1   **Adaptive radiations in natural populations of prokaryotes: innovation is key**

2

3   Michiel Vos[1*], Daniel Padfield[1,2], Christopher Quince[3,4], Rutger Vos[5,6]

4

5   1= European Centre for Environment and Human Health, University of Exeter Medical School,

6   Environment and Sustainability Institute, Penryn Campus, TR10 9FE, United Kingdom

7   2= Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall, TR10 9FE, UK

8   3= Organisms and Ecosystems, Earlham Institute, Norwich, NR4 7UZ, UK

9   4= Gut Microbes and Health, Quadram Institute, Norwich, NR4 7UQ, UK

10   5= Naturalis Biodiversity Center, Understanding Evolution, Leiden, 2333 CR, The Netherlands

11   6= Institute of Biology Leiden, Leiden University, Leiden, 2333 CR, The Netherlands

12

13   *= corresponding author m.vos@exeter.ac.uk (0044 (0)7553491614

14

15   Perspective Article

16

19

20

21

22

23

24

25   Abstract

26   Prokaryote diversity makes up most of the tree of life and is crucial to the functioning of the biosphere

27   and human health. However, the patterns and mechanisms of prokaryote diversification have received

28   relatively little attention compared to animals and plants. Adaptive radiation, the rapid diversification

29   of an ancestor species into multiple ecologically divergent species, is a fundamental process by which

30   macrobiological diversity is generated. Here, we discuss whether ecological opportunity could lead to

31   similar bursts of diversification in bacteria. We explore how adaptive radiations in prokaryotes can be

32   kickstarted by horizontally acquired key innovations allowing lineages to invade new niche space that

33   subsequently is partitioned among diversifying specialist descendants. We discuss how novel adaptive

34   zones are colonised and exploited after the evolution of a key innovation and whether certain types of

35   are more prone to adaptive radiation. Radiation into niche specialists does not necessarily lead to

36   speciation in bacteria when barriers to recombination are absent. We propose that in this scenario,

37   niche-specific genes could accumulate within a single lineage, leading to the evolution of an open

38   pan-genome.

39

40   Introduction

41   A central challenge in evolutionary biology and ecology is explaining why species richness patterns in

42   the Tree of Life vary drastically between different taxa (Scholl & Wiens, 2016) (Mooers & Heard,

43   1997). Differences in species richness are evident in many plant and animal sister clades; compare for

44   example the lone species of Hoatzin (Order Opisthocomiformes) with the 5000+ species of passerines

45   (Order Passeriformes). In eukaryotic taxa, such variation in species richness has long been

46   interrogated using analyses of phylogenetic tree shape. However, whether similar heterogeneity exists

47   in Bacteria and Archaea has received less attention (Dykhuizen, 1998). This is partly because the

48   study of bacterial biodiversity faces two major challenges. The first challenge is that most taxa are

49   under-sampled, hindering accurate estimates of species diversity (Quince *et al.*, 2008) and

50   phylogenetic reconstruction (Heath *et al.*, 2008). As a result, estimates of total bacterial diversity vary

ORIGINAL UNEDITED MANUSCRIPT

51 wildly, from ~10^4 (Mora *et al.*, 2011), via ~10^6 (Yarza *et al.*, 2014, Louca *et al.*, 2019), ~10^9 (Larsen

52 *et al.*, 2017) to ~10^{12} species (Locey & Lennon, 2016). Of course, estimates of species richness at

53 least to some extent rely on how species are defined in the first place. The second challenge is that

54 there is no one-to-one agreement between current taxonomy, species delineated based on overall

55 genomic distance, or operational taxonomic units based on clustering of 16S rRNA sequences (Parks

56 *et al.*, 2018)). Differential sampling effort and inconsistent taxonomy must mean that some of the

57 observed inter-taxon differences in bacterial species richness must be artefactual. These caveats

58 notwithstanding, it is clear that there are substantial differences in species richness when surveying

59 either named species or 16S amplicon-based Operational Taxonomic Units (OTUs) (Figure 1).

60 Numerous explanations for differences in species richness have been put forward but many of these,

61 such as the effect of trophic level, body size, geographic range, latitude, or temperature (Hutchinson,

62 1959, Rosenzweig, 1995, Dykhuizen, 1998), do not necessarily translate to prokaryotes (e.g. (Bahram

63 *et al.*, 2018)). However, reasoning from first principles, species richness, be it in animals, plants, or

64 bacteria, is ultimately the product of speciation and extinction adding and subtracting species over

65 time. Taxa with a higher net diversification rate (i.e., a higher rate of speciation than extinction) are

66 expected to have higher species richness. It is possible that different clades with identical

67 diversification rates still differ in species richness, as older clades will have had more time to

68 accumulate new species (Figure 2A).

69 Diversification can proceed at a constant rate, but can also occur in pulses (or sporadic declines).

70 Bursts in diversification ('rapid cladogenesis') are commonly ascribed to the exploitation of

71 ecological opportunity (Schluter, 2000, Gavrilets & Losos, 2009) (Figure 2B). Such adaptive

72 radiations are contingent on two main conditions: first, many niches must be available (or one large

73 niche space that can be partitioned), and second, only few lineages must be in a starting position to fill

74 them (i.e., competition must be relaxed). Laboratory experiments have demonstrated that frequency-

75 dependent competition for niche space can drive adaptive radiations in bacteria. In a seminal

76 experiment, *Pseudomonas fluorescens* predictably diversified into three types over the course of only

77 a few days when incubated in static flasks, with wrinkly spreaders inhabiting the broth-air interface,

78   fuzzy spreaders occupying the bottom of the flask, and the ancestral smooth morph residing in the

79   broth [5].

