Title:

Microevolution in the major outer membrane protein OmpA of *Acinetobacter baumannii*

Running title: *Acinetobacter baumannii* OmpA microevolution

Authors:

Alejandro M. Viale[a] and Benjamin A. Evans[b#]

Affiliations:

[a]Instituto de Biología Molecular y Celular de Rosario (IBR, CONICET) and Departamento de Microbiología, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario (UNR), 2000 Rosario, Argentina; [b]Norwich Medical School, University of East Anglia, Norwich, UK.

#Corresponding author:

Benjamin A. Evans, Norwich Medical School, University of East Anglia, Bob Champion Research and Education Building (BCRE), James Watson Road, Norwich, NR4 7UQ, UK

Email: Benjamin.evans@uea.ac.uk

Tel: 01603 591 699

Word count – abstract: 250

Word count – importance: 147

Word count – main text: 4891

Abstract

Acinetobacter baumannii is nowadays a relevant nosocomial pathogen characterized by multidrug resistance (MDR) and concomitant difficulties to treat infections. OmpA is the most abundant *A. baumannii* outer membrane (OM) protein, and is involved in virulence, host cell recognition, biofilm formation, regulation of OM stability, permeability, and antibiotic resistance. OmpA members are two-domain proteins with an N-terminal eight-stranded β-barrel domain with four external loops (ELs) interacting with the environment, and a C-terminal periplasmic domain binding non-covalently to the peptidoglycan. Here, we combined data from genome sequencing, phylogenetic, and multilocus sequence analyses from 242 strains of the *Acinetobacter calcoaceticus/Acinetobacter baumannii* complex (ACB), 222 from *A. baumannii*, to explore *ompA* microevolutionary divergence. Five major *ompA* variant groups were identified (V1 to V5) comprising 50 different alleles coding for 29 different proteins. Polymorphisms were concentrated in 5 regions corresponding to the four ELs and the C-terminal end, and provided evidence for different intra-genic recombination events. *ompA* variants were not randomly distributed across the *A. baumannii* phylogeny, with the most frequent V1a1 allele almost exclusive to clonal complex 2 (CC2) strains and the second most frequent V2a1 allele found in the majority of CC1 strains. Evidence was found for assortative exchanges of *ompA* alleles not only between different *A. baumannii* clonal lineages, but also different ACB species. Within *A. baumannii ompA* non-synonymous substitutions were concentrated in the ELs regions, but were more abundant in the transmembrane regions between different *Acinetobacter* species. The overall results have implications for *A. baumannii* evolution, epidemiology, virulence, and vaccine design.

49 **Importance**

50 *Acinetobacter baumannii* is an increasing MDR threat in nosocomial settings associated with
51 prolonged hospitalization and concomitantly increased healthcare costs. The main *A.*
52 *baumannii* OM protein, OmpA, is a multifaceted two-domain protein implicated in host cell
53 recognition and adhesion, cytotoxicity, biofilm formation, and as a slow porin for antibiotics
54 and small hydrophilic nutrients. *A. baumannii* OmpA has been proposed as a potential
55 target for anti-virulence drugs and as a vaccine candidate. Given the many interactions of
56 this protein with environmental factors including host defenses, it is certainly subjected to
57 many selective pressures. Here, we analyzed the microevolution of this OM protein in the *A.*
58 *baumannii* population to obtain clues on the extent to which selection in the clinical setting
59 has shaped this protein. The results provide relevant information on the main causes driving
60 evolution of this protein, with potential implications in *A. baumannii* epidemiology,
61 virulence, and vaccine design.

62

63 **Introduction**

64 The genus *Acinetobacter* (family *Moraxellaceae*, class *Gammaproteobacteria*) is composed
65 of Gram-negative bacterial organisms of ubiquitous environmental distribution (1). It
66 includes a number of phylogenetically closely-related species generally associated with the
67 human environment grouped into the so-called *Acinetobacter calcoaceticus*/*Acinetobacter*
68 *baumannii* complex (ACB). Of these, *A. baumannii* represents the most problematic
69 nosocomial species generally affecting critically-ill patients (1-6). Molecular typing studies
70 have indicated that infective *A. baumannii* strains belong to a limited number of globally-
71 distributed clonal groups/complexes (CCs) of recent emergence (2-9). The reasons for the
72 success of CC strains as human pathogens are still poorly understood, but are thought to

73 include the general resilience of the species, an increased ability to colonize the human
74 environment, and their propensity to acquire exogenous genetic material (1-9). Thus, CC
75 strains not only show multi-drug resistance (MDR) phenotypes from the acquisition of
76 different mobile genetic elements enriched in resistance genes , but can also rapidly modify
77 virulence-related surface structures recognized by host defenses as the result of
78 recombination events involving discrete genomic loci (2-6, 9-12). This has left very few
79 treatment options available (2-6, 10, 13) and has led to the alternative search of targets for
80 anti-virulence drugs or vaccines, and in this context a potential candidate is represented by
81 the major *A. baumannii* outer membrane (OM) protein, OmpA (6, 14-22).
82
83 *A. baumannii* OmpA belongs to the widely distributed OmpA/OprF family of OM proteins
84 from which *Escherichia coli* OmpA and *Pseudomonas* OprF are well-characterized examples
85 (23, 24). These proteins are structured in two domains: an N-terminal eight-stranded β-
86 barrel embedded in the OM exposing four loops to the environment (ELs), and a C-terminal
87 periplasmic domain that binds the peptidoglycan and contributes to cell envelope stability
88 (23-25). *E. coli* OmpA ELs play important roles in target cell recognition and bacterial survival
89 once internalized, and variations in EL2 and EL3 composition or length affects the invasive
90 potential of the bacterial cells (23). Given the many interactions of ELs with environmental
91 factors, it is not surprising that they represent the less conserved protein regions with both
92 mutation and recombination providing for the observed variabilities (23, 26, 27).
93
94 *A. baumannii* OmpA has been implicated in a multitude of functions that range from
95 structural roles, a slow porin for small hydrophilic nutrients and antimicrobials, and
96 virulence including fibronectin binding, biofilm formation, adhesion to host cells, and

97 cytotoxicity (3, 6, 18, 19, 24, 28-36). It has also been shown to interact with host defenses,
98 inducing innate immune responses, the production of antibodies both in *A. baumannii*-
99 infected patients and experimentally-infected mice, and to bind antimicrobial peptides via
100 the ELs (14-22, 37). It follows that there are different selective pressures including trade-off
101 mechanisms (23-27, 38) acting on *A. baumannii* OmpA. However, to our knowledge no
102 systematic studies have been conducted on *A. baumannii* OmpA microevolution. Here, we
103 analyzed the polymorphism of *A. baumannii* OmpA and the roles of mutation and
104 recombination in defining this variability, to analyze the extent to which selection in the
105 clinical setting has shaped this protein.
106

107 **Results**

108 *OmpA polymorphism in the A. baumannii population*

109 *ompA* sequences were extracted from 242 genomes of *Acinetobacter* species deposited in
110 databases, of which 222 were *A. baumannii* (Table S1 and Table S2). A maximum likelihood
111 phylogeny based upon the alignments of the corresponding translated sequences identified
112 five well-defined OmpA variant groups (bootstrap support ≥93 % in all cases), hereafter
113 called V1 to V5 (Figure 1). In *A. baumannii*, each of the identified variant groups was
114 composed of a different number of alleles sharing high levels of amino acid sequence
115 identity (>97%) between them (Table S2). Between these different variants the protein
116 sequence identities ranged from ca. 94 % for V1 and V2 to ca. 85 % between them and the
117 most divergent group V3 (Table S3). An unequal distribution of *ompA* variants was found,
118 with almost 89 % of the *A. baumannii* strains carrying alleles from only V1 and V2 (Table 1,
119 Table S1). Also, *A. baumannii* V1 and V4 variants each encompassed two sub-variant
120 groups designated V1_1 and V4_1, respectively (Table 1, Table S1, Table S2, Figure 2, Figures

121 S1 and S2). Each subgroup could be distinguished by the presence/absence of a 6 amino
122 acid stretch (AAAPAA) close to the C-terminal end, thus defining QEAAAPAAAQ or QQAQ as
123 alternate C-terminal sequences in the corresponding proteins. The origins of this stretch will
124 be discussed in detail below.
125
126 The OmpA V1 group was by far the most frequently observed (63 %) among the *A.*
127 *baumannii* population analyzed (Table 1). It encompassed 19 gene alleles differing by a
128 maximum of 18 nucleotides, translated to 8 different OmpA proteins differing from 1 to 4
129 amino acids. Of note, the *A. baumannii* ATCC type strains 19606 and 17978, isolated from
130 human clinical materials around 1950 (39, 40), display the V1 alleles V1a8 and V1a15,
131 respectively (Table S1). This dates the association of V1 *ompA* with *A. baumannii* clinical
132 strains prior to the emergence of the main epidemic clones in the 1960s (5, 9). A sub-variant
133 of V1, designated V1_1 whose alleles display QAAQ at their C-terminal end, were found
134 among a small percentage (1.8 %) of *A. baumannii* strains (Table 1). Two more divergent
135 V1_1 alleles, as judged by the number of amino acid substitutions as compared to V1a1,
136 were also identified in the non-*baumannii* ACB strains *A. nosocomialis* OIFC021 and *A.*
137 *seifertii* GG2 (Table S2).
138
139 V2 was the second most frequent OmpA variant found in 25.7 % of the *A. baumannii* strains
140 analyzed, with 6 different alleles in *A. baumannii* (Table 1 and Table S1). A seventh and
141 more divergent allele (V2a7) was located in *A. calcoaceticus* ANC3811 (Table S2).
142
143 The OmpA V3 group had a much lower representation (4.1 %) in *A. baumannii* (Table 1),
144 with all these strains showing V3a1 displaying the "short" QQAQ C-terminal sequence

