.,
531 Brown, C. J., Parales, R. E., Wuertz, S., & Top, E. M. (2012). Role of IncP-1β plasmids
532 pWDL7:: rfp and pNB8c in chloroaniline catabolism as determined by genomic and
533 functional analyses. *Applied and environmental microbiology, 78*(3), 828-838.

534 Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for
535 phylogenetic tree display and annotation. *Nucleic acids research, 49*(W1), W293-W296.

536 Li, X., Wang, Y., Brown, C. J., Yao, F., Jiang, Y., Top, E. M., & Li, H. (2016). Diversification
537 of broad host range plasmids correlates with the presence of antibiotic resistance genes.
538 *FEMS microbiology ecology, 92*(1), fiv151.

539 Luria, S. E., & Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus
540 resistance. *Genetics, 28*(6), 491.

541 Macartney, D. P., Williams, D. R., Stafford, T., & Thomas, C. M. (1997). Divergence and
542 conservation of the partitioning and global regulation functions in the central control region
543 of the IncP plasmids RK2 and R751. *Microbiology, 143*(7), 2167-2177.

544 Maddison, W. P. (1997). Gene trees in species trees. *Systematic biology, 46*(3), 523-536.

545 Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler,
546 A., & Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic
547 inference in the genomic era. *Molecular biology and evolution, 37*(5), 1530-1534.

548 Mostowy, R., Croucher, N. J., Andam, C. P., Corander, J., Hanage, W. P., & Marttinen, P.
549 (2017). Efficient inference of recent and ancestral recombination within bacterial
550 populations. *Molecular biology and evolution, 34*(5), 1167-1182.

551 Musovic, S., Oregaard, G., Kroer, N., & Sørensen, S. J. (2006). Cultivation-independent
552 examination of horizontal transfer and host range of an IncP-1 plasmid among gram-positive
553 and gram-negative bacteria indigenous to the barley rhizosphere. *Applied and environmental*
554 *microbiology, 72*(10), 6687-6692.

555 Norberg, P., Bergström, M., & Hermansson, M. (2014). Complete nucleotide sequence and
556 analysis of two conjugative broad host range plasmids from a marine microbial biofilm.
557 *PLoS One, 9*(3), e92321.

558 Norberg, P., Bergström, M., Jethava, V., Dubhashi, D., & Hermansson, M. (2011). The IncP-1
559 plasmid backbone adapts to different host bacterial species and evolves through homologous
560 recombination. *Nature communications, 2*(1), 268.

561 Norman, A., Hansen, L. H., & Sørensen, S. J. (2009). Conjugative plasmids: vessels of the
562 communal gene pool. *Philosophical Transactions of the Royal Society B: Biological*
563 *Sciences, 364*(1527), 2275-2289.

564 Novick, R. P. (1987). Plasmid incompatibility. *Microbiological reviews, 51*(4), 381-395.

565 Olsen, R. H., & Shipley, P. (1973). Host range and properties of the Pseudomonas aeruginosa R
566 factor R1822. *Journal of Bacteriology*, *113*(2), 772-780.

567 Pamilo, P., & Nei, M. (1988). Relationships between gene trees and species trees. *Molecular*
568 *biology and evolution, 5*(5), 568-583.

569 Pansegrau, W., Lanka, E., Barth, P. T., Figurski, D. H., Guiney, D. G., Haas, D., Helinski, D. R.,
570 Schwab, H., Stanisich, V. A., & Thomas, C. M. (1994). Complete nucleotide sequence of
571 Birmingham IncPα plasmids: compilation and comparative analysis. *Journal of molecular*
572 *biology, 239*(5), 623-663.

573 Partridge, S. R., Kwong, S. M., Firth, N., & Jensen, S. O. (2018). Mobile genetic elements
574 associated with antimicrobial resistance. *Clinical microbiology reviews, 31*(4), 10-1128.

575 Pilla, G., & Tang, C. M. (2018). Going around in circles: virulence plasmids in enteric
576 pathogens. *Nature Reviews Microbiology*, *16*(8), 484-495.

577 Popowska, M., & Krawczyk-Balska, A. (2013). Broad-host-range IncP-1 plasmids and their
578 resistance potential. *Frontiers in microbiology*, *4*, 44.

579 Rannala, B., & Yang, Z. (2003). Bayes estimation of species divergence times and ancestral
580 population sizes using DNA sequences from multiple loci. *Genetics, 164*(4), 1645-1656.

581 Schlüter, A., Heuer, H., Szczepanowski, R., Forney, L. J., Thomas, C. M., Puhler, A., & Top, E.
582 M. (2003). The 64 508 bp IncP-1 β antibiotic multiresistance plasmid pB10 isolated from a
583 waste-water treatment plant provides evidence for recombination between members of
584 different branches of the IncP-1 β group. *Microbiology, 149*(11), 3139-3153.

585 Schlüter, A., Szczepanowski, R., Pühler, A., & Top, E. M. (2007). Genomics of IncP-1 antibiotic
586 resistance plasmids isolated from wastewater treatment plants provides evidence for a widely
587 accessible drug resistance gene pool. *FEMS microbiology reviews, 31*(4), 449-477.