80   Phylogenetic methods offer ways to uncover bacterial diversification on much larger (geological)

81   timescales. They often rely on PCR amplification and sequencing of the conserved 16S ribosomal

82   marker from environmental samples serving as proxies for species or on higher-resolution

83   concatenated core genes sequenced from isolated strains. These studies indicate that bacterial

84   speciation rate is slightly higher than extinction rate (Loren *et al.*, 2014, Marin *et al.*, 2016, Louca *et*

85   *al.*, 2018) (but see (Martin *et al.*, 2004)), consistent with results for multicellular organisms where

86   turnover of taxa is high and where most diversity is now extinct (Louca *et al.*, 2018). Some studies

87   have uncovered bursts in diversification rate in (sections of) bacterial phylogenies (Morlon *et al.*,

88   2012, O'Dwyer *et al.*, 2015). As 16S-based datasets have limited power to detect diversification on

89   shallower evolutionary time scales (Louca *et al.*, 2018) and studies using higher resolution markers

90   generally survey only a relatively limited number of taxa, such burst-like evolution could be present

91   but overlooked in other studies.

92   The aim of this paper is to examine the evidence for bursts of adaptive evolution in prokaryotes and

93   their evolutionary and ecological drivers, and how these compare to those in macroscopic species. We

94   will discuss how differences in diversification rate between prokaryotes could affect other aspects of

95   bacterial biology such as the evolution of pan-genomes. Although highly insightful, lab experiments

96   are generally performed on extremely short timescales that rely solely on mutation (and seldomly

97   incorporate Horizontal Gene Transfer (HGT), a central driver of genomic and functional diversity in

98   bacteria) and are based on purely artificial selection pressures in the absence of other community

99   members. We therefore will focus on natural populations in this review and refer to other literature

100   summarising results on experimental adaptive radiations in bacteria (Travisano & Rainey, 2000, Craig

101   MacLean, 2005). We will review studies on isolates assigned traditional taxonomic labels, 16S

102   amplicons, and closely related clusters based on whole-genome sequences.

103

104   <u>Key innovations spur adaptive radiations in bacteria</u>

105   In macrobes, the open niche space that forms a prerequisite for adaptive radiations is often provided

106   by rare colonisation events of remote localities such as mountains, lakes, or islands, where competing

107   species are absent. Classic examples of such adaptive radiations include Darwin's finches in the

108   Galapagos, Cichlid fishes in East African Rift Lakes and Silversword plants in Hawaii (Schluter,

109   2000). This scenario is not likely in bacteria, as they experience little dispersal-limitation due to their

110   small size and high abundance, meaning niche specialists and niches will be efficiently matched. This

111   diminished role of biogeographical barriers and allopatry in prokaryotes (and a correspondingly

112   increased role for environmental filtering) is illustrated by many 16S-based studies (Lozupone &

113   Knight, 2005); for instance, most global soil diversity was found to be contained in an area as small as

114   Central Park in New York City (Ramirez *et al.*, 2014). A recent large-scale analysis of curated

115   genomes from around the globe found that most prokaryotic clades on Earth's surface are globally

116   distributed (Louca, 2022). Consistent with an earlier housekeeping gene-based study demonstrating

117   geographical divergence in a thermophile archaeon [10], thermophiles were found to be least

118   dispersive, which makes sense as they live in relatively small, specialised habitats that are far apart

119   (Louca, 2022). However, neither study could conclude that even extremophile species displayed

120   endemicity. There seems to be no bacterial equivalent of marsupials, and it is ecological opportunity -

121   rather than geographic isolation - that is most likely to drive bacterial diversification (Vos, 2011). The

122   oft-quoted adage "everything is everywhere, the environment selects" thus seems to be vindicated by

123   sequencing-based studies almost a century after it was first proposed (Baas Becking, 1934).

124   How could adaptive radiations occur in sympatry? One pathway to ecological innovation that is not

125   reliant on geographical isolation was developed by Miller, Mayr and Simpson in the middle of the 20th

126   century (Heard & Hauser, 1995, Schluter, 2000). These and other scientists posited that occasional

127   evolutionary 'key innovations' give rise to entirely new functional capabilities that allow the

128   colonisation of new 'adaptive zones' (Hunter, 1998, Alfaro, 2014) (Figure 2C). Such adaptations

129   could provide a release from competition and access to niche space not available before. A well-

130   known example in animals is the radiation of Notothenioid fishes in the Antarctic Ocean. The

131 evolution of antifreeze glycoproteins that lower internal freezing point in their last common ancestor

132 has allowed the invasion of comparatively empty oceanic regions with sub-zero temperatures and the

133 subsequent diversification into over 130 species (Matschiner *et al.*, 2015).

134 It could be argued that prokaryotes have an especially great potential to evolve key innovations, as

135 HGT allows the wholesale acquisition of entirely novel functional traits originating from other strains

136 and species (Lawrence, 2001, Cohan & Koeppel, 2008, Hall *et al.*, 2017). One population genomics

137 study beautifully uncovered a radiation of bacterial niche specialists driven by HGT (Hehemann *et al.*,

138 2016). In previous work, the same group had identified multiple genetically distinct *Vibrio* clusters

139 that were hypothesised to be ecologically differentiated, as they were enriched in different particle

140 size fractions in the same seawater samples (Hunt *et al.*, 2008). Subsequent genome sequencing

141 uncovered that the brown algal glycan alginate pathway had undergone extensive combinatorial

142 changes mediated by HGT within and between these clusters as well as more distantly related species,

143 leading to rapid clade diversification (Shapiro *et al.*, 2012). Subsequent growth rate experiments

144 demonstrated that variation in enzyme type, copy number and localisation (on the cell wall or

145 broadcast into the environment) translated into physiological differences, which in turn could explain

146 the differential association of different types with particle size (representing different degradable algal

147 cell wall types) and season (Hehemann *et al.*, 2016). This case bears all the hallmarks of an adaptive

148 radiation mediated via a key innovation.

