

1 **Global evolutionary epidemiology, phylogeography and resistome dynamics of**
2 *Citrobacter species, Enterobacter hormaechei, Klebsiella variicola, and Proteae clones: A*
3 *One Health analyses*

4 John Osei Sekyere# and Melese Abate Reta

5 Department of Medical Microbiology, School of Medicine, Faculty of Health Sciences, University of
6 Pretoria, 0084 Prinshof, Pretoria, South Africa

7 #Address correspondence to Dr. John Osei Sekyere, Department of Medical Microbiology, School of
8 Medicine, University of Pretoria, 0084 Prinshof, Pretoria, South Africa. jod14139@gmail.com;
9 u18392467@tuks.co.za

10 **Tweet:** Global *E. hormaechei*, *C. freundii*, *P. mirabilis*, *P. stuartii* & *P. rettgeri* strains
11 contain multiple resistance mechanism to important and reserved antibiotics in globally
12 circulating clones. These portend the dawn of pandrug resistance and a return to the pre-
13 antibiotic era.

14 **Running head:** Global phylogenomics of resistant Enterobacteriaceae

15 **Highlights/Importance**

16 *Citrobacter spp., Enterobacter hormaechei subsp., Klebsiella variicola* and *Proteae* tribe
17 members are rarely isolated Enterobacteriaceae increasingly implicated in nosocomial
18 infections. The global phylogenomics, evolution, geographical distribution and resistome
19 repertoire of these species found them to be globally distributed, being isolated from human,
20 animal, plant, and environmental sources. *E. hormaechei subsp., C. freundii, Proteus*
21 *mirabilis, Providencia stuartii, Providencia rettgeri* and *Morganella morganii* contained
22 multidrug-resistant clades that harboured resistance to clinically important reserved
23 antibiotics, portending the dawn of pandrug resistance and a potential acceleration of the
24 post-antibiotic era.

25 **Abstract**

26 **Background.** The global epidemiology and resistomes dynamics of multidrug-resistant
27 *Citrobacter spp.*, *Enterobacter hormaechei*, *Klebsiella variicola*, *morganella morganii*,
28 *Proteus mirabilis* and *Providencia spp.* have not been described, despite their importance as
29 emerging opportunistic clinical pathogens.

30 **Methods.** The genomes of the above-mentioned organisms were curated from PATRIC and
31 NCBI and used for evolutionary epidemiology, phylogeography and resistome analyses. The
32 phylogeny trees were drawn using RAXmL and edited with Figtree. The resistomes were
33 curated from GenBank and the phylogeography was manually mapped.

34 **Results and conclusion.** *Mcr-9* and other *mcr* variants were highly prevalent in *E.*
35 *hormaechei subsp.* and substantial in *C. freundii* whilst KPC, OXA-48, NDM, IMP, VIM,
36 TEM, OXA and SHV were abundant in global *E. hormaechei subsp.*, *Citrobacter freundii*, *P.*
37 *mirabilis*, *P. stuartii* and *P. rettgeri* clones/clades. Species-specific ampCs were highly
38 conserved in respective species whilst fluoroquinolones, aminoglycosides, macrolides,
39 fosfomicin, chloramphenicol, tetracycline, sulphamethoxazole and trimethoprim resistance
40 mechanisms were abundantly enriched in almost all clades of most of the species, making
41 them extensively and pandrug resistant; *K. variicola*, *C. amalonaticus* and *C. koseri* had
42 relatively few resistance genes. Vertical and horizontal resistome transmissions as well as
43 local and international dissemination of strains evolving from common ancestors were
44 observed, suggesting the anthroponotic, zoonotic, and food-/water-borne infectiousness of
45 these pathogens. There is a global risk of pandrug resistant strains escalating local and
46 international outbreaks of antibiotic-insensitive infections, initiating the dawn of a post-
47 antibiotic era.

48 **Keywords:** Resistome; epidemiology; *Citrobacter*; *Klebsiella variicola*; *Morganella*;
49 *Proteus*; *Providencia*; *Enterobacter hormaechei*

50 **Introduction**

51 Antibiotic resistance is mainly disseminated via horizontal and vertical transmission through
52 mobile genetic elements such as plasmids and transposons and through clonal and multiclonal
53 expansion of same species¹⁻⁶. Conjugative plasmids have been implicated in the transmission
54 of several resistance determinants within and across species, resulting in the presence of same
55 or very similar resistomes in same and different species and clones^{1,4,7-9}. Thus, the
56 emergence of plasmid-borne resistance genes is always a cause for concern as they help
57 breach the species barrier and shuttle resistance genes (ARGs) from commensals and non-
58 pathogenic bacteria to pathogenic ones or vice versa^{3,4,10,11}. Such has been the case with the
59 emergence and rapid spread of extended-spectrum β -lactamases (ESBLs) viz., TEM, SHV,
60 OXA and CTX-M, carbapenemases such as NDM, IMP, VIM, KPC and GES, the mobile
61 colistin resistance gene *mcr-1* (to *mcr-10*) and recently, the mobile tigecycline resistance
62 gene, *tet(X)*¹²⁻¹⁷. Thus, such conjugative plasmids influence the genomic plasticity of several
63 related and unrelated species and genera of bacteria^{8,11,18-21}.

64 Coupled with plasmid-borne dissemination of ARGs is the selection and expansion of
65 specific drug-resistant clones^{11,21}, which quickly spread under antibiotic pressure to
66 overpopulate their environments, facilitating their survival and subsequent spread to other
67 environments^{14,22,23}. In cases where such clones harbour resistance plasmids, their expansion
68 almost always lead to the concomitant replication and intra-clonal as well as inter-clonal
69 spread of such plasmids^{4,5,7,8,24,25}. Thus, as such clones are disseminated through contact,
70 food, water, farms, hospitals, and the environment, they carry with them these resistance
71 plasmids to colonize new hosts and environments^{6,26-28}. It is thus not surprising to have same

72 clones hosting the same plasmids, contain the same resistomes^{4-6,21,28}. This explains the
73 presence of multi-drug resistance (MDR) in particular international clones such as *Klebsiella*
74 *pneumoniae* ST258 and *E. coli* ST113^{21,25,29,30}.

75 Hence, tracing the phylogeography of clones and their associated resistance genes is highly
76 critical in epidemiology and public health as it provides necessary data to contain the further
77 spread of ARGs^{26,28,31}. In this work, the global evolutionary epidemiology and resistome
78 dynamics of clinically important but relatively less isolated Enterobacteriaceae pathogens are
79 described⁸. It is notable that most of the recently emerged or novel resistance genes in
80 bacteria have occurred in Enterobacteriaceae more than in any other family of bacteria,
81 making Enterobacteriaceae particularly important medically^{28,32,33}. These include ESBL-,
82 carbapenemase-, *mcr*- and *tet(X)*-producing producing Enterobacteriaceae, which have been
83 classified by the WHO as high and critical priority pathogens due to their implication in high
84 mortalities and morbidities^{32,33}. Although *Citrobacter spp.*, *Enterobacter hormaechei*,
85 *Klebsiella variicola*, *Morganella morganii*, *Proteus spp.* and *Providencia spp.* are not mostly
86 reported as *Escherichia coli*, *Klebsiella pneumoniae*, and *Salmonella enterica*, they have been
87 associated with multiple resistance and clinical fatalities^{3,7,14,25,28,34}. Due to the transferability
88 of resistance plasmids between members of the Enterobacteriaceae, the global resistome
89 epidemiology of these six genera is important as they could be eventually transferred to
90 commonly isolated Enterobacteriaceae species^{11,14,21,25,35,36}.

91 **Results**

92 ***Included genomes***

93 A total of 2,377 genomes from *C. freundii* (n=569 genomes), *C. koseri* (n=82 genomes), *C.*
94 *amalonaticus* (n=35 genomes), *E. steigerwaltii* and *E. oharae* (n=121 genomes), *E.*
95 *xiangfangensis* (n= 90 genomes), *E. hormaechei* (n=563 genomes), *K. variicola* (n=574

96 genomes), *M. morganii* (n=59 genomes), *P. mirabilis* (n= 156 genomes), and *Providencia*
97 *spp.* (n=128 genomes) were obtained from PATRIC and NCBI databases as at January 2020
98 and used for downstream analyses (Tables S1-S3). Included in *C. freundii*, *C. koseri*, *E.*
99 *xiangfangensis*, *K. variicola* and *Providencia spp.* genomes were genomes of other Gram-
100 negative bacterial species or genera that were initially classified within these respective
101 species but later reclassified by NCBI's ANI (average nucleotide identity) analysis. The
102 genomes of these reclassified species were however maintained and included in the resistome
103 analyses to serve as controls for comparison (Tables S1-S3).

104 The genomes were mainly isolated from human specimens, followed by animal (including
105 food animals), plants (including food crops) and environmental specimens. The human and
106 animal specimens used included urine, blood, stool, catheter tip, swabs etc. whilst the
107 environmental specimens used included soils, hospital environments, water, wastewater,
108 sinks etc. (Tables S1-S3). In all, these genomes were obtained from 67 countries globally,
109 with the USA having the most genomes for all species: *C. freundii/spp.* (USA=233,
110 China=34, France=14, Spain=11); *C. amalonaticus* (USA=17); *C. koseri* (USA=55); *E.*
111 *steigerwaltii/oharae* (USA=40, Japan=15); *E. hormaechei* (USA=142, China=78, Japan=72,
112 France=13, UK=10); *E. xiangfangensis* (USA=50, China=14, India=12); *K. variicola*
113 (USA=289, Germany=50, China=24, Bangladesh=23); *M. morganii* (USA=16); *P. mirabilis*
114 (USA=38, France=27, China=17); *Providencia spp.* (USA=61).

115 *C. freundii*, the only species among the species included in this analysis to have an MLST
116 scheme, had 84 different clones or sequence types (STs). ST100 (n=51), ST22 (n=51), ST62
117 (n=18), ST11 (n=14), ST299 (n=11), ST8 (n=10), ST114 (n=8), and ST98 (n=8) were the
118 commonest clones.

119 ***Species epidemiology***

120 *Citrobacter species*

121 Amongst the *C. freundii* genomes were other *Citrobacter spp.* such as *C. werkmanii*, *C.*
122 *youngae*, *C. brakii*, *C. portucalensis* etc. *C. werkmanii* were isolated from humans, sprouts,
123 and sinks whilst *C. brakii* were obtained from humans, vegetables, and hospital
124 environments. Some *C. werkmanii* and *C. brakii* strains were found in USA, Germany, and
125 India (Fig. S1A-B). *C. koseri* were mainly found from humans and from a mouse (clade B2)
126 (Fig. S1A-B). A *Citrobacter spp.* (clade A) outbreak was observed in the US (Texas, Boston)
127 and Mexico (Fig. S1A-B). *C. portucalensis* of the same clade (but different clones) were
128 found in Nigeria (vegetables) and Brazil (turtle) as well as from effluents (UK), humans
129 (China), chives, carrots and salad (Germany); these clustered in *C. freundii* clade B1. *C.*
130 *freundii* strains of different STs and countries clustered together into clades, showing the
131 wide distribution of strains from the same ancestor (or of close evolutionary distance); they
132 were isolated from humans, plants, animals, and the environment (Fig. 1 & S1A-B).

133 *C. werkmanii*, *C. brakii* and *C. youngae* clustered within *C. freundii* clade A whilst *C.*
134 *portucalensis* and *C. youngae*, were clustered in clade B; *C. koseri* clustered in clade B3 of *C.*
135 *freundii* (Fig. 1 & S1A-B). *C. freundii* clades B2 and B3 had a richer resistome than clades A
136 and B1, although *bla*_{CMY}, *qnr*, *sul1/2*, *aac*(6'/3')-like ARGs were common in all the clades;
137 *mph*(A/E), *catAB*, *dfrAB*, *aadA* and *bla*_{TEM} were common in clades B2 and B3. Important
138 ARGs such as *mcr*, *bla*_{KPC}, *bla*_{CTX-M}, *bla*_{NDM}, *bla*_{VIM}, and *bla*_{OXA-48} were relatively rare and
139 mainly found in clades B2 and B3 than in A and B1. Comparatively, other Enterobacteriaceae
140 species (*E. coli*, *K. pneumoniae*, and *S. marcescens*) had strains with richer resistomes
141 diversity than *C. freundii* clades A and B1, including chromosomal mutations and MDR
142 efflux pumps. However, Clades B2 and B3 had comparable resistome diversity and
143 abundance to the above-mentioned Enterobacteriaceae species, in which CMY was well-nigh
144 absent except in *E. coli* (Fig. 1).

145 A local outbreak of *C. braakii* was observed in the UK (in humans), with closely related
146 strains being isolated from biosolids (Canada), carrots (Germany), chicken (China) and beef
147 (Canada) (Fig. S1A-B). *C. amalonaticus* was also found in humans, animals, plants and the
148 environment (Fig. 2): *bla*_{CMY} was not common in this species, but *bla*_{SED} and *Oqx*A/B were
149 almost conserved in clades B1 and C, with clade B2 (all from France) having the richest
150 resistome repertoire that included *bla*_{TEM}, *bla*_{SHV}, *bla*_{NDM} and *mcr*-9.1 (Fig. 2). *C. koseri* had
151 a relatively limited resistome diversity, with *bla*_{MAL-1/2} (clade B2), *bla*_{CKO} (clade B3) and
152 *fos*A7 (clade B2) being the commonest ARGs. Comparatively, the other Enterobacteriaceae
153 species (e.g. *E. coli*, *K. pneumoniae*, *Enterobacter spp.*, *Serratia spp.*, and *Providencia spp.*)
154 had richer resistome diversity than *C. koseri* (Fig. 3).

155 *Enterobacter species*

156 The resistomes of *E. steigerwaltii* and *E. oharae* were extraordinarily rich and diverse,
157 although clade A of *E. steigerwaltii* had lesser resistome abundance and diversity compared
158 to clades B and C (Fig. 4). *E. oharae* only clustered in clade B of *E. steigerwaltii* and had
159 most *bla*_{CTX-M}, *bla*_{SHV}, *bla*_{NDM}, *bla*_{OXA-48} and *bla*_{VIM} ARGs than *E. steigerwaltii* clade C.
160 *bla*_{ACT} was present in almost all *Enterobacter spp.* strains whilst other AmpCs such as *bla*_{LAP},
161 *bla*_{SCO-1}, *bla*_{SFO}, *bla*_{TMB}, *bla*_{DHA}, and *bla*_{CARB} were virtually absent. *E. cloacae* strains
162 clustered closely with *E. steigerwaltii* clades B and C. overall, *E. steigerwaltii* and *E. oharae*
163 strains were richly endowed with clinically important ARGs including carbapenemases,
164 ESBLs, *aac*(6')-like, *aac*(3')-like, *aph*(3'/3'')-like, *aadA*, *dfrA*, *catA*, *fosA*, *oqxAB*,
165 *qnrA/B/S/D*, *sul1/2*, and *tet*(A/B/D). *Mcr*-9 genes were particularly abundant in clade B than
166 in clades A and C of *E. steigerwaltii/cloacae* (Fig. 4).

167 *E. cloacae* strains clustered with all *E. xiangfangensis* strains except clade A1, from which it
168 was evolutionarily distant (Fig. 5). ARGs such as *bla*_{ACT}, *catA/B*, *Oqx*A/B and *fosA* were

169 universally present in almost all the *E. xiangfangensis* and *E. cloacae* strains. Furthermore, *E.*
170 *xiangfangensis* clade A1 had the least ARGs diversity and abundance whilst *E. cloacae* A1
171 and the remaining *E. xiangfangensis* (clades A2 and B) were richly endowed with multiple
172 ARGs such as *bla*_{KPC}, *mcr-9*, *bla*_{CTX-M}, *bla*_{OXA}, *bla*_{NDM}, *bla*_{TEM}, and genes mediating
173 resistance to fluoroquinolones (*aac*(6'/3')-like, *qnrA/B/S* etc.) and aminoglycosides (*armA*,
174 *rmtC/G*, *aadA*, *aph*(4)-, *aph*(3'/3'')- and *aph*-(6')-like) as well as *tet* (A/B/D), *dfrA*, and *arr*
175 (particularly in clades A2 and B). Thus, the resistomes diversity and abundance of *E.*
176 *xiangfangensis* and *E. cloacae* were comparable to that of *E. oharae* and *E. steigerwaltii* (Fig.
177 4-5).

178 An *E. hormaechei* outbreak was observed in Germany in 2017, evolving with the spread
179 (clade B) (Fig. 6 & S2). *E. steigerwaltii* and *oharae* were mainly isolated from humans whilst
180 *E. xiangfangensis* and *E. hormaechei* were from humans, plants (*E. xiangfangensis* from rice
181 in India), animals and the environment. Strains of closely related (i.e. close evolutionary
182 distance) *E. hormaechei* in clade B were from the US, South Africa, Colombia, China,
183 Germany, Australia and Lebanon from humans, animals, and plants (Fig. 4-6). Other closely
184 related *E. hormaechei* strains from humans, animals, plants, and the environment were found
185 in different countries, showing a gradual evolution of strains emanating from a common
186 ancestor and spreading across countries through different hosts. These observations were
187 made in *E. hormaechei* clades B and C, and represented local and international outbreaks
188 spanning UK, USA, China, Serbia, Japan etc. (Fig. 6 & S2).

189 Moreover, *bla*_{ACT}, *fosA*, *OqxA/B* and *catA/B* were almost conserved in almost all the clades
190 of *E. hormaechei*. As observed with the other *E. hormaechei* subsp. (*oharae*, *steigerwaltii*
191 and *xiangfangensis*), *E. hormaechei* clades B and C were richly endowed with ARGs than
192 some clade A strains (except clade A in Fig. S2A). Specifically, clades B and C strains were
193 relatively enriched with *bla*_{TEM-1}, *bla*_{OXA}, *bla*_{CTX-M}, *aph*(3'/3'')-like, *aac*(3')-like, *aac*(6')-like,

194 *dfrA/B*, *mcr-9*, *arr*, *qnrA/B/S*, *sul-1/2*, and *tet(A/B/C/D)*. Further, clade B was mostly rich in
195 *bla_{KPC}* and *bla_{NDM}* whilst clade C was rich in *bla_{NDM}* and *ble*. A substantial number of clade
196 A strains also had *mcr-9* and *bla_{TEM}* (Fig. 6 & S2).

