Phylogenomic analyses of understudied *Neisseriaceae* species support the reclassification of the polyphyletic genera *Kingella, Simonsiella, and Alysiella*.

Daniel P. Morreale[1,2], Joseph W. St Geme III[1,2†], and Paul J Planet [1,2,3*†]

Affiliations:

[1] Perelman School of Medicine, University of Pennsylvania, Philadelphia PA.

[2]Division of Infectious Diseases, Children's Hospital of Pennsylvania, Philadelphia PA.

[3] Comparative Genomics, American Museum of Natural History, New York, New York, United States of America

*Corresponding Author: Paul J. Planet (pjplanet@chop.edu)

† Contributed equally

Keywords: ***Kingella, Simonsiella, Alysiella, Neisseriaceae,*** **Polyphyly,**

Abbreviations: ANI: Average Nucleotide identity; dDDH: digital DNA-DNA hybridization, OGRI: overall genome relatedness index, COGs: clusters of orthologous gene groups; MuLDi: Multicellular, longitudinally-dividing; DUS: DNA-uptake sequences, MALDI-TOF MS: Matrix-assisted laser desorption-ionization time of flight mass spectrometry

Abstract.

Taxonomic classification of the *Neisseriaceae* family has focused primarily on the pathogens *Neisseria meningitidis* and *Neisseria gonorrhoeae*. As a consequence, the commensal *Neisseria* species and other *Neisseriaceae* genera such as *Kingella* have been only loosely classified, resulting in several polyphyletic genera. In this study, using available 16S rRNA sequences and a phylogenetic approach, we found that the genus *Kingella* is polyphyletic, due to misclassification of *Kingella potus*. Calculation of multiple genome relatedness indices revealed that *K. potus* is more similar to *Neisseria bacilliformis* than it is to other *Kingella* species. To better understand the evolution of this clade, we examined the core genome of *Kingella* and closely related genera and discovered that the related genera *Simonsiella* and *Alysiella* form a distinct clade within the genus *Kingella* that is closely related to the pathogens *K. kingae* and *K. negevensis*. Based on these analyses, we propose incorporation of *Simonsiella* and *Alysiella* species into the genus *Kingella*.

15 **Main text.**

16 The *Neisseriaceae* family is a group of Gram-negative, β-proteobacteria that currently
17 includes 20 genera: *Neisseria*, *Wielerella*, *Vitreoscilla*, *Uruburuella*, *Stenoxybacter*, *Snodgrassela*,
18 *Simonsiella*, *Rivicola*, *Prolinoborus*, *Paralysiella*, *Morococcus*, *Kingella*, *Eikenella*, *Crenobacter*,
19 *Craterilacuibacter*, *Conchiformibius*, *Bergeriella*, *Aquella*, *Amantichitinum*, and *Alysiella* [1–3].
20 The majority of prior work devoted to understanding the taxonomic relationships between these
21 genera has focused primarily on the pathogenic *Neisseria* species, *N. meningitidis* and *N.*
22 *gonorrhoeae*. To date, the commensal *Neisseria* species and related genera have been only
23 loosely classified based on preliminary studies. The genus *Kingella* includes five taxa: *K. potus, K.*
24 *oralis, K. denitrificans, K. negevensis,* and *K. kingae*. Recently, a new species was proposed, *K.*
25 *bonacorsii* [4–10]. Each of the *Kingella* species is typically associated with the oral or
26 oropharyngeal microbiome, a common niche of other *Neisseriaceae*. Isolates of *K. kingae*, *K.*
27 *negevensis*, and *K. potus* have all been associated with bacteremia, osteoarticular infections,
28 and/or endocarditis in immunocompetent individuals [6, 8, 11]. In contrast, *K. oralis* and *K.*
29 *denitrificans* are present in dental plaque, gingival and periodontal samples, and only rarely cause
30 invasive infections, typically in immunocompromised individuals [9, 12, 13].

31 A review of published phylogenetic relationships among the commensal *Neisseria* species
32 reveals a high level of variability in the proposed taxonomic relationships, primarily owing to the
33 genes selected and method used to perform the analysis [14–17]. The genus *Kingella* is
34 polyphyletic, and the genera that interrupt a monophyletic *Kingella* clade differ between
35 published analyses. Most isolates used for defining taxonomic nomenclature in this genus were
36 collected prior to pervasive whole genome sequencing and were assigned to *Kingella* as a result

37 of phenotypic and/or minimal sequence-based analyses. Additionally, as the frequency of
38 culture-dependent and independent microbiome studies have become more prevalent, the
39 number and diversity of formally accepted species in the *Neisseriaceae* has increased [18]. With
40 this background, we sought to reassess the relatedness of the five species in *Kingella*.

