

1 **Association between lockdown and changes in**
2 **subspecies diversity with a focus on *Haemophilus***
3 ***influenzae* and *Staphylococcus aureus***
4

5 Slim Hmidi^{1,2}, Nadim Cassir^{1,3,4}, Philippe Colson^{1,3,4}, Raymond Ruimy^{5,6}, Hervé Chaudet^{1,2,4*}

6
7 ¹ IHU Méditerranée Infection, Marseille, France

8 ² Aix Marseille Université, Institut de Recherche pour le Développement (IRD), Assistance
9 Publique– Hôpitaux de Marseille (AP-HM), Service de Santé des Armées (SSA), Vecteurs –
10 Infections Tropicales et Méditerranéennes (VITROME), Marseille, France.

11 ³ Aix-Marseille Université, Institut de Recherche pour le Développement (IRD), Microbes
12 Evolution Phylogeny and Infections (MEPHI), Marseille, France

13 ⁴ Assistance Publique- Hôpitaux de Marseille (AP-HM), Marseille, France

14 ⁵ Université Côte d’Azur, Inserm, U1065, C3M, Nice, France

15 ⁶ Department of Bacteriology, Université Côte d’Azur, CHU Nice, Nice, France

16

17 * Corresponding author

18 E-mail :herve.chaudet@ap-hm.fr

19

20

21 **Abstract**

22 **Background:** The social distancing measures implemented to curb SARS-CoV-2
23 transmission provided a unique opportunity to study the association between reduced human
24 interaction and epidemiological changes related to human bacterial pathogens. While studies
25 have indicated a decrease in respiratory infections during lockdowns, further description is
26 needed regarding the changes in the incidence of bacterial populations. This study investigates
27 the changes in strain richness of community infections with two bacterial species,
28 *Haemophilus influenzae* and *Staphylococcus aureus* during the waning related to France's
29 social distancing measures, especially lockdown.

30 **Methods:** MALDI-TOF MS spectra analyses of routine clinical bacterial identifications were
31 used as proxies for genomic analyses. Spectra from lockdown and reference periods were
32 compared using unsupervised classification methods. A total of 251 main spectrum profiles of
33 *H. influenzae*, 2079 main spectrum profiles of *S. aureus* for respiratory tract and blood
34 samples, and 414 main spectrum profiles for skin samples of *S. aureus* were examined. Data
35 were analyzed using hierarchical clustering, binary discriminant analysis, and statistical tests
36 for significance.

37 **Results:** The strain mix of both bacteria during the lockdown was deeply altered, but with
38 different further evolutions. *H. influenzae* exhibited a shift in spectra composition, with a
39 subsequent return towards pre-lockdown diversity observed in 2021. In contrast, *S. aureus*
40 exhibited a persistent change in spectra composition, with a gradual return to pre-lockdown
41 patterns one year later.

42 **Conclusions:** Hindering inter-human transmission, as was done during the lockdown
43 measures, was associated with significant alterations in bacterial species compositions, with

44 differential impacts observed for *H. influenzae* and *S. aureus*. This study provides data on the
45 putative relationship between genetic diversity and transmission dynamics during a public
46 health crisis. Describing the dynamics of bacterial populations during lockdowns could
47 contribute providing information for the implementation of future strategies for infectious
48 disease control and surveillance.
49

50 Introduction

51 To hinder the spread of the SARS-CoV-2 outbreak, several restrictive measures have
52 been carried out worldwide, including population lockdowns to prevent viral transmission. In
53 France, a general lockdown of the population was enforced from March 17th to May 11th,
54 2020, for curbing the outbreak curve. This unprecedented situation can be considered as a
55 real-scale experiment to assess the changes in the epidemiology of infectious agents when
56 hindering human-to-human transmissions. This measure has been followed by less restrictive
57 ones, as restrictions on gatherings and human mobility, use of personal protective equipment.
58 However, the barrier gestures were less respected from summer 2020, allowing several
59 epidemic waves.

60 Several studies have investigated the consequences of measures against the spread of
61 SARS-CoV-2 on the incidence of other pathogens, mainly focusing on paediatric infections.
62 Rotulo *et al.* demonstrated a reduction in community-acquired respiratory infections observed
63 in a paediatric emergency department in Genoa, Italy, during the lockdown period 1. Similar
64 results were also reported by Vierucci *et al.* in Tuscany 2.

65 A recent study examining invasive infections caused by *Streptococcus pneumoniae*,
66 *Haemophilus influenzae* or *Neisseria meningitidis* across 26 countries during the early months
67 of 2020 (from January 1 to May 31) retrieved a significant and prolonged decrease in invasive
68 diseases due to these bacteria. These incidence reductions coincided with the implementation
69 of COVID-19 containment measures in each respective region 3. Similarly, a study
70 investigating the lockdown consequences on non-viral infectious agents observed significant
71 changes in the relative abundance of several bacterial species, including a decrease in
72 *Escherichia coli*, *S. pneumoniae*, and *H. influenzae* 4.

73 All these studies suggested a quantitative alteration of bacterial presence in the human
74 population, with a decrease in species abundances. Interhuman transmission events have two

75 dimensions: contact rate and inoculum size, constituting the transmission bottleneck. The
76 lockdown, aimed at stopping or hindering the interhuman transmission, might result in small
77 bottlenecks. Previous studies have shown that small bottlenecks were associated with
78 substantially different levels of bacterial diversity in the transmission target, in terms of clonal
79 richness, due to stochastic effects 5. Evaluating the epidemiological changes at the waning of
80 the lockdown on human-related bacterial populations therefore need to assess species clonal
81 richness over time.

82 Studies of bacteria subspecies optimally rely on costly genomic analyses. However,
83 MALDI-TOF spectra analyses may act as proxies of genomic analyses. Routine
84 identifications of bacterial species using MALDI-TOF (Matrix-Assisted Laser
85 Desorption/Ionization Time-of-Flight) mass spectrometry are based on the presence of
86 species' characteristic peaks 6, a kind of 'fingerprint recognition', even if characteristic peak
87 correspondence is not known 7. Identification may even delve deeper than the species level by
88 the mean of supplementary peaks capable of identifying clonal complexes, frequently used for
89 outbreak investigation and clonal characterization 8-11. We previously demonstrated 12 that
90 MALDI-TOF spectra have sufficient richness and expressivity to allow for sub-species
91 differentiations. On this basis, we have developed a set of original numerical and statistical
92 processing tools for analyzing routine identification spectra to search for possible clonal
93 complexes for epidemiologic surveillance purposes 13.

94 In this work we used these tools to address the following questions: what were the
95 changes at the waning of the lockdown-related on species' incidence and their clonal
96 richness? Were these changes similar across bacterial species and samplings? This work
97 specifically focused on *H. influenzae* and *Staphylococcus aureus*, species with enough
98 identifications.

99 **Materials and methods**

100 **Experimental design**

101 The experimental design of this study aimed to compare MALDI-TOF spectra of
102 *Haemophilus influenzae* and *Staphylococcus aureus* produced during the 8-week lockdown
103 period with 8-week reference periods before and after, for identifying alterations in spectra
104 clustering. This general framework was used to develop two protocols: the first protocol
105 compared the lockdown period to the periods immediately before and following it within the
106 same year, while the second protocol compared the lockdown period to the same one during
107 2019 and 2021.

108 **Materials**

109 The Institut Hospitalo-Universitaire Méditerranée Infection (IHUMI) has a database of
110 1.270 million MALDI-TOF MS spectra carried out since 2011 for routine bacterial
111 identification in the AP-HM, mostly from regional patient recruitment. All spectra were done
112 using three Bruker Microflex mass spectrometers following the standardized protocol
113 provided by Bruker Daltonics. A single colony or sediment was applied directly to two
114 separate spots on crushed, air-dried steel targets coated with 1 μ L of a matrix solution of α -
115 cyano-4-hydroxycinnamic acid saturated in 50% acetonitrile and 2.5% trifluoroacetic acid,
116 and air dried for 5 min. Spectrum acquisitions were controlled using FlexControl software.
117 The Bruker Bacterial Test Standard (BTS) calibration was routinely used according to Bruker
118 instructions. For each plate, BTS was used as a positive control and an uninoculated matrix
119 solution was used as a negative control. Identification of bacterial species was given using the
120 Biotyper software with species-level identification when the log (score) was ≥ 2.0 . All
121 MALDI-TOF spectra produced by the automated systems were loaded into a database, along
122 with the corresponding biological sample information (pseudonymised patient demographics

123 and stay characteristics, sample characterization, requesting unit, species identification,
124 antibiogram).

125 **Inclusion criteria**

126 We included spectra from routine clinical bacterial identification of *H. influenzae* and *S.*
127 *aureus*, corresponding to upper respiratory tract and blood culture specimens for the two
128 species and skin specimens for *S. aureus*. All samples were coming from community
129 infections. We retained spectra with existing positive plate control and belonging to the 5
130 study periods: lockdown (March 17-May 11, 2020), pre-lockdown 2020 (January 27-March
131 16, 2020), post-lockdown (May 12-July 7, 2020), baseline 2019 (March 17-May 11, 2019),
132 baseline 2021 (March 17-May 11, 2021). The spectra were deduplicated by hospital stay and
133 clinical sample kind.

134 **MALDI-TOF MS analysis**

135 The matrix gathering the intensities of the significant peaks was set up as recommended
136 by S. Gibb 14. For spectra normalization we used a process flow made of intensity
137 transformation (square root), smoothing (moving average with half window size 12), baseline
138 correction (Statistics-sensitive Non-linear Iterative Peak-clipping algorithm, 100 iterations),
139 and intensity recalibration on the peak of maximum intensity. Throughout this process, a
140 signal-to-noise ratio (SNR) of 3 was utilized. The technical replicates were merged (averaging
141 method) into main spectra profiles (msp), and species-specific peaks were removed to increase
142 the contrast between spectra, as outlined in 12.