149 Another example of an adaptive radiation driven by an HGT-acquired key innovation is offered by the

150 Thaumarchaeota, an abundant Archaeal phylum that plays a major role in the global nitrogen cycle,

151 specifically via the oxidation of ammonia. Environmental pH is a major factor affecting the

152 distribution of different Thaumarchaeota clades (Gubry-Rangin *et al.*, 2011). Phylogenetic methods

153 could show that a radiation occurred early in the evolution of the Thaumarchaeota, allowing niche

154 expansion from neutral pH environments to acidic and alkaline environments (Gubry-Rangin *et al.*,

155 2015). Interestingly, diversification rate remained high after this initial burst, which is not consistent

156 with typical adaptive radiations, where an initial high diversification is followed by a slowdown (a

157 signature also observed in adaptive radiations inferred in bacteria (Morlon *et al.*, 2012)). pH

158 adaptation in Thaumarchaeota is at least in part mediated by V-type ATPase proton pumps (Wang *et*

159 *al.*, 2019). The phylogeny of acidophile V-type-like ATPase operons in Thaumarchaeota is

160 incongruent with organismal phylogeny but is congruent with habitat, indicating that HGT is

161 responsible for ATPase-mediated niche adaptation (Wang *et al.*, 2019).

162 Ecological opportunity for adaptive radiations can be provided by abiotic factors such as resource

163 type or pH as in the case studies above. But as prokaryotes are generally embedded in highly diverse

164 and dense communities of competitors, parasites, prey, predators, hosts, symbionts and mutualists,

165 biotic factors must be highly relevant too. As different organisms can co-evolve with each other,

166 selection exerted by other organisms is not only likely to be strong, but also long lasting and

167 potentially diversifying (Van Valen, 1973). A meta-analysis on 16S diversity collected across many

168 different biomes found that the diversity of specific lineages correlated positively with whole-

169 community diversity (Madi *et al.*, 2019). This observation is consistent with more diverse

170 communities offering more available niche space through more diverse biotic interactions. It could

171 also be shown that this relationship was weaker for the most diverse communities, indicating that

172 when niches are increasingly filled, there is less opportunity for invading lineages to diversify

173 (Hehemann *et al.*, 2016, Madi *et al.*, 2019).

174

175 Entry into novel environments: adaptive zones

176 High dispersal rates mean that available niches are generally filled by the appropriate niche

177 specialists. However, it also means that there is frequent immigration of taxa that are not (well)

178 adapted to the local environment. The vast majority of such immigrants are unlikely to persist, let

179 alone diversify (Madi *et al.*, 2019). However, if an ecologically and genomically distinct migrant

180 manages to take up a niche-defining gene from the local community, it could be in a position to

181 occupy (or create) hitherto unexploited niche space and give rise to an adaptive radiation. An example

182 of one of the most drastic environmental transitions for metazoans and prokaryotes alike is that

183 between marine and terrestrial (including freshwater) environments (Cohan & Koeppel, 2008,

184 Logares *et al.*, 2009). Salinity is a major determinant structuring microbial diversity, with distinct

185 phylogenetic shifts observed over salt gradients (Dupont *et al.*, 2014, Fortunato & Crump, 2015).

186 Successful marine-terrestrial transitions require significant rewiring of central metabolism and

187 osmotic stress responses (Eiler *et al.*, 2016), which could be aided by HGT (Wisniewski-Dye *et al.*,

188 2011). Phylogenetic analyses indicate marine-terrestrial transitions occasionally occur in bacterial

189 taxa (Zhang *et al.*, 2019) and it can be argued these form an excellent model for the colonisation of

190 novel adaptive zones (Jurdzinski *et al.*, 2023).

191 Another example of the colonisation of novel adaptive zones is offered by pathogens switching host.

192 *Staphylococcus aureus* infects a wide range of vertebrates (and even invertebrates) (Matuszewska *et*

193 *al.*, 2020)). Host jumps are frequent and result in distinct genetic clusters where strains carry specific

194 host-adaptive genes, and evidence loss of host-adaptive genes associated with their previous host

195 (Matuszewska *et al.*, 2020). Specifically, different host specialists are characterised by the carriage of

196 different combinations of Mobile Genetic Elements, including genes known to target specific host

197 innate immune responses and antimicrobial resistance genes conferring resistance to antibiotics used

198 in particular husbandry regimes (Richardson *et al.*, 2018, Matuszewska *et al.*, 2020)). This further

199 exemplifies the pervasive role of HGT in opening up new niches, although it is not clear whether

200 MGEs are generally acquired just before or after host-switching events (Richardson *et al.*, 2018).

201 Major new microbial niches have originated throughout earth's history, from the emergence of

202 oxygenic habitats allowing aerobic respiration to the evolution of animal and plant hosts (Jaffe *et al.*,

203 2023). Such niches range from 'closed' with purely vertical transmission (as those occupied by

204 endosymbionts) to 'open' with mainly horizontal transmission (as those occupied by planktonic

205 marine bacteria). Some horizontal transfer needs to occur to allow the colonization of novel adaptive

206 zones, but it is not clear whether migration rates must be very high to allow rare key innovations to

207 occur, or if they need to be at some intermediate level to prevent establishment of the best currently

208 adapted species, in turn preventing the opportunity of a new best-adapted lineage to evolve).

209

210   Generalists as progenitors of adaptive radiations

211   Prokaryotes can be classified as specialists or generalists based on the broadness of their niche

212   requirements (Bell & Bell, 2021). Bacteria with larger genomes and higher metabolic versatility are

213   associated with greater niche width (Barberán *et al.*, 2014). Living in a wider range of microbiomes

214   means that such generalist species will encounter more distinct selection pressures as well as interact

215   with more species that could serve as donors of key adaptations through HGT. A large-scale meta-

216   analysis of 16S sequence data found that 16S OTUs present across a greater number of distinct

217   habitats (likely to be generalists) was found to have a 19-fold higher speciation rate than OTUs

218   present in only a single habitat (likely to be specialists) (Sriswasdi *et al.*, 2017). That generalist-to-

219   specialist transitions are more common than vice-versa, is consistent with increasing specialisation

220   resulting in the closing of doors leading to other ecological lifestyles, which is consistent with results

221   from lab experiments on bacteria (Buckling *et al.*, 2003).

222

223   Are some taxa inherently more prone to adaptively radiate?

224   Speciation rate is dependent on ecological opportunity, but also on the rate at which new niche-

225   defining traits can arise. Taxa that are more evolvable (Díaz Arenas & Cooper, 2013) thus could be

226   expected to be in a better position to radiate into novel types. Species-specific variation in factors such

227   as mutation rate, generation time and population size all influence the rate of adaptation to new

228   niches, but a high frequency of HGT specifically can be expected to facilitate the evolution of key

229   innovations (Lawrence, 2001).