197 *E. oharae* was evolutionarily closer to *E. xiangfangensis*, clustering on the same branches,
198 with a relatively few *E. oharae* strains clustering with *E. steigerwaltii*. However, *E.*
199 *steigerwaltii* was mostly distant from *E. oharae* and *E. xiangfangensis*, with a few *E.*
200 *xiangfangensis* strains (from rice in India) clustering within *E. steigerwaltii* clades A and C
201 (Fig. 2SB). Figure S2B summarises the ARGs in *E. hormaechei* and its subspecies: *bla_{ACT}*,
202 *fosA*, *OqxAB*, *qnrA/B/D/S*, and *catA/B* were conserved whilst *bla_{OXA}*, *ble*, *bla_{SHV}*, *bla_{TEM}*,
203 *bla_{KPC}*, *bla_{NDM}*, *bla_{VIM}*, *mcr*, *tet*, *mph(A)*, *sul-1/2*, *dfrA*, and fluoroquinolone and
204 aminoglycoside ARGs were richly abundant, particularly in clusters IV, V and VII.

205 *K. variicola*

206 *K. variicola* strains were isolated mainly from humans, with some being from animals, plants,
207 and the environment. On the individual branches/clades were closely related strains with very
208 close evolutionary distance but disseminated across countries, suggesting international
209 dissemination of that clade and showing little evolution during the spread from host to host,
210 e.g., clades A3, A4, A5, A6, B1, B2 and C (Fig. 7 & S3). As well, local outbreaks of closely
211 related strains in the USA, Bangladesh, Canada, and Germany were observed. A
212 reassessment of the *K. variicola* tree with genomes of *K. pneumoniae* and *K.*
213 *quasipneumoniae* largely confirmed the initial clustering of the *K. variicola* strains, with only
214 a few rearrangements of some strains within different clades (Fig. S3A and S3B).

215 Notably, *K. variicola* had fewer resistome diversity and abundance than *Citrobacter spp.*,
216 *Enterobacter spp.*, and *K. pneumoniae spp.* Conserved within the *K. variicola* genomes were
217 *emrD*, *fosA*, *OqxAB* and *bla_{LEN-2}*, whilst the other ARGs were sparse. Whilst *bla_{SHV}* was

218 conserved in *K. pneumoniae*, it was virtually absent in *K. variicola*; further, *bla*_{LEN} was
219 present in the latter but absent in the former. *K. variicola* clade B2 strains (on branch VIII),
220 specifically those from Bangladesh, had richer and more diverse resistomes than the other *K.*
221 *variicola* clades/clusters. Whilst *mcr* and *bla*_{NDM} were virtually absent in these genomes,
222 *bla*_{KPC} occurred in substantial abundance. These suggest that *K. variicola* is least likely to be
223 MDR and/or harbour ESBLs, carbapenemases, *mcr* and other clinically important ARGs
224 compared to other Enterobacteriaceae species (Fig. 7 & S3).

225 *M. morgani*

226 There were three *M. morgani* clusters/clades viz., A, B and C, which were mainly from
227 humans with a few in clades A and C being from animals (Fig. 8). Within each clade are
228 closely related strains with very close evolutionary distance that were distributed across
229 several countries; indeed, strains of the same clone were found in different countries,
230 suggesting international dissemination of the same clones. The branching order of the trees
231 within each clade shows the gradual evolution of the strains as they moved from host to host.
232 *bla*_{DHA} and *catA/B* genes were almost conserved in almost all the *M. morgani* genomes.
233 Other highly abundant ARGs in the *M. morgani* strains were *sul-1/-2*, *tet(A/B/D/Y)*, *bla*_{OXA-}
234 *1*, *ble*, *dfrA1*, *aadA*, *aac(3)*-like and *aph(3'/3'')*-like. Notably, clade B had more ARG
235 diversity and abundance than clade C and A; clade A had the least diversity and abundance of
236 ARGs (Fig. 8).

237 *P. mirabilis*

238 *P. mirabilis* clustered into three major clades and included isolates from humans, animals,
239 and the environment. Within the three clades were sub-clades consisting of closely related
240 strains from the same as well as different countries, showing local and international outbreaks
241 involving human and animal hosts, and in some cases environmental mediators (Fig. 9).

242 Specifically, local outbreaks were seen in the USA (clades A3, B2 and B3), France (clade
243 C2) and Japan (C3) whilst international dissemination of clades A2, A3, A4, B, C2 and C3
244 were observed. Clade C2 had the richest abundance of resistomes, with all the members
245 having a uniform/conserved diversity of the same ARGs, except for *bla*_{CTX-M}, *bla*_{OXA}, *bla*_{TEM},
246 and *Inu*(F/G). Clade C3 had the 2nd most abundant but more diverse ARGs than C2. Notably,
247 clade B and its subclades had lesser ARGs than clades A and C. *CatA* and *tet*(A/B/D/Y)
248 ARGs were virtually conserved in all the clades whilst *aadA*, *aac*(3')-like, *aph*(3'/3'')-like,
249 *aph*(6')-like, *dfrA7*, *sat2*, and *sul-1/2/3* were substantially prevalent in all the clades. In
250 particular, *bla*_{CARB2} and *florR* were highly conserved in clade C2; *florR* was however less
251 abundant than *bla*_{CARB}. Notably, *mcr* was almost absent except in a few strains in B3. As
252 well, *bla*_{CTX-M}, *bla*_{CMY}, *bla*_{NDM}, *bla*_{OXA-1}, *bla*_{TEM} and *ble-O/Sh* were mainly found in clades A
253 and C3, with traces in B and C1. Thus, ESBLs, carbapenemases and *mcr* genes were
254 relatively less common in *P. mirabilis* strains, compared to other Enterobacteriaceae (Fig. 9).

255 *Providencia species*

256 There were five *Providencia spp.* viz., *stuartii*, *rustigianii*, *alcalifaciens*, *heimbachae*, and
257 *rettgeri*, with *P. rettgeri* being most isolated and branching into two clades, A and B (Fig.
258 10). The other species had single clades. *Providencia spp.* were isolated from animals,
259 humans, and the environment. Highly similar strains of the same species were found across
260 countries (*P. stuartii*, *P. alcalifaciens*, *P. heimbachae* and *P. rettgeri*) and within countries (*P.*
261 *rettgeri* clade A). The distinction between the various species of *Providencia* was depicted by
262 the clustering patterns on the tree as strains of the same species clustered together;
263 *Providencia spp.* was closest evolutionarily to *S. marcescens* whilst *E. coli* was closest to
264 *Citrobacter spp.*, and *E. hormaechei*. As well, *K. pneumoniae*, *K. aerogenes*, and *K.*
265 *michiganensis* clustered together with relatively short evolutionary distance (Fig. 10).

266 The other Enterobacteriaceae species had more ARGs than *Providencia spp.* Within
267 *Providencia spp.*, *P. stuartii* and *P. rettgeri* were most endowed with richer and more diverse
268 resistomes whilst *P. alcalifaciens* and *P. rustigianii* were almost bereft of ARGs. Common
269 ARGs within *P. stuartii*, were *aac(2')-Ia*, *aph(3')*-like, *bla_{OXA}*, *bla_{TEM}*, *aadA1*, *dfrA/B*, *sul1*,
270 *fosA*, *ampC*, *tet* and *catA/B* whilst *P. rettgeri* had *aac(6')*-like, *aph(3'/3'')*-like, *bla_{SRT}*,
271 *bla_{OXA}*, *bla_{NDM}*, *aadA*, *dfrA/B*, *OqxAB*, *sul1*, *qnrA/B/D/E/S*, *ampC*, *catA/B*, *tet*, *arr*, *mph(A/E)*,
272 *cmlA*, *msr(E)*, and *Inu(F/G)*. Hence, *P. rettgeri* has the most abundant and diverse ARGs than
273 all other *Providencia spp.* (Fig. 10).

274 ***Phylogeography***

275 North America (particularly USA) and Europe (particularly Western and Southern Europe)
276 had the highest concentration of the various species, followed by South East Asia, South
277 America (particularly Brazil and Colombia) and South Africa. There were sparse reports on
278 these species from Australasia, the Middle East and Africa (except South Africa) and the
279 Caribbean (Fig. 11).

280 *C. freundii* clade A was distributed mainly in North America, Europe, and South-East Asia
281 whilst clade B was found worldwide on almost all continents. *C. amalonaticus* clade A was
282 found in North America and South Korea; clade B was found in Malawi, USA, Malaysia, and
283 France whilst clade C was only found in Malaysia, USA and Switzerland. *C. koseri* clade A
284 (USA) and clade B (USA, Spain, UK, France, Canada, China and Malaysia) were relatively
285 less reported with clade B being more widely distributed globally (Fig. 11).

286 *E. oharae* strains and *E. hormaechei* clades A, B and C were globally disseminated, with
287 clade C being most widely distributed, followed by clades B and A. *E. steigerwaltii* clade C
288 was more globally distributed than clades A (USA, Japan and Europe) and B (Argentina,
289 USA, France and Germany). Although *E. xiangfangensis* was globally disseminated, the

290 respective clades A (China, India, USA, S. Africa, and France) and B (France, USA, Egypt,
291 and China) were reported from very few countries (Fig. 11).

292 *K. variicola* strains, particularly clades B and C, were of wide geographical distribution;
293 clade A was found in relatively fewer countries globally. *M. morgani* strains were found in
294 North America, Europe (including Russia), and South-East Asia, with clades B and C being
295 found in South Africa. *P. mirabilis* clades A, B and C were found globally, with clade C
296 being reported in most countries. *P. rettgeri* strains had the widest global distribution and
297 reports from most countries among the *Providencia spp.*, followed by *P. stuartii*. *P. rettgeri*
298 clade A (China, Brazil, South Africa, Colombia, and USA) was found in fewer countries than
299 clade B. *P. stuartii* strains were of global distribution whilst *P. rustigianii* (UK and USA), *P.*
300 *alcalifaciens* (USA and India) and *P. heimbachae* (France, China, and Germany) were not
301 (Fig. 11).

302 **Frequency distribution of ARGs per species**

303 *MCR ARGs*

304 *Mcr-9.1* ARGs were the commonest *mcr* variants identified, with very few *mcr-1* and *mcr-3*
305 being found in *E. hormaechei* and *C. freundii* (*mcr-1*, -3 and -10) and a single *mcr-4.3* gene
306 being found in *P. rettgeri*. Notably, the highest prevalence of *mcr-9* was in *E. hormaechei*
307 (n=67 *mcr-9* genes), *E. steigerwaltii/oharae* (n=32 *mcr-9* genes) and *E.*
308 *xiangfangensis/cloacae* (n=19 *mcr-9* genes), followed by *C. freundii* (n=19), *C. amalonaticus*
309 (n=5) and other *Citrobacter spp.*, some of which had very few or no *mcr* genes (Fig. S4-S9).

310 *Carbapenemases*

311 One of the most prevalent carbapenemase among the species was KPC, with KPC-2 (n=123),
312 KPC-3 (n=97), KPC-4 (n=14), and KPC-6 (n=1) being common. KPC-2 was higher in all the
313 species except in *E. xiangfangensis* for which KPC-3 was more abundant (n=35) than KPC-2

314 (n=9). KPC was most prevalent in *E. hormaechei/oharae/steigerwaltii/xiangfangensis* and *C.*
315 *freundii* strains whilst *C. amalonaticus*, *C. koseri*, and *Providencia spp.* had no KPC ARGs.
316 After Ambler class A KPC, Ambler class D OXA-48-like serine carbapenemases (n=65)
317 were also very prominent in all species except *C. amalonaticus*, *E. xiangfangensis* and *K.*
318 *variicola*, with other OXA-48 variants such as OXA-181 (n=1, *M. morgani*) and OXA-396
319 (n=1, *P. rettgeri*) being relatively scarce. *C. freundii* (n=43), *E. hormaechei* (n=10), *E.*
320 *oharae/steigerwaltii* (n=9) and *C. koseri* (n=3) had OXA-48 genes whilst OXA-58 (n=2) and
321 OXA-23 (n=23) were only found in *P. mirabilis*. Other class A serine carbapenemases i.e.,
322 GES-5 (n=1, *M. morgani*) and IMI-2 (n=1, *E. hormaechei*), were also rare (Fig. S4-S9).
323 NDM was the commonest class B carbapenemase (n=159), followed by IMP (n=97) and
324 VIM (n=83). NDM-1 (n=137) was the most prevalent variant and was found in *E.*
325 *hormaechei* (n=68), *C. freundii* (n=15), *E. steigerwaltii* (n=13), *Providencia spp.* (n=12), *P.*
326 *mirabilis* (n=10), *E. xiangfangensis* (n=6), *C. amalonaticus* (n=5), *M. morgani* (n=5), and *K.*
327 *variicola* (n=3), with NDM-5 (n=16; 12 in *E. hormaechei*, 3 in *E. xiangfangensis*, and 1 in *P.*
328 *mirabilis*), NDM-7 (n=3 in *E. hormaechei*) and NDM-9 (n=3 in *K. variicola*) being less
329 prevalent. IMP-1 (n=71; 67 in *E. hormaechei* and 4 in *C. freundii*), IMP-8 (n=14; 13 in *C.*
330 *freundii*, 1 in *E. steigerwaltii*), IMP-4 (n=9; 5 in *C. freundii*, 4 in *E. hormaechei*), IMP-27
331 (n=2 in *P. mirabilis*) and IMP-13 (n=1 in *E. oharae*) were the identified variants. VIM-1
332 (n=65; 26 in *C. freundii*, 14 in *E. steigerwaltii*, 21 in *E. hormaechei*, 1 in *E. xiangfangensis*, 2
333 in *Providencia spp.*, 1 in *P. mirabilis*), VIM-4 (n=12; 5 in *E. steigerwaltii*, 5 in *E.*
334 *hormaechei*, 2 in *C. freundii*), VIM-2 (n=2 in *Providencia spp.*), VIM-31 (n=2; *E.*
335 *steigerwaltii* and *E. hormaechei*), VIM-5 (n=1 in *E. hormaechei*), and VIM-67 (n=1 in *E.*
336 *hormaechei*) were identified in the strains (Fig. S4-S9).
337 *ESBLs and ampCs*

338 TEM was the commonest ESBL and TEM-1 was the most common variant to be identified in
339 all species; particularly, TEM-1 was most abundant in *Enterobacter spp.* and *C. freundii*.
340 OXA (particularly OXA-1, -9, and -10), CTX-M (particularly CTX-M-15) and SHV
341 (particularly SHV-12) were also common in almost all species in a descending order of
342 prevalence, but were very abundant in *Enterobacter spp.*, specifically *E. hormaechei*; *C.*
343 *koseri*, *C. amalonaticus*, *P. mirabilis*, and *M. morgani* had relatively low abundance of SHV,
344 OXA and CTX-M genes. Other ESBLs genes such as *bla_{SCO}*, *bla_{LAP}*, *bla_{VEB}*, *bla_{TMB}*, *bla_{SFO}*,
345 *bla_{SMB}*, *bla_{CARB}*, *bla* etc. were also rare in the various species (Fig. S4-S9).

346 AmpC ARGs were basically strain-specific, with ACT, LEN, CMY, DHA, CKO/MAL and
347 SED being conserved in *Enterobacter spp.*, *K. variicola*, *C. freundii*, *M. morgani*, *C.koseri*
348 and *C. amalonaticus* (except clade A) respectively; FOX was rare (Fig. S4-S9).

349 *Aminoglycoside ARGs*

350 ARGs mediating resistance to aminoglycosides such as *aadA*, *aph(2'')*-like, *aph(3')*-like,
351 *aph(4')*-like, *aac(6')*-like, *aac(3)*-like, *aph(6)*-like, and *ant(2'')*-like were abundantly
352 prevalent in almost all the clades of *C. freundii*, *Enterobacter spp.*, and *P. mirabilis* and
353 sparsely abundant in the other species; *aadA*, *aph(4')*-like, *aac(6')*-like and *aac(3')*-like
354 ARGs were most common. 16S rRNA Methyltransferases such as *rmtB1* (*E. steigerwaltii* and
355 *P. mirabilis*), *rmtC/G* (*E. hormaechei*, *C. freundii* and *M. morgani*), and *armA* (*C. freundii*,
356 *E. xiangfangensis*, *P. mirabilis*, *P. stuartii*, *M. morgani* and *P. rettgeri*) were rare (Fig. S4-
357 S9).

358 *Fluoroquinolone ARGs*

359 *Aac(6')*-like, *OqxAB*, *QnrA/B/D/S* and *qepA* ARGs were identified, albeit *qepA* (*M.*
360 *morgani*) was rare and *OqxA* was less prevalent than *OqxB* in all but one species. Notably,

361 *OqxAB* were virtually absent in *C. freundii* clades. whilst chromosomal mutations in *gyrAB*
362 and *parCE* were only observed in *K. variicola* (Fig. S4-S9).

363 *Other ARGs*

364 Chloramphenicol ARGs, *cmlA*, *catA* and *catB* were found in all the species, albeit *cmlA* was
365 relatively rare in all the species whilst *catA/B* were conserved in *Enterobacter spp.* and *M.*
366 *morganii*; *catA* was more prevalent than *catB*. *catA/B* were also abundant in *Citrobacter spp.*,
367 *P. mirabilis* and *Providencia spp.* (Fig. S4-S9). Sulphamethoxazole-trimethoprim ARGs, *sul-*
368 *1/2/3* and *dfrA*, were enriched in *Providencia spp.*, *P. mirabilis*, *M. morganii*, *Enterobacter*
369 *spp.*, *C. koseri* (*dfrA* was virtually absent), *C. amalonaticus* clade B2, and *C. freundii*. *Sul1*
370 was more prevalent than *Sul2*, with *Sul3* being relatively rare whilst the *dfrA* variants were
371 remarkably diverse. Indeed, both *sul1*, *sul2* and/or *sul3* as well as several *dfrA* variants were
372 present concurrently in some single strains (Fig. S4-S9; Tables S2).