145 (Figures S1 and S2). Five other V3 alleles were detected among various non-*baumannii* ACB
146 species (Table S2). Notably, a subgroup of V3 alleles (designated V3_1) was also found
147 among *A. pittii*, *A. calcoaceticus*, and *A. oleivorans* strains (Table S2).
148
149 The OmpA V4 group had a low representation in *A. baumannii* (3.6 %, Table 1 and Table S1).
150 It was composed of five alleles displaying the "long" QEAAAPAAAQ C-terminus differing by a
151 maximum of 14 nucleotides translated to only 3 different proteins (Table S2). A V4_1 sub-
152 variant displaying the shorter QQAQ C-terminal end was found in the *A. baumannii*
153 DSM30011 strain (Table 1 and Table S2). DSM30011 was isolated before 1944 from a desert
154 plant source and, in agreement with an environmental origin, is susceptible to most
155 clinically-employed antimicrobials including folate synthesis inhibitors (41).
156
157 Finally, V5 had a very low representation in *A. baumannii* (1.35 %, only 3 strains all showing
158 the same V5a1 allele, Table 1). A second V5 allele (V5a2) was found in an *A. pittii* strain
159 (Table S2).
160
161 *A. baumannii OmpA polymorphism*
162 A Shannon entropy plot of all aligned *A. baumannii* OmpA variant alleles (Figure 2A) showed
163 that approximately 60 from a total of almost 357 amino acid positions (17 %) are
164 polymorphic. This polymorphism was concentrated in five well-defined regions, four of
165 them overlapping with the four extracellular loops located at the N-terminal domain (EL1 –
166 EL4) and the fifth at the C-terminal end (Figure 2A). Amino acid (Figure 2A) and nucleotide
167 (Figure 2B) alignments show that the differences in length among variants are due to indels
168 located in EL1, EL2, and EL3, and in the C-terminal end.

169

170 *Selection acting on ompA alleles*

171 Given that V1/V1_1 contains the largest number of *ompA* alleles, and that they are found
172 both in *A. baumannii* and non-*baumannii* strains (Table S2), we decided to examine in detail
173 the nature of the mutations occurring both within *A. baumannii* V1 and V1_1 alleles (19 V1
174 plus 2 V1_1) and also between *Acinetobacter baumannii* and the two non-*baumannii* V1_1
175 alleles (one from *A. seifertii* GG2 and the other from *A. nosocomialis* OIFC 021) (Table S2).
176 Within *A. baumannii* V1/V1_1 alleles (intra-species changes) the total number of
177 synonymous substitutions was around 4-fold higher than that of non-synonymous ones
178 (78% versus 22%) (Table 2, Figure S3), with 6/7 of the non-synonymous substitutions (86%)
179 concentrated in the EL regions. Providing that the ELs encompass around 20% of the total
180 length of the protein, there is a clear bias towards the EL regions carrying non-synonymous
181 substitutions. A detailed analysis of the amino acid changes resulting from these non-
182 synonymous substitutions indicated only conservative amino acid changes (Figure S3).
183 Concerning inter-species changes (*A. baumannii* V1a1 *versus A. seifertii* GG2 V1_1 or *A.*
184 *nosocomialis* OIFC021 V1_1, Table 2), the total number of nucleotide substitutions as
185 compared to intra-*A. baumannii* changes increased a little depending on the non-*baumannii*
186 species considered (40 for *A. seifertii* GG2 V1_1 and 32 for *A. nosocomialis* OIFC021 V1_1, as
187 compared to 31 intra-species). The number of non-synonymous changes inter-species in the
188 EL regions was similar to that found intra-species, except that one amino acid change at EL4
189 (Phe177 for Ala177, Figure S3) could have been positively selected. These comparisons also
190 disclosed a substantial increase inter-species in non-synonymous substitutions in the non-
191 exposed OmpA regions (7 in both cases as compared to none intra-species, Table 2).
192 Moreover, 6 of these changes occurred in the TM regions, and two among them (Phe119 for

193 Ile119 at TM5, and Tyr159 for Trp159 at TM6) are considered disfavored (Figure S3) and
194 could have been positively selected.
195 With the majority of total sequence variation within the different *A. baumannii ompA*
196 variants located within the ELs, we hypothesized that selection may be acting differently
197 upon these regions as compared to the rest of the protein. To test this we investigated the
198 variation in substitution patterns for each of the V groups except for V3 and V5, as these
199 only contained one allele each in *A. baumannii* (Table S2). For V1, V2 and V4 groups when
200 the complete allele alignments are used, a substitution model where parameters are
201 allowed to vary between the exposed EL regions and the non-exposed rest of the protein
202 sequence fits the data significantly better than a fixed model where it is assumed
203 substitution patterns are uniform across the sequence (Table S5). However, when the
204 analysis is repeated using an alignment with recombinant regions removed, only the V2
205 group retains evidence for differences in substitution patterns. This suggests that, while
206 within each V group the EL and non-exposed regions have different substitution patterns
207 indicative of a different evolutionary history, this variation has largely been introduced by
208 recombination.
209
210
211 *Intra-genic recombination at the ompA locus*
212 Detailed comparative analyses of the amino acid and corresponding nucleotide sequences
213 spanning the *A. baumannii* OmpA variant's variable regions (Fig. 2B) suggested a history of
214 intra-genic recombination. The use of RDP4 (42) found evidence for intra-genic
215 recombinatorial exchanges between V3 and V4 at EL1, and also at a longer region between
216 V4 and V5 encompassing the entire EL3 and part of EL4 (Fig. 2B and Fig. S2). Visual

217 inspection additionally detected substantial sequence identity between V4 and V5 along the
218 entire EL2 and between V2 and V4 along EL4 (Fig. 2B). These observations are suggestive of
219 several intra-genic recombinatorial exchanges during the early stages of evolution of the
220 different *ompA* variants.
221
222 Intra-genic recombination may also explain the existence of two alternate C-terminal ends
223 among otherwise very similar OmpA variants (see above). The main difference between
224 them resides in the presence/absence of a hydrophobic tract of 6 amino acids composed
225 mostly by alanine residues close to the C-terminal end (Figure 2B). We noted that the DNA
226 sequence coding this tract, 5′-GCGGCAGCTCCTGCAGCA-3′, shows a higher GC content (72
227 %) when compared to the rest of the *A. baumannii ompA* gene (42 %) or even to the
228 average GC content of *Acinetobacter* genomes (1-6). This suggests that the long
229 QEAAAPAAAQ C-terminal region may have derived from a non-homologous (illegitimate)
230 recombination event involving the insertion of a fragment originating in a high-GC DNA
231 source outside the *Acinetobacter* genus (43-46). Moreover, this event is likely to have been
232 selected recently in *A. baumannii* or in a related ACB complex species, as judged by the
233 presence of OmpAs bearing QEAAAPAAAQ as C-terminal ends only among *A. baumannii*
234 and other ACB species (Table S2) but not the other ecologically-differentiated clades (Table
235 S4) into which the *Acinetobacter* genus was recently divided (1).
236
237 We analyzed possible donor(s) of the 18 nucleotide high-GC fragment among GenBank
238 bacterial genomes by conducting a BLASTN search adjusted for short sequences
239 identification (https://blast.ncbi.nlm.nih.gov). Among *Acinetobacter* sequences this stretch
240 was found only in genomes of *A. baumannii* and some other ACB species, always inserted at

241 the same position near the 3´-end of the *ompA* gene (Fig. 2A, Table S4). Remarkably
242 however, identical or highly similar sequences(17 out of 18 nucleotide identity), which were
243 not linked to *ompA*/*oprF*-related genes, were also found in the genomes of several bacterial
244 species displaying GC contents ≥60% and assigned to different phyla including the
245 *Actinobacteria*, the *Bacteroidetes*, and different subdivisions of the *Proteobacteria* (Table
246 S6). Of note, the above taxa provide not only for the predominant microorganisms in soils
247 but also for the bacterial community associated with hand palms in humans (47). These
248 interconnecting niches, also shared by *Moraxellaceae* members including *A. baumannii* and
249 other ACB species (1, 47), comprise a gene exchange community from which the above
250 high-GC fragment may have been acquired by *A. baumannii* (or other ACB member) *ompA*
251 by illegitimate recombination (44-46). The presence of identical high-GC stretches at similar
252 gene positions in different *A. baumannii ompA* variants and subvariants (Fig. 2A) not only
253 supports the notion that, once acquired by *ompA*, the further dissemination of this
254 particular fragment proceeded by lateral gene transfer and fine-scale homologous
255 recombination (43), but also that it confers some fitness advantage(s) in the human-
256 associated niches in which *A. baumannii* and other ACB species thrive (1).
257
258 *Intra- and inter-species exchange of ompA variants*
259 We investigated the possibility of *ompA* exchange within different lineages of the *A.*
260 *baumannii* population (intra-species recombination) and also between different ACB species
261 (inter-species recombination). The overall observations for a number of representative
262 strains are summarized in Figure 3 (see Table S1 and Table S2 for further details). In
263 agreement with other studies using different sets of core genes and phylogenetic strategies
264 (1, 4, 8, 41), the *A. baumannii* strain cluster is clearly demarcated in the ML tree of Fig. 3