41 To examine the *Kingella* species, we employed a phylogenetic approach. 840 of the
42 available 16S rRNA gene sequences for isolates from the *Neisseriaceae* and in the RefNR database
43 were downloaded from the SILVA r138.1 database on September 08, 2022 [19, 20]. 16S rRNA
44 sequences were aligned with MAFFT, and a phylogeny was reconstructed using RAxML v. 8.4.2
45 with the general time reversible model of nucleotide substitution and gamma model of rate
46 heterogeneity, and bootstrapped 100 times [21, 22]. The resulting tree was annotated using
47 Figtree v.1.4.4 and ITOL and is shown in Fig. 1 [23]. By 16S rRNA relatedness alone, it is clear that
48 a major driver of the polyphyletic structure of *Kingella* is *K. potus.* This species was recovered
49 from an infected wound caused by an animal bite and is more closely related to *Neisseria*
50 *bacilliformis* than it is to any currently described *Kingella* species [8, 24]. *N. bacilliformis* was
51 identified shortly after *K. potus* was defined, is a part of the oral microbiome of non-human
52 mammals, and is associated with bacteremia in humans after animal bites [24].

53 To better understand the relationship between the *K. potus* and *N. bacilliformis*, we
54 calculated average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) between
55 the type strains of these species using FastANI v.1.33 and the Type Strain Genome Server (TYGS),
56 respectively [25–27]. The ANI between *K. potus* and *N. bacilliformis* was calculated at 86.75%,
57 below the established species cut-off of 95% [28]. However, the ANI between *K. potus and N.*
58 *bacilliformis* is appreciably greater than the ANI calculated between *K. potus* and other *Kingella*

59 species (Table 1). Similarly, while the calculated dDDH is below species thresholds (Table 2) [29],
60 it is clear that *K. potus* is far more similar to *N. bacilliformis* than it is to other species within the
61 genus *Kingella.*

62 To further define the relationships between *Neisseriaceae*, whole genome sequences for
63 select species in this family were downloaded from NCBI (on September 09, 2022). These
64 analyses were limited to only those species most closely related to *Kingella, by* 16S rRNA
65 similarity and include *Neisseria elongata, Neisseria sheyganii, N. bacilliformis, Eikenella*
66 *corrodens, Alysiella filiformis, Alysiella crassa, Simonsiella muelleri, Conchiformibius steedae, K.*
67 *oralis, K. bonacorsii, K. potus, K. denitrificans, K. kingae,* and *K. negevensis* [7, 30]. In total, 134
68 genomes were reassembled with Prokka v1.14.6, using default settings for bacterial sequences
69 [31]. Subsequent assemblies were analyzed with Roary v3.13.0. to identify the core genome [32].
70 Core genes were clustered at the 75%, 80%, 85%, 90%, and 95% sequence similarity level and
71 were limited to only those clusters of orthologous groups (COGs) found in all genomes.
72 Nucleotide sequences of core genes were compared in a codon-aware alignment by PRANK
73 within Roary, and core gene alignments were used to reconstruct phylogenies using RAxML as
74 above. Each phylogeny was overwhelmingly congruent with the others, and we therefore
75 proceeded with the lowest similarity cutoff for detailed analysis (Fig. 2A). These phylogenies
76 further support the close relationship between *K. potus* and *N. bacilliformis*.

77 The phylogenomic analyses also supported a close evolutionary relationship between
78 *Kingella*, *Alysiella*, and *Simonsiella*. *Alysiella* and *Simonsiella* are commensal genera associated
79 with the oral microbiome of mammals [14, 33, 34]. These two genera, along with the genus
80 *Conchiformibius,* are distinctive for being multicellular, longitudinally-dividing (MuLDi) species

81 that comprise the only known animal multicellular bacterial symbionts [15]. Based on a whole
82 genome approach, *Alysiella* and *Simonsiella* diverged from the rest of the *Kingella* species more
83 recently than *K. denitrificans* or *K. oralis* diverged. ANI, dDDH, and G+C% were calculated for type
84 strains of *Kingella*, *Alysiella*, and *Simonsiella* and are shown in Tables 1 and 3. While there is no
85 universally accepted ANI cutoff for demarcating a bacterial genus, ANI scores greater than 73%
86 have been proposed [35]. All of these overall genome relatedness indices (OGRI) support the
87 close relationship between these three genera, and the ANI falls above the proposed 73% genus
88 cutoff score. Moreover, Nyongesa and co-workers [15] hypothesize that the MuLDi phenotype
89 evolved twice in the *Neisseriaceae,* first in *Conchiformibius* species and then in the common
90 ancestor of *Simonsiella and* Alysiella, a hypothesis that is strongly supported by our phylogenetic
91 analyses.