373 Several tetracycline ARGs such as *tet(A)*, *tet(B)*, *tet(C)*, *tet(D)*, *tet(G)*, *tet(J)*, *tet(S)*, *tet(Y)*,
374 and *tet(41)*, were present in all the species. Notably, *tet(A)*, *tet(B)*, and *tet(D)*, were highly
375 enriched in the various genomes with a descending order of frequency; *tet(J)* was most
376 prevalent in *P. mirabilis* (Fig. S4-S9). *fosA* variants were present in all the species except *M.*
377 *morganii*, but were most enriched and conserved in *Enterobacter spp.* *K. variicola*, *P. stuartii*
378 and *P. rettgeri*. As well, *ere(A)*, *emr(D)*, *erm(B)*, *msr(E)*, *mph(A)* and *mph(E)* macrolide
379 ARGs were common in the various species, with *mph(A)* being richly abundant; *mph(E)* and
380 *msr(E)* were enriched in *C. freundii*, *Providencia spp.* and *P. mirabilis*, *emr(D)* was abundant
381 in *K. variicola* whilst *ere(A)* was abundantly enriched in *E. hormaechei* and *E.*
382 *xiangfangensis*. Rifamycin ARG, *arr*, was identified in the various species, represented by
383 *arr-3* *arr-2* and *arr-1* in all the species (Fig. S4-S9).

384 **Discussion**

385 Among Gram-negative bacterial species, *Pseudomonas aeruginosa*, *Acinetobacter*
386 *baumannii*, and Enterobacteriaceae such as *K. pneumoniae*, *E. coli*, *S. enterica*, *Vibrio*
387 *cholerae*, and *Shigella spp.* are commonly isolated and implicated in nosocomial
388 (anthroponotic), zoonotic and water/food-borne infections with MDR, extensively and
389 pandrug resistant (XDR and PDR respectively) phenomes^{4,26,28,37,38}. However, other less
390 isolated Enterobacteriaceae species such as *Citrobacter spp.*, *Enterobacter hormaechei*
391 *subsp.*, *K. variicola*, *P. mirabilis*, *M. morgani* and *Providencia spp.*, are increasingly being
392 implicated in MDR, XDR and PDR infections globally as opportunistic pathogens^{7,8,28,34,39–}
393 ⁴¹. We show herein, that *C. freundii*, *Enterobacter hormaechei subsp. hormaechei*,
394 *xiangfangensis*, *oharae* and *steigerwaltii*, and Proteaceae strains harbour multiple resistance
395 mechanisms that can make them MDR, XDR and PDR pathogens. More concerning is the
396 global distribution and multiple (human, animal, plants and environmental) specimen sources
397 of these strains, which suggest that they can cause anthroponotic, zoonotic and food- and
398 water-borne infections⁴².

399 Hence, these opportunistic pathogens demand more attention than they have been given
400 hitherto as the rich resistome repertoire identified in their genomes makes them reservoirs of
401 ARGs^{11,14,21,43}. Moreover, being intestinal denizens and commensals, they can easily share
402 these ARGs with facultative and obligate pathogens of humans and animals^{22,44–46}. Further,
403 their presence on plants and the environment further suggests that they can share their ARGs
404 with food-borne and water-borne pathogens^{44–46}. Fortunately, the *E. xiangfangensis* strains
405 found in rice from India had very few ARGs, albeit a few had multiple ARGs (Fig. S2B).
406 Of greater concern is the rich resistome repertoire and abundance of globally distributed *E.*
407 *hormaechei subsp.* strains. Specifically, *E. hormaechei subsp.* contained clinically important
408 ARGs such as *mcr-9*, carbapenemases, and ESBLs, alongside fluoroquinolones,
409 aminoglycoside, tetracycline, macrolide, fosfomycin, chloramphenicol, rifamycin and

410 sulphomethoxazole-trimethoprim resistance mechanisms. This resistome repertoire was also
411 seen in *C. freundii*, *P. rettgeri*, *P. mirabilis*, *P. stuartii*, and *M. morganii*, to a relatively lesser
412 degree in a descending order. This is a worrying observation as colistin, carbapenems and
413 tigecycline are last resort antibiotics used to treat fatal bacterial infections^{47,48}. In most cases,
414 these antibiotics are used in combination with fosfomycin, fluoroquinolones and
415 aminoglycosides to treat carbapenem-resistant Enterobacteriaceae (CRE) infections^{47,48}.
416 Evidently, the presence of all these resistance mechanisms to these antibiotics, could make
417 these species PDR and automatically qualify them as critical priority pathogens per the WHO
418 criteria^{32,33}.

419 As well, members of the tribe Proteaceae viz., *M. morganii*, *P. mirabilis*, and *Providencia spp.*,
420 are known to have intrinsic resistance to colistin, tigecycline, aminopenicillins, amikacin,
421 tobramycin, lincosamides, macrolides, fosfomycin and first- and second-generation
422 cephalosporins^{1,7,49,50}. Thus, the presence of additional resistance determinants in this tribe is
423 especially worrying. Already, there are increasing reports on the isolation of Proteaceae species
424 in recurrent urinary tract infections (UTIs) infections, which is facilitated by the increasing
425 use of colistin to treat MDR infections; their broad intrinsic resistance mechanisms enable
426 them to flourish during antibiotic chemotherapy^{1,7,8,49,51,52}. These observations evince the
427 growing threat of antimicrobial resistance globally and its associated after-effects, supporting
428 the need for efficient antibiotic stewardship to safeguard current antibiotic arsenals as well as
429 discover novel ones^{3,14}.

430 As shown in the phylogenomic and phylogeographic analyses, local and international
431 transmission, or outbreaks of strains within the same clone, clade and subclade i.e., of very
432 close evolutionary distance, were observed. Notably, these closely related strains were
433 isolated from humans, animals, plants, and the environment and they harboured important
434 resistance determinants as described above. The phylogenomics showed the gradual

435 evolution of a single strain during dissemination from host to host and depict the fact that
436 antibiotic resistance respects no boundaries. Notably, a large part of *E. hormaechei* clade C
437 consisted of strains from human stools/urine in Japan; these were closely related strains
438 evolving from Spain, Taiwan, China, and Greece (Fig. 6 & S2). A similar observation was
439 made with strains of close evolutionary distance from humans in Nigeria, France, Spain,
440 Portugal, and Lebanon as well as with strains from humans, animals, and the environment in
441 several countries such as France, Germany, USA, Pakistan, Morocco, Lebanon, and Poland in
442 clade C. *E. hormaechei* strains were also found in desert sands in Morocco, showing their
443 broad and diverse niches (Fig. 6 & S2).

444 Uniform and non-uniform resistome patterns were seen between strains of the same
445 clade/clone in almost all the species. For instance, the same resistome was seen in *C.*
446 *amalonaticus* clade B2 (Fig. 2), *E. cloacae* clade B (Fig. 5) and *P. mirabilis* clade C2 (Fig. 9)
447 whilst differing resistome patterns were observed in the other species and clades. This
448 observation supports two phenomena: firstly, the clonal expansion of strains harbouring the
449 same resistome repertoire on both chromosomes and plasmids and secondly, the horizontal
450 transmission of genetic elements bearing the same of different ARGs across clones and
451 species. During clonal expansion of strains, there is a concomitant replication of resistance
452 plasmids alongside chromosomal replication, leading to daughter cells with the same
453 resistome diversity^{11,12,21,53}. As well, horizontal gene transfer of ARGs between bacteria can
454 alter the resistome diversity and composition of daughter clones emanating from the same
455 ancestor^{11,12,21,53}. In this case, both phenomena are being observed, showing that both vertical
456 and horizontal transmission of ARGs are ongoing during the evolutionary epidemiology of
457 the various clades and species across the globe.

458 The presence of multiple ARGs in a single strain might not necessarily mean they are all
459 being expressed in the bacteria's phenome as antibiotic-susceptible strains have been found to

460 harbour ARGs. For instance, colistin- and fosfomocycin-sensitive Enterobacteriaceae have
461 have been reported in strains harbouring the *mcr-9* and *fosA* genes^{1,5,7,8}. Nevertheless, the
462 ability of these ARGs to be expressed in the presence of strong promoter or transferred to
463 another host with a stronger promoter for subsequent expression cannot be gainsaid⁵⁴.
464 Indeed, antibiotic abuse could serve as an inducer to trigger the transfer and expression of
465 these vast resistomes in the microbial phenomes^{49,55}, necessitating the importance for
466 judicious antibiotic use.

467 It is revealing to note that *K. variicola*, *C. amalonaticus* and *C. koseri* strains had very few
468 ARGs except for *K. variicola* clade B2 and *C. amalonaticus* clade B2, despite the global
469 distribution of *K. variicola* (Fig. 2-3, 7, & S3). Notably, *C. amalonaticus* clade B2, which
470 were all from France, were remarkably enriched with ARGs including *bla*_{NDM}, representing a
471 local outbreak of XDR *C. amalonaticus* strains (Fig. 2). Thus, even in species with fewer
472 resistome diversity and abundance, there are MDR, XDR and PDR strains causing local
473 outbreaks.

474 Notably, most of the genomes included in this analysis were from the USA, Europe, and
475 South East Asia. This may be due to the fact that these regions have higher prevalence and
476 incidence of infections resulting from these pathogens or that these areas have enough
477 financial and technical means to undertake genomic sequencing of these isolates in periodic
478 surveillance studies. Specifically, genomes of these species were relatively scarce from a
479 large part of Russia, Middle and North-West Asia, Africa, the Caribbean and parts of South
480 America and Canada. Given the alarming resistome diversity and composition realised in this
481 analyses, it is incumbent for all nations to intensify and adopt genome-based epidemiological
482 studies to quickly identify the sources and reservoirs of ARGs to pre-empt outbreaks of
483 MDR, XDR and PDR pathogens.

484 **Conclusion**

485 In conclusion, less described Enterobacteriaceae species viz., *Enterobacter hormaechei*
486 *subsp. hormaechei*, *xiangfangensis*, *steigerwaltii* and *oharae*, *C. freundii*, *M. morgani*, *P.*
487 *mirabilis*, *P. stuartii* and *P. rettgeri*, contain globally distributed MDR, XDR and PDR strains
488 that cause local and international outbreaks, transmitting through humans, animals, food,
489 water and other environmental media or sources. Notably, the resistome repertoire of these
490 relatively rare species were equally or more abundant and diverse as those of commonly
491 isolated species. Hence, intensified efforts should be made to increase education on antibiotic
492 stewardship to safeguard the potency of available antibiotics and reduce the selection and
493 dissemination antibiotic-resistant Enterobacteriaceae. Infection prevention and control as well
494 periodic genomic surveillance of communities, hospitals, farms, water bodies and the general
495 environment (One Health) should be undertaken to pre-empt outbreaks of MDR strains and
496 inform infection control measures.

497 Notwithstanding the revealing details obtained in this study, strains with clinically important
498 ARGs whose genomes are not deposited in NCBI/PATRIC or whose genomes are not
499 sequenced before January 2020 will be missed; hence, the information contained herein are
500 true up to January 2020. Nevertheless, the global phylogeography and resistome
501 epidemiology of these emerging opportunistic pathogens provide an important picture of the
502 ARGs, sources and transmission patterns of these species.

503 **Methods**

504 *Included genomes*

505 Genomes of *Citrobacter spp.* (including *amalonaticus*, *freundii*, *koseri*, *werkmanii*, *brakii*,
506 *portucalensis* and *youngae*), *Enterobacter hormaechei subsp. hormaechei*, *xiangfangensis*,
507 *steigerwaltii*, and *oharae*, *Providencia spp.* (including *alcalifaciens*, *burhodogranariea*,

508 *heimbachae*, *rettgeri*, *rustigianii*, and *stuartii*) and *Proteus mirabilis* deposited at GenBank
509 (<https://www.ncbi.nlm.nih.gov/genbank/>) and PATRIC (<https://www.patricbrc.org/>) up to
510 January 2020 were pooled and filtered to remove poor genome sequences. These were used
511 for the downstream phylogenetics, phylogeography and resistome analyses.

512 *Phylogenetics and evolutionary epidemiology analyses*

513 Briefly, plasmid sequences, phages, poor genomes i.e., genomes with sizes below the average
514 genome size of each species viz., 3-4Mb, and genomes of strains that could not share at least
515 1000 core protein genes with all the included genomes were removed. The remaining
516 genomes were aligned and run through RAXmL in batches of 200 genomes to draw
517 phylogenetic trees using the maximum-likelihood method. A minimum of 1000 genes were
518 used for the alignment and a bootstrap resampling of 1000x was used. The trees were
519 annotated using Figtree to show their sequence type (ST), host (species), country and year of
520 isolation. The various clades and sub-clades within each species or genera were visually
521 identified based on their clustering distance and accordingly labelled.

522 *Phylogeography*

523 The various clades and subclades per species were manually drawn onto maps to show their
524 phylogeographic distribution using Paint 3D. Different colour codes were used to distinguish
525 between the various species and clades.

526 *Resistome analyses*

527 The resistomes of the included genomes were individually obtained from the NCBI Pathogen
528 Detection database (<https://www.ncbi.nlm.nih.gov/pathogens/isolates#/search/>). The
529 resistomes were aligned per strain and colour-coded per clade or species to show their
530 association per species, clone, or clade. These were then associated with the phylogenomic
531 trees to ascertain the resistome dynamics per clone, clade, species, and geographical location.

532 References

- 533 1. Mbelle, N. *et al.* Genomic analysis of two drug-resistant clinical *Morganella morganii* strains isolated
534 from UTI patients in Pretoria, South Africa. *Lett. Appl. Microbiol.* **70**, 21–28 (2020).
- 535 2. Osei Sekyere, J. Genomic insights into nitrofurantoin resistance mechanisms and epidemiology in
536 clinical Enterobacteriaceae. *Futur. Sci. OA* **4**, FSO293 (2018).
- 537 3. Osei Sekyere, J. & Asante, J. Emerging mechanisms of antimicrobial resistance in bacteria and fungi:
538 advances in the era of genomics. *Future Microbiol.* **13**, 241–262 (2018).
- 539 4. Pedersen, T. *et al.* Spread of Plasmid-Encoded NDM-1 and GES-5 Carbapenemases among Extensively
540 Drug-Resistant and Pandrug-Resistant Clinical Enterobacteriaceae in Durban, South Africa. *Antimicrob.*
541 *Agents Chemother.* **62**, e02178-17 (2018).
- 542 5. Mbelle, N. M. *et al.* Pathogenomics and Evolutionary Epidemiology of Multi-Drug Resistant Clinical
543 *Klebsiella pneumoniae* Isolated from Pretoria, South Africa. *Sci. Rep.* **10**, 1–17 (2020).
- 544 6. Osei Sekyere, J. & Mensah, E. Molecular epidemiology and mechanisms of antibiotic resistance in
545 *Enterococcus* spp., *Staphylococcus* spp., and *Streptococcus* spp. in Africa: a systematic review from a
546 One Health perspective. *Ann. N. Y. Acad. Sci.* **1465**, 29–58 (2020).
- 547 7. Mbelle, N. M. *et al.* Genomic analysis of a multidrug-resistant clinical *Providencia rettgeri* (PR002)
548 strain with the novel integron In1483 and an A/C plasmid replicon. *Ann. N. Y. Acad. Sci.* **1462**, 92–103
549 (2019).
- 550 8. Osei Sekyere, J., Maningi, N. E., Modipane, L. & Mbelle, N. M. Emergence of *mcr-9.1* in ESBL-
551 producing Clinical Enterobacteriaceae in Pretoria, South Africa: Global Evolutionary Phylogenomics,
552 Resistome and Mobilome. *John. mSystems* **5**, e00148-20 (2020).
- 553 9. Mbelle, N. M. *et al.* The Resistome, Mobilome, Virulome and Phylogenomics of Multidrug-Resistant
554 *Escherichia coli* Clinical Isolates from Pretoria, South Africa. *Sci. Rep.* **9**, 1–43 (2019).
- 555 10. Chirindze, L. M. *et al.* Faecal colonization of *E. coli* and *Klebsiella* spp. producing extended-spectrum
556 beta-lactamases and plasmid-mediated AmpC in Mozambican university students. *BMC Infect. Dis.* **18**,
557 244 (2018).
- 558 11. Decano, A. G. *et al.* Plasmids shape the diverse accessory resistomes of *Escherichia coli* ST131. *bioRxiv*
559 2020.05.07.081380 (2020). doi:10.1101/2020.05.07.081380
- 560 12. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements Associated with
561 Antimicrobial Resistance. *Clin. Microbiol. Rev.* **31**, 1–61 (2018).
- 562 13. Böhm, M.-E., Razavi, M., Marathe, N. P., Flach, C.-F. & Larsson, D. G. J. Discovery of a novel
563 integron-borne aminoglycoside resistance gene present in clinical pathogens by screening environmental
564 bacterial communities. *Microbiome* **8**, 41 (2020).
- 565 14. Asante, J. & Osei Sekyere, J. Understanding antimicrobial discovery and resistance from a metagenomic
566 and metatranscriptomic perspective: Advances and applications. *Environ. Microbiol. Rep.* **11**, 62–86
567 (2019).
- 568 15. Van Boeckel, T. P. *et al.* Global trends in antimicrobial resistance in animals in low- and middle-income
569 countries. *Science (80-.)*. **365**, eaaw1944 (2019).
- 570 16. He, T. *et al.* Emergence of plasmid-mediated high-level tigecycline resistance genes in animals and
571 humans. *Nat. Microbiol.* **4**, 1450–1456 (2019).
- 572 17. Li, J. *et al.* Inter-host Transmission of Carbapenemase-Producing *Escherichia coli* among Humans and
573 Backyard Animals. *Environ. Health Perspect.* **127**, 107009 (2019).
- 574 18. Jaidane, N. *et al.* Genomic analysis of in vivo acquired resistance to colistin and rifampicin in
575 *Acinetobacter baumannii*. *Int. J. Antimicrob. Agents* **51**, 266–269 (2018).
- 576 19. Olaitan, A. O., Diene, S. M., Assous, M. V. & Rolain, J.-M. Genomic Plasticity of Multidrug-Resistant
577 NDM-1 Positive Clinical Isolate of *Providencia rettgeri*. *Genome Biol. Evol.* **8**, 723–8 (2016).