265 from those composed of its phylogenetically closely-related ACB species including *A.*
266 *nosocomialis*, *A. seifertii*, *A calcoaceticus*, *A. oleivorans*, and *A. pittii*, which are in turn well-
267 differentiated between them. Moreover, the *A. baumannii* strains comprising the major
268 global clonal complexes CC1, CC2, and CC3 formed well-differentiated subclusters in this
269 tree, as did other CCs or STs of lower representation.
270 Several additional inferences were derived from the analysis shown in Fig. 3. First, and in
271 sharp contrast to non-*baumannii* ACB species, *ompA* V1 alleles predate the different
272 branches of the *A. baumannii* cluster except for a few epidemiologically-related lineages (e.
273 g., CC3, ST25, CC515 and ST417). This, added to the elevated frequency of V1 alleles among
274 *A. baumannii* clinical strains (Table 1), and their presence already in the ATCC clinical strains
275 17978 and 19606 disclosed above, supports the notion that OmpA V1 variants represent
276 long important determinants in the interaction of *A. baumannii* with the human host.
277 Second, the distribution of *ompA* variants between different *A. baumannii* subclusters
278 revealed non-random associations of particular variants (or variant alleles) and certain CCs
279 (Fig. 3, see also Table S1). For instance, almost all CC2 strains (100 from 101) displayed
280 *ompA* V1a1 with the only exception, ABNIH3, bearing V1a2 differing from V1a1 by a single
281 synonymous substitution (Table S2). This indicated a strong pressure for the preservation of
282 the OmpA protein codified by V1a1 during the global dissemination/expansion of CC2. In
283 addition, V1a1 was found in two other *A. baumannii* strains branching separately from the
284 CC2 cluster: AB_2007-09-110-01-7 and NIPH70 (Fig. 3) thus suggesting their acquisition by
285 lateral gene transfer. AB_2007-09-110-01-7, in particular, is epidemiologically related to
286 three other strains bearing V1_1a2 alleles (CC428, Fig. 3). This additionally supported the
287 notion that this strain acquired V1a1 via a recombinatorial *ompA* exchange in which the
288 donor was probably a CC2 strain. Third, the majority of the CC1 strains analyzed (26 out of

289 31) displayed the V2a1 allele (Fig. 3 and Table S1). These V2a1-containing strains belong to
290 the main (L1) sublineage of CC1, estimated to have separated from the CC1 founding
291 genotype early during the 1970s (9). This also indicates that this V2 allele was strongly
292 preserved during the global dissemination of this CC1 sub-lineage. Two other CC1 strains
293 (clustering together in Fig. 3) showed the V2a6 allele, differing from V2a1 by one nucleotide
294 change translated to a conservative amino acid substitution at EL2 (Table S2). It follows that
295 V2, and in particular V2a1, represents the predominant *ompA* variant associated with CC1.
296 Still, the remaining three CC1 strains (which clustered together in the ML tree of Fig. 3) all
297 showed an allele of a different variant: V1a8. Moreover, these V1a8-containing strains
298 belong to the L2 sub-lineage of CC1 (9) (Table S1), supporting the idea of a complete *ompA*
299 variant replacement most likely occurring very early during CC1 evolution. Fourth, similar
300 links could be observed for other *A. baumannii* clonal lineages, as is the case of V2a1
301 associated with CC3, this V2 allele also with ST417 and ST417-related strains; V1a13 with
302 CC10; V1_1a2 with ST428 and ST428-related strains; and V3a1 with ST25 and ST25-related
303 strains (Fig. 3 and Table S1). In the case of V3a1, its presence also in strains not affiliated to
304 ST25 such as WC-348 and EGD-HP18 (Fig. 3) provided further evidence for the existence of
305 assortative exchanges of *ompA* variants between separate *A. baumannii* lineages.
306
307 Concerning inter-species *ompA* reassortments, evidence for exchanges between *A.*
308 *baumannii* and non-*baumannii* ACB species could also be inferred from the above analysis
309 (Fig. 3 and Table S2). For instance, V1_1 alleles were also detected in *A. nosocomialis* and *A.*
310 *seifertii* strains, a V2a7 allele was found in a strain of *A. calcoaceticus*, V3 alleles were
311 detected in *A. nosocomialis*, *A. calcoaceticus* and *A. pittii* strains, and a V5a2 allele was
312 found in *A. pittii*. On the contrary, V1, V4 and V4_1 alleles were found among *A. baumannii*

313 but not among non-*baumannii* strains and, conversely, V3_1 alleles were found in *A.*
314 *oleivorans*, *A. calcoaceticus*, and *A. pittii* strains but not in *A. baumannii*. These differences
315 may be related to differential role(s) of these OmpA variants in the occupation of the
316 different niches in which the different ACB members thrive (1).
317
318 We also noted that the length of the pre-OmpA signal peptides vary between different ACB
319 members (Table S4). Thus, pre-OmpAs of *A. baumannii*, *A. nosocomialis*, and *A. seifertii* all
320 show identical signal peptides of 22 amino acid residues, while those of *A. calcoaceticus*, *A.*
321 *oleivorans*, and *A. pittii* contain 21 amino acids. Although this could hinder in principle
322 successful exchanges of *ompA* variants between ACB species in which the length of the
323 signal peptide of the pre-protein differs, this seems not to be a critical limitation as judged
324 by the presence of V2 and V3 alleles in *A. baumannii* and also in *A. calcoaceticus*, and V5
325 alleles in *A. baumannii* and also in *A. pittii* (Fig. 3, Table S2). These exchanges certainly
326 required the preservation of the corresponding *ompA* signal peptide coding region in the
327 recipient organism, and provide a further example of fine-scale intragenic recombination
328 operating at the level of the *Acinetobacter ompA* locus, in this case following lateral gene
329 transfer between different ACB species.
330
331 **Discussion**
332 The study of *A. baumannii ompA* microevolution conducted here provides clues on the
333 extent to which selection in the clinical habitat has shaped evolution of this OmpA/OprF
334 family member. We detected five well-defined polymorphic OmpA variants in the
335 *Acinetobacter* population analyzed, with most of the variation occurring within the exposed
336 EL regions at the N-terminal β-barrel domain of these proteins, as well as in the C-terminal

337 region of these proteins (Fig. 2). The EL sectors of surface-exposed bacterial proteins are
338 subjected to strong selective pressures by many environmental factors including host
339 defenses, and thus represent the less conserved regions with both mutation and
340 recombination providing for the observed variabilities (23, 26, 27). Our analyses indicated
341 that lateral gene transfer and recombination, both assortative and fine-scale, appeared as
342 the main agents driving evolution of *A. baumannii* OmpA. We found signs of different fine-
343 scale intra-genic recombination events in *A. baumannii ompA* variants, some encompassing
344 the EL regions which appeared to have occurred during their early stages of divergence from
345 their recent common ancestor. Other events, which appear more subtle and more recent,
346 resulted in different C-terminal sequences in the otherwise highly conserved C-terminal
347 periplasmic domain of OmpA (Fig. 2). The existence of two alternate C-terminal ends
348 differing in an alanine-rich AAAPAA tract not only among different OmpA variants but also
349 within a given variant, as in the case of V1 and V4 (Fig. 2), strongly point to lateral gene
350 transfer and fine-scale homologous recombination affecting the conserved *ompA* 3´-region
351 as the likely cause of this variability. Our analysis additionally suggested that the high-GC
352 DNA fragment encoding this tract represents a recent acquisition resulting from illegitimate
353 recombination occurring in an *A. baumannii* (or a closely-related ACB species) member
354 adapting to human-associated environments. *Acinetobacter* members can take up DNA
355 from environmental sources including soil microcosms, and incorporate short fragments
356 into the genome by homologous as well as non-homologous recombination (44-46). In fact,
357 natural transformation with highly fragmented (≥ 20 bp) and even partially damaged DNA
358 has been shown to be much more efficient than spontaneous mutation to generate
359 adjacent nucleotide polymorphism in the *Acinetobacter* genome (46). The OmpA
360 periplasmic domain contributes to the stability of the cell envelope by interacting non-

361 covalently with the peptidoglycan (25), and its deletion in *A. baumannii* resulted in an
362 increased cell susceptibility to different antimicrobials including cell wall synthesis inhibitors
363 (50). This region of OmpA has been found to associate with different OM and periplasmic
364 proteins including an acquired OXA-23 β-lactamase in the carbapenem-resistant *A.*
365 *baumannii* strain AB5075 (36). It is then tempting to speculate that the insertion of the
366 hydrophobic Ala tract at the OmpA C-terminal end may have resulted in an increased
367 envelope stability towards the various damaging substances used in the nosocomial
368 environment (29, 48, 49), thus providing the driving force for its selection during adaptation
369 of *A. baumannii* to this particular niche.
370
371 Analyses of the substitution patterns in the V1 and V1_1 *ompA* alleles provided further
372 evidence for selection acting upon the protein (Table 2). When looking within *A. baumannii*,
373 non-synonymous substitutions were concentrated in the EL regions of the sequence,
374 suggesting adaptation to the extracellular environment. However, when comparing
375 between *A. baumannii* and other ACB species, more non-synonymous changes were found
376 in the TM regions. If we consider that non-synonymous substitutions provide a measure of
377 adaptive changes to a given environment, then recombinatorial exchanges of V1/V1_1
378 *ompA* alleles between *A. baumannii* and non-*baumannii* species were apparently followed
379 by the rapid selection of adaptive mutations mostly at the level of the TM sectors in the
380 receiving species. This may have resulted from different physicochemical properties of the
381 OM between *A. baumannii* and non-*baumannii* species, and offers another example that the
382 boundaries between ACB species may limit, but not obliterate, inter-species lateral transfer
383 of *ompA* alleles.