143 Our analysis was based on unsupervised classification, aiming at finding possible links
144 between spectra compositions and their corresponding periods. Following the protocol we
145 have described for epidemiological surveillance 13, we employed an ascendent hierarchical
146 classification using Bray-Curtis distances between spectra, with Ward agglomeration and

147 Gruveaus-Wainer ordination. The leaves of the dendrograms were color-coded to denote
148 period membership (green: before the lockdown period, red: during, black: after, regardless of
149 the protocol). During dendrogram analysis, the optimal number of clusters was determined
150 using the Rousseeuw's silhouette method 15, the homogeneity of period distribution among
151 the clusters was tested using the chi2 or Fisher test, the grouping of lockdown strains was
152 assessed using the Wald-Wolfowitz Runs Test (WWRT) 16, and we have systematically
153 tested the proximity of lockdown and post-lockdown strains with Unconditional Barnard's
154 unilateral Exact Test (BET) on nearest neighbour counts, which is particularly suited to 2x2
155 table of counts 17. The most discriminating peaks between clusters were identified using the
156 Gibb and Strimmer's binary discriminant analysis 18.

157 The analysis was done using R and the packages MALDIquant v1.16.2, Seriation v1.2-2
158 for dendrogram ordering, BinDA v1.0.3 for binary discriminant analysis, and Vegan for
159 distance metrics.

160 **Ethical approval statement**

161 The surveillance system is in conformance with the Regulation (EU) 2016/679 of the
162 European Parliament and of the Council of 27 April 2016 on the protection of natural persons
163 with regard to the processing of personal data and on the free movement of such data, and
164 repealing Directive 95/46/EC (General Data Protection Regulation), declared on the
165 RGD/AP-HM register with number 2019-73 ("Epidemiological monitoring of potentially
166 pathogenic microorganisms"). The study SpectraSurv was approved by the *Commission*
167 *Nationale Informatique et Libertés* (decision DR-2018-177). The data were accessed for
168 research purposes during June 2022, and authors had no access to information that could
169 identify individual participants during or after data collection.

170 RESULTS

171 *Haemophilus Influenzae*

172 Comparison with reference periods from the same year

173 We analysed 283 *H. influenzae* spectra grouped into 140 msp, and distributed as
174 follows: 72 before lockdown (51.4% of all msp), 17 during lockdown (12.1%), 51 after
175 lockdown (36.5%) (Fig 1a, colouring in accordance with the schema presented in the
176 methods). According Rousseeuw's methods, the dendrogram was partitioned into 5 clusters,
177 heterogeneous in their composition regarding the periods, and contrasted by the presence of
178 most of the lockdown samples in the lower subtree (Fisher test: $p=0.0055$). There was no
179 grouping of lockdown strains together (WWRT=-1.96, $p=0.051$), but we observed a closer
180 proximity of lockdown spectra with post-lockdown spectra (BET=-2.02, $p=0.028$)

181
182 **Fig 1: Results of the analysis of *H. influenzae* spectra from the same year.** a) Dendrogram of the
183 140 main spectra profiles. The colour scale shows the distribution of the leaves over time: before
184 lockdown (green), during lockdown (red) and after lockdown (black). The dotted line corresponds to
185 the cut height for the optimal number of clusters. b) Binary discriminant analysis showing the 40
186 highest ranked peaks contrasting the lockdown period with the 8 weeks periods before and after
187 lockdown. Peaks are indicated by their m/z. For each selected peak, the entropic t-score ranking is
188 shown, positive when the peak is present in the group.

189
190 A binary discriminant analysis directly contrasting the spectra from the three periods
191 showed an opposition of lockdown and pre-lockdown most discriminant peaks, with an
192 important similarity of lockdown and post-lockdown periods (Fig 1b). This finding was in
193 accordance with the proximity of spectra belonging to these two periods observed on the

194 dendrogram and suggests that clonal complexes were replaced during the lockdown by new
195 ones, persisting after the lockdown.

196 **Comparison with years 2019 and 2021**

197 We studied 260 spectra grouped into 128 msp, 73 (58.1%) from 2019, 19 (11.2%) from
198 lockdown (2020), and 36 (30.7%) from 2021. The 3 clusters of the dendrogram (Fig 2a) were
199 heterogeneous according to the period distributions with a single cluster gathering 13/19
200 lockdown strains ($\chi^2=12.015$, $p=0.017$), and showed a grouping of the samples from the
201 lockdown period ($WWRT=-4.02$, $p=5.86 \cdot 10^{-5}$). However, there was no privileged tying of
202 lockdown strains with post-lockdown strains ($BET=-1.37$, $p=0.097$).

203

204 **Fig 2: Results of the analysis of *H. influenza* spectra from the same period of the 3 years. a)**
205 Dendrogram of the 128 main spectra profiles. The colour scale shows the distribution of leaves over
206 time: 2019 (green), 2020 (red) and 2021 (black). The dotted line corresponds to the cut height for the
207 optimal number of clusters. **b)** Binary discriminant analysis showing the 40 highest ranked peaks
208 contrasting the three yearly periods during 2019, 2020 (lockdown), and 2021. Peaks are indicated by
209 their m/z. For each selected peak, the entropic t-score ranking is shown, positive when the peak is
210 present in the group.

211

212 The binary discriminant analysis of the peaks contrasting these 3 periods (Fig 2b)
213 showed that the most discriminating peaks appeared during the lockdown, but with a tendency
214 during 2021 to return to the 2019 situation with an intermediate configuration. This result,
215 associated with the absence of tying in the dendrogram, suggests that the clonal complex
216 replacement observed during year 2020 was transient, with a slow and partial return to the
217 pre-lockdown state. The two binary discriminant analyses shared 19/40 most discriminant
218 peaks, showing a relative stability of the peaks supporting the discrimination.

219 *Staphylococcus aureus*

220 For this species we also examined skin samples in addition to respiratory samples, due
221 to the importance of the transmission mode by direct contact.

222 **Respiratory tract and blood samples**

223 **Comparison with reference periods from the same year**

224 We analysed 2147 spectra grouped into 1019 msp of *S. aureus* distributed as follows:
225 410 msp before lockdown represented in green (44.57% of the spectra), 200 msp during
226 lockdown represented in red (18.37 %), 409 msp after lockdown represented in black
227 (37.06%) (Fig 3a). Silhouette analysis of this dendrogram selected 6 clusters, heterogeneous
228 according to the period distribution, with a single cluster having 32% of its strains belonging
229 to the lockdown period and gathering 46% of these strains ($\text{Chi}^2=64.87$, $p=4.3 \cdot 10^{-10}$).
230 Lockdown strains were particularly grouped according to their spectra similarity ($\text{WWRT}=-$
231 10.685 , $p<2.2 \cdot 10^{-16}$). However, there was no privileged tying of lockdown strains with post-
232 lockdown strains ($\text{BET}=0.10$, $p=0.48$).

233
234 **Fig 3: Results of the analysis of *S. aureus* spectra from respiratory tract and blood samples of the**
235 **same year. a)** Dendrogram of the 1019 main spectra profiles. The colour scale shows the distribution
236 of the leaves over time: before lockdown (green), during lockdown (red) and after lockdown (black).
237 The dotted line corresponds to the cut height for the optimal number of clusters. **b)** Binary
238 discriminant analysis of main spectra profiles showing the 40 highest ranked peaks contrasting the
239 lockdown period with the 8 weeks periods before and after lockdown. Peaks are indicated by their
240 m/z. For each selected peak, the entropic t-score ranking is shown, positive when the peak is present in
241 the group.

242
243 Like *H. influenzae*, the binary discriminant analysis comparing the spectra
244 corresponding to the three periods (before, during, after lockdown) showed an opposition of

245 the most discriminating peaks between the pre-lockdown and the lockdown period (Fig 3b).
246 However, this opposition no more persisted during the post-lockdown period, with a partial
247 evolution toward the pre-lockdown spectra characteristics.

248 **Comparison with years 2019 and 2021**

249 We analysed 2653 spectra grouped into 1260 msp of *S. aureus* distributed as follows:
250 475 (37.7%) msp from 2019, 200 (15.9%) msp from the lockdown period, 585 (46.4%) msp
251 from 2021. Among the 3 dendrogram's clusters confirmed by the silhouette analysis (Fig 4a),
252 the middle one is characterised by a predominance of lockdown and post-lockdown strains at
253 the expense of pre-lockdown ones ($\text{Chi}^2=79.801$, $p<2.2 \cdot 10^{-16}$). As for the previous
254 comparison, the lockdown strains were particularly grouped according to their spectra
255 similarity ($\text{WWRT}=-13.782$, $p<2.2 \cdot 10^{-16}$), and we found a tying of lockdown and post-
256 lockdown strains ($\text{BET}=2.315$, $p=0.011$).

257

258 **Fig 4: Results of the analysis of *S. aureus* spectra from respiratory tract and blood samples of the**
259 **same period over the 3 years. a)** Dendrogram of the 1260 main spectra profiles. The colour scale
260 shows the distribution of leaves over time: 2019 (green), 2020 (red) and 2021 (black). The dotted line
261 corresponds to the cut height for the optimal number of clusters. **b)** Binary discriminant analysis of
262 main spectra profiles showing the 40 highest ranked peaks contrasting the three yearly periods during
263 2019, 2020 (lockdown), and 2021. Peaks are indicated by their m/z. For each selected peak, the
264 entropic t-score ranking is shown, positive when the peak is present in the group.