230   High rates of HGT mediated by Gene Transfer Agents (GTAs; exapted bacteriophages that function

231   to secrete host DNA) have been implicated in a well-documented case of a bacterial adaptive radiation

232   (Guy *et al.*, 2013). *Bartonella* are vectorborne, intracellular pathogens of mammals comprising

233   multiple species-level clades. Two clades with similar host range display evidence of increased

234   diversification, and both could be shown to have independently taken up the VirB type IV secretion

235   system (T4SS) which acts to inject virulence factors into host cells (Engel *et al.*, 2011). All ancestral

236 strains harboured a GTA capable of *in vitro* gene transfer (Guy *et al.*, 2013); interestingly, the GTA is

237 co-located with the T4SS genes which results in a higher-than-average chance of being secreted and

238 taken up by other cells (Tamarit *et al.*, 2018). It has been proposed that his coupling of niche-defining

239 genes and genes increasing recombination has allowed the successful diversification of this pathogen

240 genus (Guy *et al.*, 2013).

241 It is important to stress that HGT transfers do not necessarily lead to adaptive radiations when they do

242 not increase functional diversity or when ecological opportunity is absent. For instance, hybridisation

243 events where donor DNA replaces up to 20% of the recipient genome have been observed in a variety

244 of human pathogens (Chen *et al.*, 2014, Croucher & Klugman, 2014) without concomitant

245 diversification. Moreover, it is possible that high rates of HGT could impede, rather than promote

246 adaptive radiations. One of the very few studies that has discussed the concept of key adaptations in

247 the context of prokaryotes has argued that HGT hinders adaptive radiations, because it could result in

248 key adaptations being transferred to many lineages rather than just a single one (Martin *et al.*, 2004).

249

250 <u>Adaptive radiations with and without speciation: implications for pan genome evolution</u>

251 HGT in bacteria, like meiotic sex in eukaryotes, is a double-edged sword: on one hand it is central to

252 creating genetic diversity, but on the other hand it can impede genetic divergence of nascent niche

253 specialists. Without some ecological or genetic barrier to HGT, diversification cannot proceed to the

254 species-level (Shapiro & Polz, 2014). It is possible that many adaptive radiations in prokaryotes are

255 'stuck' on the strain-level because there are no barriers to recombination allowing speciation to occur

256 (Figure 2D) (Shapiro & Polz, 2014). As a consequence, there could be unappreciated links between

257 ecological diversification, recombination barriers, and the evolution of pan genomes (Figure 3).

258 The evolution and ecology of pan genomes, the total complement of genes within a species which is

259 usually much larger than the number of genes in any individual genome, is a topic of great interest in

260 evolutionary microbiology (Bobay, 2020, Domingo-Sananes & McInerney, 2021). Several distinct,

261 non-mutually exclusive hypotheses have been put forward to explain the existence of pan genomes.

262 Some explanations invoke adaptive benefits where different gene repertoires correspond to

263 differential niche specialisation (Domingo-Sananes & McInerney, 2021). Other explanations invoke

264 neutral processes; some species might be more prone to take up genes by HGT because their genomes

265 are more accommodating to novel genetic diversity or because they are surrounded by a higher

266 diversity of community members (Brockhurst *et al.*, 2019). Greater effective population size is

267 expected to result in greater pan genome diversity (Andreani *et al.*, 2017), specifically via retainment

268 of accessory genes with near-neutral fitness effects (Bobay & Ochman, 2018).

269 However, there is another potential explanation of why pan genome size can vary among species,

270 which is directly linked to diversification. Every time a new niche specialist evolves and

271 recombination with the ancestor ceases, the niche specialists start with a 'minimal' pan genome

272 (Figure 3A). Although this pan genome will increase in size during the lifetime of a species through

273 adaptive processes (e.g., diversifying selection), non-adaptive processes (e.g., the uptake of parasitic

274 mobile genetic elements) and neutral processes, it will be small initially. In contrast, if recombination

275 barriers are absent, for instance when different genotypes remain in close physical contact, new niche

276 specialists still evolve, but their core genes will remain tied together through continued recombination

277 (Shapiro & Polz, 2014). Clade-specific accessory genes will remain part of the pan genome, which

278 will grow with the evolution of each new niche specialist (Figure 3B). *Escherichia* might fit this latter

279 scenario: species numbers in this genus are low and *E. coli* has a famously large pan genome. In this

280 scenario, *E. coli* displays an evolutionary 'shallow' adaptive radiation where niche specialists are

281 unable to evolve into species (Figure 2D)).

282

283 <u>Discussion and Conclusions</u>

284 Adaptive radiations have been implicated in bursts of species richness in animals and plants, and

285 multiple high quality case studies have demonstrated that they also occur in bacteria. However, the

286 study of adaptive radiations in prokaryotes is still in its infancy and many questions remain to be

287 answered. For instance, are certain genetic (e.g., restriction/modification systems) or ecological

288 characteristics (e.g. type of metabolism or microbiome) especially conducive to the radiation of

289 lineages? Are particular traits unlikely to give rise to adaptive radiations because they are especially

290 prone to horizontal spread and unlikely to transfer to a single lineage? Do key adaptations come as

291 single genes or operons or can they be more complex, involving many genes, such as in the evolution

292 of cell walls (Cohan & Koeppel, 2008)? Could some radiations be started by purely mutational

293 processes rather than HGT, as has been shown experimentally with the mutational evolution of citrate

294 metabolism in *E. coli* (Blount *et al.*, 2008)? Are some clades species-rich because they are old rather

295 than having undergone burst-like evolution?