384 The different OmpA variants were unequally distributed among the *A. baumannii*
385 population, with V1 alleles being found in around two thirds of the strains analyzed and V2
386 alleles in around one fourth of them (Table 1). We found a strong pressure for the
387 preservation of particular alleles of these two variants during the dissemination and
388 expansion of the major *A. baumannii* global clones CC1 and CC2. Thus, V1a1 *ompA* was
389 found in almost all CC2 strains, while V2a1 was predominant among the main L1 lineage of
390 CC1 (Fig. 3). This suggests that the characteristics evolved by these two OmpA variants
391 confer significant fitness advantages to the strains that compose these epidemic *A.*
392 *baumannii* clonal complexes. Significant genetic diversity is seemingly a hallmark of *A.*
393 *baumannii*, both between epidemiologically- and phylogenetically-closely related strains (2,
394 3, 5, 9, 11, 12). In both CC2 and CC1 strains different whole-genome sequencing studies
395 have indicated that most of this diversity is driven by recombination, with genetic material
396 derived from different co-existing *A. baumannii* lineages (5, 9, 11, 12). Many of these
397 exchanges induced loss or swapping-out of exposed structures potentially recognized by
398 host defenses, such as the capsular polysaccharide, the outer core lipooligosaccharide, iron
399 capture components, efflux pumps, and OM proteins such as CarO (9-12). It is generally
400 assumed that this ongoing process contributes to immune evasion facilitating long-term
401 patient colonization and spread of the *A. baumannii* epidemic clones (9-12). *A. baumannii*
402 OmpA has been found capable of eliciting immune responses including the production of
403 anti-OmpA antibodies both in infected patients (17, 37) and experimentally infected mice
404 (15, 20). It is then noteworthy that, in sharp contrast with the other surface-exposed
405 macromolecules mentioned above, particular *ompA* V1 and V2 alleles have been strongly
406 preserved during the dissemination and expansion of the main *A. baumannii* global clones
407 CC1 and CC2. This preservation supports the notion that these two OmpA variants play

408 pivotal roles in *A. baumannii* pathogenicity (3, 6, 18, 19, 24, 28-36). Their escape from host
409 defenses could result, at least in part, from the recently disclosed shielding by capsule
410 polysaccharide production during the course of *A. baumannii* infection (50). Thus, the
411 observation that particular *ompA* variants, variant alleles, or even *ompA* gene sectors can be
412 effectively exchanged by either assortative or fine-scale recombination not only among
413 different *A. baumannii* lineages but also between different ACB species (Fig. 3) represents a
414 worrying scenario, providing that any innovation evolved by *ompA* affecting host specificity
415 or persistence may be rapidly disseminated among all components of this expanded
416 *Acinetobacter* community.
417
418 Finally, there has been interest in developing a vaccine against *A. baumannii* using OmpA as
419 immunogen (14-20, 37). However, the recombinatorial potential of *ompA* variants shown in
420 this work may pose severe limitations to the effectiveness of a vaccine based solely on a
421 single OmpA variant. It remains to be tested whether polyvalent OmpA vaccines generated
422 through directed protein evolution (51) using the analysis provided in this work could
423 prevent infections due to *A. baumannii* and other ACB strains bearing different *ompA*
424 variants. Further investigations involving functional studies of different OmpA
425 variants/variant alleles may shed light upon the relative strengths of selection acting upon
426 these OM proteins and help the development of new strategies for combating *A. baumannii*
427 and other ACB infections.
428
429 **Materials and Methods**
430 *Sequence retrieval, alignment and phylogeny estimation*

431 The nucleotide sequence of *ompA* from *A. baumannii* ACICU (accession number
432 NC_017162.1, locus tag ABK1_RS15935) was used to query the JGI SIMG/M database
433 (https://img.jgi.doe.gov/) for available *ompA* sequences from *Acinetobacter* genomes.
434 Nucleotide sequences were translated, aligned, and separated into groups of similarity.
435 Based on these groupings, an alignment of translated *ompA* sequences of the 21 identified
436 *A. baumannii* OmpA protein alleles was used to estimate an unrooted maximum-likelihood
437 (ML) phylogenetic tree using PhyML (http://phylogeny.lirmm.fr/phylo_cgi/index.cgi). The
438 assembled *Acinetobacter* strain genomes used were obtained from the NCBI database, and
439 their MLST types (Pasteur scheme) were retrieved from
440 https://pubmlst.org/bigsdb?db=pubmlst_abaumannii_isolates&page=query. A core gene
441 alignment was obtained using Roary (52), and the ML phylogeny was estimated using
442 FastTree v2.1.10, with support estimated from 1,000 resamples (53). For *Acinetobacter*
443 strains where the species designation was uncertain, *rpoB* sequences were used in a BLAST
444 search to confirm the species to which they belong (54).
445
446 *Protein sequence diversity*
447 Variation between *A. baumannii* OmpA variants was determined by calculating the Shannon
448 entropy from alignments of the 21 alleles described here
449 (https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy.html), as described
450 previously for *A. baumannii* CarO (10).
451
452 *Intragenic recombination detection*
453 An alignment containing one of each *A. baumannii ompA* allele was loaded into the
454 Recombination Detection Program v4 (RDP4) (42). An automated recombination analysis

455 was run using the RDP (55), GENCONV (56) and Chimera (57) methods. Regions in which
456 significant evidence for recombination was detected were removed using the 'Save
457 alignment with recombinant regions removed' feature in RDP, and used in subsequent
458 analyses.
459
460 *Tests for selection*
461 To investigate whether natural selection is acting similarly in the *A. baumannii ompA* EL
462 regions as compared to the rest of the protein, we used a partition model in the program
463 codeml implemented in PAML v4.9d (58, 59). Separate alignments for alleles from groups
464 V1, V2 and V4 either including all data, or with recombinant regions removed, were used to
465 estimate trees using PhyML, which then had distance information manually removed for use
466 as input to codeml. Input alignment files were partitioned into sequences corresponding to
467 exposed EL regions and those corresponding to non-exposed TM and periplasmic domains.
468 Two models were fitted to the data in codeml, a null model where parameters values are
469 assumed to be equal between the exposed and non-exposed regions of the sequence
470 (model = 0, NSsites = 0, Mgene = 0), and an alternative model where κ
471 (transition/transversion rate ratio), ω (d_N/d_S), πs (codon frequencies) and cs (proportional
472 branch lengths) are allowed to vary between the two regions (model = 0, NSsites = 0, Mgene
473 = 4). The resulting log likelihood values from the two models were compared using χ^2 tests
474 to determine whether the alternative models better fit the data than the null models. V3
475 and V5 were excluded from this analysis as they only contain one allele each in *A.*
476 *baumannii*.
477
478 **Funding Information**

BAE was funded by the University of East Anglia, and AMV by Agencia Nacional de Promoción Científica y Tecnológica (PICT-2015-1072), Consejo Nacional de Investigaciones Científicas y Tecnicas (CONICET), and Ministerio de Salud, Provincia de Santa Fe, Argentina. AMV is a staff member of CONICET and Professor at the National University of Rosario (UNR), Argentina.

References

1. Garcia-Garcera M, Touchon M, Brisse S, Rocha EPC. 2017. Metagenomic assessment of the interplay between the environment and the genetic diversification of *Acinetobacter*. Environ Microbiol. 19(12):5010-5024. doi: 10.1111/1462-2920.13949.
2. Peleg AY, Seifert H, Paterson DL. 2008. *Acinetobacter baumannii*: emergence of a successful pathogen. Clin Microbiol Rev 21:538-82.
3. Roca I, Espinal P, Vila-Farres X, Vila J. 2012. The *Acinetobacter baumannii* Oxymoron: Commensal Hospital Dweller Turned Pan-Drug-Resistant Menace. Front Microbiol 3:148.
4. Evans BA, Hamouda A, Amyes SG. 2013. The Rise of Carbapenem-Resistant *Acinetobacter baumannii*. Curr Pharm Des 19:223-38.
5. Antunes LC, Visca P, Towner KJ. 2014. *Acinetobacter baumannii*: evolution of a global pathogen. Pathog Dis. 71(3):292-301. doi: 10.1111/2049-632X.12125.
6. Lee CR, Lee JH, Park M, Park KS, Bae IK, Kim YB, Cha CJ, Jeong BC, Lee SH. 2017. Biology of *Acinetobacter baumannii*: Pathogenesis, antibiotic resistance mechanisms, and prospective treatment options. Front Cell Infect Microbiol. 7 (55), 1-35. doi: 10.3389/fcimb.2017.00055
7. Bartual SG, Seifert H, Hippler C, Luzon MA, Wisplinghoff H, Rodriguez-Valera F. 2005. Development of a multilocus sequence typing scheme for characterization of clinical isolates of *Acinetobacter baumannii*. J Clin Microbiol 43:4382-90.
8. Diancourt L, Passet V, Nemec A, Dijkshoorn L, Brisse S. 2010. The population structure of *Acinetobacter baumannii*: expanding multiresistant clones from an ancestral susceptible genetic pool. PLoS One. 5(4):e10034. doi: 10.1371/journal.pone.0010034.
9. Holt K, Kenyon JJ, Hamidian M, Schultz MB, Pickard DJ, Dougan G, Hall R. 2016. Five decades of genome evolution in the globally distributed, extensively antibiotic-resistant *Acinetobacter baumannii* global clone 1. Microb Genom. 2(2):e000052. doi: 10.1099/mgen.0.000052.
10. Mussi MA, Limansky AS, Relling V, Ravasi P, Arakaki A, Actis LA, Viale AM. 2011. Horizontal gene transfer and assortative recombination within the *Acinetobacter baumannii* clinical population provide genetic diversity at the single *carO* gene, encoding a major outer membrane protein channel. J Bacteriol 193:4736-48. doi: 10.1128/JB.01533-10. Epub 2011 Jul 15.