- 578 20. Conlan, S. *et al.* Plasmid Dynamics in KPC-Positive *Klebsiella pneumoniae* during Long-Term Patient
579 Colonization. *MBio* **7**, (2016).
- 580 21. Decano, A. G. *et al.* Complete Assembly of *Escherichia coli* Sequence Type 131 Genomes Using Long
581 Reads Demonstrates Antibiotic Resistance Gene Variation within Diverse Plasmid and Chromosomal
582 Contexts. *mSphere* **4**, 130–19 (2019).
- 583 22. Osei Sekyere, J., Maningi, N. E. & Fourie, P. B. Mycobacterium tuberculosis, antimicrobials,
584 immunity, and lung–gut microbiota crosstalk: current updates and emerging advances. *Ann. N. Y. Acad.*
585 *Sci.* **1467**, 21–47 (2020).
- 586 23. Liu, B.-T. *et al.* Characteristics of Carbapenem-Resistant Enterobacteriaceae in Ready-to-Eat
587 Vegetables in China. *Front. Microbiol.* **9**, 1147 (2018).
- 588 24. Tshisevhe, V. S., Lekalakala, M. R., Tshuma, N., Janse van Rensburg, S. & Mbelle, N. Outbreak of
589 carbapenem-resistant *Providencia rettgeri* in a tertiary hospital. *S. Afr. Med. J.* **107**, 31–33 (2016).
- 590 25. Kopotsa, K., Osei Sekyere, J. & Mbelle, N. M. Plasmid evolution in carbapenemase-producing
591 Enterobacteriaceae: a review. *Ann. N. Y. Acad. Sci.* **1457**, 61–91 (2019).
- 592 26. Weill, F. X. *et al.* Global phylogeography and evolutionary history of *Shigella dysenteriae* type 1. *Nat.*
593 *Microbiol.* **1**, (2016).
- 594 27. Hargreaves, M. L. *et al.* Clonal Dissemination of *Enterobacter cloacae* Harboring bla KPC-3 in the
595 Upper Midwestern United States. *Antimicrob. Agents Chemother.* **59**, 7723–7734 (2015).
- 596 28. Osei Sekyere, J. & Reta, M. A. Phylogeography and Resistome Epidemiology of Gram-Negative
597 Bacteria in Africa: A Systematic Review and Genomic Meta-Analysis from a One-Health Perspective.
598 *medRxiv* 2020.04.09.20059766 (2020). doi:10.1101/2020.04.09.20059766
- 599 29. Cerqueira, G. C. *et al.* Multi-institute analysis of carbapenem resistance reveals remarkable diversity,
600 unexplained mechanisms, and limited clonal outbreaks. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 1135–1140
601 (2017).
- 602 30. Tijet, N. *et al.* Lateral dissemination and inter-patient transmission of blaKPC-3: role of a conjugative
603 plasmid in spreading carbapenem resistance. *J. Antimicrob. Chemother.* **71**, 344–347 (2016).
- 604 31. Petty, N. K. *et al.* Global dissemination of a multidrug resistant *Escherichia coli* clone. *Proc. Natl. Acad.*
605 *Sci. U. S. A.* **111**, 5694–5699 (2014).
- 606 32. Rello, J. *et al.* A global priority list of the TOp TEn resistant Microorganisms (TOTEM) study at
607 intensive care: a prioritization exercise based on multi-criteria decision analysis. *Eur. J. Clin. Microbiol.*
608 *Infect. Dis.* **38**, 319–323 (2019).
- 609 33. Asokan, G. V., Ramadhan, T., Ahmed, E. & Sanad, H. WHO Global Priority Pathogens List: A
610 Bibliometric Analysis of Medline-PubMed for Knowledge Mobilization to Infection Prevention and
611 Control Practices in Bahrain. *Oman Med. J.* **34**, 184–193 (2019).
- 612 34. Mbelle, N. *et al.* Genomic analysis of two drug-resistant clinical *Morganella morganii* strains isolated
613 from UTI patients in Pretoria, South Africa. *Lett. Appl. Microbiol.* **70**, 21–28 (2020).
- 614 35. Lv, L. *et al.* Emergence of a Plasmid-Encoded Resistance-Nodulation-Division Efflux Pump Conferring
615 Resistance to Multiple Drugs, Including Tigecycline, in *Klebsiella pneumoniae*. *MBio* **11**, (2020).
- 616 36. Ellem, J. A. *et al.* Locally Acquired mcr-1 in *Escherichia coli*, Australia, 2011 and 2013. *Emerg. Infect.*
617 *Dis.* **23**, 1160–1163 (2017).
- 618 37. Perez, A. *et al.* High incidence of MDR and XDR *Pseudomonas aeruginosa* isolates obtained from
619 patients with ventilator-associated pneumonia in Greece, Italy and Spain as part of the MagicBullet
620 clinical trial. *J. Antimicrob. Chemother.* **74**, 1244–1252 (2019).
- 621 38. D’Onofrio, V. *et al.* Epidemiology of colistin-resistant, carbapenemase-producing Enterobacteriaceae
622 and *Acinetobacter baumannii* in Croatia. *Infect. Genet. Evol.* **81**, 104263 (2020).
- 623 39. Lob, S. H. *et al.* In vitro activity of imipenem-relebactam against resistant phenotypes of
624 Enterobacteriaceae and *Pseudomonas aeruginosa* isolated from intraabdominal and urinary tract

- 625 infection samples - SMART Surveillance Europe 2015-2017. *J. Med. Microbiol.* **69**, 207–217 (2020).
- 626 40. Lu, Y., Feng, Y., McNally, A. & Zong, Z. Occurrence of colistin-resistant hypervirulent *Klebsiella*
627 *variicola*. *J. Antimicrob. Chemother.* (2018). doi:10.1093/jac/dky301
- 628 41. Rodriguez-Medina, N., Barrios-Camacho, H., Duran-Bedolla, J. & Garza-Ramos, U. *Klebsiella*
629 *variicola*: an emerging pathogen in humans. *Emerg. Microbes Infect.* **8**, 973–988 (2019).
- 630 42. Exner, M. *et al.* Antibiotic resistance: What is so special about multidrug-resistant Gram-negative
631 bacteria? *GMS Hyg. Infect. Control* **12**, Doc05 (2017).
- 632 43. Yang, D. *et al.* The Occurrence of the Colistin Resistance Gene *mcr-1* in the Haihe River (China). *Int. J.*
633 *Environ. Res. Public Health* **14**, (2017).
- 634 44. Lerner, A., Matthias, T. & Aminov, R. Potential effects of horizontal gene exchange in the human gut.
635 *Front. Immunol.* **8**, (2017).
- 636 45. Boto, L., Pineda, M. & Pineda, R. Potential impacts of horizontal gene transfer on human health and
637 physiology and how anthropogenic activity can affect it. *FEBS J.* **286**, 3959–3967 (2019).
- 638 46. Anderson, M. T. & Seifert, H. S. Opportunity and means: Horizontal gene transfer from the human host
639 to a bacterial pathogen. *MBio* **2**, (2011).
- 640 47. Osei Sekyere, J., Govinden, U., Bester, L. A. & Essack, S. Y. Colistin and tigecycline resistance in
641 carbapenemase-producing Gram-negative bacteria: emerging resistance mechanisms and detection
642 methods. *J. Appl. Microbiol.* **121**, 601–617 (2016).
- 643 48. Osei Sekyere, J., Sephofane, A. K. & Mbelle, N. M. Comparative Evaluation of CHROMagar COL-
644 APSE, MicroScan Walkaway, ComASP Colistin, and Colistin MAC Test in Detecting Colistin-resistant
645 Gram-Negative Bacteria. *Sci. Rep.* **10**, 1–13 (2020).
- 646 49. Liu, H., Zhu, J., Hu, Q. & Rao, X. *Morganella morganii*, a non-negligent opportunistic pathogen.
647 *International Journal of Infectious Diseases* **50**, 10–17 (2016).
- 648 50. Hope, R. *et al.* Tigecycline activity: low resistance rates but problematic disc breakpoints revealed by a
649 multicentre sentinel survey in the UK. *J Antimicrob Chemother* **65**, 2602–2609 (2010).
- 650 51. Hayakawa, K. *et al.* Growing prevalence of *Providencia stuartii* associated with the increased usage of
651 colistin at a tertiary health care center. *Int. J. Infect. Dis.* **16**, e646-8 (2012).
- 652 52. Gajdács, M. & Urbán, E. Comparative epidemiology and resistance trends of Proteae in urinary tract
653 infections of inpatients and outpatients: A 10-year retrospective study. *Antibiotics* **8**, (2019).
- 654 53. Carattoli, A. Plasmids in Gram negatives: Molecular typing of resistance plasmids. *Int. J. Med.*
655 *Microbiol.* **301**, 654–658 (2011).
- 656 54. Carroll, L. M. *et al.* Identification of novel mobilized colistin resistance gene *mcr-9* in a multidrug-
657 resistant, colistin-susceptible salmonella enterica serotype typhimurium isolate. *MBio* **10**, e00853-19
658 (2019).
- 659 55. Liu, G., Bogaj, K., Bortolaia, V., Olsen, J. E. & Thomsen, L. E. Antibiotic-Induced, Increased
660 Conjugative Transfer Is Common to Diverse Naturally Occurring ESBL Plasmids in *Escherichia coli*.
661 *Front. Microbiol.* **10**, 2119 (2019).

662

663

664

665

666

667

668 **Acknowledgements:** None

669 **Funding:** None

670 **Transparency declaration:** None (Authors declare no conflict of interest)

671 **Figure 1.** Evolutionary epidemiology and resistome of global *Citrobacter freundii* isolates.

672 *C. freundii* clustered into four main clades (A, B1, B2 and B3), highlighted with distinct

673 colours. Clade B3 had the most resistome abundance and diversity. Strains from humans

674 (blue labels), animals (red labels), plants (purple/mauve labels) and the environment (green

675 labels) were found in the same clade/cluster. *Bla_{CMY}* was conserved in these genomes.

676 **Figure 2.** Evolutionary epidemiology and resistome of global *Citrobacter amalonaticus*

677 isolates. *C. amalonaticus* strains clustered into clades A (red highlight), B1 (green highlight),

678 B2 (blue highlight) and C (mauve highlight); clade B2 had very rich resistome repertoire and

679 were all from France, but the other clades had very few resistance genes. Strains from

680 humans (blue labels), animals (red labels), plants (purple/mauve labels) and the environment

681 (green labels) were found in the same clade/cluster. *Bla_{SED}* and *oqxAB* were almost conserved

682 in these genomes.

683 **Figure 3.** Evolutionary epidemiology and resistome of global *Citrobacter koseri* isolates. *C.*

684 *koseri* strains clustered into clades A (grey highlight), B1 (light blue highlight), B2 (orange

685 highlight) and B3 (mauve highlight). Strains from humans (blue labels) and animals (red

686 labels) were found in the same clade/cluster. *Bla_{CKO}* and *bla_{MAL}* were almost conserved in

687 these genomes.

688 **Figure 4.** Evolutionary epidemiology and resistome of global *Enterobacter*

689 *steigerwaltii/oharae* isolates. *E. steigerwaltii/oharae* isolates clustered into clades A, B and

690 C, which were all obtained from humans from countries distributed across the globe. Clades

691 B and C had very rich resistome repertoire; *bla_{ACT}* was conserved in these genomes..

692 **Figure 5.** Evolutionary epidemiology and resistome of global *Enterobacter xiangfangensis*
693 isolates. The *E. xiangfangensis* contained *E. cloacae* genomes and they clustered into clades
694 A1, A2 and B, with clades A2 and B having rich and diverse resistome repertoire; these
695 clades were distributed globally from humans (blue labels) and animals (red labels). *bla_{ACT}*
696 was conserved in these genomes.

697 **Figure 6.** Evolutionary epidemiology and resistome of global *Enterobacter hormaechei*
698 isolates. The *E. hormaechei* isolates clustered into three main clades A, B and C (with distinct
699 highlights) that contained strains distributed globally from humans (blue labels), and animals
700 (red labels), plants (purple/mauve labels) and the environment (green labels). Clades B and C
701 contained diverse and rich resistome repertoire. *bla_{ACT}* was conserved in these genomes.

702 **Figure 7.** Evolutionary epidemiology and resistome of global *Klebsiella variicola* isolates.
703 The *K. variicola* strains clustered into nine clades viz., A1, A2, A3, A4, A5, A6, B1, B2 and
704 C, which were highlighted with distinct colours and were isolated from countries around the
705 globe. The clades contained strains distributed globally from humans (blue labels), animals
706 (red labels), plants (purple/mauve labels) and the environment (green labels). Besides a few
707 strains in clade B2, the other strains contained very few resistance genes. *bla_{LEN}* was
708 conserved in these genomes.

709 **Figure 8.** Evolutionary epidemiology and resistome of global *Morganella morganii* isolates.
710 The *M. morganii* strains clustered into three clades, A (red highlight), B (light blue highlight)
711 and C (yellow/gold highlight), containing isolates obtained globally from humans (blue
712 labels) and animals (red labels).

713 **Figure 9.** Evolutionary epidemiology and resistome of global *Proteus mirabilis* isolates. The
714 *P. mirabilis* isolates clustered into 10 clades, A-A3, B1-B3, and C1-C3 (shown with different
715 highlights), which contained diverse and abundant resistomes with conserved *catA* and *tet*

716 genes. The clades contained strains distributed globally from humans (blue labels), animals
717 (red labels), plants (purple/mauve labels) and the environment (green labels).

718 **Figure 10.** Evolutionary epidemiology and resistome of global *Providencia* spp. isolates. The
719 *Providencia* spp. clustered into 12 branches, with *P. stuartii*, *P. rustigianii*, *P. alcalifaciens*,
720 *P. heimbachei* and *P. rettgeri* clustering into branch VIII to XII respectively. *P. rettgeri*
721 clustered into clades A and B, consisting of globally distributed isolates. *P. stuartii* and *P.*
722 *rettgeri* contained richer and more abundant resistomes than the other *Providencia* strains and
723 contained strains distributed globally from humans (blue labels), animals (red labels), plants
724 (purple/mauve labels) and the environment (green labels).

725 **Figure 11.** Global phylogeography of *Citrobacter freundii*, *Citrobacter amalonaticus*,
726 *Citrobacter koseri*, *Enterobacter hormaechei* subsp. *hormaechei*, *xiangfangensis*,
727 *steigerwaltii* and *oharae*, *Klbesiella variicola*, *Morganella morgani*, *Proteus mirabilis*, and
728 *Providencia* spp. Most of these genomes were obtained from USA, Europe, South-East Asia
729 and South America in a descending order of frequency. *C. freundii*, *Enterobacter* spp., *K.*
730 *variicola*, and *P. mirabilis* had more diverse distribution across the globe. Each species is
731 designated with a different colour code.

732 **Supplemental dataset 1.** Raw metadata of downloaded genomes from PATRIC containing
733 all the data associated with each genome.

734 **Supplemental dataset 2.** Species by species tabulation and analyses of the resistomes,
735 specimen sources, country of isolation, MLST, Biosample accession number, and strain name
736 of all the genomes according to their order on the phylogeny trees.

737 **Supplemental dataset 3.** Colour-coded species by species tabulation of the resistomes,
738 specimen sources, country of isolation, MLST, Biosample accession number, and strain name
739 of all the genomes according to their order on the phylogeny trees.

740 **Figure S1 (A and B).** Evolutionary epidemiology and resistome of global *Citrobacter*
741 *freundii* isolates, A and B. *C. freundii* clustered into four main clades (A, B1, B2 and B3),
742 highlighted with distinct colours. Clade B3 had the most resistome abundance and diversity.
743 Strains from humans (blue labels), animals (red labels), plants (purple/mauve labels) and the
744 environment (green labels) were found in the same clade/cluster. *Bla_{CMY}* was conserved in
745 these genomes.

746 **Figure S2 (A and B).** Evolutionary epidemiology and resistome of global *Enterobacter*
747 *hormaechei* isolates, A and B. The *E. hormaechei* isolates clustered into three main clades A,
748 B and C (with distinct highlights) that contained strains distributed globally from humans
749 (blue labels), and animals (red labels), plants (purple/mauve labels) and the environment
750 (green labels). Clades B and C contained diverse and rich resistome repertoire. *bla_{ACT}* was
751 conserved in these genomes.

752 **Figure S3 (A and B).** Evolutionary epidemiology and resistome of global *Klebsiella*
753 *variicola* isolates, A and B. The *K. variicola* strains clustered into nine clades viz., A1, A2,
754 A3, A4, A5, A6, B1, B2 and C, which were highlighted with distinct colours and were
755 isolated from countries around the globe. The clades contained strains distributed globally
756 from humans (blue labels), animals (red labels), plants (purple/mauve labels) and the
757 environment (green labels). Besides a few strains in clade B2, the other strains contained very
758 few resistance genes. *bla_{LEN}* was conserved in these genomes.

759 **Figure S4.** Frequency distribution of antibiotic resistance genes in *Citrobacter freundii* (A),
760 and *Citrobacter species* (B and C).

761 **Figure S5.** Frequency distribution of antibiotic resistance genes in *Citrobacter amalonaticus*
762 (A), and *Citrobacter koseri* (B).

763 **Figure S6.** Frequency distribution of antibiotic resistance genes in *Enterobacter steigerwaltii*
764 and *oharae* (A), and *Enterobacter xiangfangensis* (B).