588 Schmidhauser, T. J., & Helinski, D. R. (1985). Regions of broad-host-range plasmid RK2
589 involved in replication and stable maintenance in nine species of gram-negative bacteria.
590 *Journal of bacteriology, 164*(1), 446-455.

591 Sen, D., Brown, C. J., Top, E. M., & Sullivan, J. (2013). Inferring the evolutionary history of
592 IncP-1 plasmids despite incongruence among backbone gene trees. *Molecular biology and*
593 *evolution, 30*(1), 154-166.

594 Sen, D., Van der Auwera, G. A., Rogers, L. M., Thomas, C. M., Brown, C. J., & Top, E. M.
595 (2011). Broad-host-range plasmids from agricultural soils have IncP-1 backbones with
596 diverse accessory genes. *Applied and environmental microbiology, 77*(22), 7975-7983.

597 Sen, D., Yano, H., Suzuki, H., Król, J. E., Rogers, L., Brown, C. J., & Top, E. M. (2010).
598 Comparative genomics of pAKD4, the prototype IncP-1δ plasmid with a complete backbone.
599 *Plasmid, 63*(2), 98-107.

600 Shintani, M., Takahashi, Y., Yamane, H., & Nojiri, H. (2010). The behavior and significance of
601 degradative plasmids belonging to Inc groups in Pseudomonas within natural environments
602 and microcosms. *Microbes and environments, 25*(4), 253-265.

603 Sota, M., Tsuda, M., Yano, H., Suzuki, H., Forney, L. J., & Top, E. M. (2007). Region-specific
604 insertion of transposons in combination with selection for high plasmid transferability and
605 stability accounts for the structural similarity of IncP-1 plasmids. *Journal of bacteriology,*
606 *189*(8), 3091-3098.

607 Sota, M., Yano, H., M Hughes, J., Daughdrill, G. W., Abdo, Z., Forney, L. J., & Top, E. M.
608 (2010). Shifts in the host range of a promiscuous plasmid through parallel evolution of its
609 replication initiation protein. *The ISME journal, 4*(12), 1568-1580.

610 Soucy, S. M., Huang, J., & Gogarten, J. P. (2015). Horizontal gene transfer: building the web of
611 life. Nature Reviews *Genetics, 16*(8), 472-482.

612 Stalder, T., Rogers, L. M., Renfrow, C., Yano, H., Smith, Z., & Top, E. M. (2017). Emerging
613 patterns of plasmid-host coevolution that stabilize antibiotic resistance. *Scientific Reports,*
614 *7*(1), 4853.

615 Swenson, N. G. (2009). Phylogenetic resolution and quantifying the phylogenetic diversity and
616 dispersion of communities. *PloS one, 4*(2), e4390.

617 Thomas, C. M., & Smith, C. A. (1987). Incompatibility group P plasmids: genetics, evolution,
618 and use in genetic manipulation. *Annual Reviews in Microbiology*, *41*(1), 77-101.

619 Thorsted, P. B., Shah, D. S., Macartney, D., Kostelidou, K., & Thomas, C. M. (1996).
620 Conservation of the genetic switch between replication and transfer genes of IncP plasmids
621 but divergence of the replication functions which are major host-range determinants.
622 *Plasmid, 36*(2), 95-111.

623 Vedler, E., Vahter, M., & Heinaru, A. (2004). The completely sequenced plasmid pEST4011
624 contains a novel IncP1 backbone and a catabolic transposon harboring tfd genes for 2, 4-
625 dichlorophenoxyacetic acid degradation. *Journal of bacteriology, 186*(21), 7161-7174.

626 Yakimov, M. M., Crisafi, F., Messina, E., Smedile, F., Lopatina, A., Denaro, R., Pieper, D. H.,
627 Galyshin, P. N., & Giuliano, L. (2016). Analysis of defence systems and a conjugative IncP-
628 1 plasmid in the marine polyaromatic hydrocarbons-degrading bacterium *Cycloclasticus sp*.
629 78-ME. *Environmental Microbiology Reports, 8*(4), 508-519.

630 Yano, H., Deckert, G. E., Rogers, L. M., & Top, E. M. (2012). Roles of long and short
631 replication initiation proteins in the fate of IncP-1 plasmids. *Journal of bacteriology, 194*(6),
632 1533-1543.

633 Yano, H., Wegrzyn, K., Loftie-Eaton, W., Johnson, J., Deckert, G. E., Rogers, L. M., Konieczny,
634 I. K., & Top, E. M. (2016). Evolved plasmid-host interactions reduce plasmid interference
635 cost. *Molecular microbiology, 101*(5), 743-756.

636 Yano, H., Rogers, L. M., Knox, M. G., Heuer, H., Smalla, K., Brown, C. J., & Top, E. M.
637 (2013). Host range diversification within the IncP-1 plasmid group. *Microbiology,*
638 *159*(Pt_11), 2303-2315.

639 Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species
640 tree reconstruction from partially resolved gene trees. *BMC bioinformatics, 19*, 15-30.