296 Generalisation of patterns and processes between very different organisms and lifestyles is a main

297 challenge (Gillespie *et al.*, 2020)). We would argue that bacterial diversification does not differ

298 qualitatively from that in macrobes but only quantitatively. In other words: 'prokaryotes also disperse,

299 adapt, recombine and speciate, just to different extents'. HGT allows the uptake of complete operons

300 from different species and could increase the likelihood of key innovations. This effect is likely much

301 more pervasive but not wholly different from hybridisation events preceding adaptive radiations in

302 eukaryote species (Seehausen, 2004). When genome-wide HGT remains ongoing between

303 differentially adapted lineages this means that adaptive radiations cannot proceed and will not result

304 in increased species richness, but rather highly diverse 'strain flocks'. The same process has been

305 observed in sticklebacks, where speciation also occurs along a continuum, including repeated and

306 reversible specialisation and reproductive isolation (Hendry *et al.*, 2009)). Arguably the most

307 pronounced difference between prokaryotes and multicellular organisms is that environmental

308 filtering is much more important than dispersal limitation.

309 In-depth genomic and ecological knowledge on species and ecotypes will be necessary to identify

310 patterns of increased diversification, links to ecological niches, barriers to recombination and specific

311 key innovations. As in all fields of microbiology, the way we study bacterial diversification depends

312 greatly on technological advances. Increasing sequencing power allows for the routine use of

313 Metagenome Assembled Genomes (MAGs) (Parks, et al. 2017; Bickhart, et al. 2020). Ancient DNA

314 (Wibowo *et al.*, 2021), HGT transfers (Davín *et al.*, 2018) and bacteria-eukaryote associations (Wang

315 & Luo, 2021) all can help explicitly date radiations and aid in the reconstruction of ancestral states.

316 Despite technical and computational challenges, it could be argued it is actually easier to study

317 adaptive radiations in bacteria, as vicariance is less important relative to selection. In addition,

318 genomic diversification is more tractable compared to macrobes and experiments can be designed to

319 test the ecological function of genes under controlled lab conditions. Experimental evolution studies

320 incorporating multiple species and allowing HGT (e.g. (Hall *et al.*, 2016)) are a crucial way forward

321 to study diversification. We look forward to more high-resolution genomic studies of natural

322 populations examining the interplay between ecology, evolution and genetics that ultimately leads to

323 diversification of clades, genomes, pan genomes and microbial communities.

324

328

329 **References**

330 Alfaro ME. Key evolutionary innovations. *The Princeton guide to evolution (eds DA Baum, DJ*

331   *Futuyma, HE Hoekstra, RE Lenski, AJ Moore, CL Peichel, D Schluter, MC Whitlock)* 2014:

332   594-600.

333 Andreani NA, Hesse E, Vos M. Prokaryote genome fluidity is dependent on effective population size.

334   *ISME J* 2017; **11**: 1719-1721.

335 Baas Becking LGM. Geobiologie of inleiding tot de milieukunde. *WP Van Stockum & Zoon NV, Den*

336   *Haag, The Netherlands* 1934.

337 Bahram M, Hildebrand F, Forslund SK *et al.* Structure and function of the global topsoil microbiome.

338   *Nature* 2018;**560**: 233-7.

339 Barberán A, Ramirez KS, Leff JW *et al.* Why are some microbes more ubiquitous than others?
340   Predicting the habitat breadth of soil bacteria. *Ecol Lett* 2014;**17**: 794-802.

341 Bell TH, Bell T. Many roads to bacterial generalism. *FEMS Microbiol Ecol* 2021;**97**: fiaa240.

342 Blount ZD, Borland CZ, Lenski RE. Historical contingency and the evolution of a key innovation in
343   an experimental population of *Escherichia coli*. *Proc Natl Acad Sci USA* 2008;**105**: 7899-906.

344 Bobay L-M. The prokaryotic species concept and challenges. *The Pangenome* 2020: 21-49.

345 Bobay L-M, Ochman H. Factors driving effective population size and pan-genome evolution in
346   bacteria. *BMC Evol Biol* 2018;**18**: 1-12.

347 Brockhurst MA, Harrison E, Hall JP *et al.* The ecology and evolution of pangenomes. *Curr Biol*
348   2019;**29**: R1094-R103.

349 Buckling A, Wills MA, Colegrave N. Adaptation limits diversification of experimental bacterial
350   populations. *Science* 2003;**302**: 2107-9.

351 Chen L, Mathema B, Pitout JD *et al.* Epidemic *Klebsiella pneumoniae* ST258 is a hybrid strain. *mBio*
352   2014;**5**: e01355-14.

353 Cohan FM, Koeppel AF. The origins of ecological diversity in prokaryotes. *Curr Biol* 2008;**18**:
354   R1024-R34.

355 Croucher NJ, Klugman KP. The emergence of bacterial "hopeful monsters". *mBio* 2014;**5**: e01550-14.

356 Davín AA, Tannier E, Williams TA *et al.* Gene transfers can date the tree of life. *Nature Ecol Evol*
357   2018;**2**: 904-9.

358 Díaz Arenas C, Cooper TF. Mechanisms and selection of evolvability: experimental evidence. *FEMS*
359   *Microbiol Rev* 2013;**37**: 572-82.

360 Domingo-Sananes MR, McInerney JO. Mechanisms that shape microbial pangenomes. *Trends*
361   *Microbiol* 2021.

362 Dupont CL, Larsson J, Yooseph S *et al.* Functional tradeoffs underpin salinity-driven divergence in

363   microbial community composition. *PLoS One* 2014;**9**: e89549.

364 Dykhuizen DE. Santa Rosalia revisited: why are there so many species of bacteria? *Antonie Van*

365   *Leeuwenhoek* 1998;**73**: 25-33.

366 Eiler A, Mondav R, Sinclair L *et al.* Tuning fresh: radiation through rewiring of central metabolism in

367   streamlined bacteria. *ISME J* 2016;**10**: 1902.

368 Engel P, Salzburger W, Liesch M *et al.* Parallel evolution of a type IV secretion system in radiating

369   lineages of the host-restricted bacterial pathogen Bartonella. *PLoS Genet* 2011;**7**: e1001296.

370 Fortunato CS, Crump BC. Microbial gene abundance and expression patterns across a river to ocean

371   salinity gradient. *PLoS One* 2015;**10**: e0140578.

372 Gavrilets S, Losos JB. Adaptive radiation: contrasting theory with data. *Science* 2009;**323**: 732-7.

373 Gillespie RG, Bennett GM, De Meester L *et al.* Comparing adaptive radiations across space, time, and

374   taxa. *J Hered* 2020;**111**: 1-20.