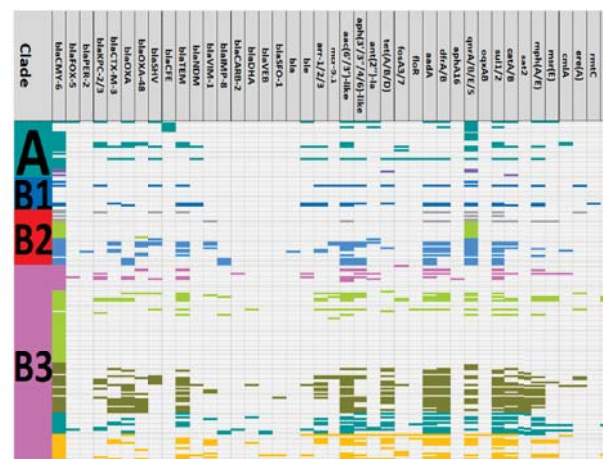
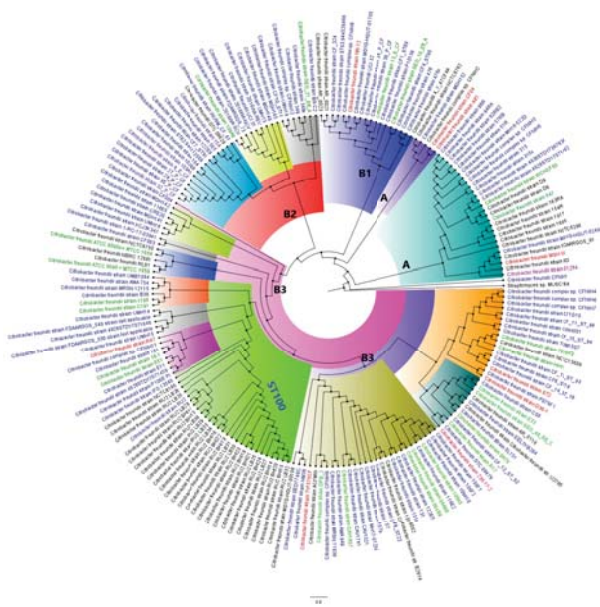
765 **Figure S7.** Frequency distribution of antibiotic resistance genes in *Enterobacter hormaechei*
766 (A, B and C).

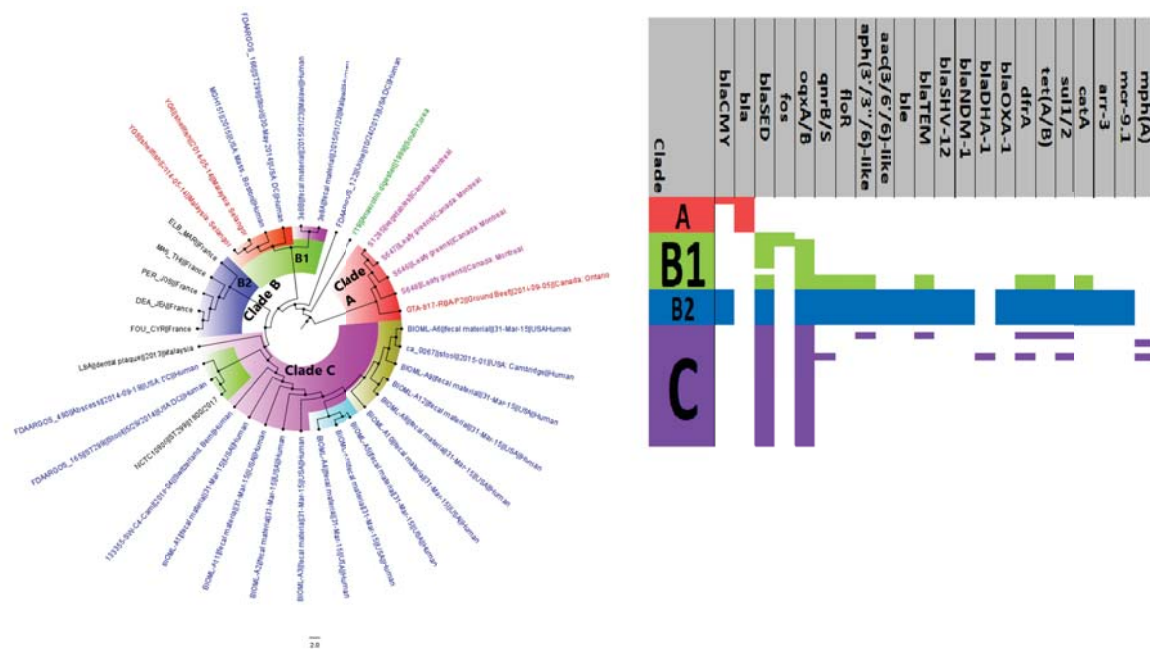
767 **Figure S8.** Frequency distribution of antibiotic resistance genes in *Klebsiella variicola* (A, B
768 and C).

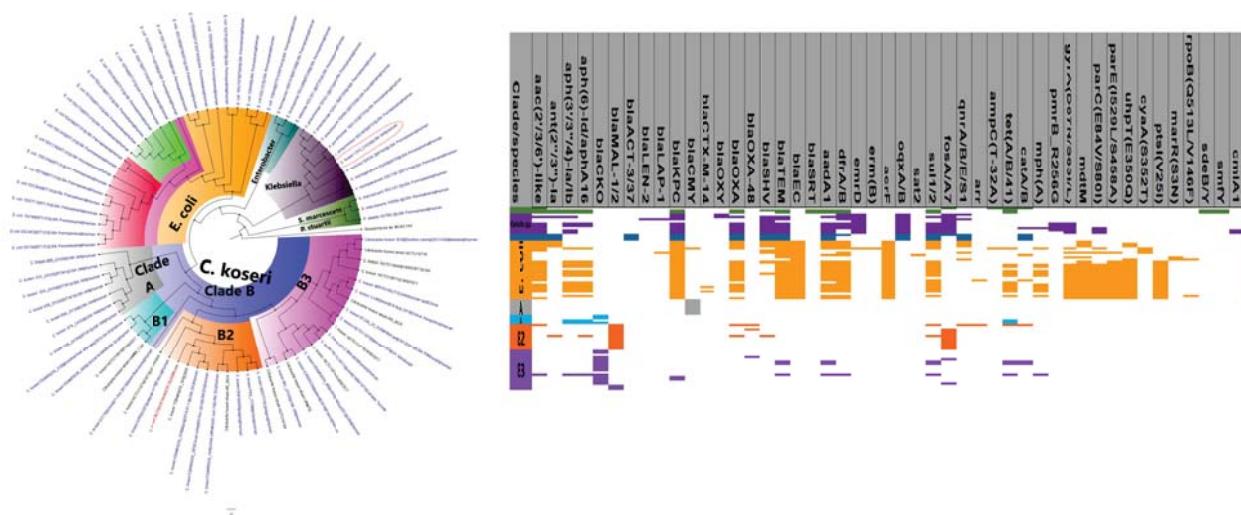
769 **Figure S9.** Frequency distribution of antibiotic resistance genes in *Morganella morganii* (A),
770 *Proteus mirabilis* (B) and *Providencia species* (C).

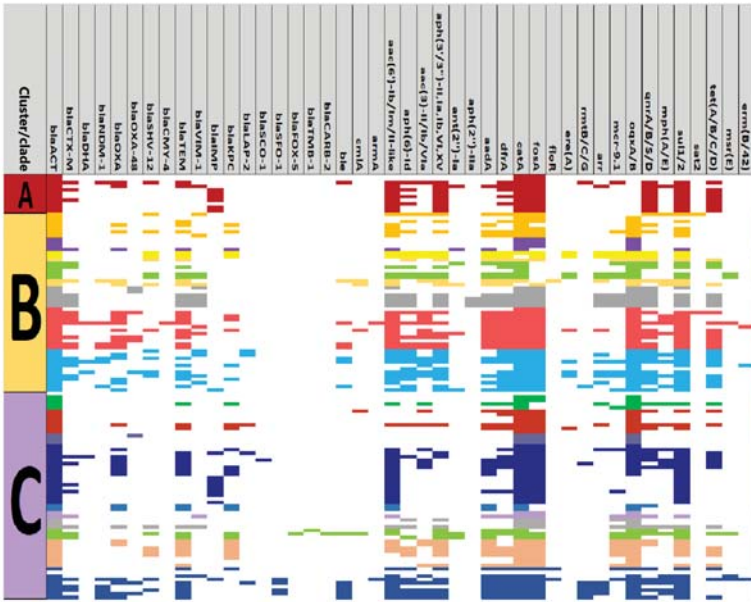
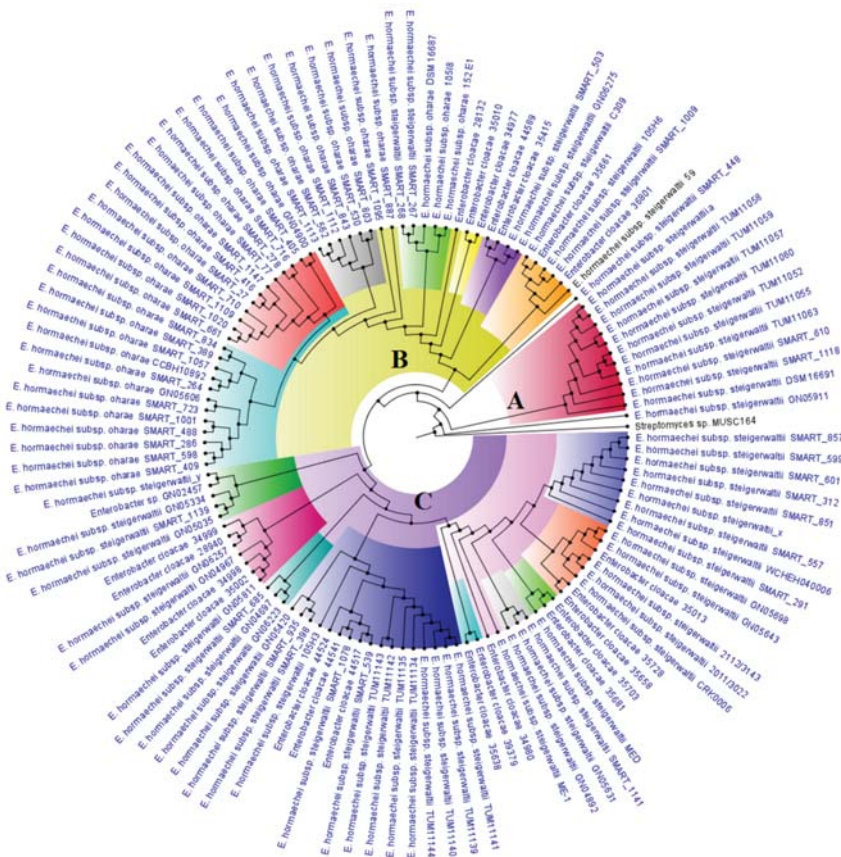
771

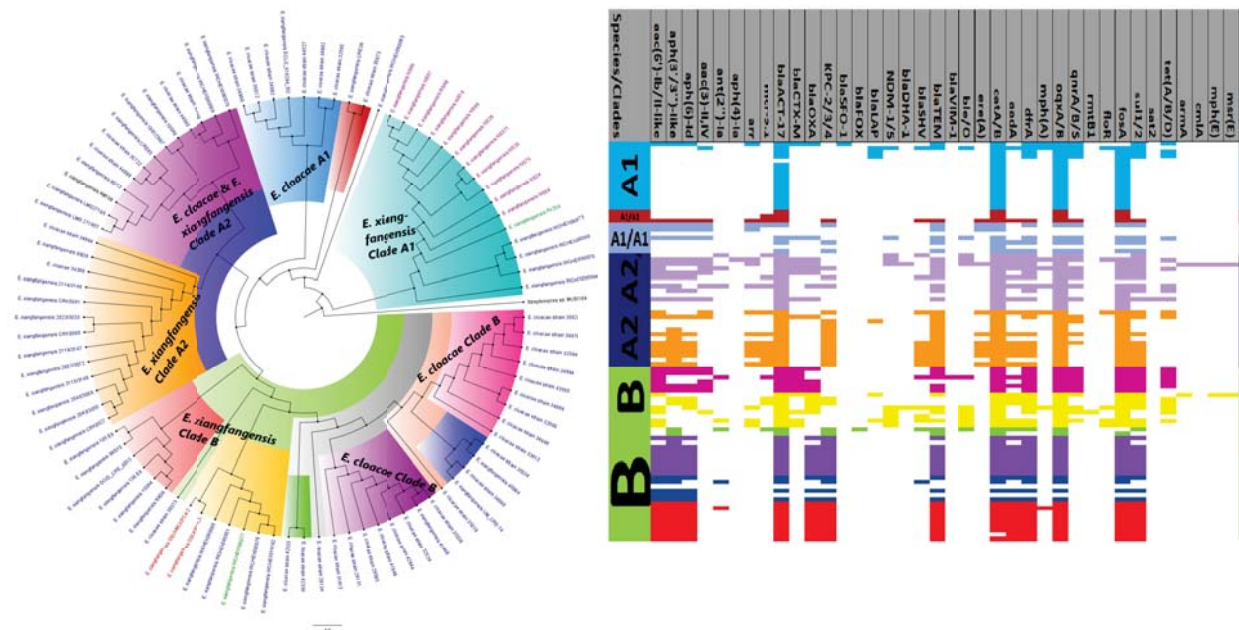
772

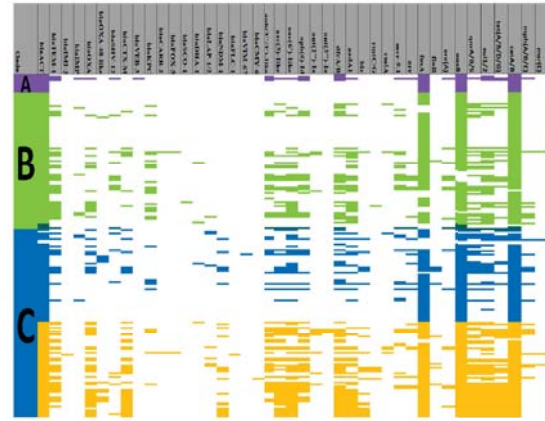
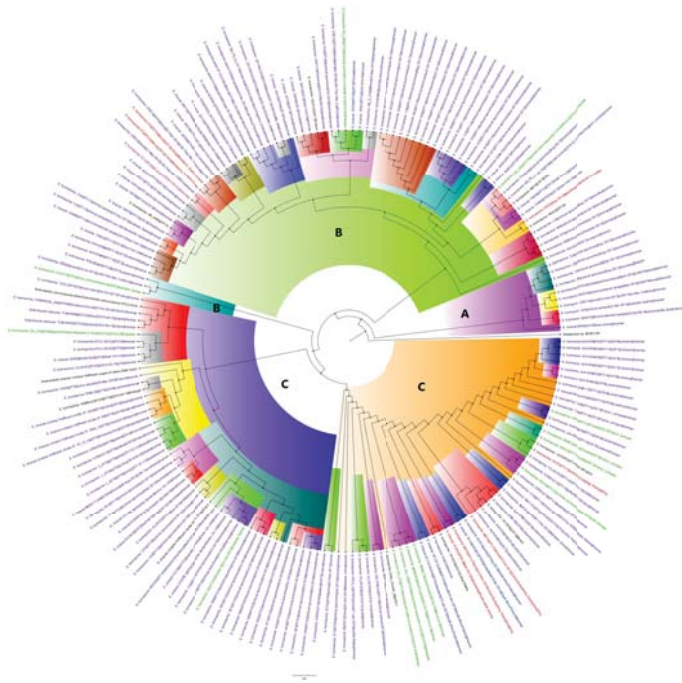


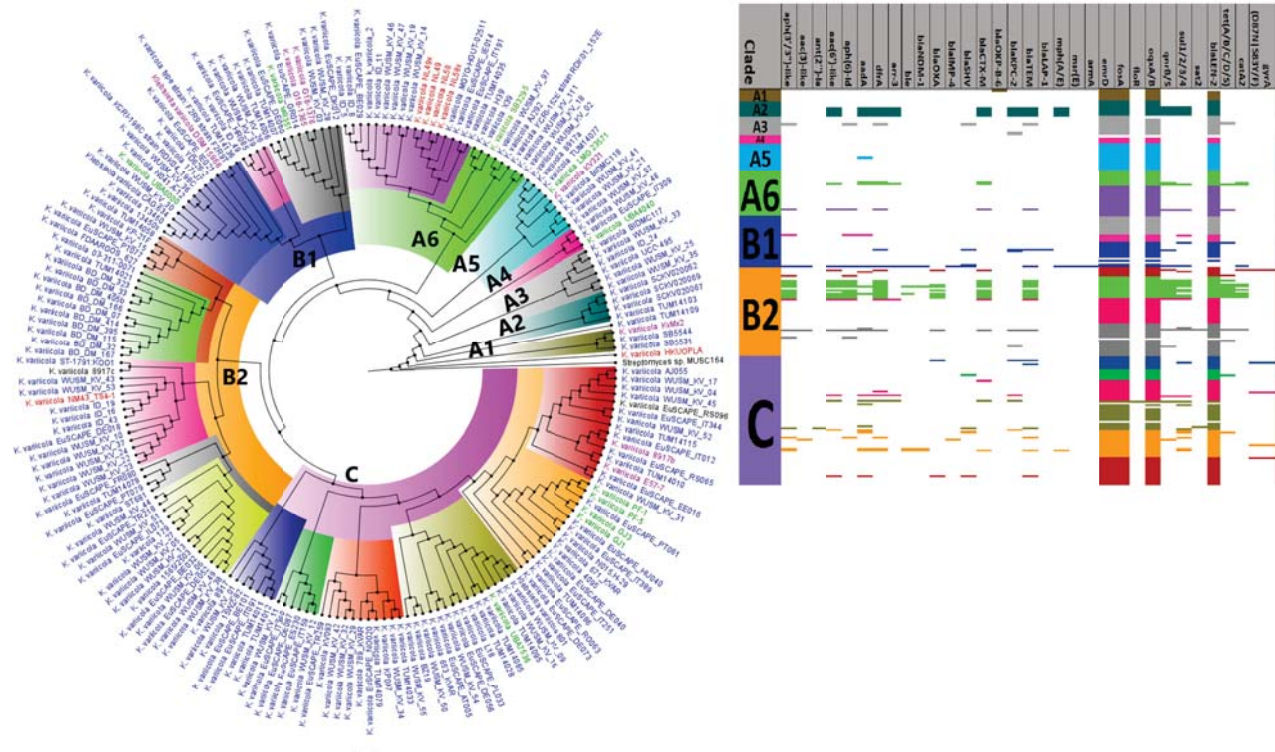


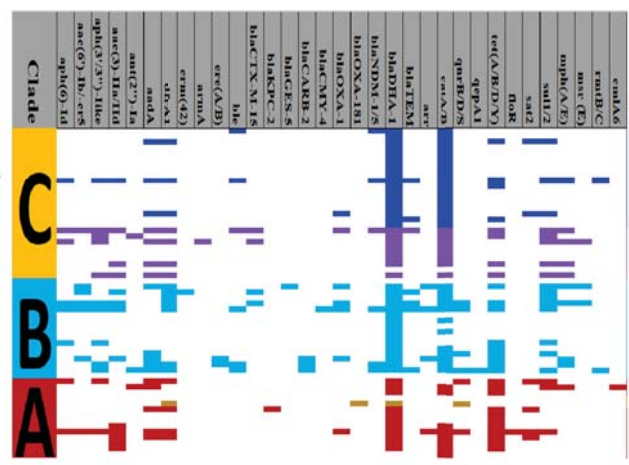
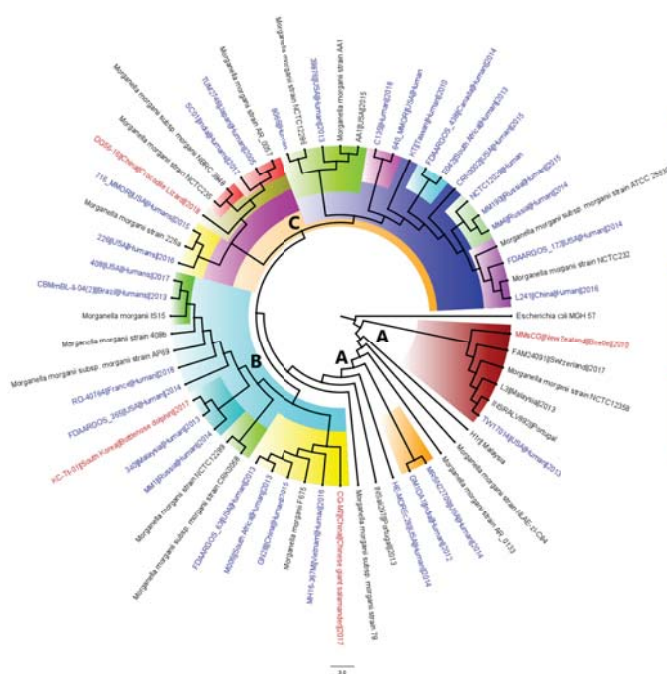