265

266 The overall binary discriminant analysis (Fig 4b) showed a similarity between
267 lockdown and 2021 and a difference from 2019, suggesting a persistence of the changes
268 observed during the lockdown on the protein composition of *S. aureus*, which seems to agree
269 with the results of the comparison between before and after it during the same year. The two

270 binary discriminant analyses shared 21/40 most discriminant peaks, showing a stability of the
271 peaks supporting the discrimination.

272 **Skin Samples**

273 **Comparison with reference periods from the same year**

274 This study analysed 367 msp including 177 msp before lockdown represented in green
275 (48.2% of the spectra), 51 msp during lockdown represented in red (13.9 %), 139 msp after
276 lockdown represented in black (37.9%). The silhouette analysis found three optimal clusters
277 (Fig 5a) in the dendrogram, with a middle cluster having few lockdown strains ($\text{Chi}^2=10.949$,
278 $p=0.027$). Samples from the lockdown period were grouped by the clustering process of their
279 spectra ($\text{WWRT}=-8.506$, $p<2.2 \cdot 10^{-16}$). We also noticed that spectra from pre-lockdown and
280 post-lockdown samples were distributed over the dendrogram, without protein proximity
281 between lockdown and post-lockdown periods ($\text{BT}=-1.509$, $p=0.07$).

282

283 **Fig 5: Results of the analysis of *S. aureus* spectra from skin samples of the same year. a)**

284 Dendrogram of the 367 main spectra profiles. The colour scale shows the distribution of the leaves
285 over time: before lockdown (green), during lockdown (red) and after lockdown (black). The dotted
286 line corresponds to the cut height for the optimal number of clusters. **b)** Binary discriminant analysis
287 of main spectra profiles showing the 40 highest ranked peaks contrasting the lockdown period with the
288 8 weeks periods before and after lockdown. Peaks are indicated by their m/z. For each selected peak,
289 the entropic t-score ranking is shown, positive when the peak is present in the group.

290

291 Unlike the previous comparison, the binary discriminant analysis (Fig 5b) showed a
292 similarity of the distribution of the most discriminant peaks between pre and post lockdown
293 spectra, contrasting with the lockdown period, which indicates that the changes during
294 lockdown on the spectra composition of *S. aureus* were transitory.

295 **Comparison with years 2019 and 2021**

296 We analysed 605 spectra grouped into 288 msp of *S. aureus* distributed as follows: 221
297 (76.8%) msp from 2019, 51 (17.7%) msp from 2020, 16 (5.5%) msp from 2021. The two
298 optimal clusters of the dendrogram (Fig 6a) presented a homogeneous distribution of the
299 periods (Fisher Exact Test: $p=0.656$), but with a strong grouping of the lockdown strains
300 (WWRT=-10.362, $p<2.2 \cdot 10^{-16}$) and a statistical association between lockdown and post-
301 lockdown spectra despite the small amount of the last ones (BET=-1.715, $p=0.048$).

302
303 **Fig 6: Results of the analysis of *S. aureus* spectra from skin samples of the same period over the**
304 **3 years. a)** Dendrogram of the 1019 main spectra profiles. The colour scale shows the distribution of
305 leaves over time: 2019 (green), 2020 (red) and 2021 (black). The dotted line corresponds to the cut
306 height for the optimal number of clusters. **b)** Binary discriminant analysis of main spectra profiles
307 showing the 40 highest ranked peaks contrasting the three yearly periods during 2019, 2020
308 (lockdown), and 2021. Peaks are indicated by their m/z. For each selected peak, the entropic t-score
309 ranking is shown, positive when the peak is present in the group..

310
311 The overall binary analysis (Fig 6b) showed a similarity between lockdown and 2021
312 and a difference from 2019, suggesting a persistence of changes that occurred during the
313 waning due to the lockdown on the protein composition of *S. aureus*, which seems to
314 contradict the results of the comparison between before and after and during the same year,
315 but agrees with the results of comparison with reference period from 2019 and 2021 for the
316 respiratory tract and blood samples. The two binary discriminant analyses shared 14/40 most
317 discriminant peaks, showing a change in the peaks supporting the discrimination.

318 Discussion

319 Our objective was to evaluate the changes in bacterial species incidence during the
320 waning due to the lockdown and associated social distancing measures, and their association
321 with their clonal richness. This study focused on community strains of *Haemophilus*
322 *influenzae* and *Staphylococcus aureus*, two species for which a notable and prolonged
323 decrease was observed throughout this period. Bacterial identifications from blood and
324 respiratory tract were pooled, as previous studies indicated that strains coming from these two
325 kinds of samples could not be distinguished in most cases 19. Additionally, identifications of
326 *S. aureus* from cutaneous samples were explored for comparison, as they are influenced by
327 non-human transmissions 20. To our knowledge, no similar study, focusing on bacteria
328 subspecies diversity, has been conducted thus far. For this purpose, we used MALDI-TOF
329 MS, which is used for routine identification of bacterial species, and which is cost-effective
330 compared to classic genotyping methods 21. This was possible because, in addition to
331 species-specific peaks, MALDI-TOF MS spectra exhibit clusters of frequently co-occurring
332 peaks, enabling subspecies differentiation 12.

333 This study included 251 msp of *H. Influenzae* and 2683 msp (2079 from respiratory
334 sample, 604 from skin samples) of *S. aureus*, representing all identifications of these two
335 species made during the included periods. All identifications were obtained from the routine
336 work of the clinical microbiology laboratories of the four university hospitals of the
337 Assistance Publique-Hôpital de Marseille. However, these results are limited to the AP-HM
338 recruitment, which is primarily regional.

339 All six analyses of the two species revealed a dramatic change in the MALDI-TOF MS
340 proteome composition during the lockdown, suggesting a shift in their strain composition.
341 Regarding *H. influenzae*, this change persisted immediately after the lockdown, with a partial
342 return in 2021 to a strain mix like the pre-lockdown situation. The frequent proximity in the

343 dendrogram of strains coming from the lockdown and immediate post-lockdown periods
344 suggests that subsequently encountered strains may have been selected during the lockdown,
345 with a progressive increase in subspecies diversity over time. The grouping of lockdown
346 spectra in the year comparisons, even if they are spread across several dendrogram clusters,
347 supports the idea that they differ from the spectra of the other periods. Conversely, regardless
348 of sample types, *S. aureus* displayed a change in spectral composition during the lockdown,
349 but without a continuation during the immediate post-lockdown period. Similarities in spectra
350 profiles with the lockdown period emerged later during the 2021 period, confirmed by
351 statistically significant closer proximity of these spectra. While 21 out of 40 most
352 discriminant peaks were shared by the two discriminant analyses of the respiratory and blood
353 samples, indicating a resemblance of distinctive protein profile immediately around the
354 lockdown period with the samples one year later, only 14 out of 40 most discriminant peaks
355 were shared by the skin samples, showing a looser likeliness of spectral profiles one year
356 later.

357 Interpreting the results requires consideration that the apparent population diversity
358 conveyed by MALDI-TOF spectra reflects the species' strain diversity, which is a trade-off
359 between within-host diversities, mutations, inoculum strain heterogeneity, inoculum
360 frequency, environmental selection, and transmission bottleneck. Previous simulation studies
361 showed that small bottlenecks rapidly result in a single dominant strain carried and
362 transmitted in the infected population 5. Our results indicated a decrease in strain diversity for
363 both studied bacteria. An hypothesis is that the lockdown and preventive measures could have
364 been associated with a reduction in inoculum frequency and an increase in the bottleneck
365 effect due to reduced social interaction, and could have contributed to a reduction in genetic
366 diversity within the bacterial population. Additionally, strict hygiene measures, such as
367 frequent handwashing, use of disinfectants, and wearing masks, all distancing measures used

368 during the end of 2020, were reported to reduce bacteria diversity 22. Household transmission
369 could have become the primary mode of transmission, at the expense of community
370 transmission.

371 *H. influenzae* is an obligate commensal of the human respiratory tract and non-typeable
372 *H. influenzae* (NTHi) are currently the predominant strains 23. NTHi strains are more diverse
373 and influenced by recombination compared to encapsulated *H. influenzae* 24. These genetic
374 mechanisms, altering the gene expression or content, generate strains with fitness
375 characteristics aimed at surviving under various selective pressure within a same kind of host
376 25. In particular, NTHi are able to switch between genetic variants at high frequency (phase
377 variation mechanism) 26. This capability of dynamic recombination could explain why the
378 change in *H. influenzae* strains' diversity was only transitory for *H. influenzae*.

379 On the contrary, *S. aureus* is found in warm-blood animals and in the environment,
380 constituting an important reservoir 20. It shows a slow but detectable genetic evolution of its
381 population, at a rate of 2.72 mutations per megabase per year 19, in association with a fast
382 replication rate 27. Twenty to twenty-five percent of the human population present a
383 persistent carriage of *S. aureus* 28, which is rarely polyclonal (about 6.6 %) 29, mainly
384 influenced by household transmission 20, possibly potentialized by host genetic factors 30.
385 Epidemiologically isolated populations have shown very low diversity of their carriage 30.
386 The rapid return to the pre-lockdown spectral configuration, followed by a progression toward
387 the lockdown situation one year later, may be explained by a first influence of the reservoir
388 multiplicity, followed by a return to the predominance of household transmissions, strengthen
389 by the post-lockdown preventive measures.

390 This study showed that, for our human population and the studied bacteria species,
391 lockdown was associated with a dramatic change of species compositions, but with different
392 further evolution. We observed a transitory persistence of the strain composition that occurred

393 during the lockdown for a human-dependant bacteria species, with a high frequency of
394 genetic variant. Conversely, we observed a long-term evolution toward a return to the
395 lockdown composition for a bacterial species with multiple reservoirs, a slow mutation rate,
396 and an important household-dependant transmission.