92 Additionally, we calculated pairwise ANI scores for each of the genomes listed above. The
93 ANI scores confirm the majority of the proposed species classifications with one exception.
94 Recently, an unclassified *Kingella* isolate was proposed to be a new species and was named *K.*
95 *bonacorsii* [4]. This isolate was recovered from a gingival sample, and the authors report a 16S
96 rRNA similarity of 98.7%, ANI of 95.83%, and dDDH of 63.6% relative to *K. oralis* UB-38, as well as
97 non-identification by MALDI-TOF mass-spectrometry [4]. We replicated each of the genomic
98 calculations reported, as well as additional dDDH calculations using alternative formulae (Tables
99 1 and 4) [28, 36]. The 16S rRNA similarity and ANI score are both above the threshold for
100 classification of *K. bonacorsii* and *K. oralis* as the same species, consistent with the very similar
101 G+C% and their common niche. Importantly, we note that the algorithm used to calculate dDDH
102 is critical in this instance. Analysis with either *d0* or *d6* yields dDDH above the threshold of 70%.

103 Only if dDDH is calculated with *d4,* which is considered most appropriate for highly fragmented
104 or gapped genomes, the dDDH falls below the accepted threshold of 70% (Table 4). Very few
105 isolates of *K. oralis* are publicly available for additional comparisons, and without additional
106 isolates and sequencing data, it is difficult to appreciate the diversity of *K. oralis* or account for
107 bias in strain selection. Based on the available data, we conclude that *K. bonacorsii* is likely not a
108 unique species in the genus *Kingella* and instead may be a subspecies of *K. oralis.*

109 Finally, another recent evolutionary study of these taxa has used less stringent similarity
110 and gene presence cutoff to define the core genome for species in the *Neisseriaceae* for
111 subsequent phylogenetic analyses [15]. We replicated this analysis by reanalyzing the data using
112 a sequence similarity cut off of 50% for genes present in 99% or 80% of analyzed genomes. By
113 lowering the threshold for a gene to be considered part of the core genome, we increase the
114 number of genes considered from 109 at the 75% similarity level up to 620 genes. Notably, as we
115 increase the number of genes considered part of the core genome, the phylogenetic architecture
116 remains constant (Fig. 3A).

117 The reproducibility of the core genome phylogenetic structure very strongly supports the
118 inclusion of *Kingella*, *Alysiella*, and *Simonsiella* in the same genus, and the exclusion of *K. potus*.
119 We hypothesized that prior misclassification could be due to high levels of horizontal gene
120 transfer of accessory genes between *Kingella* (*sensu stricto*) genomes. To test this, we generated
121 a presence/absence matrix generated by Roary at 75% sequence similarity level, a technique that
122 emphasizes accessory genes over core genes, and we reconstructed a phylogeny using the RAxML
123 binary model with gamma model of rate heterogeneity, and bootstrapped 100 times (Fig. 3B).
124 This was the only analysis in which we observed a monophyletic *Kingella* genus, based on current

125 taxonomy, though *Kingella potus* was still excluded. *Simonsiella*, *K. oralis*, and *K. kingae* utilize
126 specific DNA-uptake sequences (DUS) to limit horizontal gene transfer (HGT) events, with
127 *Simonsiella*, *K. oralis* seemingly sharing the same DUS [37]. Given that members of the
128 *Neisseriaceae* are highly recombinogenic and undergo significant HGT of accessory genes, we
129 assert that the core genome is a higher fidelity evolutionary record for this family [38].

130 Based on these results, we propose several changes to the genera *Kingella, Alysiella, and*
131 *Simonsiella.* As *K. potus* is more related to *N. bacilliformis* than other *Kingella* species by 16S
132 rRNA, OGRI, and whole genome analyses, we propose revising the taxonomic nomenclature of
133 this isolate to reflect this close relationship. Additionally, as phylogenomic analysis of the core
134 genomes of *Alysiella* and *Simonsiella* supports their close relationship to the genus *Kingella,* we
135 propose incorporating these genera into the genus *Kingella*, as *K. crassa*, *K. filiformis*, and *K.*
136 *muelleri*. These changes would adjust the accepted taxonomic nomenclature to better reflect the
137 phylogenomic relationships between these isolates and result in a monophyletic *Kingella* genus.
138 Finally, further characterization will be required for confirm if *K. bonacorsii* constitutes its own
139 species independent of *K. oralis* or if it would be better classified as a subspecies (*K. oralis* subsp.
140 *bonacorsii*).

141 **Acknowledgements.**

142 We thank the members of the Planet and St. Geme Labs for their useful feedback and discussion.

143 **Funding.**

144 This work was supported by the National Institute of Allergy and Infectious Diseases under the
145 training grant in Microbial Pathogenesis and Genomics (5T32141493) to DPM and under award
146 1R01AI121015 to JWS.

Conflicts of Interest.

The authors report no conflicts of interest.

150 **References.**

151 1. **Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, Göker M**. List of prokaryotic names
152 with standing in nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol*
153 2020;70:5607–5612.

154 2. **Parker CT, Tindall BJ, Garrity GM**. International code of nomenclature of Prokaryotes. *Int*
155 *J Syst Evol Microbiol* 2019;69:S1.

156 3. **Oren A, Garrity G**. Notification of changes in taxonomic opinion previously published
157 outside the IJSEM. List of changes in taxonomic opinion no. 36. *Int J Syst Evol Microbiol*
158 2022;72:005413.