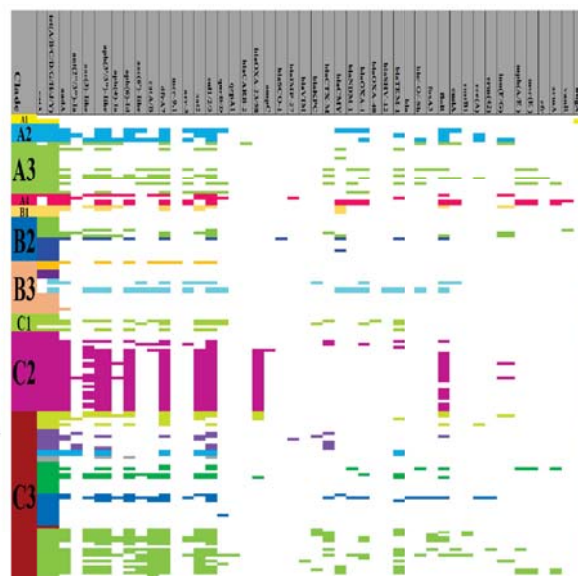
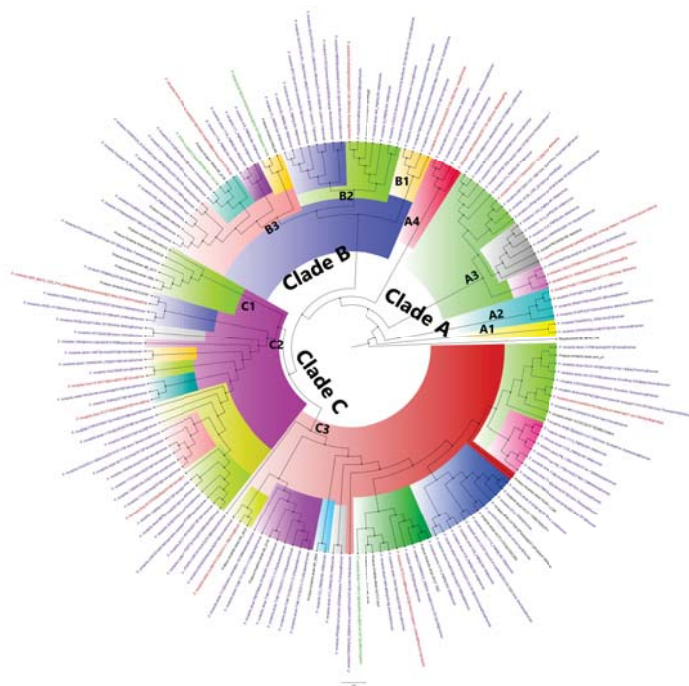


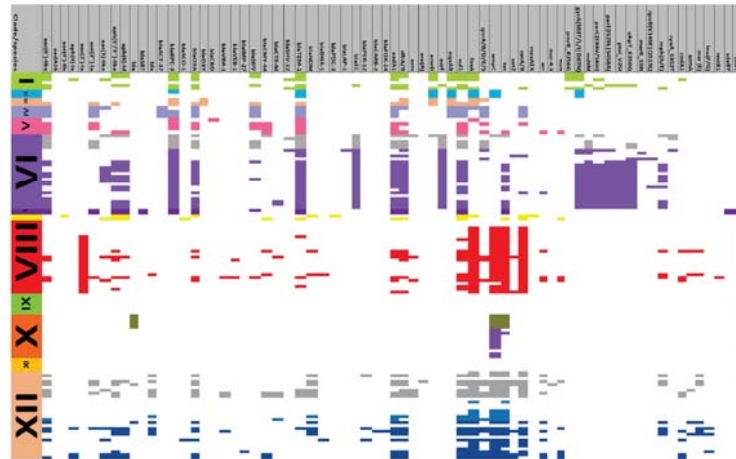
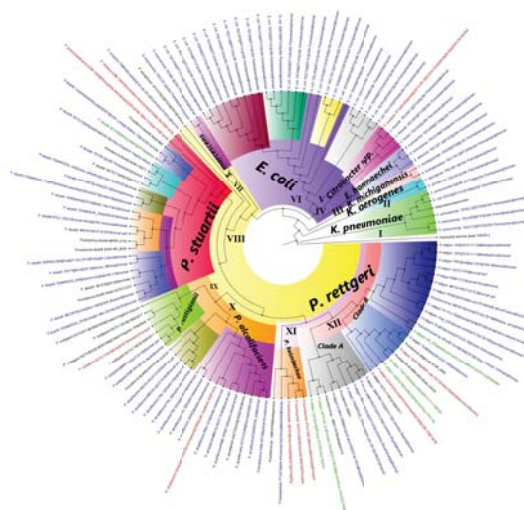


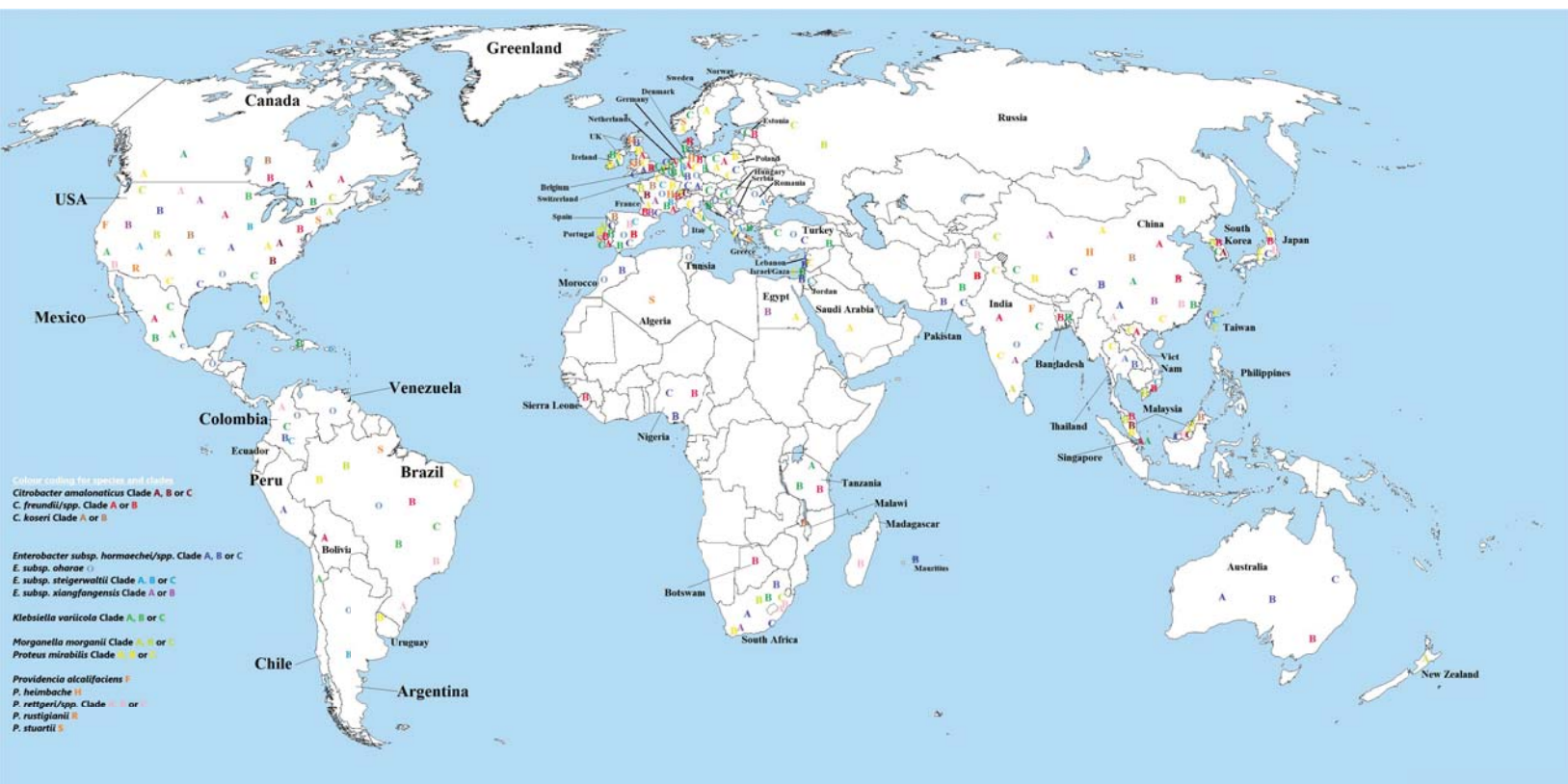




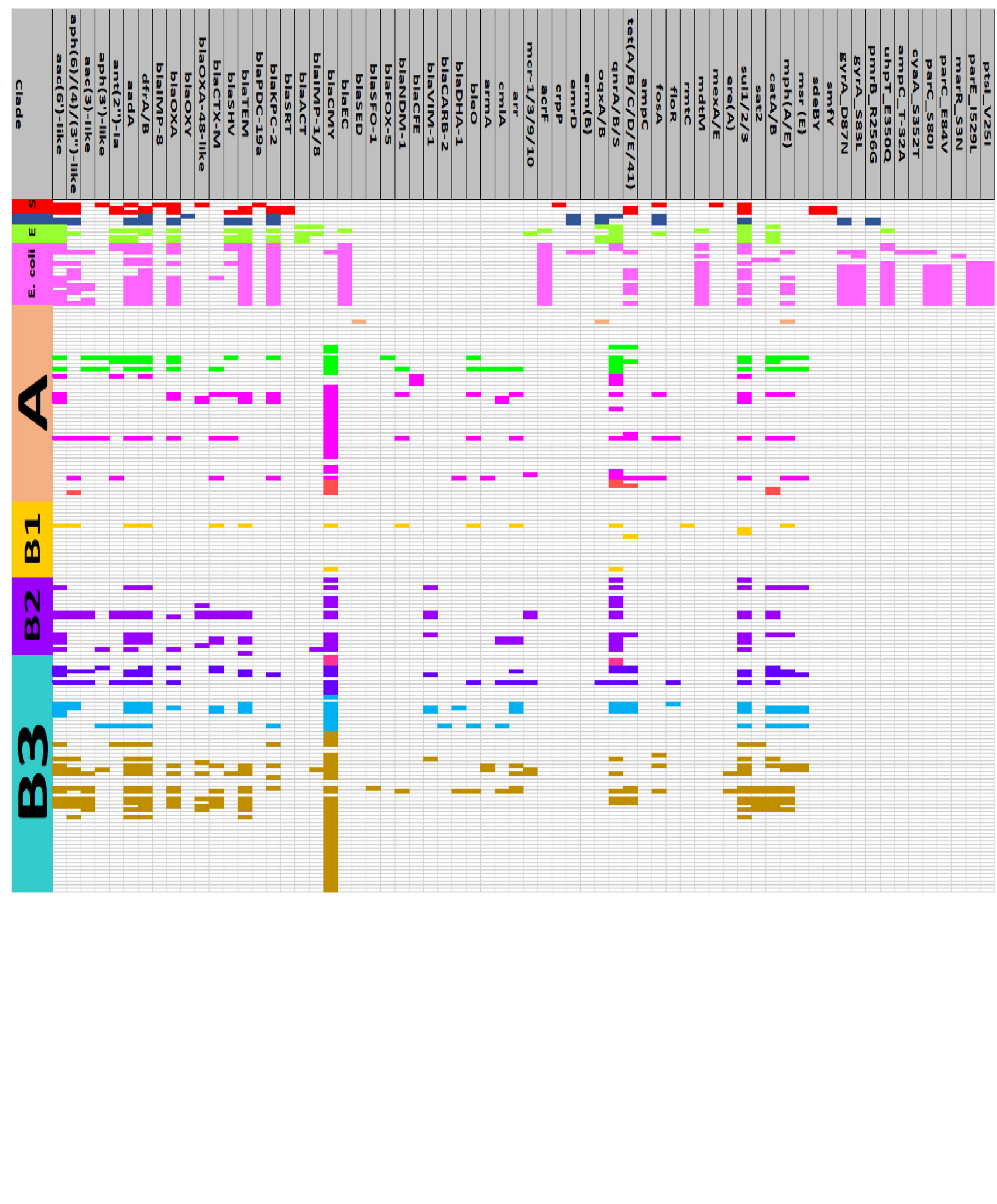
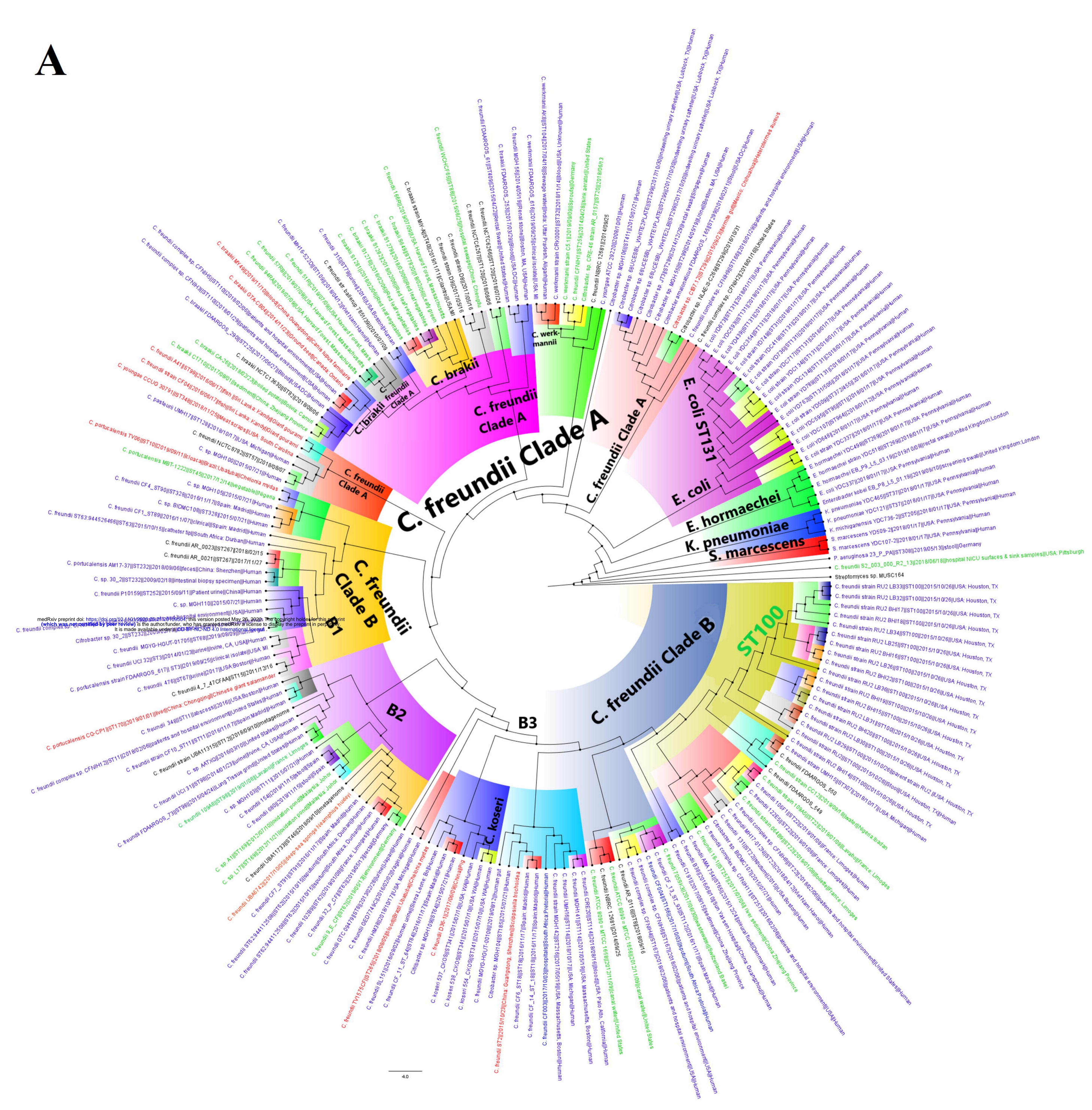