518 11. Snitkin ES, Zelazny AM, Montero CI, Stock F, Mijares L; NISC Comparative Sequence
519 Program, Murray PR, Segre JA. 2011. Genome-wide recombination drives
520 diversification of epidemic strains of *Acinetobacter baumannii*. Proc Natl Acad Sci U S
521 A. 108(33):13758-63. doi: 10.1073/pnas.1104404108.
522 12. Wright MS, Haft DH, Harkins DM, Perez F, Hujer KM, Bajaksouzian S, Benard MF,
523 Jacobs MR, Bonomo RA, Adams MD. 2014. New insights into dissemination and
524 variation of the health care-associated pathogen *Acinetobacter baumannii* from
525 genomic analysis. MBio 5(1):e00963-13. doi: 10.1128/mBio.00963-13.
526 13. Evans BA, Amyes SG. 2014. OXA beta-lactamases. Clin Microbiol Rev 27:241-63. doi:
527 10.1128/CMR.00117-13.
528 14. Kim SA, Yoo SM, Hyun SH, Choi CH, Yang SY, Kim HJ, Jang BC, Suh SI, Lee JC. 2008.
529 Global gene expression patterns and induction of innate immune response in human
530 laryngeal epithelial cells in response to *Acinetobacter baumannii* outer membrane
531 protein A. FEMS Immunol Med Microbiol. 54(1):45-52. doi: 10.1111/j.1574-
532 695X.2008.00446.x.
533 15. Luo G, Lin L, Ibrahim AS, Baquir B, Pantapalangkoor P, Bonomo RA, Doi Y, Adams MD,
534 Russo TA, Spellberg B. 2012. Active and passive immunization protects against lethal,
535 extreme drug resistant-*Acinetobacter baumannii* infection. PLoS One 7:e29446. doi:
536 10.1371/journal.pone.0029446.
537 16. Moriel DG, Beatson SA, Wurpel DJ, Lipman J, Nimmo GR, Paterson DL, Schembri MA.
538 2013. Identification of novel vaccine candidates against multidrug-resistant
539 *Acinetobacter baumannii*. PLoS One 8:e77631. doi: 10.1371/journal.pone.0077631.
540 17. Fajardo Bonin R, Chapeaurouge A, Perales J, da Silva JG, Jr., do Nascimento HJ,
541 D'Alincourt Carvalho Assef AP, Moreno Senna JP. 2014. Identification of
542 immunogenic proteins of the bacterium *Acinetobacter baumannii* using a proteomic
543 approach. Proteomics Clin Appl 8:916-23. doi: 10.1002/prca.201300133.
544 18. Wang N, Ozer EA, Mandel MJ, Hauser AR. 2014. Genome-wide identification of
545 *Acinetobacter baumannii* genes necessary for persistence in the lung. MBio
546 5:e01163-14. doi: 10.1128/mBio.01163-14.
547 19. Mendez JA, Mateos J, Beceiro A, Lopez M, Tomas M, Poza M, Bou G. 2015.
548 Quantitative proteomic analysis of host--pathogen interactions: a study of
549 *Acinetobacter baumannii* responses to host airways. BMC Genomics 16:422. doi:
550 10.1186/s12864-015-1608-z.
551 20. Zhang X, Yang T, Cao J, Sun J, Dai W, Zhang L. 2016. Mucosal immunization with
552 purified OmpA elicited protective immunity against infections caused by multidrug-
553 resistant *Acinetobacter baumannii*. Microb Pathog. 2016 96:20-5. doi:
554 10.1016/j.micpath.2016.04.019.
555 21. Lin MF, Tsai PW, Chen JY, Lin YY, Lan CY. 2015. OmpA binding mediates the effect of
556 antimicrobial peptide LL-37 on *Acinetobacter baumannii*. PLoS One 10:e0141107.
557 doi: 10.1371/journal.pone.0141107.
558 22. Vila-Farrés X, Parra-Millán R, Sánchez-Encinales V, Varese M, Ayerbe-Algaba R, Bayó
559 N, Guardiola S, Pachón-Ibáñez ME, Kotev M, García J, Teixidó M, Vila J, Pachón J,
560 Giralt E, Smani Y. 2017. Combating virulence of Gram-negative bacilli by OmpA
561 inhibition. Sci Rep. 7(1):14683. doi: 10.1038/s41598-017-14972-y.
562 23. Smith SG, Mahon V, Lambert MA, Fagan RP. 2007. A molecular Swiss army knife:
563 OmpA structure, function and expression. FEMS Microbiol Lett 273:1-11.

564 24. Confer AW, Ayalew S. 2013. The OmpA family of proteins: roles in bacterial
565 pathogenesis and immunity. Vet Microbiol 163:207-22. doi:
566 10.1016/j.vetmic.2012.08.019.
567 25. Grizot S, Buchanan SK. 2004. Structure of the OmpA-like domain of RmpM from
568 *Neisseria meningitidis*. Mol Microbiol. 51:1027-1037.
569 26. Pirnay JP, De Vos D, Mossialos D, Vanderkelen A, Cornelis P, Zizi M. 2002. Analysis of
570 the *Pseudomonas aeruginosa oprD* gene from clinical and environmental isolates.
571 Environ Microbiol 4:872-82.
572 27. Hao W. 2013. Unrecognized fine-scale recombination can mimic the effects of
573 adaptive radiation. Gene. 518(2):483-8. doi: 10.1016/j.gene.2012.12.107.
574 28. Sugawara E, Nikaido H. 2012. OmpA is the principal nonspecific slow porin of
575 *Acinetobacter baumannii*. J Bacteriol 194:4089-96. doi: 10.1128/JB.00435-12.
576 29. Iyer R, Moussa SH, Durand-Réville TF, Tommasi R, Miller A. 2018. *Acinetobacter*
577 *baumannii* OmpA is a selective antibiotic permeant porin. ACS Infect Dis. 4(3):373-
578 381. doi: 10.1021/acsinfecdis.7b00168
579 30. Choi CH, Lee JS, Lee YC, Park TI, Lee JC. 2008. *Acinetobacter baumannii* invades
580 epithelial cells and outer membrane protein A mediates interactions with epithelial
581 cells. BMC Microbiol 8:216. doi:10.1186/1471-2180-8-216.
582 31. Choi CH, Hyun SH, Lee JY, Lee JS, Lee YS, Kim SA, Chae JP, Yoo SM, Lee JC. 2008.
583 *Acinetobacter baumannii* outer membrane protein A targets the nucleus and induces
584 cytotoxicity. Cell Microbiol 10:309-19.
585 32. Gaddy JA, Tomaras AP, Actis LA. 2009. The *Acinetobacter baumannii* 19606 OmpA
586 protein plays a role in biofilm formation on abiotic surfaces and in the interaction of
587 this pathogen with eukaryotic cells. Infect Immun 77:3150-60. doi:
588 10.1128/IAI.00096-09. Epub 2009 May 26.
589 33. Lee JS, Choi CH, Kim JW, Lee JC. 2010. *Acinetobacter baumannii* outer membrane
590 protein A induces dendritic cell death through mitochondrial targeting. J Microbiol
591 48:387-92. doi: 10.1007/s12275-010-0155-1. Epub 2010 Jun 23.
592 34. Sánchez-Encinales V, Álvarez-Marín R, Pachón-Ibáñez ME, Fernández-Cuenca F,
593 Pascual A, Garnacho-Montero J, Martínez-Martínez L, Vila J, Tomás MM, Cisneros
594 JM, Bou G, Rodríguez-Baño J, Pachón J, Smani Y. 2017. Overproduction of outer
595 membrane protein A by *Acinetobacter baumannii* as a risk factor for nosocomial
596 pneumonia, bacteremia, and mortality rate increase. J Infect Dis. 215(6):966-974.
597 doi: 10.1093/infdis/jix010.
598 35. Moon DC, Choi CH, Lee JH, Choi CW, Kim HY, Park JS, Kim SI, Lee JC. 2012.
599 *Acinetobacter baumannii* outer membrane protein A modulates the biogenesis of
600 outer membrane vesicles. J Microbiol 50:155-60. doi: 10.1007/s12275-012-1589-4.
601 36. Wu X, Chavez JD, Schweppe D K, Zheng C, Weisbrod CR, Eng JK, Murali A, Lee SA,
602 Ramage E, Gallagher LA, Kulasekara HD, Edrozo ME, Kamischke CN, Brittnacher MJ,
603 Miller SI, Singh PK, Manoil C, Bruce JE. 2016. In vivo protein interaction network
604 analysis reveals porin-localized antibiotic inactivation in *Acinetobacter baumannii*
605 strain AB5075. Nature communications, 7, 13414. doi:10.1038/ncomms13414.
606 37. Nafarieh T, Bandehpour M, Hashemi A, Taheri S, Yardel V, Jamaati H, Moosavi SM,
607 Mosaffa N. 2017. Identification of antigens from nosocomial *Acinetobacter*
608 *baumannii* clinical isolates in sera from ICU staff and infected patients using the
609 antigenome technique. World J Microbiol Biotechnol. 33(10):189. doi:
610 10.1007/s11274-017-2355-x.