159 4. **Antezack A, Boxberger M, Rolland C, Monnet-Corti V, La Scola B**. Isolation and
160 Characterization of Kingella bonacorsii sp. nov., A Novel Kingella Species Detected in a
161 Stable Periodontitis Subject. *Pathogens* 2021;10:1–13.

162 5. **Henricksen SD, Bovre K**. Transfer of *Moraxella kingae* Henriksen and Bovre to the genus
163 *Kingella* gen. nov. in the family *Neisseriacea*e. *Int J Syst Bacteriol* 1976;26:447–450.

164 6. **El Houmami N, Bakour S, Bzdrenga J, Rathored J, Seligmann H, *et al.*** Isolation and
165 characterization of *Kingella negevensis* sp. nov., a novel *Kingella* species detected in a
166 healthy paediatric population. *Int J Syst Evol Microbiol* 2017;67:2370–2376.

167 7. **Dewhirst FE, Chen CKC, Paster BJ, Zambon JJ**. Phylogeny of species in the family
168 Neisseriaceae isolated from human dental plaque and description of *Kingella oralis* sp.
169 nov [corrected]. *Int J Syst Bacteriol* 1993;43:490–499.

170 8. **Lawson PA, Malnick H, Collins MD, Shah JJ, Chattaway MA, *et al.*** Description of Kingella
171 potus sp. nov., an Organism Isolated from a Wound Caused by an Animal Bite. *J Clin*

172 *Microbiol* 2005;43:3526.

173 9. **Snell JJS, Lapage SP**. Transfer of some saccharolytic *Moraxella* species to *Kingella*
174 Henriksen and Bovre 1976, with descriptions of *Kingella indologenes* sp. nov. and
175 *Kingella denitrificans* sp. nov. *Int J Syst Bacteriol* 1976;26:451–458.

176 10. **Hollis DG, Weaver RE, Riley PS**. Emended description of *Kingella denitrificans* (Snell and
177 Lapage 1976): correction of the maltose reaction. *J Clin Microbiol* 1983;18:1174.

178 11. **Yagupsky P**. *Kingella kingae*: Carriage, transmission, and disease. *Clinical Microbiology*
179 *Reviews* 2015;28:54–79.

180 12. **Chen C, Chen C**. Distribution of a newly described species, *Kingella oralis*, in the human
181 oral cavity. *Oral Microbiol Immunol* 1996;11:425–427.

182 13. **Kopyt N, Kumar A, Agrawal V**. A Case of Suppurative Peritonitis by a Commensal Oral
183 Organism, *Kingella denitrificans*, in an Adult Peritoneal Dialysis Patient. *Perit Dial Int*
184 2015;35:105.

185 14. **Adeolu M, Gupta RS**. Phylogenomics and molecular signatures for the order Neisseriales:
186 proposal for division of the order Neisseriales into the emended family Neisseriaceae and
187 Chromobacteriaceae fam. nov. *Antonie Van Leeuwenhoek* 2013;104:1–24.

188 15. **Nyongesa S, Weber PM, Bernet È, Pulido F, Nieves C, *et al.*** Evolution of longitudinal
189 division in multicellular bacteria of the Neisseriaceae family. *Nat Commun 2022 131*
190 2022;13:1–18.

191 16. **Garrity GM, Bell JA, Lilburn T**. Class II. Betaproteobacteria class. nov. *Bergey's Manual®*
192 *Syst Bacteriol* 2005;575–922.

