

1 **Genomic epidemiology and evolution of rhinovirus in western Washington State, 2021-22**

2

3 Stephanie Goya¹, Seffir T. Wendm¹, Hong Xie¹, Tien V. Nguyen¹, Sarina Barnes¹, Rohit R.
4 Shankar¹, Jaydee Sereewit¹, Kurtis Cruz¹, Ailyn C. Pérez-Osorio¹, Margaret G. Mills¹, Alexander
5 L Greninger^{1,2,#}.

6

7 ¹Department of Laboratory Medicine and Pathology, University of Washington Medical Center,
8 Seattle, Washington, USA

9 ² Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle,
10 Washington, USA

11

12

13 **Abstract**

14 **Background:** Human rhinoviruses (RV) primarily cause the common cold, but infection
15 outcomes vary from subclinical to severe cases, including asthma exacerbations and fatal
16 pneumonia in immunocompromised individuals. To date, therapeutic strategies have been
17 hindered by the high diversity of serotypes. Global surveillance efforts have traditionally focused
18 on sequencing VP1 or VP2/VP4 genetic regions, leaving gaps in understanding RV true
19 genomic diversity.

20 **Methods:** We sequenced 1,003 RV genomes from nasal swabs of symptomatic and
21 asymptomatic individuals to explore viral evolution during two epidemiologically distinct periods
22 in Washington State: when the COVID-19 pandemic affected the circulation of other seasonal
23 respiratory viruses except for RV (February – July 2021), and when the seasonal viruses
24 reemerged with the severe RSV and influenza outbreak (November-December 2022). We
25 constructed maximum likelihood and BEAST-phylogenetic trees to characterize intra-genotype
26 evolution.

27 **Results:** We detected 100 of 168 known genotypes, identified two new genotypes (A111 and
28 C59), and observed inter-genotypic recombination and genotype cluster swapping from 2021 to
29 2022. We found a significant association between the presence of symptoms and viral load, but
30 not with RV species or genotype. Phylodynamic trees, polyprotein selection pressure, and
31 Shannon diversity revealed co-circulation of divergent clades within genotypes with high amino
32 acid constraints throughout polyprotein.

33 **Discussion:** Our study underscores the dynamic nature of RV genomic epidemiology within a
34 localized geographic region, as more than 20% of existing genotypes within each RV species
35 co-circulated each month. Our findings also emphasize the importance of investigating
36 correlations between rhinovirus genotypes and serotypes to understand long-term immunity and
37 cross-protection.

38

39

40 **Background**

41 Human rhinoviruses (RV) are a primary etiology of the common cold, causing 8.9 – 19 % of
42 total acute upper respiratory infections (ARI) and 4 – 6.1% of asymptomatic respiratory
43 infections [1–4]. A recent study of RV socioeconomical burden in New York City found that
44 23.5% of the individuals with ARIs associated with RV missed at least 1 day of work or school
45 (1). RV can also cause lower respiratory tract infections in immunocompromised patients with
46 high mortality rate [5,6], and are a major contributor of exacerbations of asthma [7–9], cystic
47 fibrosis [10] and chronic obstructive pulmonary disease [11,12]. While 100 different RV
48 serotypes were described by 1987 [13], and the first complete RV genome was published in
49 1984 [14], fewer than 500 complete RV genomes were available in public databases at time of
50 initiation of this study (January 2021), meaning most RV genotypes had fewer than 5 genomes
51 available. Despite its more widespread disease burden, RV genomic data pale in comparison to
52 the 163,000 influenza virus genomes or 11,500 respiratory syncytial virus (RSV) genomes
53 currently available in GISAID or INSDC databases.

54

55 RV are non-enveloped viruses with a single-stranded positive-sense genome of around 7,300
56 nucleotides in length. The RV genome is organized into a single polyprotein that undergoes
57 autocatalytically-based and protease-based cleavage to produce 4 capsid proteins (VP1 – VP4)
58 and 7 non-structural proteins [15]. Global RV molecular epidemiology is based on phylogenetic
59 association of VP1 genetic sequence, which allowed to classify three RV species (RV-A, RV-B,
60 and RV-C), subclassified into genotypes (81 RV-A, 32 RV-B, and 55 RV-C genotypes) [16]. The
61 co-circulation of many highly diverse genotypes, considered antigenically distinct, hampers the
62 development of an effective broadly reactive RV vaccine [17].

63

64 Here, using a set of more than one thousand newly generated RV genomes constituting more
65 than half of publicly available RV genomes, we describe the epidemiology and evolution of RV
66 infections from symptomatic and asymptomatic individuals in Puget Sound region (Washington
67 State, USA) during two distinct epidemiological time periods of the COVID-19 pandemic [18–
68 22].

69

70 **Results**

71 *Demographic and clinical characteristics of RV infections*

72 Our RV screening was based on a subset of SARS-CoV-2 negative nasal swab
73 specimens collected at the UW COVID-19 community testing sites taken during February-July
74 2021 and November-December 2022. A total of 1,083 out of 10,050 nasal swab specimens
75 (10.8%) were RV positive from February-July 2021, and 820 out of 10,656 samples (7.8%) from
76 November-December 2022. The RV positive population consisted of 49% male persons and
77 57% between the ages of 18 to 65 years (Table 1).