611 38. Ferenci T, Phan K. 2015. How porin heterogeneity and trade-offs affect the antibiotic
612 susceptibility of Gram-negative bacteria. Genes (Basel) 6:1113-24. doi:
613 10.3390/genes6041113.
614 39. Schaub, IG, Hauber FD. 1948. A biochemical and serological study of a group of
615 identical unidentifiable gram-negative bacilli from human sources. J. Bacteriol.
616 56:379-385.
617 40. Piechaud M, Second L. 1951. Studies of 26 strains of *Moraxella lwoffi*. Ann. Inst.
618 Pasteur (Paris) 80:97-99.
619 41. Repizo GD, Viale AM, Borges V, Cameranesi MM, Taib N, Espariz M, Brochier-
620 Armanet C, Gomes JP, Salcedo SP. 2017. The environmental *Acinetobacter*
621 *baumannii* isolate DSM30011 reveals clues into the preantibiotic era genome
622 diversity, virulence potential, and niche range of a predominant nosocomial
623 pathogen. Genome Biol Evol. 9(9):2292-2307. doi: 10.1093/gbe/evx162.
624 42. Martin DP, Murrell B, Khoosal A, Muhire B. 2017. Detecting and analyzing genetic
625 recombination using RDP4. Methods Mol Biol 1525:433-460.
626 43. Chan CX, Beiko RG, Darling AE, Ragan MA. 2013. Lateral transfer of genes and gene
627 fragments in prokaryotes. Genome Biol Evol. 1:429-38. doi: 10.1093/gbe/evp044.
628 44. Nielsen KM, Smalla K, van Elsas JD. 2000. Natural transformation of *Acinetobacter* sp.
629 strain BD413 with cell lysates of *Acinetobacter* sp., *Pseudomonas fluorescens*, and
630 *Burkholderia cepacia* in soil microcosms. Appl Environ Microbiol. 66(1):206-12.
631 45. Hülter N, Wackernagel W. 2008. Double illegitimate recombination events integrate
632 DNA segments through two different mechanisms during natural transformation of
633 *Acinetobacter baylyi*. Mol Microbiol. 67(5):984-95. doi: 10.1111/j.1365-
634 2958.2007.06096.x.
635 46. Overballe-Petersen S, Harms K, Orlando LA, Mayar JV, Rasmussen S, Dahl TW, Rosing
636 MT, Poole AM, Sicheritz-Ponten T, Brunak S, Inselmann S, de Vries J, Wackernagel W,
637 Pybus OG, Nielsen R, Johnsen PJ, Nielsen KM, Willerslev E. 2013. Bacterial natural
638 transformation by highly fragmented and damaged DNA. Proc Natl Acad Sci U S A.
639 110(49):19860-5. doi: 10.1073/pnas.1315278110.
640 47. Fierer N, Hamady M, Lauber CL, Knight R. 2008. The influence of sex, handedness,
641 and washing on the diversity of hand surface bacteria. Proc Natl Acad Sci U S A.
642 105(46):17994-9. doi: 10.1073/pnas.0807920105.
643 48. Fernandez-Cuenca F, Tomas M, Caballero-Moyano FJ, Bou G, Martinez-Martinez L,
644 Vila J, Pachon J, Cisneros JM, Rodriguez-Bano J, Pascual A. 2015. Reduced
645 susceptibility to biocides in *Acinetobacter baumannii*: association with resistance to
646 antimicrobials, epidemiological behaviour, biological cost and effect on the
647 expression of genes encoding porins and efflux pumps. J Antimicrob Chemother
648 70(12):3222-9. doi: 10.1093/jac/dkv262.
649 49. Kwon HI, Kim S, Oh MH, Na SH, Kim YJ, Jeon YH, Lee JC. 2017. Outer membrane
650 protein A contributes to antimicrobial resistance of *Acinetobacter baumannii*
651 through the OmpA-like domain. J Antimicrob Chemother. 72(11):3012-3015. doi:
652 10.1093/jac/dkx257.
653 50. Wang-Lin SX, Olson R, Beanan JM, MacDonald U, Balthasar JP, Russo TA. 2017. The
654 capsular polysaccharide of *Acinetobacter baumannii* is an obstacle for therapeutic
655 passive immunization strategies. Infect Immun. 85(12): e00591-17. doi:
656 10.1128/IAI.00591-17.51.

657 51. Wang SN, Cheng ZX, Ling XP, Chu X, Peng XX, Li H. 2018. Construction, immune
658 protection and innate immune response of shuffled polyvalent ompAs vaccines. Fish
659 Shellfish Immunol 34:325-331. doi: 10.1016/j.fsi.2017.12.048
660 52. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D,
661 Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis.
662 Bioinformatics 31(22): 3691-3.
663 53. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2--approximately maximum-likelihood
664 trees for large alignments. PLoS One 5:e9490.
665 54. Gundi VAKB, Dijkshoorn L, Burignat S, Raoult D , La Scola B. 2009. Validation of
666 partial *rpoB* gene sequence analysis for the identification of clinically important and
667 emerging *Acinetobacter* species. Microbiology 155:2333-41.
668 55. Martin D, Rybicki E. 2000. RDP: detection of recombination amongst aligned
669 sequences. Bioinformatics 16:562-3.
670 56. Padidam M, Sawyer S, Fauquet CM. 1999. Possible emergence of new geminiviruses
671 by frequent recombination. Virology 265:218-25.
672 57. Posada D, Crandall KA. 2001. Evaluation of methods for detecting recombination
673 from DNA sequences: computer simulations. Proc Natl Acad Sci U S A 98:13757-62.
674 58. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol
675 24:1586-91.
676 59. Yang Z, Swanson WJ. 2002. Codon-substitution models to detect adaptive evolution
677 that account for heterogeneous selective pressures among site classes. Mol Biol Evol
678 19:49-57.

679

680

681

682

683

684
685
686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703 Table 1: Characteristics of the different OmpA variant (V) groups.

704

V group	No. alleles (genes/ proteins)	No. of *Acinetobacter* isolates (total = 242)	Gene length (nt)	Protein precursor length (aa)	Mature protein length (aa)	C-terminus sequence	% in *A. baumannii* isolates (total = 222)
V1	19/8	140	1,071	356	334	QEAAAPAAAQ	63.0 %
V1_1	4/4	6	1,053	350	328	QQAQ	1.8 %
V2	7/3	58	1,062	353	331	QEAAAPAAAQ	25.7 %
V3	6/5	17	1,029	342	320	QQAQ	4.1 %
V3_1	6/3	7	1,044	347	326	QEAAAPAAAQ	0
V4	5/3	8	1,050	349	327	QEAAAPAAAQ	3.6 %

V4_1	1/1	1	1,032	343	321	QQAQ	0.45 %
V5	2/2	5	1,044	347	325	QQAQ	1.35 %

705

706

707

708

709

710

711

712

713 Table 2: Summary of synonymous and non-synonymous substitutions at polymorphic sites

714 detected between V1/V1_1 *ompA* alleles in the *Acinetobacter* population analyzed. The first

715 1,041 nucleotide positions (equivalent to the first 347 codons) of the different V1/V1_1

716 *ompA* alleles were aligned, and the mutations resulting in synonymous (syn) and non-

717 synonymous (non-syn) changes were calculated. The numbers between brackets denote the

718 corresponding percentages among the 347 codons in each case.

Gene regions	**Total changes**	**Synonymous changes**	**% syn**	**Non-synonymous changes**	**% non-syn**
a)Intra-species changes (19 V1 plus 2 V1_1 *A. baumannii ompA* alleles)					
All	31 (8.9)	24 (6.9)	78	7 (2.0)	22
EL regions	16 (4.6)	10 (2.9)	33	6 (1.7)	19
Non-exposed regions (total)	15 (4.3)	14 (4.0)	45	1 (0.3)	3
TM regions only	8 (2.3)	8 (2.3)	26	0	0
b) Inter-species changes					
b1) A. seifertii GG2 V1_1 *versus A. baumannii* V1a1 *ompA*					
All	40 (11.6)	27 (7.9)	68	13 (3.7)	32
EL regions	16 (4.6)	12 (3.4)	29	4 (1.2)	10

Non-exposed regions	24 (6.9)	15 (4.3)	37	9 (2.6)	22
TM regions only	16 (4.6)	9 (2.6)	22	7 (2.0)	17
b2) A. nosocomialis OIFC021 V1_1 *versus A. baumannii* V1a1 *ompA*					
All	32 (9.2)	20 (5.8)	63	12 (3.4)	37
EL regions	16 (4.6)	11 (3.2)	35	5 (1.4)	15
Non-exposed regions	16 (4.6)	9 (2.6)	28	7 (2.0)	22
TM regions only	15 (4.3)	8 (2.3)	25	7 (2.0)	22

719

720

721

722

723

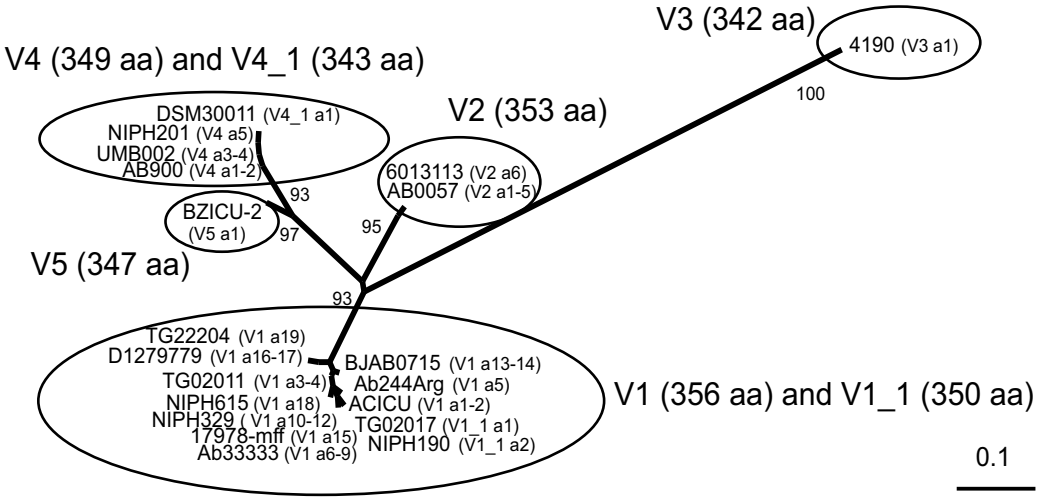
724

725 Figure 1: Co-existence of different OmpA variants and sub-variants in the *A. baumannii*
726 population. An unrooted maximum-likelihood (ML) phylogenetic tree was constructed using
727 PhyML (http://phylogeny.lirmm.fr/phylo_cgi/index.cgi) using alignments of translated *ompA*
728 sequences of the 21 indicated *A. baumannii* isolates representing all OmpA protein alleles
729 identified in this species. The lengths of the branches are proportional to the evolutionary
730 distance, with the scale bar (estimated changes per site) shown at the bottom right.
731 Bootstrap support (percentages of 100 resamplings) for the different clusters are indicated
732 at the corresponding branches. The analysis shows that OmpA proteins comprise 5 well-
733 defined similarity groups of alleles (V1 – V5, indicated by ovals) each defining a particular
734 variant. In addition, V1 and V4 encompass sub-variant groups (designated V1_1 and V4_1,
735 respectively) differing from the corresponding V proteins at their C-terminal motifs. The
736 length in amino acid residues of the corresponding variant and sub-variant proteins is
737 indicated for each cluster.