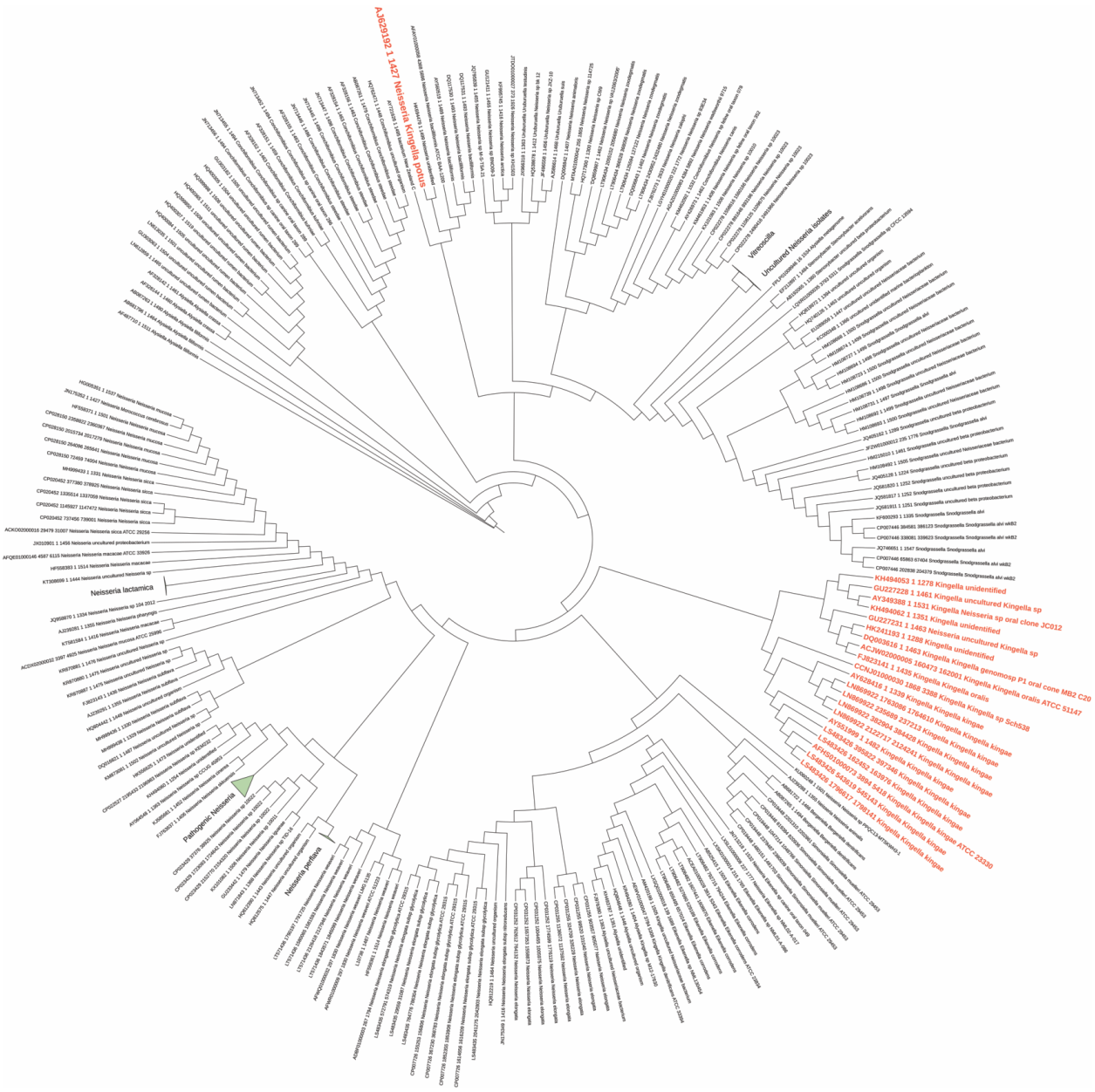
193 17. **Enright MC, Carter PE, MacLean IA, McKenzie H**. Phylogenetic relationships between

194 some members of the genera *Neisseria*, *Acinetobacter, Moraxella*, and *Kingella* based on
195 partial 16S ribosomal DNA sequence analysis. *Int J Syst Bacteriol* 1994;44:387–391.
196 18. **Maiden MCJ, Harrison OB**. Population and functional genomics of *Neisseria* revealed
197 with gene-by-gene approaches. *J Clin Microbiol* 2016;54:1949–1955.
198 19. **Glöckner FO, Yilmaz P, Quast C, Gerken J, Beccati A, *et al.*** 25 years of serving the
199 community with ribosomal RNA gene reference databases and tools. *J Biotechnol*
200 2017;261:169–176.
201 20. **Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, *et al.*** The SILVA ribosomal RNA gene
202 database project: improved data processing and web-based tools. *Nucleic Acids Res*
203 2013;41:D590–D596.
204 21. **Stamatakis A**. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
205 phylogenies. *Bioinformatics* 2014;30:1312.
206 22. **Katoh K, Standley DM**. MAFFT Multiple Sequence Alignment Software Version 7:
207 Improvements in Performance and Usability. *Mol Biol Evol* 2013;30:772.
208 23. **Letunic I, Bork P**. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree
209 display and annotation. *Nucleic Acids Res* 2021;49:W293–W296.
210 24. **Han XY, Hong T, Falsen E**. *Neisseria bacilliformis* sp. nov. Isolated from Human Infections.
211 *J Clin Microbiol* 2006;44:474.
212 25. **Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S**. High throughput ANI
213 analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun 2018*
214 *91* 2018;9:1–8.
215 26. **Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M**. TYGS and LPSN: a database