78 Community testing site prevalence data correlated with that of Seattle hospital-based
79 respiratory virus PCR surveillance data, which showed a RV monthly positivity rate of 13-25%
80 from February-July 2021 and 5-10% in November-December 2022 (Figure 1). These two
81 periods were epidemiologically distinct for respiratory virus circulation, as hospital-based
82 surveillance data during February-July 2021 yielded a 4% SARS-CoV-2 positivity rate with no
83 detection of other seasonal respiratory viruses. During November-December 2022, SARS-CoV-
84 2 had a 10% average positivity rate, while seasonal respiratory viruses strongly re-emerged with
85 a 34% positivity rate for RSV in November 2022, and 30% for influenza virus in both November
86 and December.

87 Among RV positives, persons aged 5 to 17 years exhibited significantly lower Ct values
88 compared to individuals older than 18 years (Wilcoxon test, $p\text{-adj}=1.0\times 10^{-4}$, Figure 2A).
89 Symptomatic individuals comprised 66% of the RV positive population, including cough (35%),

90 sore throat (23%), runny nose (17%), headache (13%), congestion (11%), and fever (9%) being
91 the most common. Less commonly reported symptoms (<5% of cases) included sneezing, body
92 ache, fatigue, difficulty breathing, and vomiting. Individuals reporting respiratory symptoms had
93 lower Ct values than the asymptomatic individuals (Kruskal-Wallis test, $p=2.1 \times 10^{-11}$, Figure 2B).
94 RV positives were slightly more likely to be symptomatic in 2021 compared to 2022 (70% vs.
95 63%) (Table 1). No association was found between the age of the individuals and the presence
96 of symptoms (Figure 2C).

97 Older individuals were associated with lower viral loads when adjusted for the presence
98 of symptoms (ANOVA, $p=0.03$) (Figure 2D). Individuals with RV Ct value <25 had a significantly
99 higher likelihood of reporting respiratory symptoms than those with Ct value >25 (Wald odds
100 ratio= 1.63, 95% confidence limits: 1.03-2.59).

101 102 *Epidemiological and geographical characterization of RV species*

103 We first characterized RV epidemiology at the species level. Our metagenomic sequencing
104 effort resulted in 1,003 complete or almost complete RV genomes (883 from 2021 and 120 from
105 2022), consisting of 588 RV-A, 119 RV-B, and 296 RV-C. Seasonal analysis revealed an overall
106 shift from RV-C to RV-A prevalence from February to July 2021. Combined with the higher
107 prevalence of RV-C during November and December 2022, our data indicate a higher relative
108 prevalence of RV-A in spring and summer with RV-C circulating at higher levels in autumn and
109 winter (Figure 3A) [23]. RV-B exhibited 9.7% of the total cases in 2021 and 27.5% in 2022.

110 No difference in the presence of symptoms was observed by RV species (G-test of
111 independence, p value>0.05). Children under 5 years old were more likely to be infected by RV-
112 A or RV-C than RV-B (G-test of independence, p value = 2.02×10^{-4} ; CMLE odds ratio [95%
113 confidence] RV-C vs RV-A=2.02 [1.19-3.47], RV-C vs RV-B=17.61 [2.86-723.92], RV-A vs RV-
114 B=8.70 [1.39-361.16]). No tendency was found when analyzing RV species by non-stratified
115 age.

116 We investigated if RV species distribution differed based on the location of COVID-19
117 community testing sites where specimens were collected. We found that the co-circulation,
118 prevalence, and replacement of RV-A to RV-C was not geographically distinct during 2021,
119 while no clear prevalence nor replacement was found in 2022, consistent with the overall
120 seasonality (Figure 4).

121

122 *RV genotype diversity in the Puget Sound region*

123 We next characterized the RV diversity at the genotype level. Datasets of 1,327 RV-A,
124 314 RV-B, and 587 RV-C sequences were analyzed comprising all whole genomes available in
125 NCBI Virus (October 2023), and VP1 reference sequences were included to represent RV-A
126 and RV-C genotypes without available genomes. Within this dataset, 60% of sequences were
127 collected in Washington State. To evaluate the continuity or reemergence of the genotype
128 circulation in the Puget Sound region, 17 positive RV samples from the Seattle hospital-based
129 respiratory virus surveillance, randomly sectioned from January to June 2023, were sequenced
130 and included in the phylogenetic analyses. All complete genomes were trimmed to the VP1
131 region to proceed with the globally established genotyping classification [16].

132 We detected 47 RV-A genotypes, 14 RV-B genotypes, and 36 RV-C genotypes (Figure
133 3B-D, Figure 5A-B). A limited group of genotypes – A1B, A25, A39, A78, A101 in RV-A; B6 in
134 RV-B; and C3, C11, C15, C17, C20 in RV-C – exhibited frequencies surpassing 1% of total
135 cases though their prevalence varied throughout the studied time (Figure 5B). High RV-C
136 prevalence from February to April 2021 was driven by four genotypes with limited change over
137 time (C3, C11, C17, and C20), while RV-A dominance from May to July 2021 was accompanied
138 by the replacement of A101 by A1B (Figure 3B-3D, Figure 5A). At the same time, B6 was the
139 prevalent RV-B genotype during 2021. Remarkably, fifteen months later, most of the genotypes
140 with high frequencies were not detected (Figure 5A).