397

398

399 **References**

- 400 1 Rotulo GA, Percivale B, Molteni M, et al. The impact of COVID-19 lockdown on
401 infectious diseases epidemiology: The experience of a tertiary Italian Pediatric
402 Emergency Department. *The Am J Emerg Med.* 2021;43:115–7.
- 403 2 Vierucci F, Bacci C, Mucaria C, Dini F, Federico G, Maielli M, Vaccaro A. How
404 COVID-19 Pandemic Changed Children and Adolescents Use of the Emergency
405 Department: the Experience of a Secondary Care Pediatric Unit in Central Italy. *SN*
406 *Comprehensive Clinical Medicine.* 2020;2:1959–1969.
- 407 3 Brueggemann AB, Jansen van Rensburg MJ, Shaw D, McCarthy ND. Changes in the
408 incidence of invasive disease due to *Streptococcus pneumoniae*, *Haemophilus*
409 *influenzae*, and *Neisseria meningitidis* during the COVID-19 pandemic in 26 countries
410 and territories in the Invasive Respiratory Infection Surveillance Initiative: a
411 prospective analysis of surveillance data. *Lancet Digit Health.* 2021 Jun;3(6):e360-
412 e370.
- 413 4 Kaba L, Giraud-Gatineau A, Jimeno MT, Rolain JM, Colson P, Raoult D, et al.
414 Consequences of the COVID-19 Outbreak Lockdown on Non-Viral Infectious Agents
415 as Reported by a Laboratory-Based Surveillance System at the IHU Méditerranée
416 Infection, Marseille, France. *Journal of Clinical Medicine.* 21 juill 2021;10(15):3210.
- 417 5 Worby CJ, Lipsitch M, Hanage WP. Within-host bacterial diversity hinders accurate
418 reconstruction of transmission networks from genomic distance data. *PLoS Comput*
419 *Biol.* 2014 Mar 27;10(3):e1003549. doi: 10.1371/journal.pcbi.1003549. PMID:
420 24675511; PMCID: PMC3967931.
- 421 6 Seng P, Rolain JM, Fournier PE, La Scola B, Drancourt M, Raoult D. MALDI-TOF-
422 mass spectrometry applications in clinical microbiology. *Future Microbiology.* nov
423 2010;5(11):1733-54.

- 424 7 Freiwald A, Sauer S. Phylogenetic classification and identification of bacteria by mass
425 spectrometry. *Nat Protoc.* mai 2009;4(5):732-42.
- 426 8 Månsson V, Resman F, Kostrzewa M, Nilson B, Riesbeck K. Identification of
427 *Haemophilus influenzae* Type b Isolates by Use of Matrix-Assisted Laser Desorption
428 Ionization–Time of Flight Mass Spectrometry. McAdam AJ, éditeur. *J Clin Microbiol.*
429 juill 2015;53(7):2215-24.
- 430 9 Josten M, Reif M, Szekat C, Al-Sabti N, Roemer T, Sparbier K, et al. Analysis of the
431 Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrum of
432 *Staphylococcus aureus* Identifies Mutations That Allow Differentiation of the Main
433 Clonal Lineages. *J Clin Microbiol.* juin 2013;51(6):1809-17.
- 434 10 Ueda O, Tanaka S, Nagasawa Z, Hanaki H, Shobuike T, Miyamoto H. Development
435 of a novel matrix-assisted laser desorption/ionization time-of-flight mass spectrum
436 (MALDI-TOF-MS)-based typing method to identify meticillin-resistant
437 *Staphylococcus aureus* clones. *Journal of Hospital Infection.* juin 2015;90(2):147-55.
- 438 11 Sauget M, van der Mee-Marquet N, Bertrand X, Hocquet D. Matrix-assisted laser
439 desorption ionization-time of flight Mass spectrometry can detect *Staphylococcus*
440 *aureus* clonal complex 398. *Journal of Microbiological Methods.* août 2016;127:20-3.
- 441 12 Giraud-Gatineau A, Texier G, Garnotel E, Raoult D, Chaudet H. Insights Into
442 Subspecies Discrimination Potentiality From Bacteria MALDI-TOF Mass Spectra by
443 Using Data Mining and Diversity Studies. *Frontiers in Microbiology* [Internet]. 13
444 août 2020 [cité 23 juin 2022];11. Disponible sur:
445 <https://www.frontiersin.org/article/10.3389/fmicb.2020.01931/full>
- 446 13 Giraud-Gatineau A, Texier G, Fournier PE, Raoult D, Chaudet H. Using MALDI-TOF
447 spectra in epidemiological surveillance for the detection of bacterial subgroups with a
448 possible epidemic potential. *BMC Infect Dis.* déc 2021;21(1):1109.

- 449 14 Gibb S, Strimmer K. MALDIquant: a versatile R package for the analysis of mass
450 spectrometry data. *Bioinformatics*. 1 sept 2012;28(17):2270-1.
- 451 15 Rousseeuw P.J. Silhouettes: a graphical aid to the interpretation and validation of
452 cluster analysis. *Computational and Applied Mathematics* 1987; 20: 53–65
- 453 16 Wald and J. Wolfowitz, "On a test whether two samples are from the same
454 population" (*Annals of Math. Stat.*, Vol. XI, No. 2, June (1940) pp. 147
- 455 17 Barnard GA (1945). A new test for 2×2 tables. *Nature* 156:177
- 456 18 Gibb S, Strimmer K. Differential protein expression and peak selection in mass
457 spectrometry data by binary discriminant analysis. *Bioinformatics*. 1 oct
458 2015;31(19):3156-62.
- 459 19 Young BC, Golubchik T, Batty EM, Fung R, Larner-Svensson H, Votintseva AA,
460 Miller RR, Godwin H, Knox K, Everitt RG, Iqbal Z, Rimmer AJ, Cule M, Ip CL,
461 Didelot X, Harding RM, Donnelly P, Peto TE, Crook DW, Bowden R, Wilson DJ.
462 Evolutionary dynamics of *Staphylococcus aureus* during progression from carriage to
463 disease. *Proc Natl Acad Sci U S A*. 2012 Mar 20;109(12):4550-5.
- 464 20 Knox J, Uhlemann AC, Lowy FD. *Staphylococcus aureus* infections: transmission
465 within households and the community. *Trends in Microbiology*. juill
466 2015;23(7):437-44.
- 467 21 Seng P, Abat C, Rolain JM, Colson P, Lagier JC, Gouriet F, Fournier PE, Drancourt
468 M, La Scola B, Raoult D. Identification of rare pathogenic bacteria in a clinical
469 microbiology laboratory: impact of matrix-assisted laser desorption ionization-time of
470 flight mass spectrometry. *J Clin Microbiol*. 2013 Jul;51(7):2182-94.
- 471 22 Finlay BB, Amato KR, Azad M, Blaser MJ, Bosch TCG, Chu H, Dominguez-Bello
472 MG, Ehrlich SD, Elinav E, Geva-Zatorsky N, Gros P, Guillemin K, Keck F, Korem T,
473 McFall-Ngai MJ, Melby MK, Nichter M, Pettersson S, Poinar H, Rees T, Tropini C,

- 474 Zhao L, Giles-Vernick T. The hygiene hypothesis, the COVID pandemic, and
475 consequences for the human microbiome. *Proc Natl Acad Sci U S A*. 2021 Feb
476 9;118(6):e2010217118. doi: 10.1073/pnas.2010217118. Erratum in: *Proc Natl Acad*
477 *Sci U S A*. 2021 Mar 16;118(11): PMID: 33472859; PMCID: PMC8017729.
- 478 23 De Chiara M, Hood D, Muzzi A, Pickard DJ, Perkins T, Pizza M, Dougan G,
479 Rappuoli R, Moxon ER, Soriani M, Donati C. Genome sequencing of disease and
480 carriage isolates of nontypeable *Haemophilus influenzae* identifies discrete population
481 structure. *Proc Natl Acad Sci U S A*. 2014 Apr 8;111(14):5439-44.
- 482 24 Meats E, Feil EJ, Stringer S, Cody AJ, Goldstein R, Kroll JS, Popovic T, Spratt BG.
483 Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and
484 determination of phylogenetic relationships by multilocus sequence typing. *J Clin*
485 *Microbiol*. 2003 Apr;41(4):1623-36.
- 486 25 Gilsdorf JR, Marrs CF, Foxman B. *Haemophilus influenzae*: genetic variability and
487 natural selection to identify virulence factors. *Infect Immun*. 2004 May;72(5):2457-61.
- 488 26 Bayliss CD, Sweetman WA, Moxon ER. Mutations in *Haemophilus influenzae*
489 mismatch repair genes increase mutation rates of dinucleotide repeat tracts but not
490 dinucleotide repeat-driven pilin phase variation rates. *J Bacteriol*. 2004
491 May;186(10):2928-35.
- 492 27 Szafrńska AK, Junker V, Steglich M, Nübel U. Rapid cell division of *Staphylococcus*
493 *aureus* during colonization of the human nose. *BMC Genomics*. 2019 Mar
494 20;20(1):229.
- 495 28 Sollid JU, Furberg AS, Hanssen AM, Johannessen M. *Staphylococcus aureus*:
496 determinants of human carriage. *Infect Genet Evol*. 2014 Jan;21:531-41.