216 tandem for fast and reliable genome-based classification and nomenclature of
217 prokaryotes. *Nucleic Acids Res* 2022;50:D801–D807.
218 27. **Meier-Kolthoff JP, Göker M**. TYGS is an automated high-throughput platform for state-
219 of-the-art genome-based taxonomy. *Nat Commun 2019 101* 2019;10:1–10.
220 28. **Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M**. Genome sequence-based species
221 delimitation with confidence intervals and improved distance functions. *BMC*
222 *Bioinformatics* 2013;14:1–14.
223 29. **Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, *et al.*** Report of the Ad Hoc
224 Committee on Reconciliation of Approaches to Bacterial Systematics. *Int J Syst Evol*
225 *Microbiol* 1987;37:463–464.
226 30. **Dewhirst FE, Paster BJ, Bright PL**. *Chromobacterium, Eikenella, Kingella, Neisseria,*
227 *Simensiella*, and *Vitreoscilla* species comprise a major branch of the beta group
228 Proteobacteria by 16S ribosomal ribonucleic acid sequence comparison: Transfer of
229 *Eikenella* and *Simonsiella* to the family Neisseriaceae (emend.). *Int J Syst Bacteriol*
230 1989;39:258–266.
231 31. **Seemann T**. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–
232 2069.
233 32. **Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, *et al.*** Roary: rapid large-scale
234 prokaryote pan genome analysis. *Bioinformatics* 2015;31:3691–3693.
235 33. **Hedlund BP, Tønjum T**. *Simonsiella*. *Bergey's Man Syst Archaea Bact* 2015;1–12.
236 34. **Xie CH, Yokota A**. Phylogenetic analysis of *Alysiella* and related genera of Neisseriaceae:
237 proposal of *Alysiella crassa* comb. nov., *Conchiformibium steedae* gen. nov., comb. nov.,

238 *Conchiformibium kuhniae* sp. nov. and *Bergeriella denitrificans* gen. nov., comb. nov. *J*
239 *Gen Appl Microbiol* 2005;51:1–10.
240 35. **Barco RA, Garrity GM, Scott JJ, Amend JP, Nealson KH, *et al.*** A genus definition for
241 bacteria and archaea based on a standard genome relatedness index. *MBio*;11. Epub
242 ahead of print 2020. DOI: 10.1128/MBIO.02475-19/SUPPL_FILE/MBIO.02475-19-
243 S0001.DOCX.
244 36. **Auch AF, von Jan M, Klenk HP, Göker M**. Digital DNA-DNA hybridization for microbial
245 species delineation by means of genome-to-genome sequence comparison. *Stand*
246 *Genomic Sci* 2010;2:117.
247 37. **Frye SA, Nilsen M, Tønjum T, Ambur OH**. Dialects of the DNA Uptake Sequence in
248 Neisseriaceae. *PLOS Genet* 2013;9:e1003458.
249 38. **Hanage WP, Fraser C, Spratt BG**. Fuzzy species among recombinogenic bacteria. *BMC*
250 *Biol* 2005;3:1–7.
251
252

253 **Figures.**



254

255 **Figure 1.** High quality 16S rRNA sequences were downloaded from the SILVA database for the

256 family Neisseriaceae (taxid: 408). Sequences were aligned with MAFFT, and a maximum

257 likelihood phylogeny was constructed using RAxML. Species that belong to the genus *Kingella* are

258 colored red. By 16S rRNA relatedness, *Kingella* forms a polyphyletic clade with *K. potus* falling

259 outside of the rest of the genus. *K. potus* forms a clade with the oral-associated *Neisseria*

260 *bacilliformis*. Bootstrap values greater than 70% are shown.