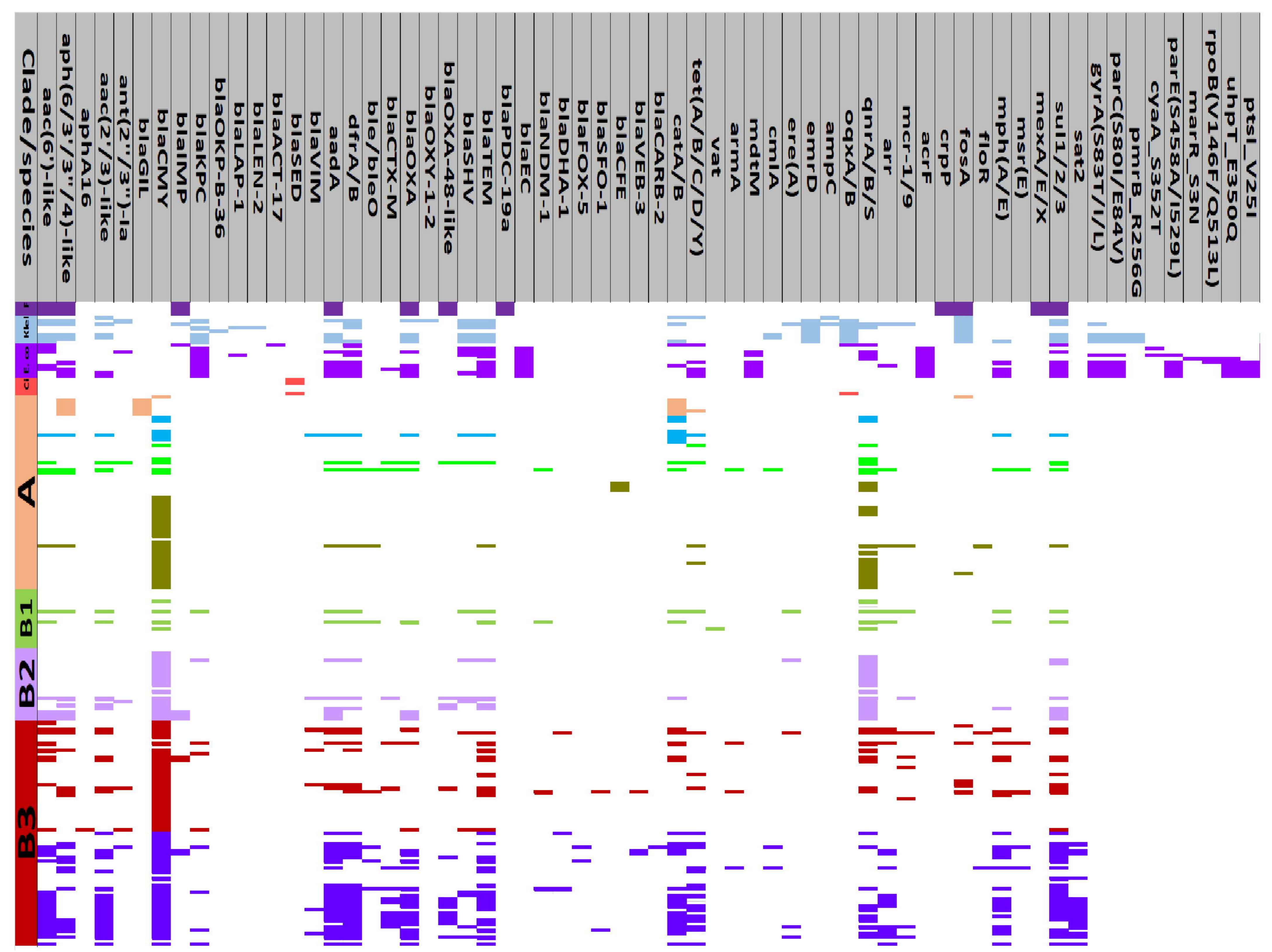
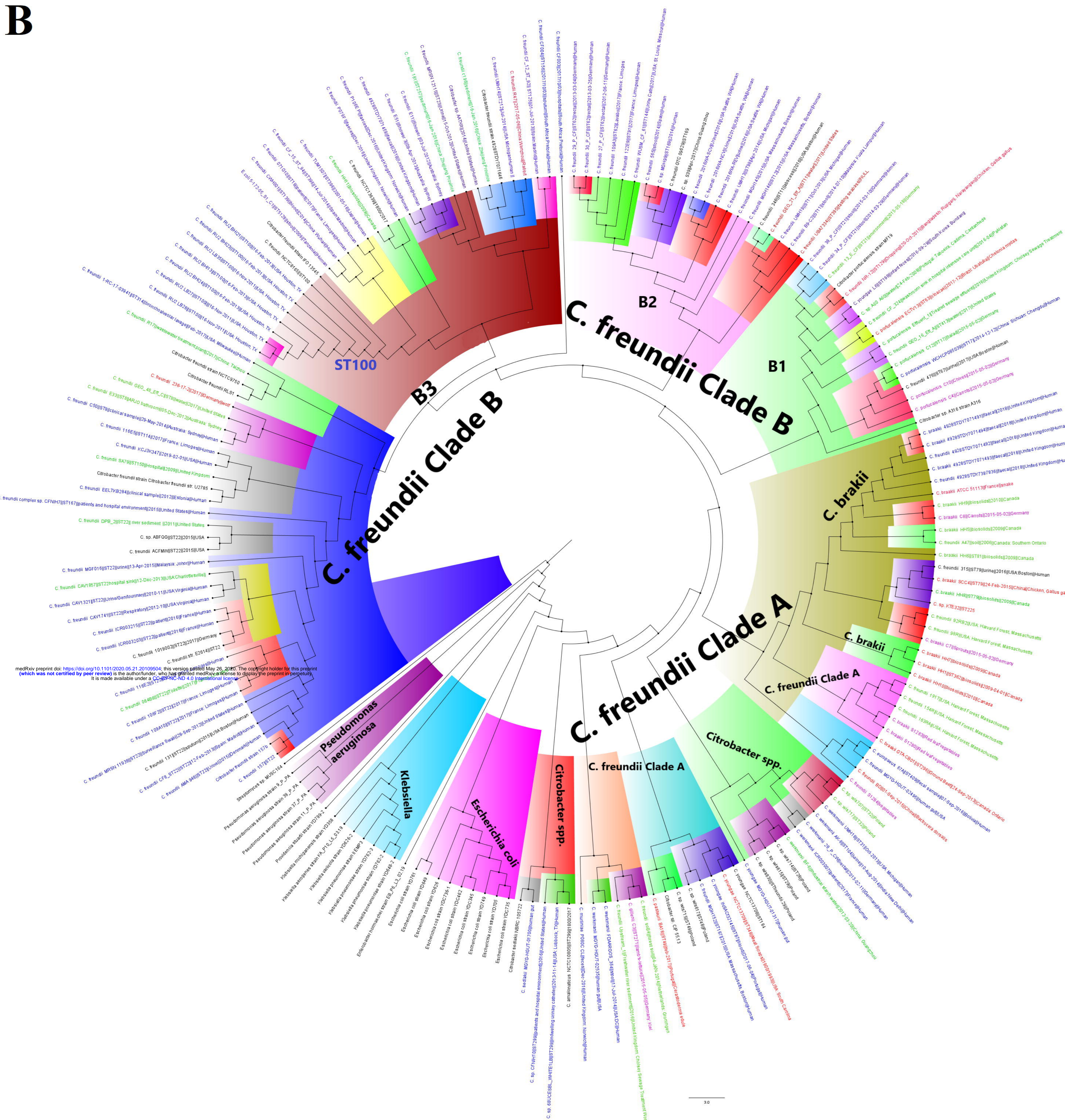




A

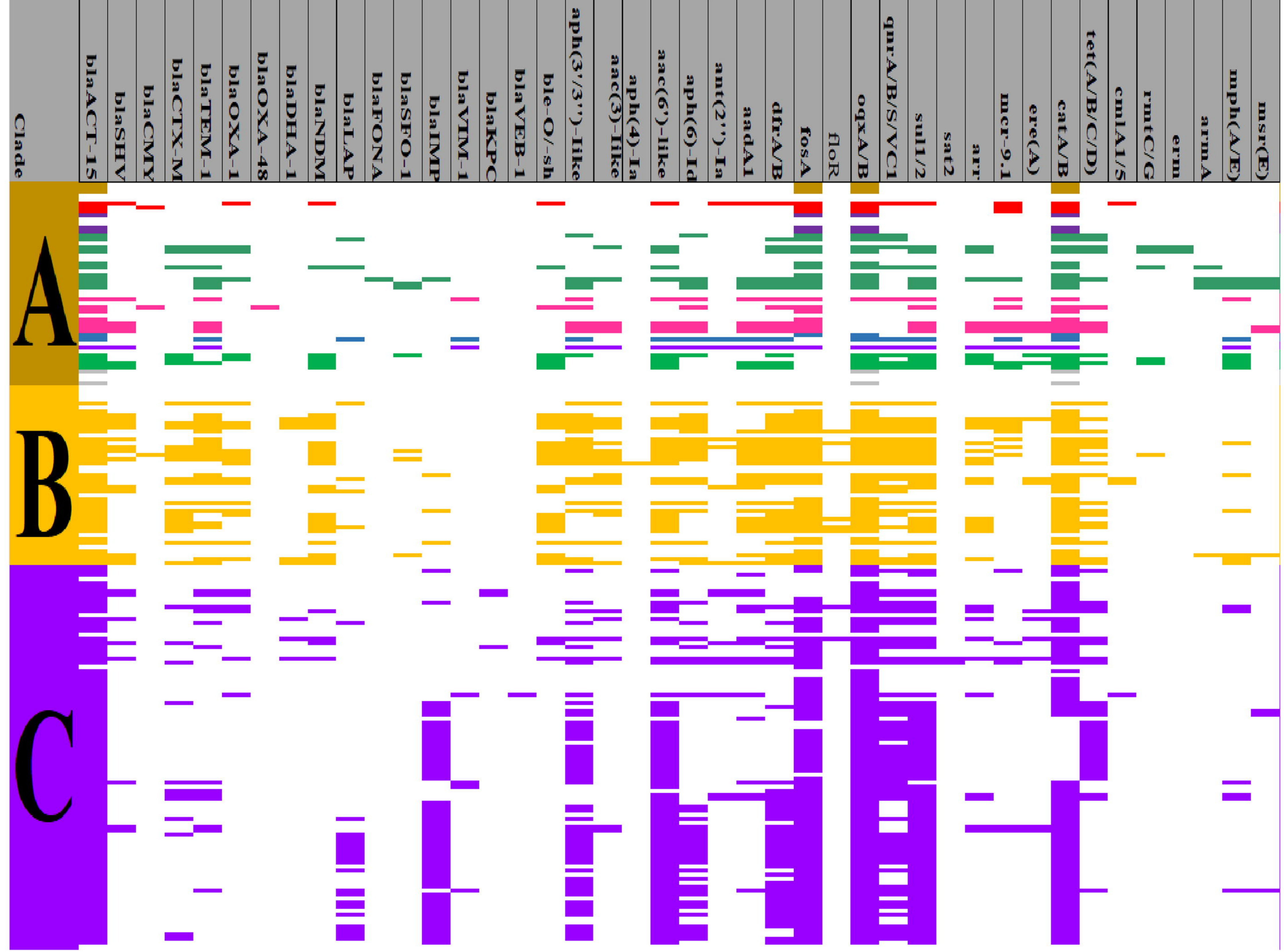
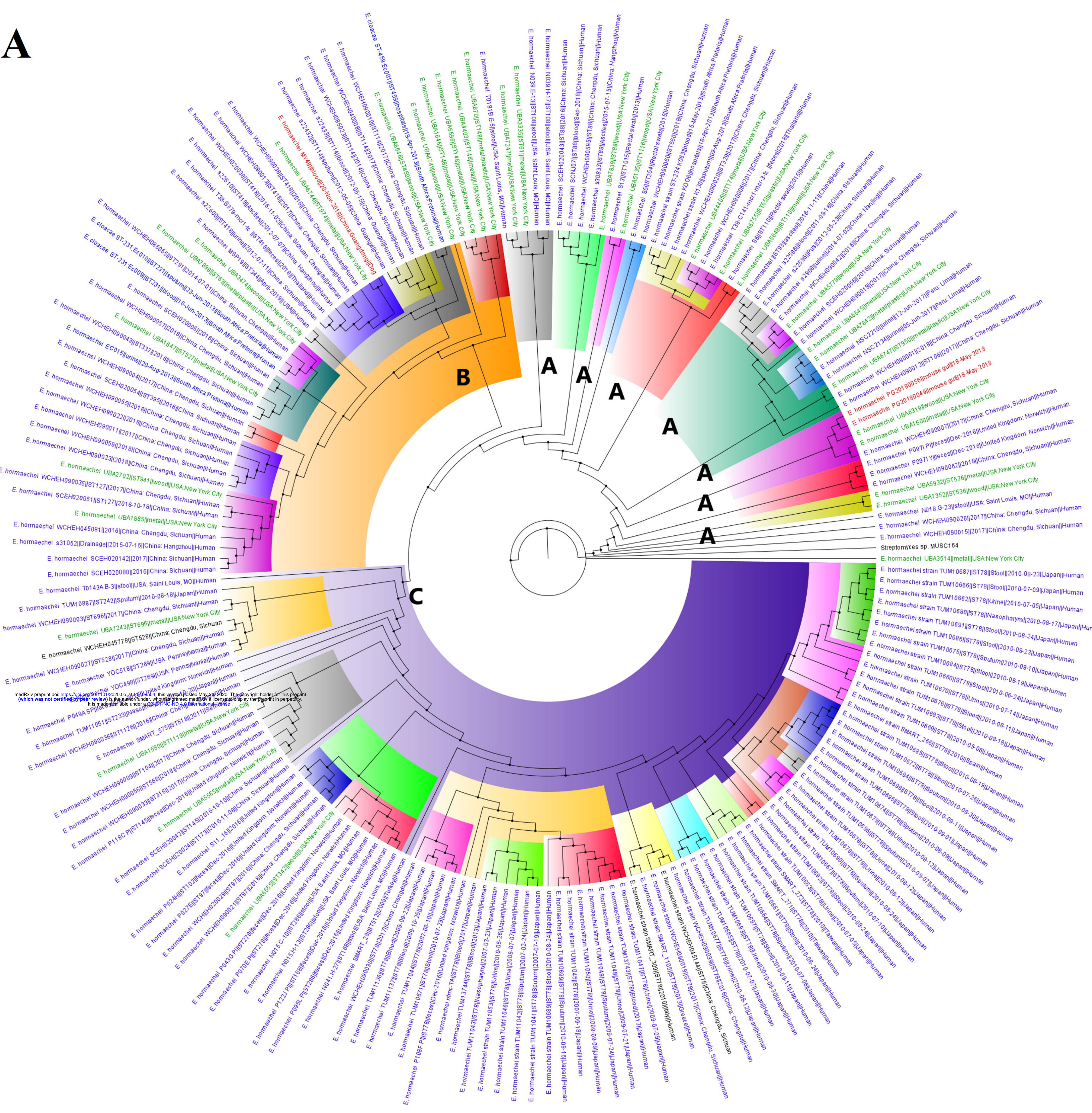