375 Gubry-Rangin C, Hai B, Quince C *et al.* Niche specialization of terrestrial archaeal ammonia

376   oxidizers. *Proc Natl Acad Sci USA* 2011;**108**: 21206-11.

377 Gubry-Rangin C, Kratsch C, Williams TA *et al.* Coupling of diversification and pH adaptation during

378   the evolution of terrestrial Thaumarchaeota. *Proc Natl Acad Sci USA* 2015;**112**: 9370-5.

379 Guy L, Nystedt B, Toft C *et al.* A gene transfer agent and a dynamic repertoire of secretion systems

380   hold the keys to the explosive radiation of the emerging pathogen Bartonella. *PLoS Genet*

381   2013;**9**: e1003393.

382 Hall JP, Brockhurst MA, Harrison E. Sampling the mobile gene pool: innovation via horizontal gene

383   transfer in bacteria. *Philos Trans R Soc Lond B: Biol Sci* 2017;**372**: 20160424.

384 Hall JP, Wood AJ, Harrison E & Brockhurst MA. Source–sink plasmid transfer dynamics maintain

385   gene mobility in soil bacterial communities. *Proc Natl Acad Sci USA* 2016; **113**: 8260-8265.

386 Heard SB, Hauser DL. Key evolutionary innovations and their ecological mechanisms. *Hist Biol*
387   1995;**10**: 151-73.

388 Heath TA, Hedtke SM, Hillis DM. Taxon sampling and the accuracy of phylogenetic analyses. *J Syst*
389   *Evol* 2008;**46**: 239-57.

390 Hehemann J-H, Arevalo P, Datta MS *et al.* Adaptive radiation by waves of gene transfer leads to fine-
391   scale resource partitioning in marine microbes. *Nat Comms* 2016;**7**.

392 Hendry AP, Bolnick DI, Berner D *et al.* Along the speciation continuum in sticklebacks. *J Fish Biol*
393   2009;**75**: 2000-36.

394 Hunt DE, David LA, Gevers D *et al.* Resource partitioning and sympatric differentiation among
395   closely related bacterioplankton. *Science* 2008;**320**: 1081-5.

396 Hunter JP. Key innovations and the ecology of macroevolution. *Trends Ecol Evol* 1998;**13**: 31-6.

397 Hutchinson GE. Homage to Santa Rosalia or why are there so many kinds of animals? *Am Nat*
398   1959;**93**: 145-59.

399 Jaffe AL, Castelle CJ, Banfield JF. Habitat transition in the evolution of bacteria and archaea. *Annu*
400   *Rev Microbiol* 2023;**77**.

401 Jurdzinski KT, Mehrshad M, Delgado LF *et al.* Large-scale phylogenomics of aquatic bacteria reveal
402   molecular mechanisms for adaptation to salinity. *Sci Adv* 2023;**9**: eadg2059.

403 Larsen BB, Miller EC, Rhodes MK *et al.* Inordinate fondness multiplied and redistributed: the number
404   of species on earth and the new pie of life. *Q Rev Biol* 2017;**92**: 229-65.

405 Lawrence JG. Catalyzing bacterial speciation: correlating lateral transfer with genetic headroom. *Syst*
406   *Biol* 2001;**50**: 479-96.

407 Locey KJ, Lennon JT. Scaling laws predict global microbial diversity *Proc Natl Acad Sci USA*
408   2016;**113**: 5970-5.

409   Logares R, Bråte J, Bertilsson S *et al.* Infrequent marine–freshwater transitions in the microbial

410   world. *Trends Microbiol* 2009;**17**: 414-22.

411   Loren JG, Farfan M, Fuste MC. Molecular phylogenetics and temporal diversification in the genus

412   Aeromonas based on the sequences of five housekeeping genes. *PLoS One* 2014;**9**: e88805.

413   Louca S. The rates of global bacterial and archaeal dispersal. *ISME J* 2022;**16**: 159-67.

414   Louca S, Mazel F, Doebeli M *et al.* A census-based estimate of Earth's bacterial and archaeal

415   diversity. *Plos Biol* 2019;**17**: e3000106.

416   Louca S, Shih PM, Pennell MW *et al.* Bacterial diversification through geological time. *Nature Ecol*

417   *Evol* 2018;**2**: 1458-67.

418   Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities.

419   *Appl Env Microbiol* 2005;**71**: 8228-35.

420   MacLean CR. Adaptive radiation in microbial microcosms. *J Evol Biol* 2005;**18**: 1376-86.

421   Madi N, Vos M, Legendre P *et al.* Diversity begets diversity in microbiomes. *eLife* 2020; **9***:e58999.*

422   Marin J, Battistuzzi FU, Brown AC *et al.* The timetree of prokaryotes: new insights into their

423   evolution and speciation. *Mol Biol Evol* 2016;**34**: 437-46.

424   Martin AP, Costello EK, Meyer AF *et al.* The rate and pattern of cladogenesis in microbes. *Evolution*

425   2004;**58**: 946-55.

426   Matschiner M, Colombo M, Damerau M *et al.* The adaptive radiation of notothenioid fishes in the

427   waters of Antarctica *Extremophile fishes*: Springer, 2015, 35-57.

428   Matuszewska M, Murray GG, Harrison EM *et al.* The evolutionary genomics of host specificity in

429   Staphylococcus aureus. *Trends Microbiol* 2020;**28**: 465-77.

430   Mooers AO, Heard SB. Inferring evolutionary process from phylogenetic tree shape. *Q Rev Biol*

431   1997;**72**: 31-54.

432   Mora C, Tittensor DP, Adl S *et al.* How many species are there on Earth and in the ocean? *Plos Biol*

433   2011;**9**: e1001127.

434   Morlon H, Kemps BD, Plotkin JB *et al.* Explosive radiation of a bacterial species group. *Evolution*

435   2012;**66**: 2577-86.

436   O'Dwyer JP, Kembel SW, Sharpton TJ. Backbones of evolutionary history test biodiversity theory for

437   microbes. *Proc Natl Acad Sci USA* 2015;**112**: 8356-61.