- 497 29 Cespedes C, Said-Salim B, Miller M, Lo SH, Kreiswirth BN, Gordon RJ, Vavagiakis
498 P, Klein RS, Lowy FD. The clonality of Staphylococcus aureus nasal carriage. J Infect
499 Dis. 2005 Feb 1;191(3):444-52.
- 500 30 Ruimy R, Angebault C, Djossou F, Dupont C, Epelboin L, Jarraud S, Lefevre LA, Bes
501 M, Lixandru BE, Bertine M, El Miniai A, Renard M, Bettinger RM, Lescat M,
502 Clermont O, Peroz G, Lina G, Tavakol M, Vandenesch F, van Belkum A, Rousset F,
503 Andremont A. Are host genetics the predominant determinant of persistent nasal
504 Staphylococcus aureus carriage in humans? J Infect Dis. 2010 Sep 15;202(6):924-34.
505

medRxiv preprint doi: <https://doi.org/10.1101/2024.05.08.24307047>; this version posted May 9, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

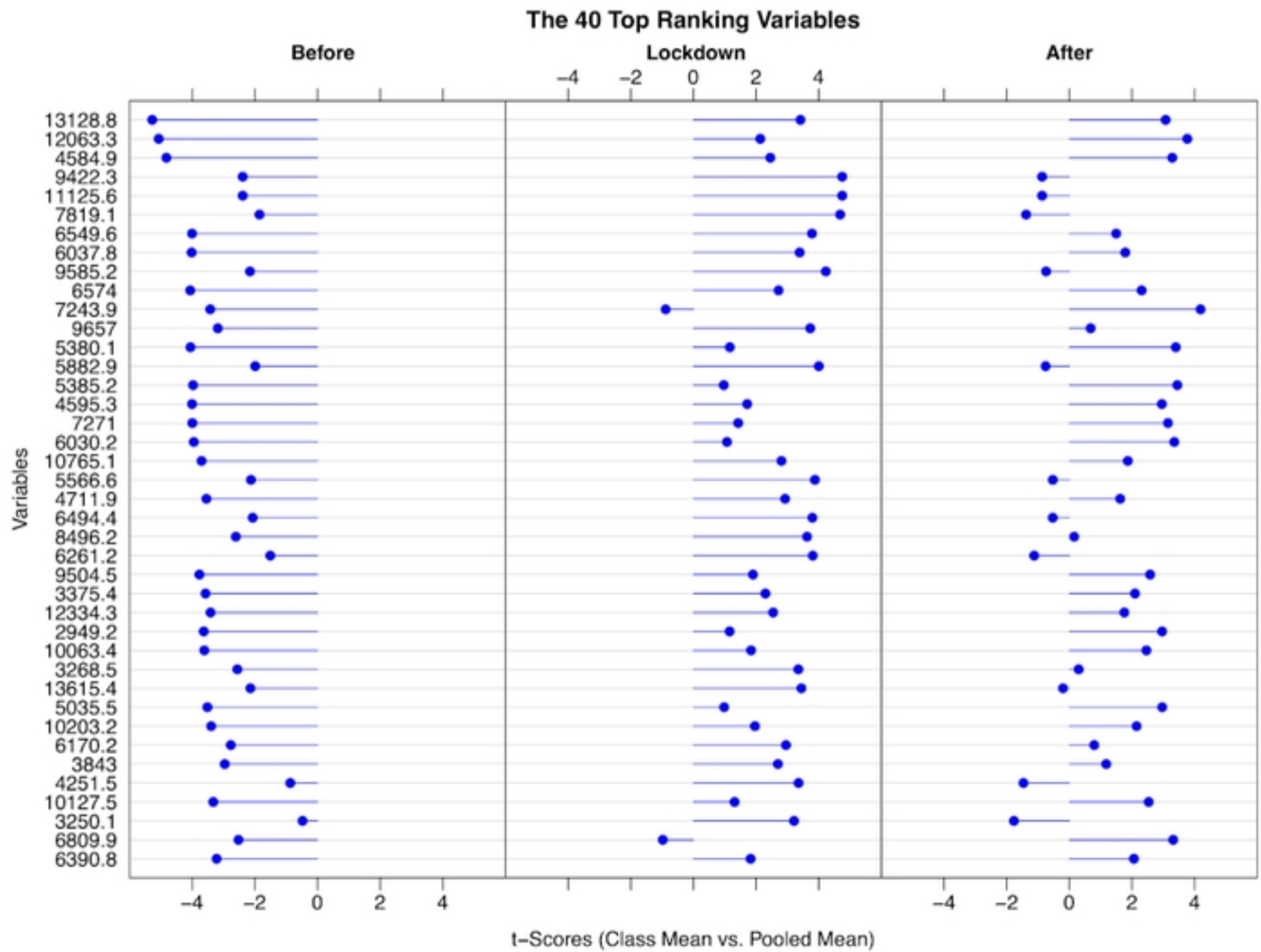
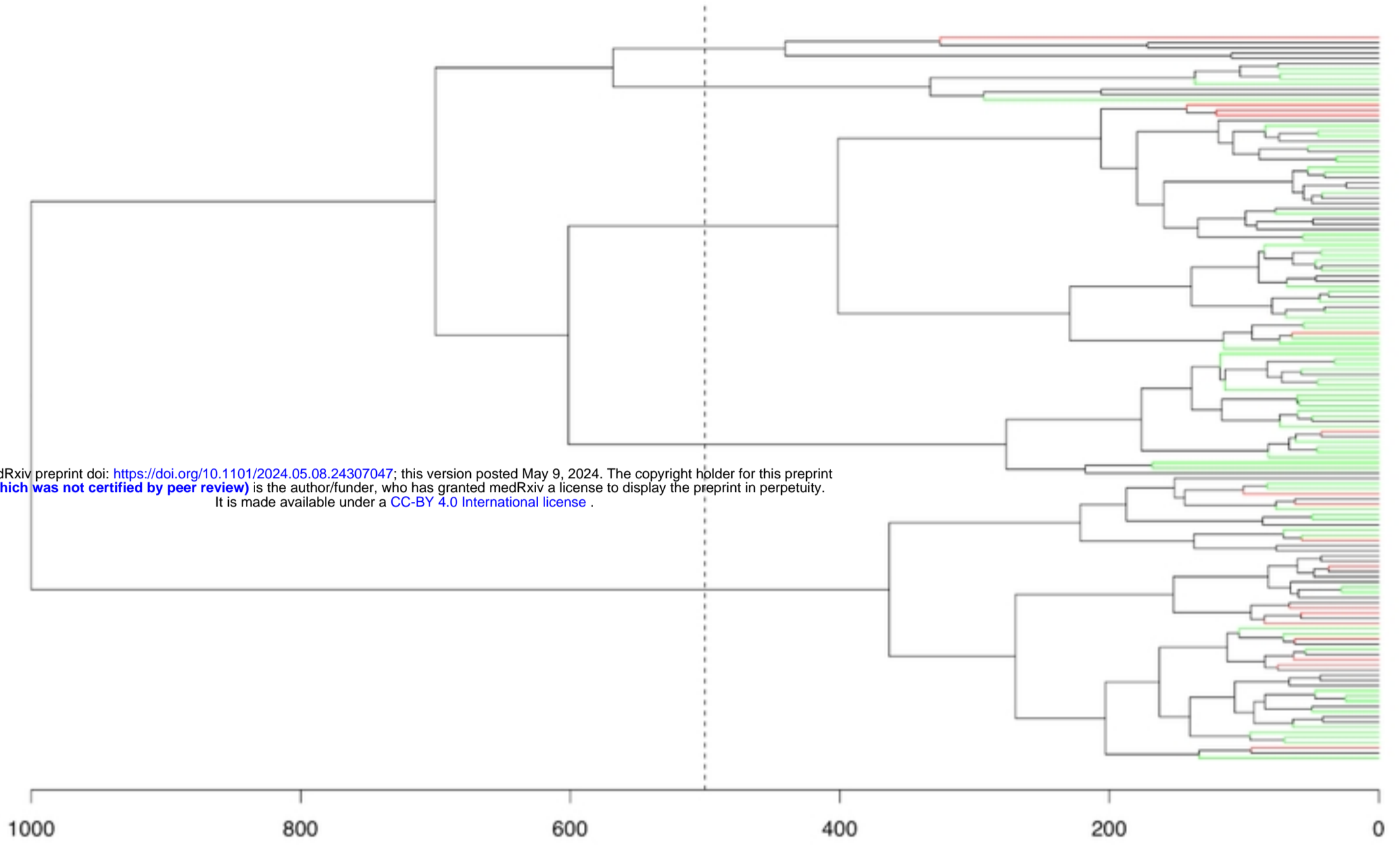


Figure1

medRxiv preprint doi: <https://doi.org/10.1101/2024.05.08.24307047>; this version posted May 9, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