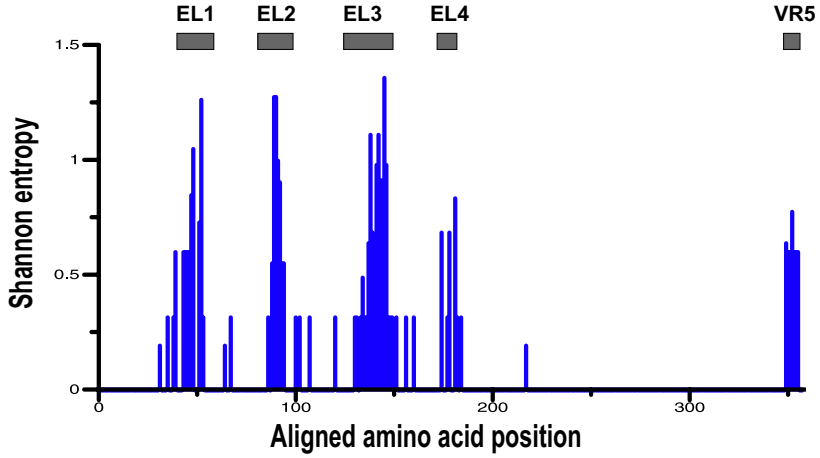
738

739 Figure 2: Sequence variability between *A. baumannii* OmpA variants and sub-variants. (A)
740 Shannon entropy variation along the amino acid alignments of the OmpA variant and sub-
741 variant representative alleles shown in Figure 1. The program available at
742 http://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy.html was used for entropy
743 computations. The heights of the bars are proportional to the degree of amino acid
744 variation at a particular location in the alignments. The span of the predicted external loops
745 EL1 to EL4 at the N-terminal domain and the variable C-terminal motif (VR5) are indicated
746 by closed bars above the figure. Alignments are numbered from the corresponding N-
747 terminal regions including the signal peptide. (B) Amino acid (left columns) and
748 corresponding nucleotide alignments (right columns) of the EL regions and the C-terminal
749 domains of representative *omp*A variant alleles. The amino acid and nucleotide positions
750 encompassing the EL1 to EL4 regions indicated above the sequences are those
751 corresponding to ACICU (V1) OmpA. In VR5 the numbers at the end indicate the position of
752 the last amino acid or nucleotide in the corresponding columns, and the high-GC 5′-
753 GCGGCAGCTCCTGCAGCA-3′ insertion coding for the extra AAAPAA stretch found at the C-
754 terminus in some variant sequences is shadowed in light gray. Uppercase letters in
755 nucleotide alignments indicate the same base at a given position in all sequences. RDP4 (ref.
756 42) detected evidence for recombination between V3 and V4 variants at EL1 and between
757 V4 and V5 variants at a gene region encompassing the entire EL3 and part of EL4 (shadowed
758 in light gray). In addition, visual inspection detected sequence identity between V4 and V5
759 sequences along the entire EL2 (shadowed in light gray), as well as between V2 and V4 along
760 the complete EL4 region (shadowed in dark gray only for V2). The complete amino acid and
761 nucleotide alignments, as well as the topology predictions, are shown in Figs. S1 and S2.
762

763 Figure 3: Evolutionary relatedness and *ompA* alleles of the *Acinetobacter* strains analyzed in
764 this work. An approximately-maximum-likelihood core gene relationships was constructed
765 using 901 core genes (present in ≥99% of strains) with the strain labels colored by species,
766 as shown in the key. The tree is shown without scaling so that the relationships between
767 strains can be seen, and is rooted on the branch separating the *A. baumannii* strains from
768 the other *Acinetobacter* ACB species. Bootstrap support values ≥99% on relevant branches
769 are indicated by red dots. The *ompA* gene alleles are indicated by the two rings of shapes
770 surrounding the tree. The inner ring of shapes indicate the major V-group that the *ompA*
771 allele belongs to (V1, V1_1, V2, V3, V3_1, V4, V4_1 or V5), and the outer ring indicates the
772 specific gene allele. V1 alleles are coloured red, orange and yellow, V2 alleles are blue, V3
773 alleles are green, V4 alleles are pink and purple, and V5 alleles are grey. The outermost
774 circle indicates the clusters of *A. baumannii* strains assigned to the main clonal complexes
775 CC1, CC2, and CC3, as well as other less common clonal complexes and sequence types
776 according to the Pasteur MLST scheme (see Table S1 for details). These CCs evolved after
777 the middle of last century from the clonal expansion each of a highly adaptive independent
778 genotype, and they are generally thought to represent adaptations to specific microniches
779 and be robust with respect to gene choice (5-9).


V3 (342 aa)
4190 (V3 a1)
100
V4 (349 aa) and V4_1 (343 aa)
DSM30011 (V4_1 a1)
NIPH201 (V4 a5)
UMB002 (V4 a3-4)
AB900 (V4 a1-2)
V2 (353 aa)
6013113 (V2 a6)
AB0057 (V2 a1-5)
93
BZICU-2
(V5 a1)
97
95
V5 (347 aa)
93
TG22204 (V1 a19)
D1279779 (V1 a16-17)
BJAB0715 (V1 a13-14)
TG02011 (V1 a3-4)
Ab244Arg (V1 a5)
NIPH615 (V1 a18)
ACICU (V1 a1-2)
NIPH329 (V1 a10-12)
TG02017 (V1_1 a1)
17978-mff (V1 a15)
NIPH190 (V1_1 a2)
Ab33333 (V1 a6-9)
V1 (356 aa) and V1_1 (350 aa)
0.1