438   Parks DH, Chuvochina M, Waite DW *et al.* A standardized bacterial taxonomy based on genome

439   phylogeny substantially revises the tree of life. *Nat Biotechnol* 2018;**36**: 996-1004.

440   Quince C, Curtis TP, Sloan WT. The rational exploration of microbial diversity. *ISME J* 2008;**2**: 997-

441   1006.

442   Ramirez KS, Leff JW, Barberán A *et al.* Biogeographic patterns in below-ground diversity in New

443   York City's Central Park are similar to those observed globally. *Proc R Soc B: Biol Sci*

444   2014;**281**: 20141988.

445   Richardson EJ, Bacigalupe R, Harrison EM *et al.* Gene exchange drives the ecological success of a

446   multi-host bacterial pathogen. *Nat Ecol Evol* 2018;**2**: 1468-78.

447   Rosenzweig ML. Species Diversity in Space and Time. *Cambridge University Press, Cambridge*

448   1995.

449   Schluter D. *The ecology of adaptive radiation*: OUP Oxford, 2000.

450   Scholl JP, Wiens JJ. Diversification rates and species richness across the Tree of Life. *Proc R Soc B:*

451   *Biol Sci* 2016;**283**: 20161334.

452   Seehausen O. Hybridization and adaptive radiation. *Trends Ecol Evol* 2004;**19**: 198-207.

453   Shapiro BJ, Friedman J, Cordero OX *et al.* Population genomics of early events in the ecological

454   differentiation of bacteria. *Science* 2012;**336**: 48-51.

455   Shapiro BJ, Polz MF. Ordering microbial diversity into ecologically and genetically cohesive units.

456   *Trends Microbiol* 2014;**22**: 235-47.

457   Sriswasdi S, Yang C-c, Iwasaki W. Generalist species drive microbial dispersion and evolution. *Nat*

458   *Comms* 2017;**8**: 1-8.

459   Tamarit D, Neuvonen M-M, Engel P *et al.* Origin and evolution of the Bartonella gene transfer agent.

460   *Mol Biol Evol* 2018;**35**: 451-64.

461   Travisano M, Rainey PB. Studies of adaptive radiation using model microbial systems. *Am Nat*

462   2000;**156**: S35-S44.

463   Van Valen L. A new evolutionary law. *Evol Theory* 1973;**1**: 1-30.

464   Vos M. A species concept for bacteria based on adaptive divergence. *Trends Microbiol* 2011;**19**: 1-7.

465   Wang B, Qin W, Ren Y *et al.* Expansion of Thaumarchaeota habitat range is correlated with

466   horizontal transfer of ATPase operons. *ISME J* 2019;**13**: 3067-79.

467   Wang S, Luo H. Dating Alphaproteobacteria evolution with eukaryotic fossils. *Nat Comms* 2021;**12**:

468   1-9.

469   Wibowo MC, Yang Z, Borry M *et al.* Reconstruction of ancient microbial genomes from the human

470   gut. *Nature* 2021;**594**: 234-9.

471   Wisniewski-Dye F, Borziak K, Khalsa-Moyers G *et al.* Azospirillum genomes reveal transition of

472   bacteria from aquatic to terrestrial environments. *PLoS Genet* 2011;**7**: e1002430.

473   Yarza P, Yilmaz P, Pruesse E *et al.* Uniting the classification of cultured and uncultured bacteria and

474   archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014;**12**: 635-45.

475   Zhang H, Yoshizawa S, Sun Y *et al.* Repeated evolutionary transitions of flavobacteria from marine

476   to non-marine habitats. *Env Microbiol* 2019;**21**: 648-66.

477

478
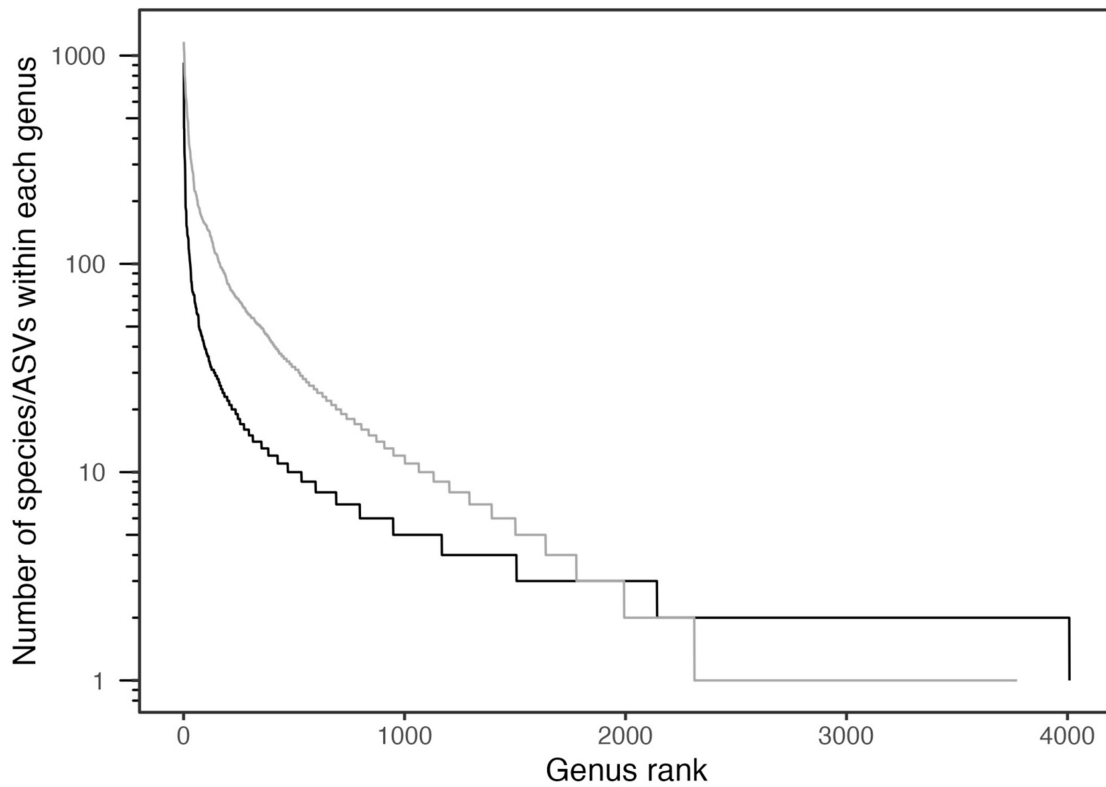
479

480   Figure 1. Variation in species diversity among bacterial genera. Rank abundance curves of total

481   species diversity from all taxonomically recognised species (LPSN; https://www.bacterio.net/) (black