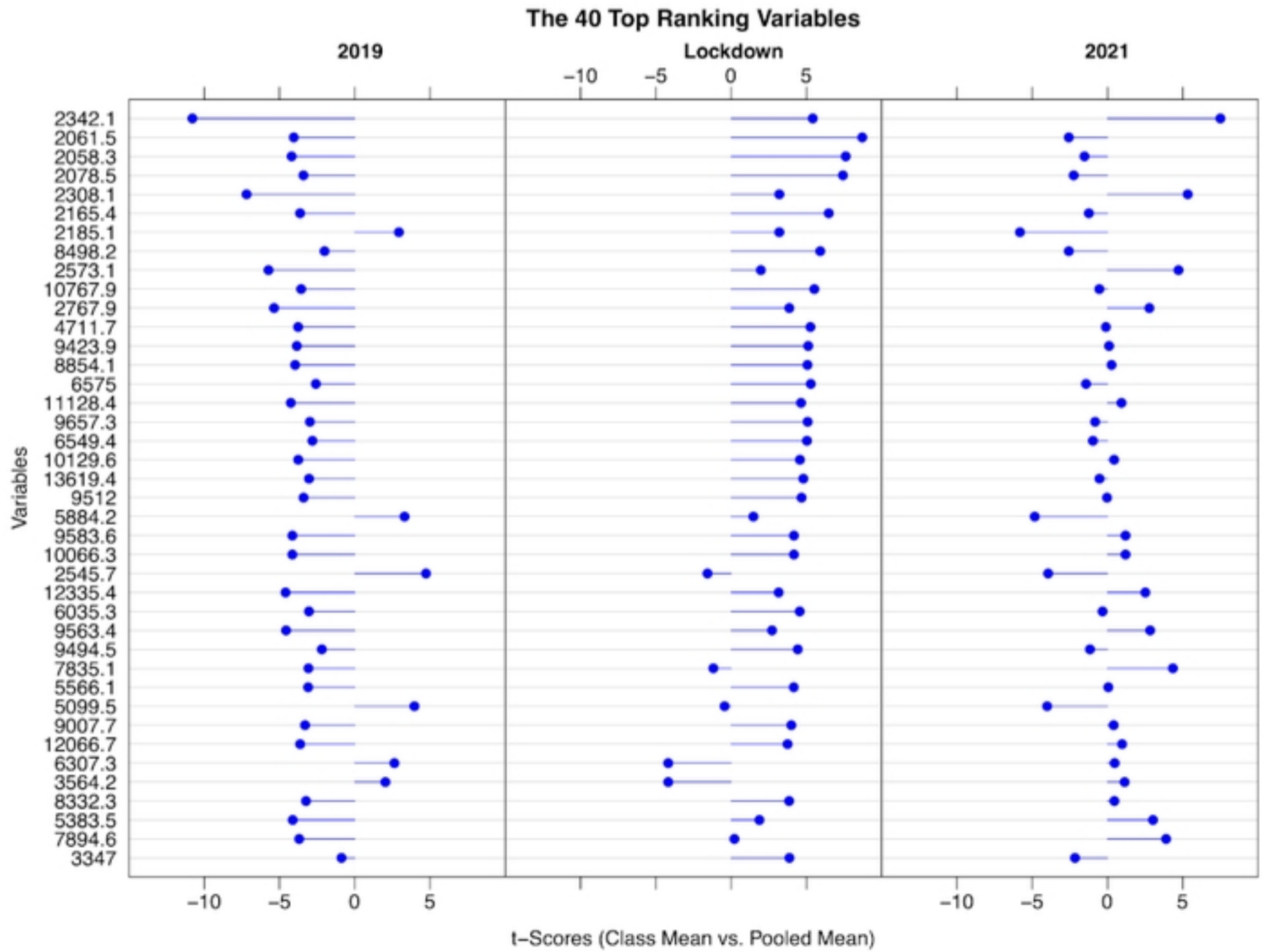
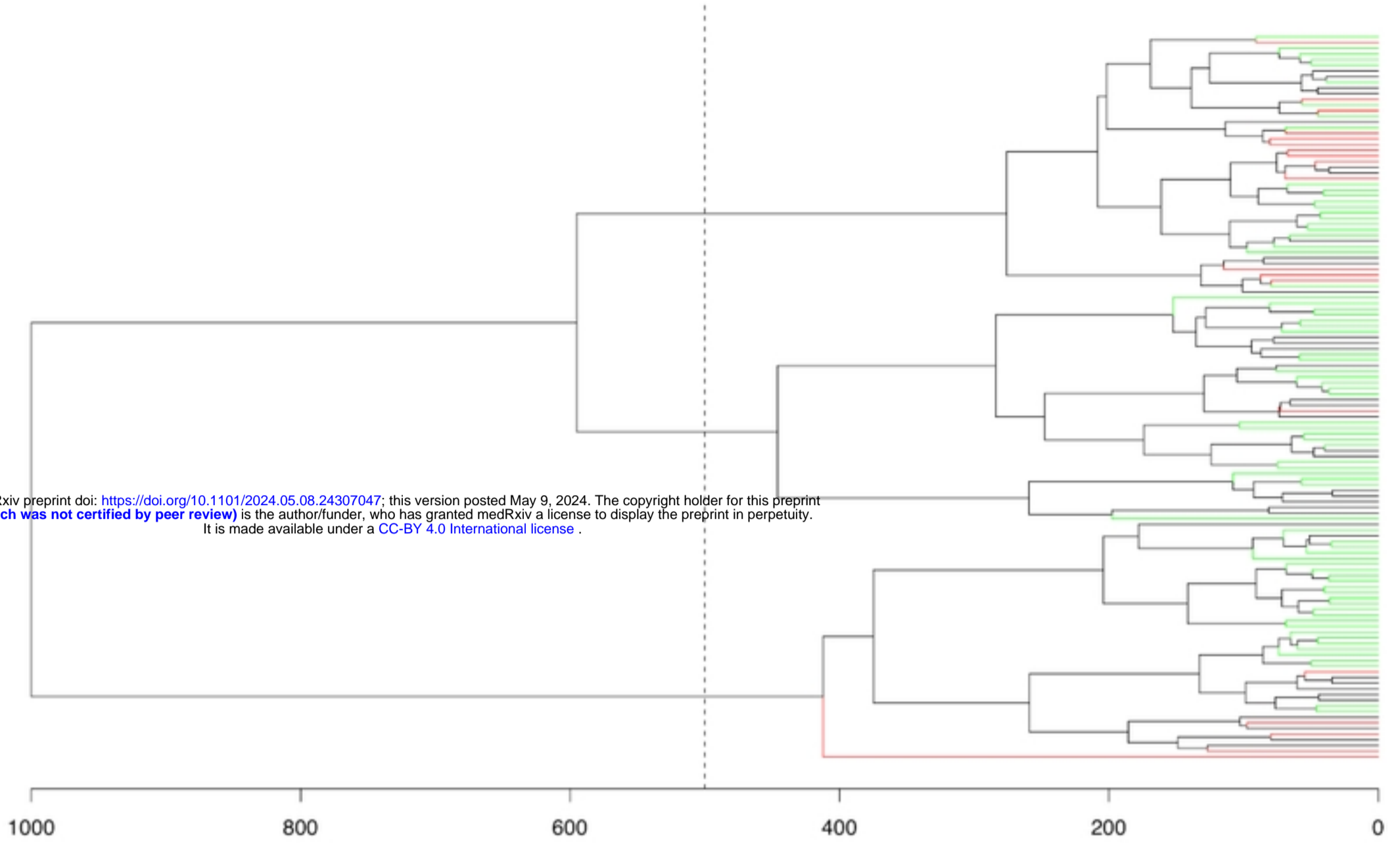