261

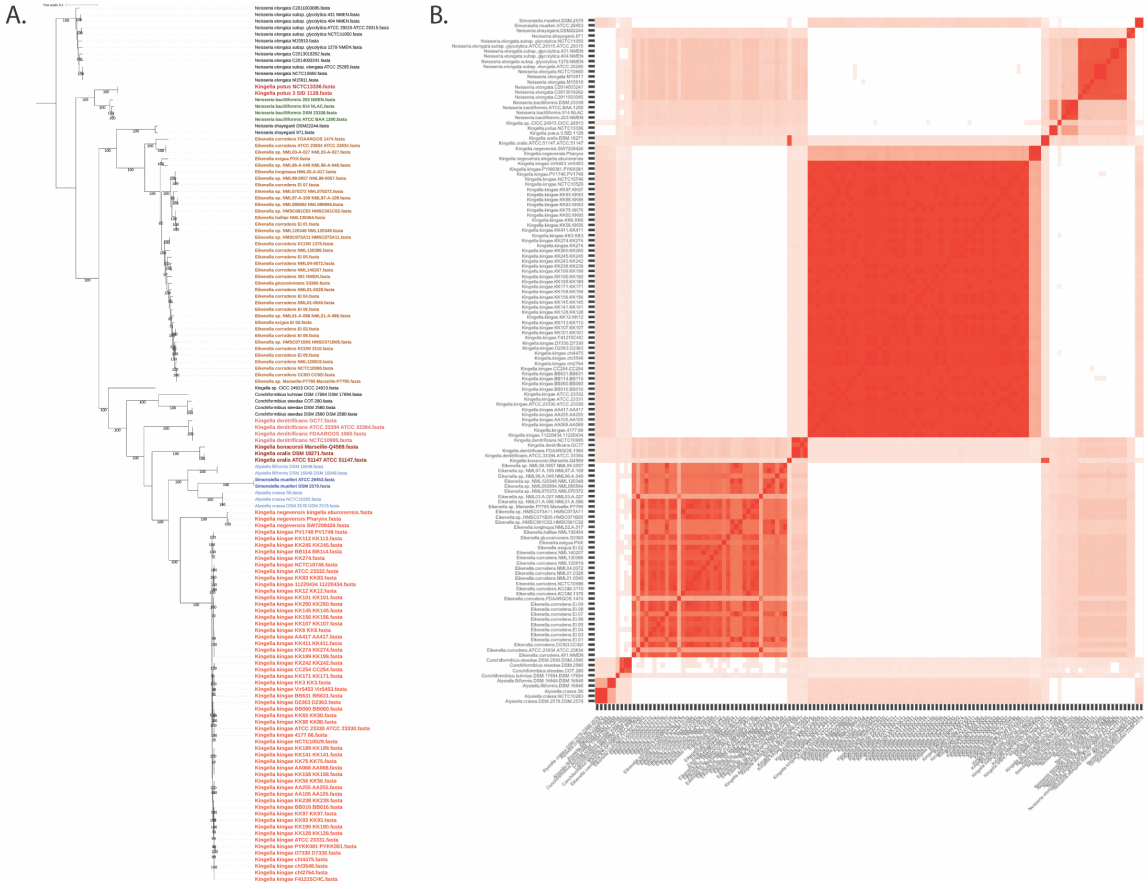


Figure 2. A. The core genome of the strains from the 13 species most closely related to *Kingella kingae* was determined using Roary with a minimum homology cutoff of 75%. The core genomes, which include the sequences of 109 genes total, which were aligned with PRANK in Roary and RAxML was used to reconstruct a maximum likelihood phylogeny. *K. potus* is only distantly related to the rest of the species in the genus *Kingella.* The genera *Alysiella* and *Simonsiella* fall within the *Kingella* clade, suggesting that the core genome of the species in each of these genera is very closely related to that of *Kingella.* Additionally, *K. bonacorsii* is very closely related to *K. oralis,* and these two species form a well-supported clade. The tree is rooted by *N. elongata,* and bootstraps greater than 70% are shown. **B.** ANI score between each isolate is shown on the tree.

272 The ANI score when comparing *K. oralis* and *K. bonacorsii* suggests that differentiating between

273 these isolates as two species may not reflect speciation events.

274

Figure 3. A. The core genome of the strains from the 13 species most closely related to *Kingella kingae* was determined using Roary with a minimum protein similarity cutoff of 50%, effectively reducing the stringency for genes to be considered a part of the core genome. The core genomes, which include the sequences of 620 CDS total, were aligned with MAFFT, and RAxML was used to reconstruct a maximum likelihood phylogeny with 100 bootstraps. Bootstrap values >70% are shown, and the tree is rooted by *N. elongata*. Unlike previous analyses with

282 similar datasets, *Alysiella* and *Simonsiella* are still more closely related to *K. kingae* than to
283 either *K. oralis* or *K. denitrificans*. **B.** The pangenome of the species closely related to *Kingella*
284 was determined using Roary, at the 75% similar level and ignoring genes that occur in all
285 isolates. RAxML was used to reconstruct a phylogeny of gene presence/absence. Bootstrap
286 values >70% are shown, and the tree is rooted by *N. elongata*. This phylogeny is different than
287 the 75% sequence similarity tree shown in panel A, suggesting that the pangenome gene
288 content may heavily influence the phylogenetic relation relationship between these species,
289 clouding the evolutionary history demonstrated by core genes.
290

291 **Table 1. Average ANI scores calculated among taxa of interest.** ANI scores were calculated using
292 FastANI and are shown only if the ANI >75%. Empty cells denote data excluded from the table.

	N. bacilliformis	*K. denitrificans*	*K. oralis*	*K. negevensis*	*K. kingae*	*N. elongata*	*A. crassa*	*A. filiformis*	*S. muelleri*
K. potus	86.75	78.88	78.74	<75	<75	80.7	-	-	-
K. bonacorsii	-	79.84	96.41	78.06	77.9	78.76	77.98	77.87	77.14
A. crassa	-	77.81	77.96	79.37	79.74	<75	100	85.43	79.79
A. filiformis	-	77.81	77.96	79.37	79.74	<75	85.43	100	78.75
S. muelleri	-	<75	76.9	79.16	79.96	<75	79.79	78.75	100

293

294

295 **Table 2. dDDH values of *K. potus* calculated against closely related taxa.** dDDH was calculated
296 using TYGS using each of the possible algorithms, as well as the difference in G+C content
297 between *K. potus* and each of the subject type strains.