A



B

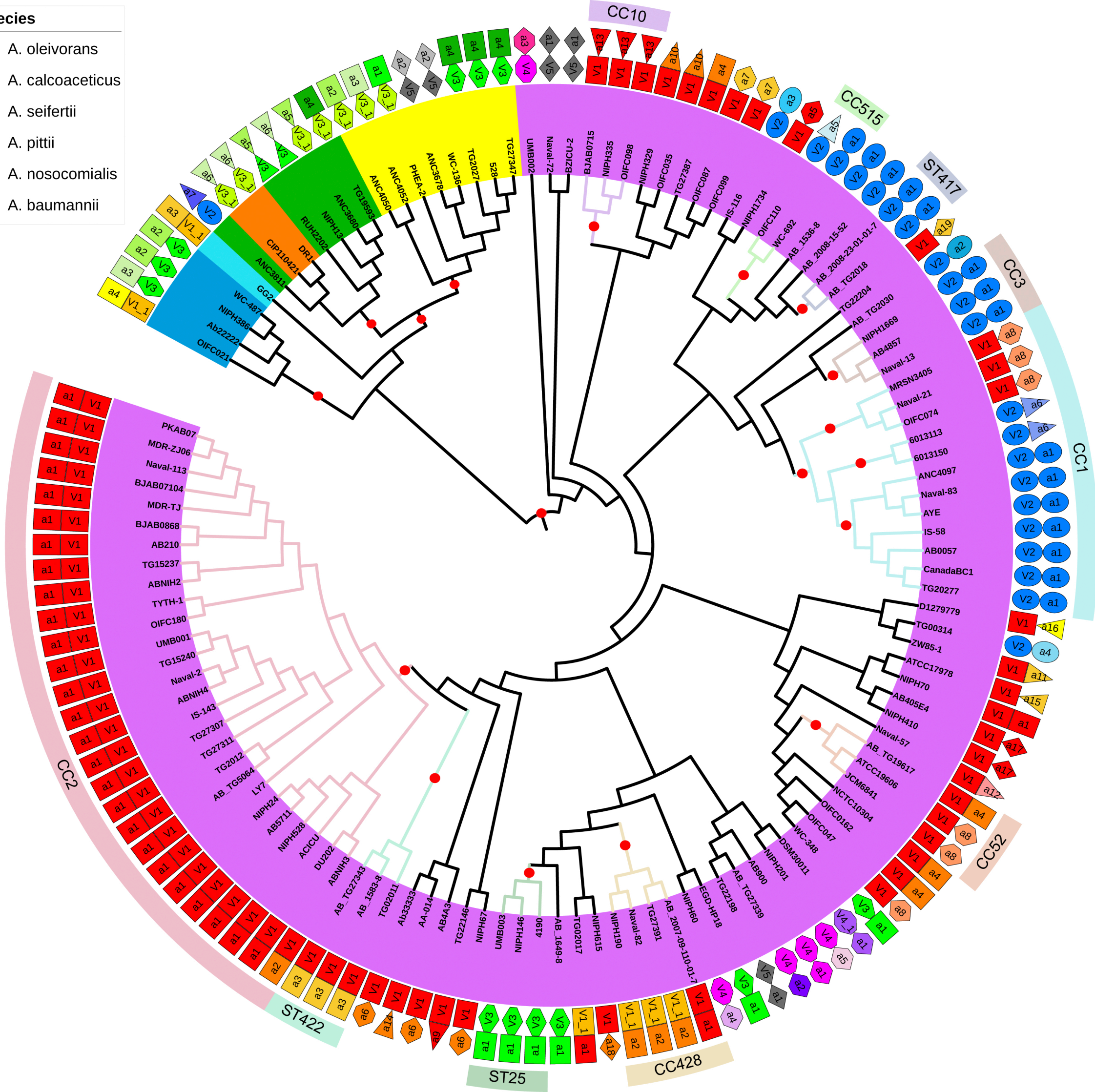
```
EL1                 36                      58   106                                                                    174
ACICU_V1_a1          QDSQHNNGGKDGNLTNSPELQDD      CAAGAcagccAaCAcAACAAtggcggtaaagatggtaacTTaACtaAttcacCtGAgtTaCAAGAcGAT
TG02017_V1_1_a1      QDSQHNNGGKDGNLTNAPELQDD      CAAGAcagccAaCAcAACAAtggcggtaaagatggtaacTTaACtaAtgcacCtGAgtTaCAAGAcGAT
AB0057_V2_a1         QDTQHNNGGKDGELTNGPELQDD      CAAGAcactcAaCAcAACAAtggcggtaaagatggcgagTTaACtaAcggacCtGAatTaCAAGAcGAT
4190_V3_a1           QDSEHNN----NKLTDHAELQDD      CAAGAcagcgAgCAcAACAAc------------aataaaTTaACtgAtcatgCtGAatTgCAAGAtGAT
AB900_V4_a1          QDSEHNN----HKLTDSPELQDD      CAAGAttctgAaCAtAACAAc------------cataaaTTaACtgAtagccCaGAgcTaCAAGAtGAT
DSM30011_V4_1_a1     QDSEHNN----HKLTDSPELQDD      CAAGAttctgAaCAtAACAAc------------cataaaTTaACtgAtagccCaGAgcTaCAAGAtGAT
BZICU-2_V5_a1        QDSQHNNGGKDGSLTNGPELQDD      CAAGAcagccAaCAcAACAAtggcggtaaagatggtagcTTgACaaAtggtcCtGAgtTaCAAGAcGAT

EL2                 80                    100     238                                                                 297
ACICU_V1_a1          NQVKGDVDGASA-GAEYKQKQ        AAcCAAGTtAAAGGcGAcgTaGAcggcgcttctgct---ggtGcTGAaTAtAAaCAaAAAacaa
TG02017_V1_1_a1      NQVKGDVDGASA-GAEYKQKQ        AAcCAAGTtAAAGGtGAcgTaGAcggcgcttctgct---ggtGcTGAaTAtAAaCAaAAAacaa
AB0057_V2_a1         NQVKGDVDGLAA-GAEYKQKQ        AAcCAAGTtAAAGGtGAtgTaGAcggtcttgcagct---ggcGcTGAaTAcAAgCAaAAAacaa
4190_V3_a1           NQVKGDLDGTGVQGAEYKQKT        AAtCAAGTaAAAGGtGAtcTtGAtggtactggcgttcaaggcGcTGAgTAcAAaCAgAAgact
AB900_V4_a1          NQVKGDVD---TNYGEYKQKQ        AAcCAAGTtAAAGGtGAtgTaGAc---------acgaactatGgTGAaTAtAAaCAgAAacaa
DSM30011_V4_1_a1     NQVKGDVD---PNYGEYKQKQ        AAcCAAGTtAAAGGtGAtgTaGAt---------ccaaactacGgTGAaTAtAAgCAgAAacaa
BZICU-2_V5_a1        NQVKGDVD---TNYGEYKQKQ        AAcCAAGTtAAAGGtGAtgTaGAc---------acgaactatGgTGAaTAcAAaCAaAAAacaa

EL3                 127                     151 379                                                                 453
ACICU_V1_a1          GAGHYKYDFDGVNRGTRGTSEEGTL    GGTGCtGGTCActatAAAtatGActTTgaTGgcgtaaatcgtggtacacgtggtacttctgAagAAGgTACTtTa
TG02017_V1_1_a1      GAGHYKYDFDGVNRGTRGTSEEGTL    GGTGCtGGTCActatAAAtacGActTTgaTGgcgtaaaccgtggtacacgtggtacttctgAagAAGgTACTtTa
AB0057_V2_a1         GAGHYKYEIPDL---SYHNDEEGTL    GGTGCtGGTCActacAAAtacGAgaTTccTGacctt---------tcttatcacaacgatgAagAAGgTACTtTa
4190_V3_a1           GAGQTKTEFDGI-----YEDKKDTI    GGTGCgGGTCAaactAAAactGAgtTTgaTGgtatc---------------tatgaagacaAgaAAGaTACTaTc
AB900_V4_a1          GAGHYKYDFDDA-RLAYHDGEEGTL    GGTGCaGGTCAttatAAAtatGAttTTgaTGatgct---cgtttagcttaccatgatggtgAagAAGgTACTtTa
DSM30011_V4_1_a1     GAGHYKYDFDDA-RLAYHDGEEGTL    GGTGCaGGTCAttatAAAtatGAttTTgaTGatgct---cgtttagcttaccatgatggtgAagAAGgTACTtTa
BZICU-2_V5_a1        GAGHYKYDFDDA-RLAYHDGEEGTL    GGTGCaGGTCAttatAAAtatGAttTTgaTGatgct---cgtttagcttaccatgatggtgAagAAGgTACTtTa

EL4                 176       183               526                     549
ACICU_V1_a1          NADEEFWN                     aAtgcTGAtGAAgAgTtCTGGaaC
TG02017_V1_1_a1      NADEEFWN                     aAtgcTGAtGAAgAgTtCTGGaaC
AB0057_V2_a1         NFDEKFWN                     aActtTGAcGAAaAaTtCTGGaaC
4190_V3_a1           NFDEKYWN                     gAtttTGAtGAAaAaTaCTGGcgC
AB900_V4_a1          NFDEKFWN                     aActtTGAcGAAaAaTtCTGGaaC
DSM30011_V4_1_a1     NFDEKFWN                     aActtTGAcGAAaAaTtCTGGaaC
BZICU-2_V5_a1        NFDEQFWN                     aActtTGAtGAAcAaTtCTGGaaC

VR5: C-terminal region
ACICU_V1_a1          TVVVQPGQEAAAPAAAQ 356         ACTGTAGTAGTTCAACCTGGTCAAgAAgcggcagctcctgcagcaGCTCAATAA 1071
TG02017_V1_1_a1      TVVVQPGQQ------AQ 350         ACTGTAGTAGTTCAACCTGGTCAAcAA------------------GCTCAATAA 1053
AB0057_V2_a1         TVVVQPGQEAAAPAAAQ 353         ACTGTAGTAGTTCAACCTGGTCAAgAAgcggcagctcctgcagcaGCTCAATAA 1062
4190_V3_a1           TVVVQPGQQ------AQ 342         ACTGTAGTAGTTCAACCTGGTCAAcAA------------------GCTCAATAA 1029
AB900_V4_a1          TVVVQPGQEAAAPAAAQ 349         ACTGTAGTAGTTCAACCTGGTCAAgAAgcggcagctcctgcagcaGCTCAATAA 1050
DSM30011_V4_1_a1     TVVVQPGQQ------AQ 343         ACTGTAGTAGTTCAACCTGGTCAAcAA------------------GCTCAATAA 1032
BZICU-2_V5_a1        TVVVQPGQQ------AQ 347         ACTGTAGTAGTTCAACCTGGTCAAcAA------------------GCTCAATAA 1044
```


Species
A. oleivorans
A. calcoaceticus
A. seifertii
A. pittii
A. nosocomialis
A. baumannii
CC10
CC515
ST417
CC3
CC1
CC52
CC428
ST25
ST422
CC2
UMB002
Naval-72
BZICU-2
BJAB0715
NIPH335
OIFC098
NIPH329
OIFC035
TG27387
OIFC087
OIFC099
IS-116
NIPH1734
OIFC110
WC-692
AB_1536-8
AB_2008-15-52
AB_2008-23-01-01-7
AB_TG2018
TG22204
AB_TG2030
NIPH1669
AB4857
Naval-13
MRSN3405
Naval-21
OIFC074
6013113
6013150
ANC4097
Naval-83
AYE
IS-58
AB0057
CanadaBC1
TG20277
D1279779
TG00314
ZW85-1
ATCC17978
NIPH70
AB405E4
NIPH410
Naval-57
AB_TG19617
ATCC19606
JCM6841
NCTC10304
OIFC0162
OIFC047
WC-348
DSM30011
NIPH201
AB900
AB_TG27339
TG22198
EGD-HP18
NIPH60
AB_2007-09-110-01-7
TG27391
Naval-82
NIPH190
NIPH615
TG02017
AB_1649-8
4190
NIPH146
UMB003
NIPH67
TG22146
AB4A3
AA-014
Ab333333
TG02011
AB_1583-8
AB_TG27343
ABNIH3
DU202
ACICU
NIPH528
AB5711
NIPH24
LY7
AB_TG5064
TG2012
TG27311
TG27307
IS-143
ABNIH4
Naval-2
TG15240
UMB001
OIFC180
TYTH-1
ABNIH2
TG15237
AB210
BJAB0868
MDR-TJ
BJAB07104
Naval-113
MDR-ZJ06
PKAB07
TG27347
528
TG2027
WC-136
ANC3678
PHEA-2
ANC4052
ANC4050
TG19593
ANC3680
NIPH113
RUH2202
DR1
CIP110421
ANC3811
GG2
WC-487
NIPH386
Ab22222
OIFC021