Figure2

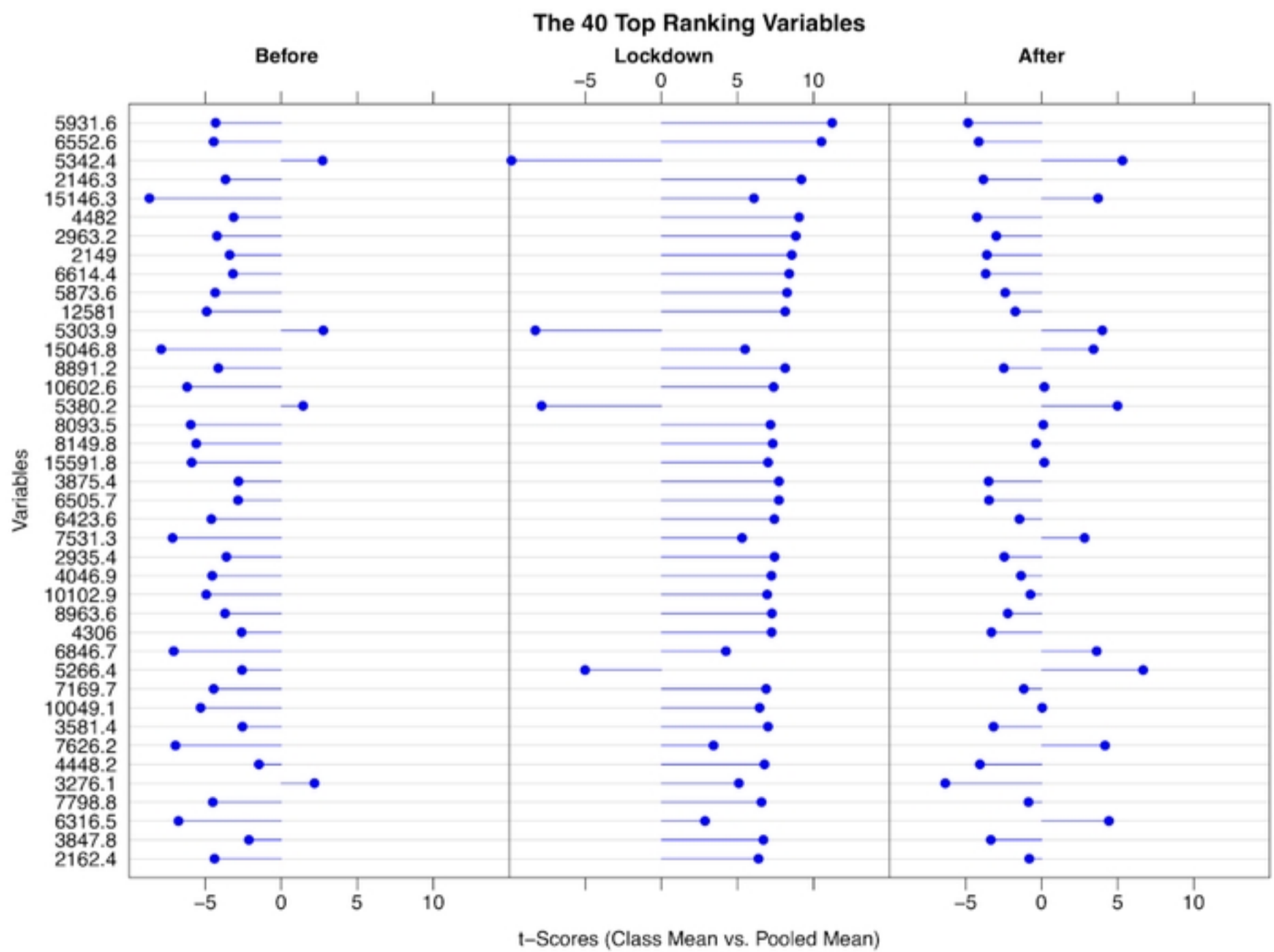
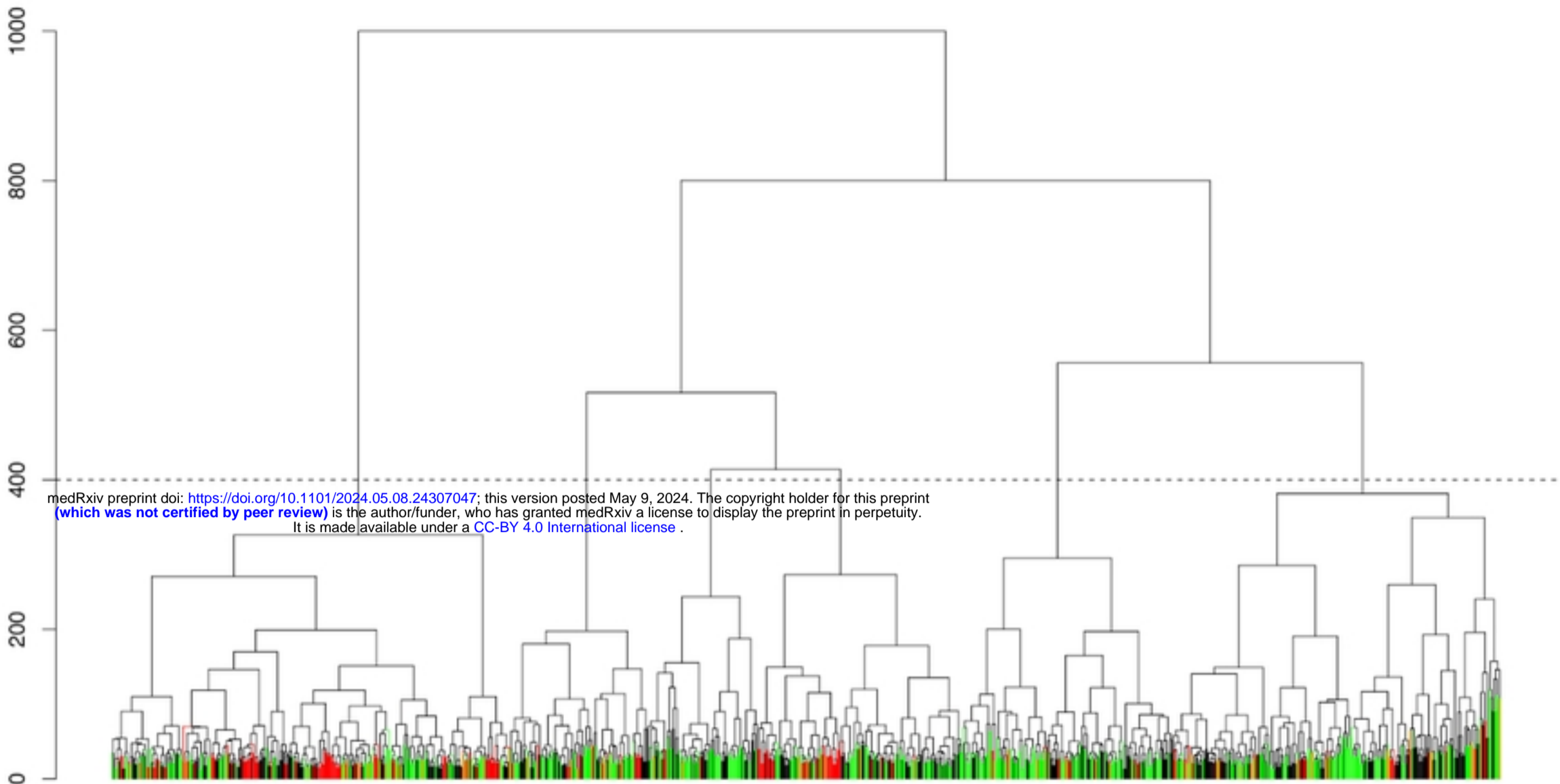