Query strain	Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G+C content difference (in %)
K. potus NCTC13336	*N. bacilliformis* ATCC BAA-1200	45.6	[42.2 - 49.0]	30	[27.6 - 32.5]	41	[38.0 - 44.0]	1.86
K. potus NCTC13336	*N. elongata* subsp. *glycolytica* ATCC 29315	20.3	[17.1 - 23.9]	22.5	[20.2 - 25.0]	19.7	[17.0 - 22.8]	3.71
K. potus NCTC13336	*K. negevensis* Sch538	12.9	[10.2 - 16.2]	23.5	[21.2 - 25.9]	13.3	[10.9 - 16.0]	12.24

298

299

300 **Table 3. OGRI of *Kingella*, *Alysiella*, and *Simonsiella*.** dDDH was calculated using TYGS using each
301 of the possible algorithms, as well as the difference in G+C content between type strains of
302 *Alysiella* or *Simonsiella* species and each of the closest related subject type strains.

Query strain	Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G+C content difference (in %)
A. crassa NCTC10283	*A. filiformis* DSM 16848	21.7	[18.5 - 25.3]	34.4	[31.9 - 36.9]	21.9	[19.1 - 25.0]	1.29
A. crassa NCTC10283	*W. bovis* CCUG 44465T	18.4	[15.3 - 22.0]	33.4	[31.0 - 35.9]	18.8	[16.1 - 21.8]	2.67
A. crassa NCTC10283	*S. muelleri* ATCC 29453	16.9	[13.8 - 20.4]	23.2	[20.9 - 25.6]	16.9	[14.3 - 19.8]	3.83
A. crassa NCTC10283	*K. kingae* ATCC 23330	16.3	[13.3 - 19.8]	23.8	[21.5 - 26.3]	16.4	[13.8 - 19.3]	1.45
A. crassa NCTC10283	*K. negevensis* Sch538	15.6	[12.7 - 19.0]	23.3	[21.0 - 25.8]	15.7	[13.2 - 18.6]	0.17
A. crassa NCTC10283	*K. bonacorsii* Marseille-Q4569	14.2	[11.4 - 17.6]	22.7	[20.5 - 25.2]	14.5	[12.0 - 17.3]	8.71
A. crassa NCTC10283	*K. oralis* ATCC 51147	14.2	[11.4 - 17.6]	22	[19.7 - 24.4]	14.5	[12.1 - 17.3]	9.01
A. filiformis DSM 16848	*A. crassa* NCTC10283	21.8	[18.6 - 25.5]	34.1	[31.7 - 36.6]	22	[19.2 - 25.1]	1.25
A. filiformis DSM 16848	*W. bovis* CCUG 44465T	17.2	[14.2 - 20.8]	30.5	[28.1 - 33.0]	17.5	[14.9 - 20.5]	3.95
A. filiformis DSM 16848	*K. kingae* ATCC 23330	16.5	[13.5 - 20.0]	24	[21.7 - 26.5]	16.5	[14.0 - 19.5]	0.17
A. filiformis DSM 16848	*K. negevensis* Sch538	15.1	[12.2 - 18.6]	24.1	[21.8 - 26.6]	15.3	[12.8 - 18.2]	1.11
A. filiformis DSM 16848	*S. muelleri* ATCC 29453	15.1	[12.2 - 18.6]	23.7	[21.4 - 26.1]	15.3	[12.8 - 18.2]	5.12
A. filiformis DSM 16848	*C. steedae* DSM 2580	15.1	[12.2 - 18.5]	22.6	[20.3 - 25.0]	15.3	[12.8 - 18.2]	4.17
A. filiformis DSM 16848	*K. bonacorsii* Marseille-Q4569	14.6	[11.7 - 18.0]	22.8	[20.5 - 25.2]	14.8	[12.3 - 17.6]	7.43
A. filiformis DSM 16848	*K. oralis* ATCC 51147	14.6	[11.8 - 18.0]	22.2	[19.9 - 24.7]	14.8	[12.4 - 17.7]	7.72

303

304

305 **Table 4. dDDH calculations for the proposed species *K. bonacorsii.*** dDDH was calculated using
306 TYGS using each of the possible algorithms, as well as the difference in G+C content between *K.*
307 *bonacorsii* and each of the subject type strains.

Query strain	Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G+C content difference (in %)
K. bonacorsii Marseille-Q4569	*K. oralis* ATCC 51147	84.3	[80.5 - 87.4]	63.6	[60.7 - 66.4]	83.1	[79.8 - 86.0]	0.29
K. bonacorsii Marseille-Q4569	*K. denitrificans* ATCC 33394	15.9	[13.0 - 19.4]	23.2	[20.9 - 25.7]	16	[13.5 - 18.9]	0.06
K. bonacorsii Marseille-Q4569	*N. bacilliformis* ATCC BAA-1200	15.1	[12.2 - 18.5]	25.3	[23.0 - 27.8]	15.3	[12.8 - 18.2]	5.55
K. bonacorsii Marseille-Q4569	*K. potus* NCTC13336	14.4	[11.6 - 17.8]	23.2	[20.9 - 25.7]	14.7	[12.2 - 17.5]	3.84
K. bonacorsii Marseille-Q4569	*K. negevensis* Sch538	14.4	[11.6 - 17.8]	22	[19.7 - 24.4]	14.7	[12.2 - 17.5]	8.54

308