B

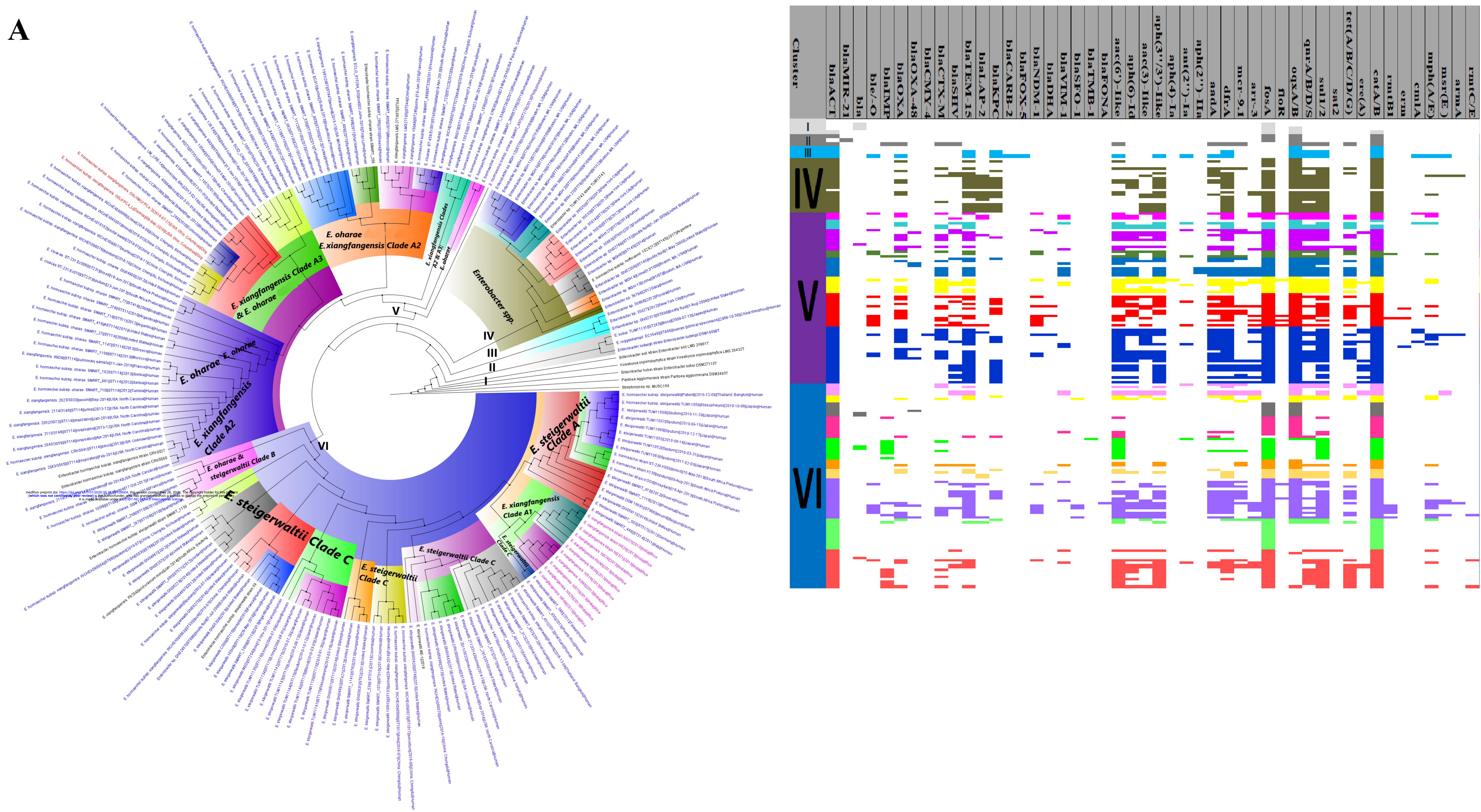


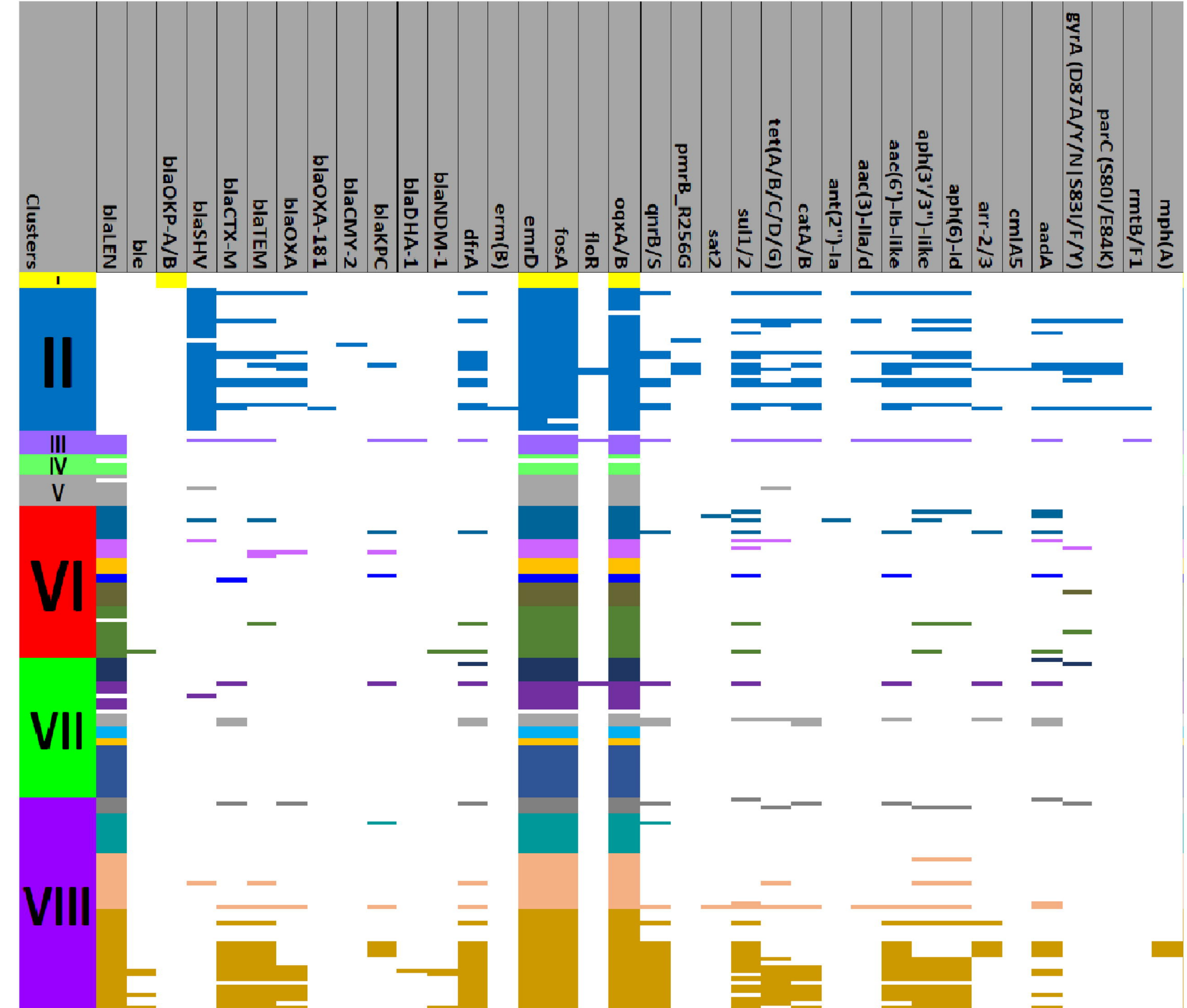
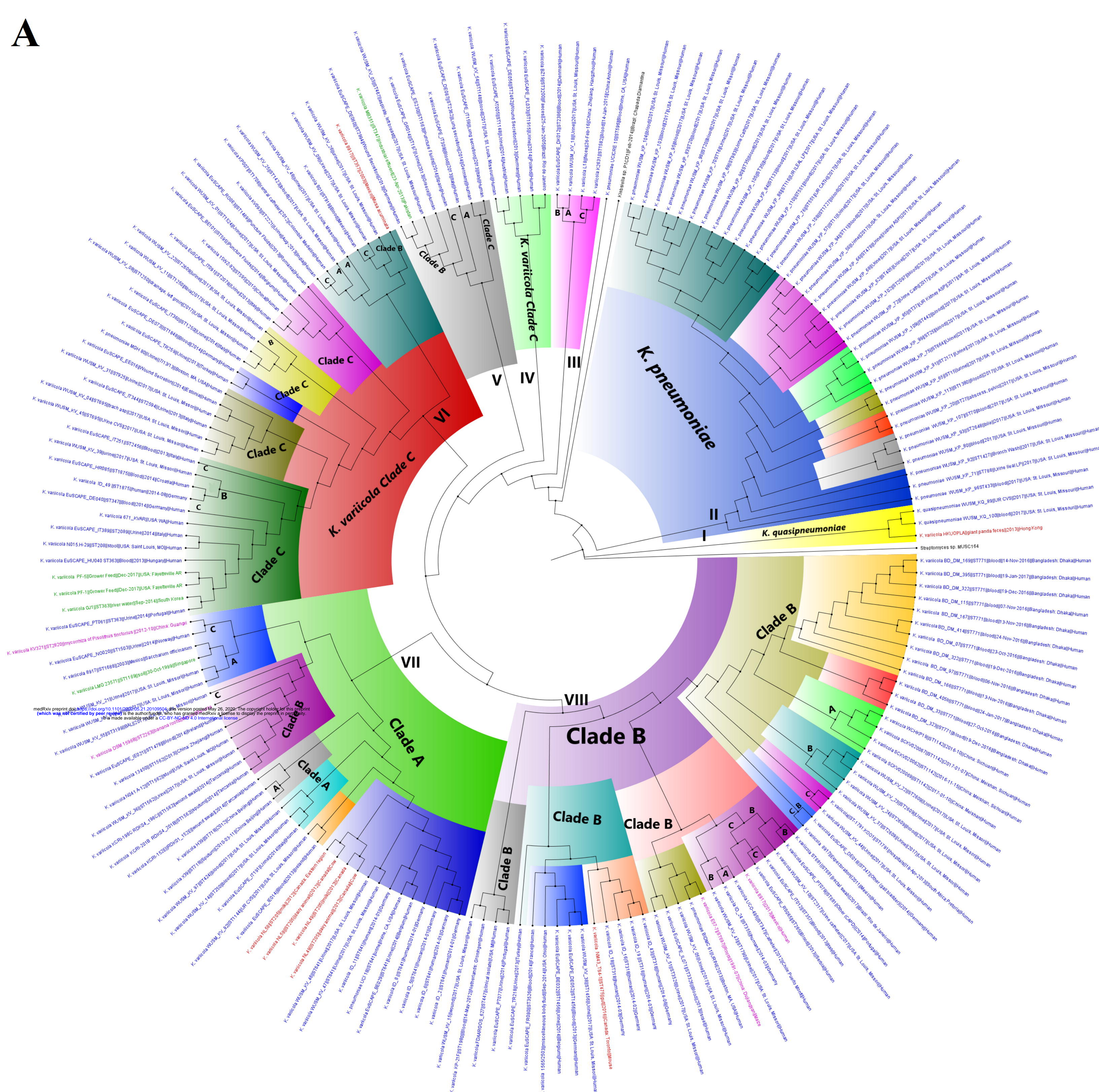
medRxiv preprint doi: <https://doi.org/10.1101/2020.05.21.20109504>; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

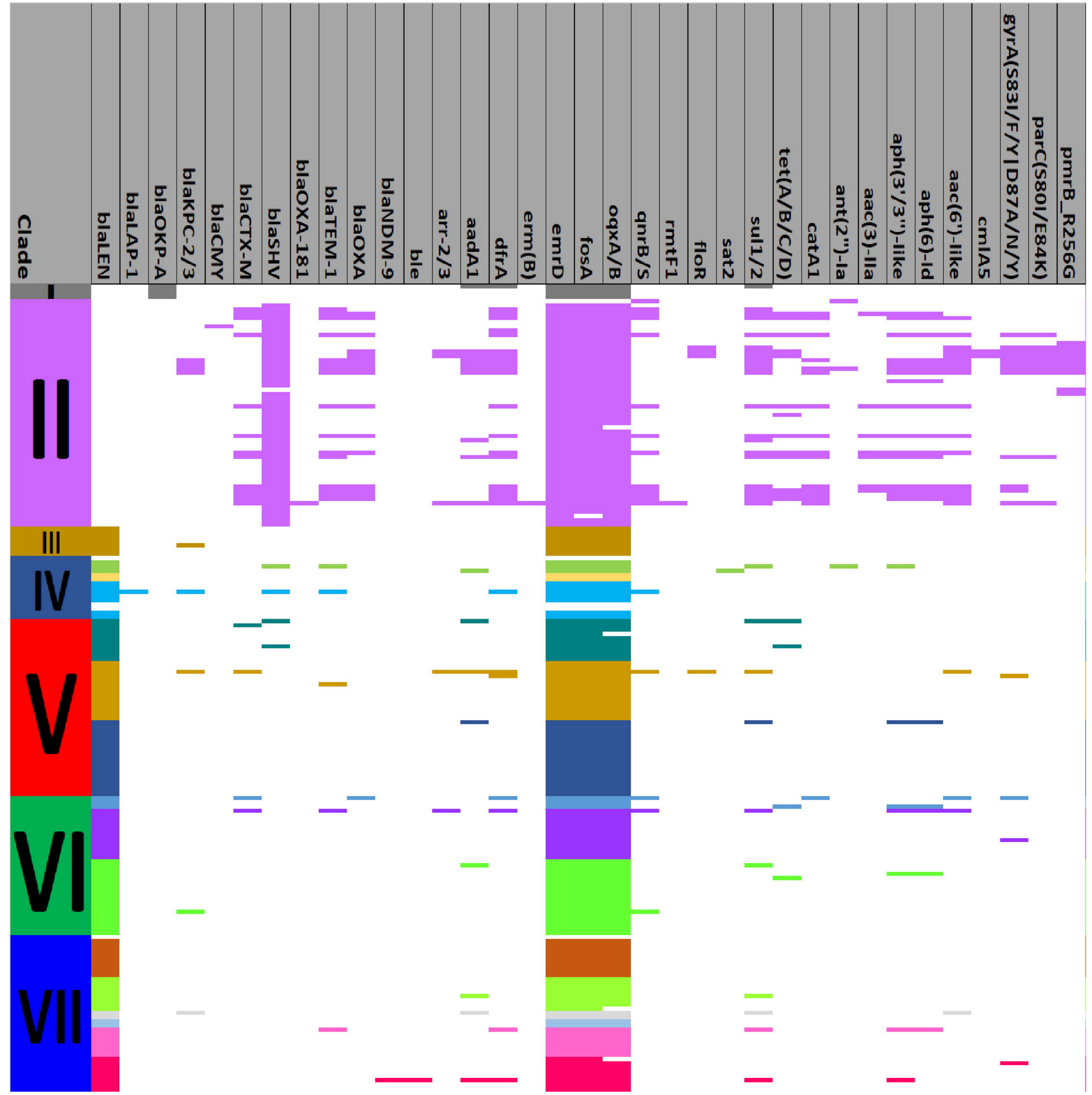
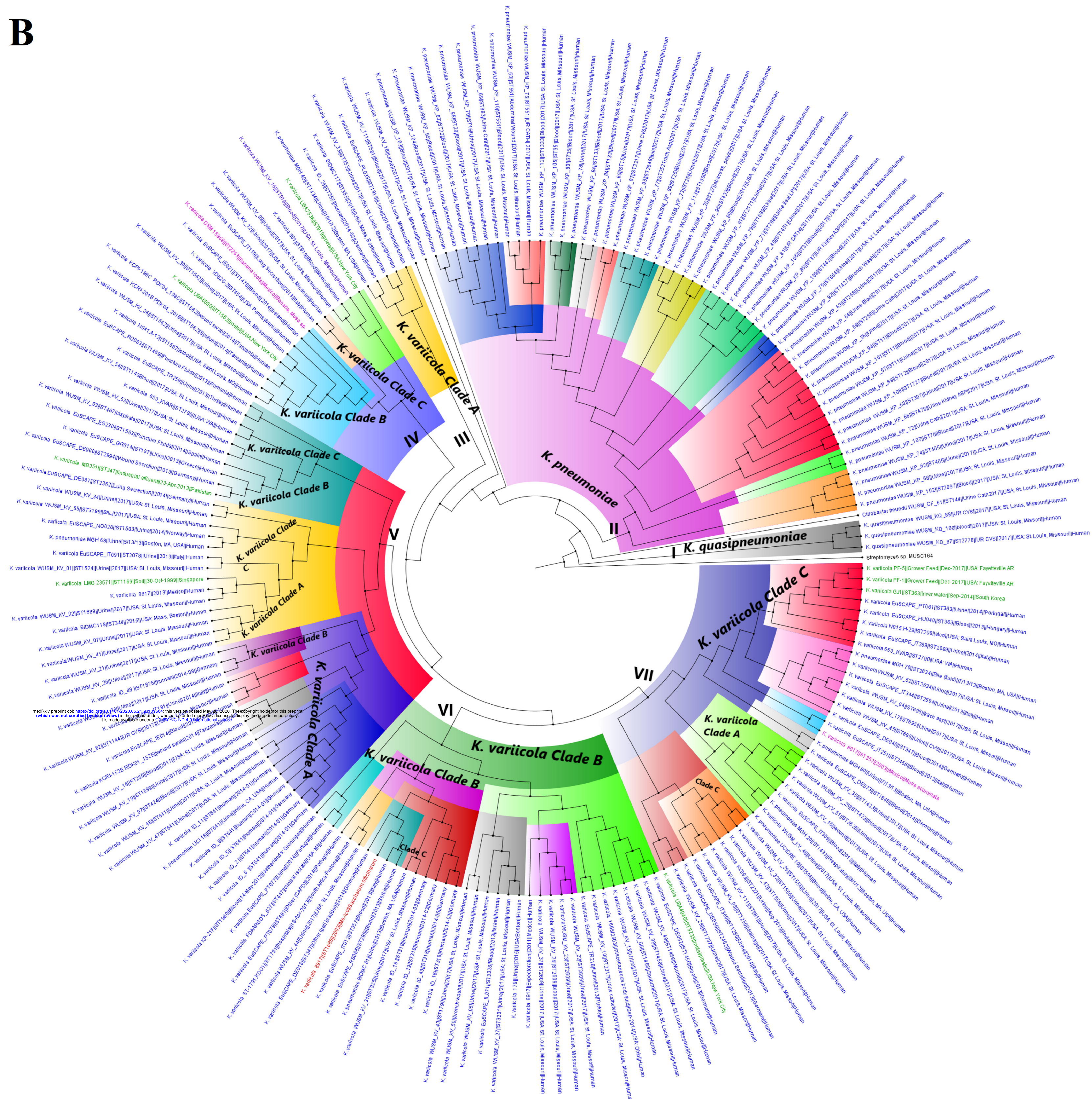
A



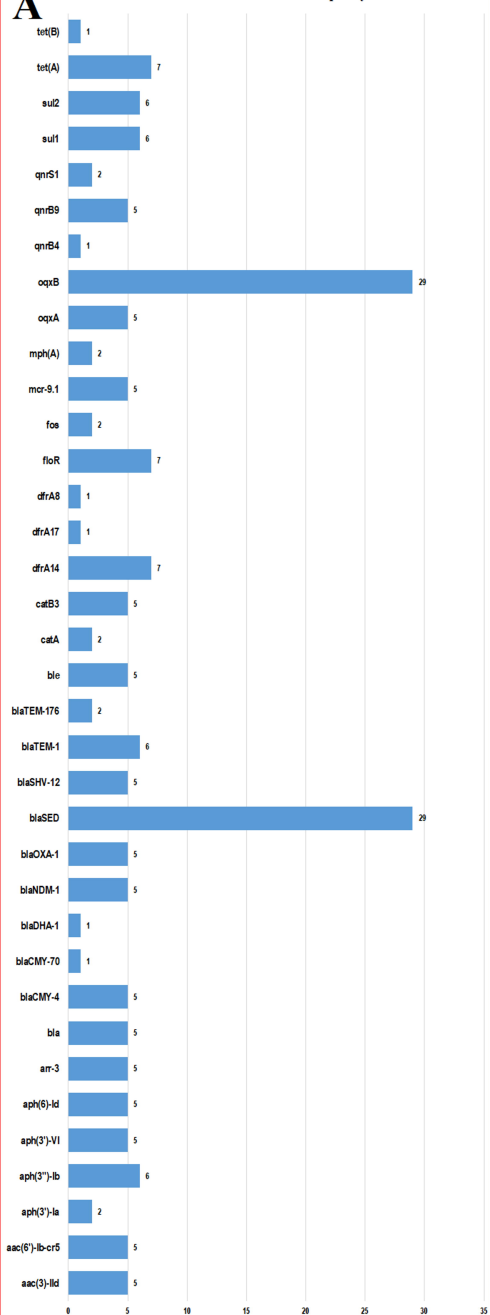
medRxiv preprint doi: <https://doi.org/10.1101/2020.04.24.20091111>; this version posted May 26, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

A

A

B

medRxiv preprint doi: <https://doi.org/10.1101/2020.05.21.20100000>; this version posted May 21, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Citrobacter amalonaticus AMR frequency*C. koseri* AMR frequency