Figure3

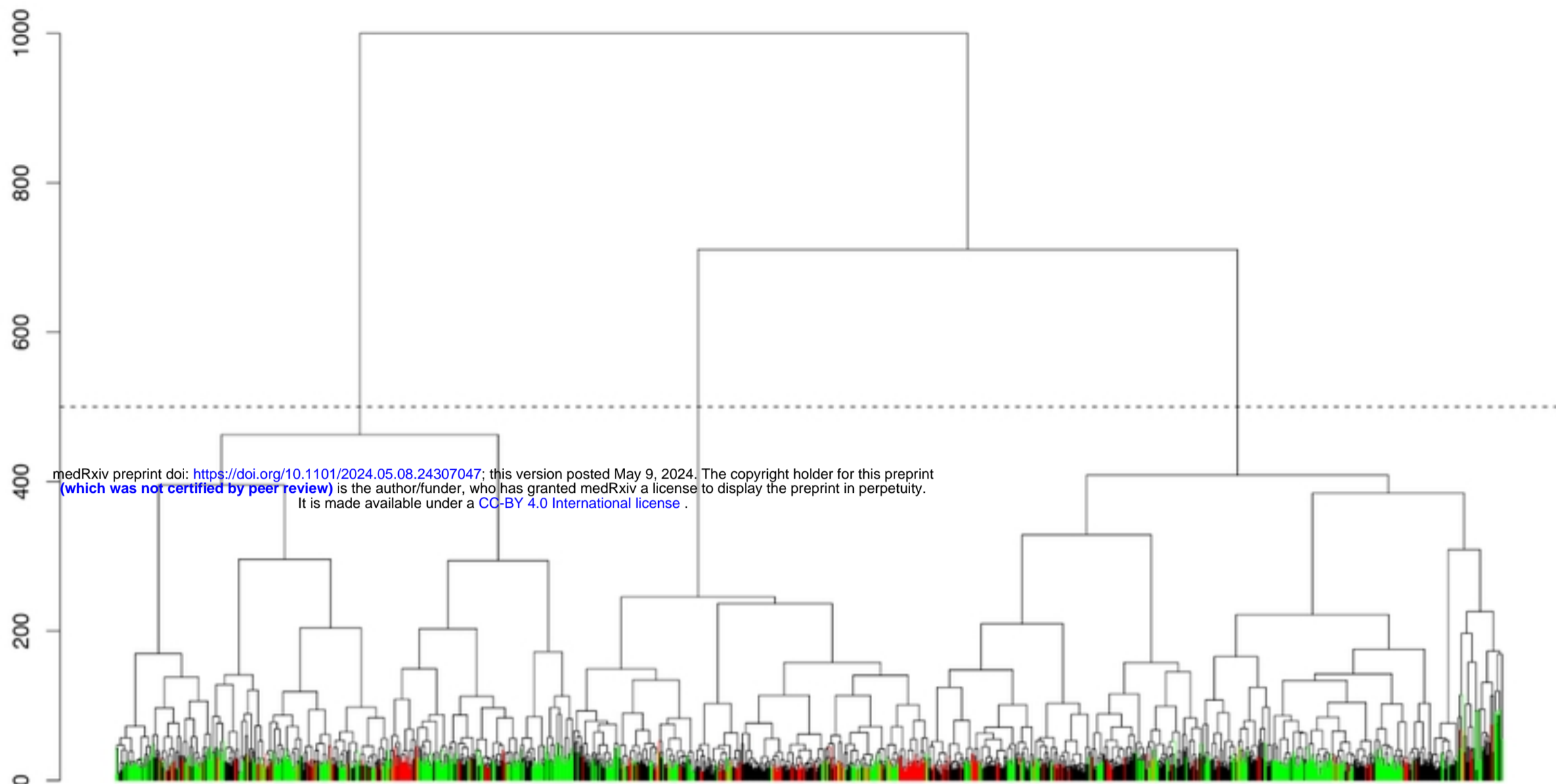


Figure4

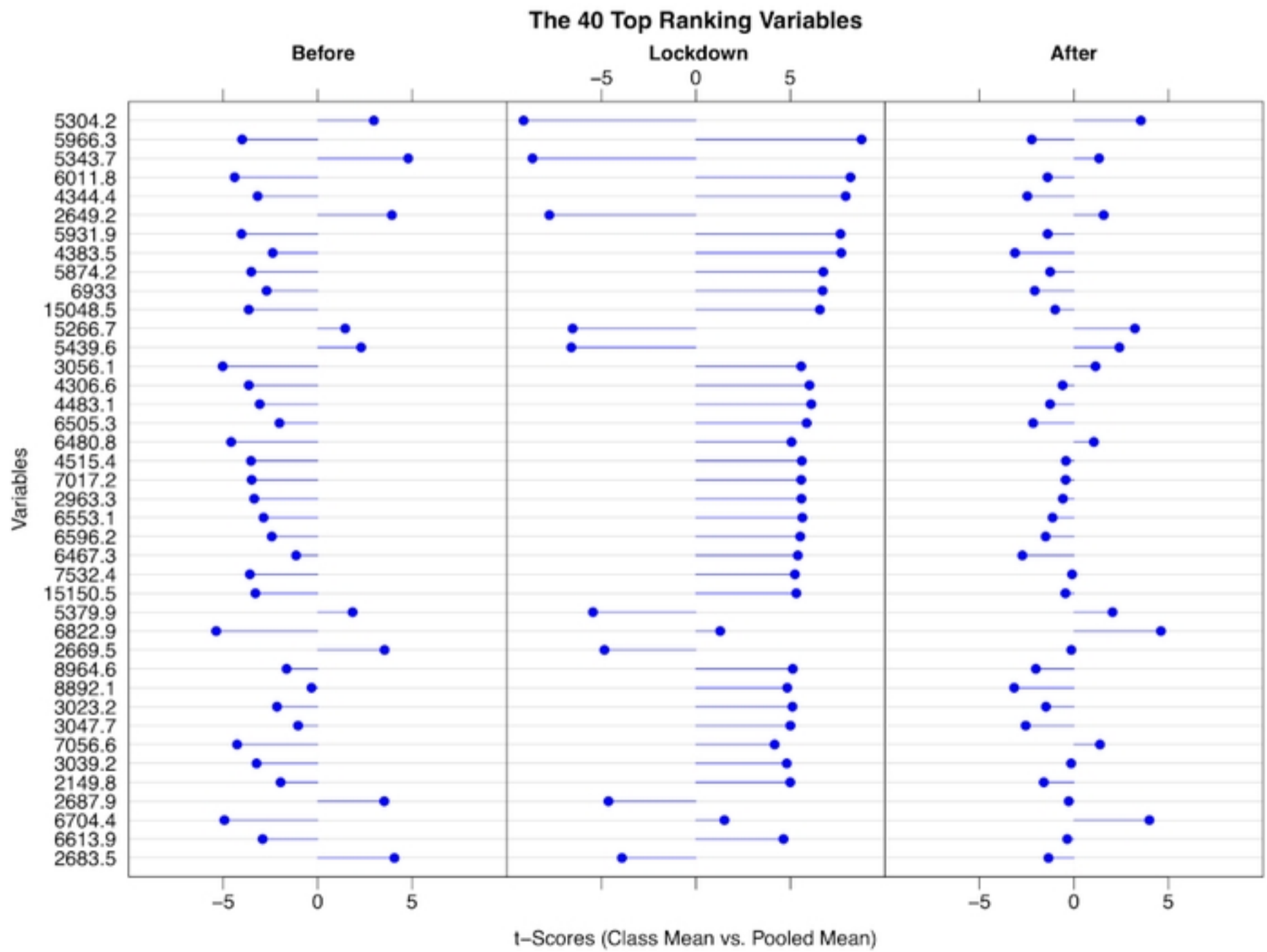
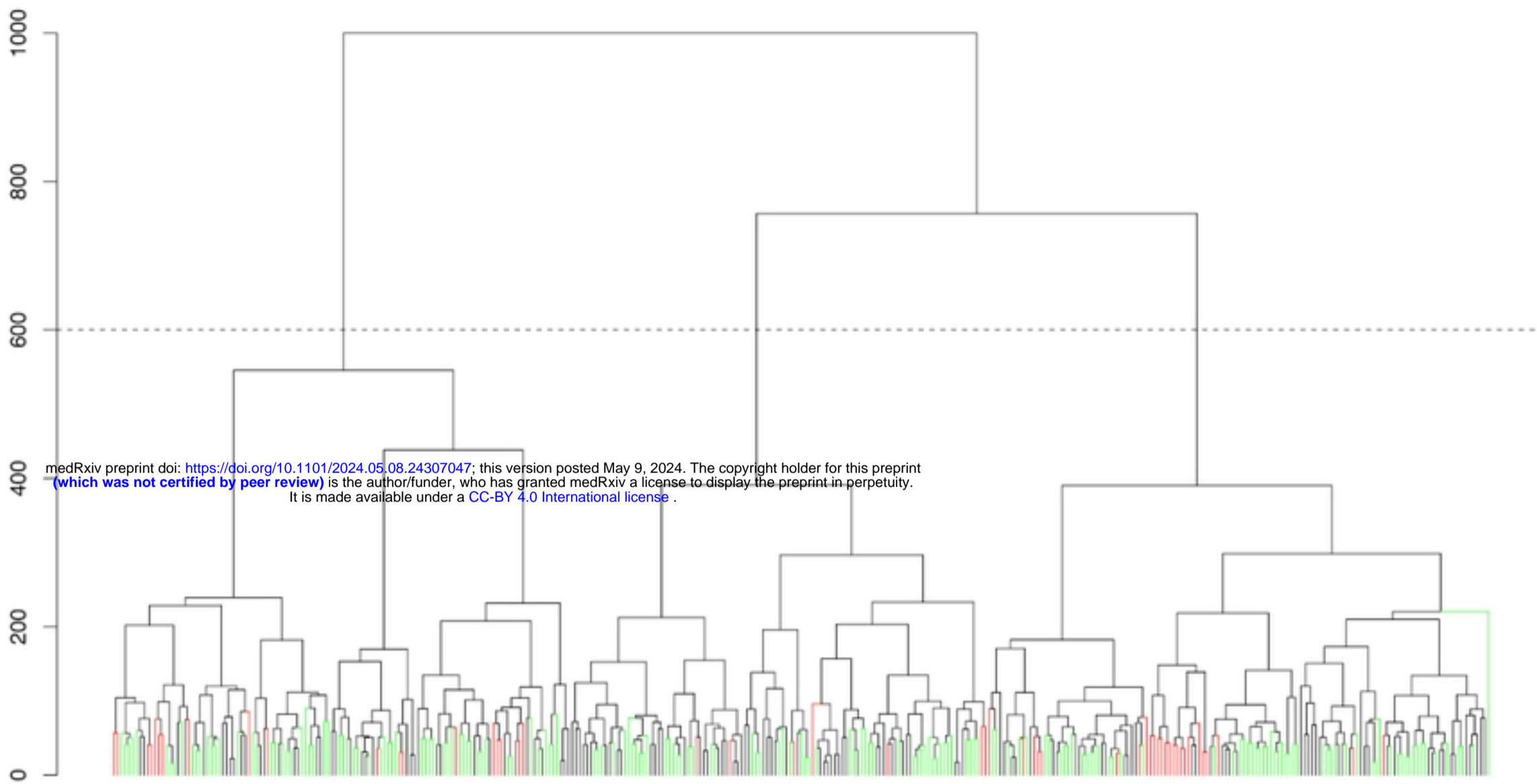


Figure5

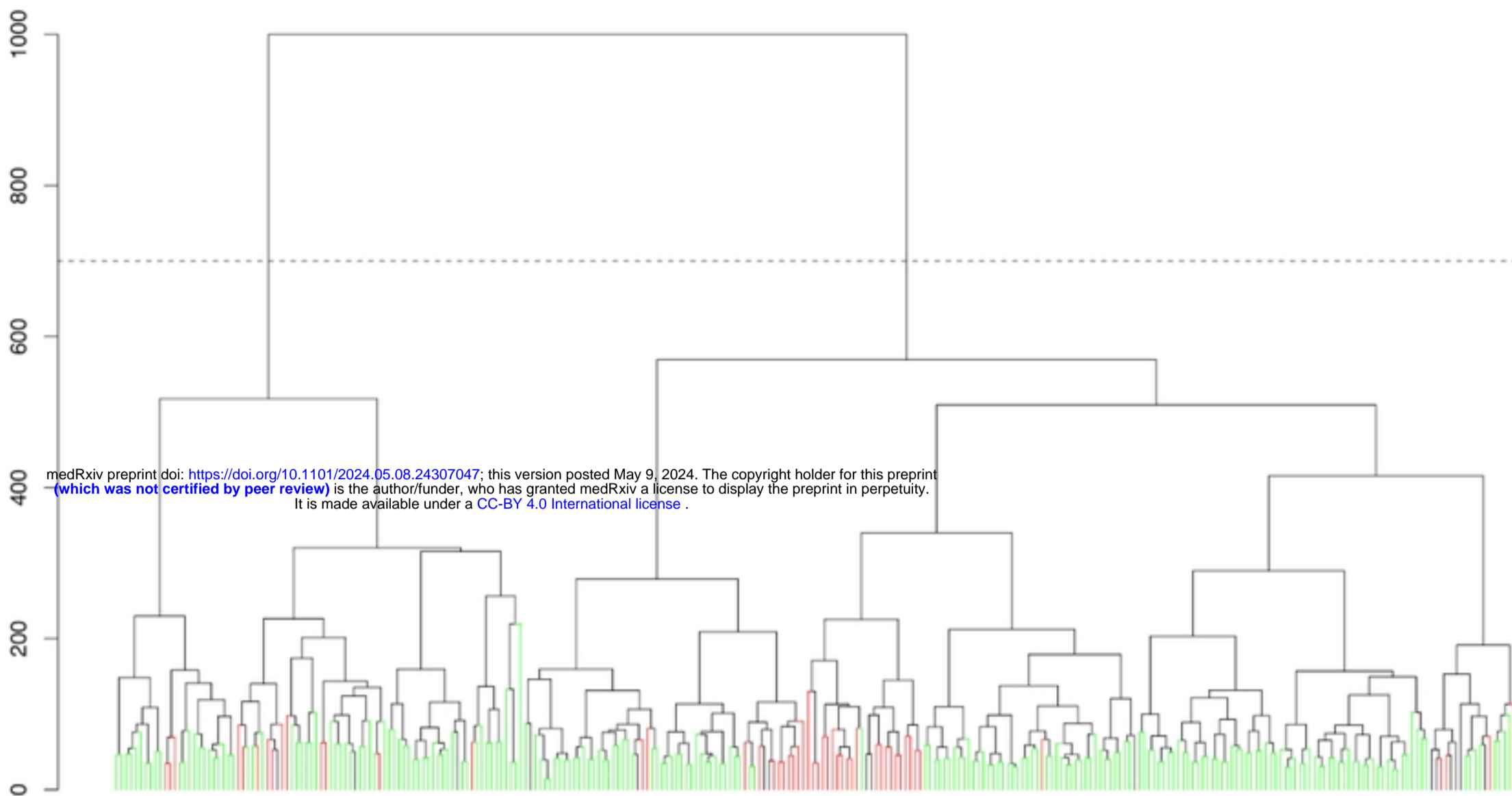


Figure6