482   line) and 16S rRNA-based ASVs (Amplicon Sequence Variants) assigned to genera in the Earth

483   Microbiome Project (grey line).

484

485

486
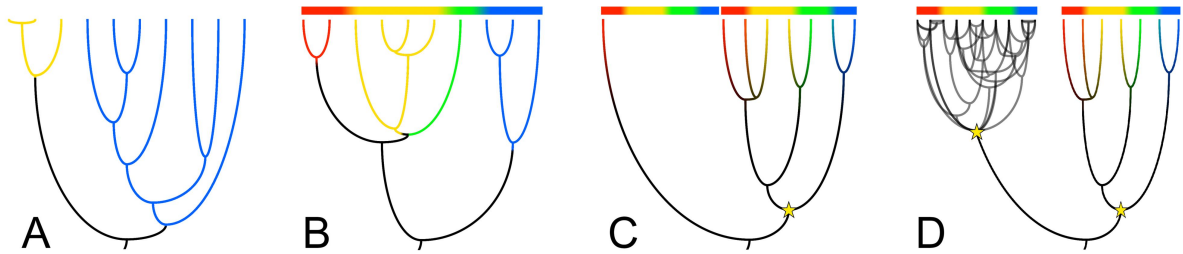
Figure 2. Four scenarios leading to differences in species richness between taxa. A: all else being equal, older clades should be of larger size. The root ages for both sister clades are different, such that the blue clade has had more time to diversify than the yellow clade. B: clades might diversify when faced with multiple potential niches to exploit, demonstrated by partitioning and subsequent diversification into red, yellow, green and blue niches. C: the capacity for diversification into multiple lineages might be mediated by the presence (or absence) of a key innovation, here indicated by the star. The clade on the right has acquired the capacity to exploit multiple niches into which it diversifies, while the branch on the left does not. D: Adaptive radiations caused by key adaptations (star symbols) in the presence of recombination barriers, allowing new niche specialists to evolve into distinct species (deep branches, right clade), or in the absence of recombination barriers, leading to the evolution of many niche specialists that do not evolve into species 'proper', with a shared core genome (shallow intermingled branches, left clade).
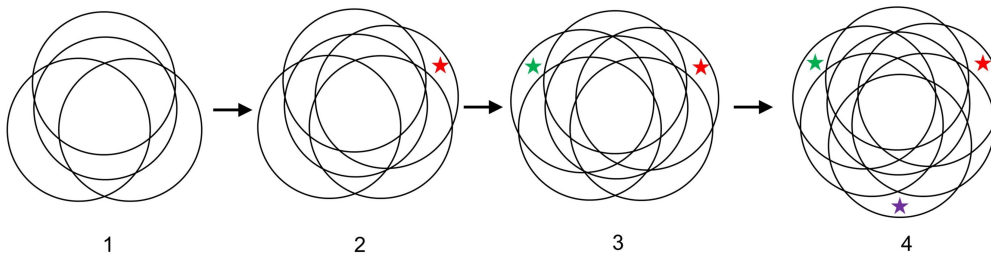
499

500

A: pan genome size growth is restricted when diversification is coupled to speciation
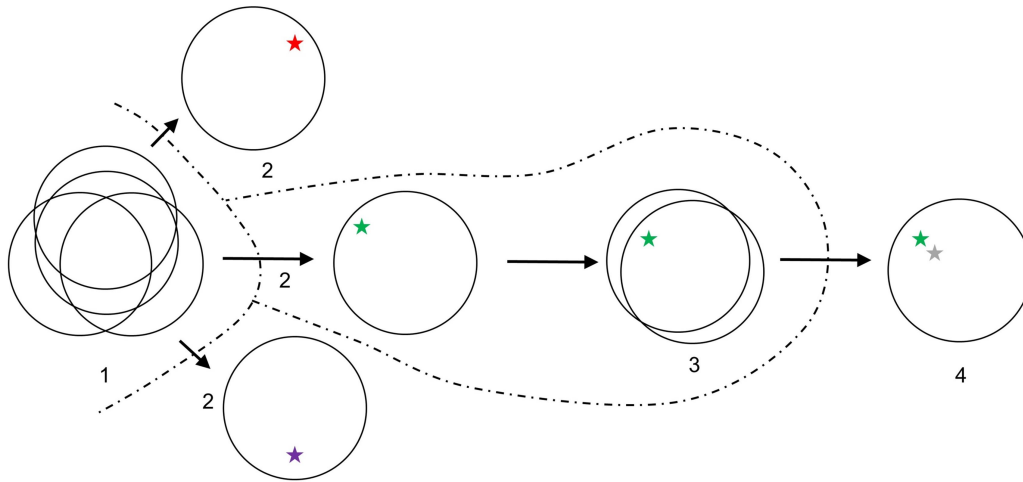
B: pan genomes expand when diversification is uncoupled from speciation

501

Figure 3. Pan genome diversification with and without barriers to recombination. A: diversification coupled to speciation in a species with barriers to recombination. The ancestral pan genome (1) acquires different key innovations (2); each uniquely adapted lineage ceases to recombine with the ancestor or with other newly evolved niche specialists because of recombination barriers (dashed lines). New niche specialists subsequently grow their pangenome through adaptive and non-adaptive processes (3). When a new key innovation occurs (4), the process is repeated. B: Diversification of a species without barriers to recombination. The ancestral pangenome (1) grows progressively with each key innovation, depicted by red (2), green (3) and purple (4) stars as well as non-adaptive gene additions.