141 Rarefaction analysis indicated that the genotype richness was well characterized during
142 2021, while in 2022 genotypic coverage lagged despite our sampling efforts (Appendix,
143 Supplementary Figure 1). Nonetheless, we covered more than 75% of estimated genotypic
144 richness (Appendix). Based on the observed genotypic richness, Chao1 index calculation
145 indicated the presence of up to 24 potentially unsampled genotypes in RV-A, 4 in RV-B and 17
146 in RV-C (Appendix).

147

148 *Novel RV-A and RV-C genotypes detected*

149 We next examined the overall RV phylogenies for monophyletic clusters suggestive of novel
150 genotypes (Supplementary Figure 2). Two monophyletic clusters exhibited distinctive patristic
151 distances and further analysis indicated they had an average inter-genotype p-distance above
152 the cutoff value of 0.13 (Supplementary Figure 2) [16]. We proposed them as genotypes A111
153 and C59. Genotype A111 comprised exclusively sequences from Washington State from 2017
154 to 2023, and C59 included sequences from Wisconsin from 2014 and Washington State from
155 2018 and 2023.

156

157 *Association of clusters of genotypes with clinical and demographic characteristics*

158 The low number of cases for many of the detected RV genotypes impedes statistical association
159 studies of clinical and epidemiological characteristics. Therefore, we evaluated whether
160 phylogenetically-related groups of genotypes might demonstrate any such association. We
161 analyzed principal component analysis (PCA) based on a polyprotein-based pairwise genetic
162 distance matrix applying unsupervised hierarchical clustering. Genotypes were grouped into 4
163 clusters in RV-A, 5 clusters in RV-B, and 7 clusters in RV-C (Appendix, Supplementary Figure
164 3). While clusters showed no association with symptom, sex, or geographic location, they were
165 associated with the age of the individuals. Specifically, in HRV-B, cluster 2 (genotypes B3, B72,
166 B83, B92 and B100) was more related with older individuals (median 42 years old, interquartile

167 range: 33-56 years old) cluster 1 (genotype B6, median 20 years old, interquartile range=10-32
168 years old) (Kruskal-Wallis test, $p=0.04$). In RV-A, cluster 4 (genotypes A7, A12, A20, A28, A45,
169 A46, A58, A68, A78) was more associated with the group of individuals <5 and >65 years old,
170 while cluster 2 (genotypes A2, A9, A11, A16, A18, A22-25, A29-34, A39, A47, A49, A54, A60-
171 64, A66, A67, A73, A85, A94, A105) and cluster 3 (genotype A1B) were associated with the
172 group individuals aged 5 to 64. While not statistically supported, RV-C cluster 4 (genotype C15)
173 was more related with older individuals (mean 42 years old) compared to the other clusters
174 (range of means: 21 – 29 years).

175 We also evaluated the seasonality of the cluster of genotypes and found that RV-A and
176 RV-B clusters were clearly distinguished by the year of sample collection ($p<0.05$). Specifically,
177 cluster 1 (A101) and 3 (A1B) in RV-A, and cluster 1 (B6) and 4 (B27) in RV-B were associated
178 with detections in 2021 compared to other clusters. In the case of RV-C, the comparison by
179 clusters did not result in a single supported association, however cluster 3 (C4, C9, C13, C19,
180 C23, C25-27, C29, C33, C35, C36, C40, C46, C53, C55) and 4 (C15) were mainly detected in
181 2022, while cluster 7 (C20, C34) was mainly detected in 2021.

182

183 *Recombination between A105 and A21 genotypes within 3C protease gene*

184 We next analyzed the complete genomes of the RV circulating in Puget Sound to examine for
185 non-capsid changes. Interestingly, topological comparison of VP1 and 3D/RdRp phylogenetic
186 trees revealed inconsistencies in a monophyletic clade comprised of four sequences from
187 Washington State dating from 2016 and 2021. The clade was associated with A105 VP1 capsid
188 sequence and A21 3D polymerase sequence (Figure 6A). Further analysis identified nucleotide
189 5,250 (95% confident interval: 5,216 – 5,271, GenBank accession number MZ268661),
190 corresponding to amino acids 37-55 of the 3C protease, as the likely recombination breakpoint
191 and was confirmed by all algorithms tested with RDP software ($p<10^{-13}$) (Figure 6B).

192

193 *Intra-genotypic RV evolution and variability*

194 We reconstructed genome-based phylodynamic trees with the genomes available for the highly
195 prevalent genotypes in this study (A1B, A25, A39, A78, A101 in RV-A; B6 in RV-B; and C3,
196 C11, C15, C17, C20 in RV-C). Genotype A25 showed no temporal signal, and A39 failed to
197 converge in BEAST analysis. The rest of genotypes showed consistently similar intra-genotype
198 evolutionary (range: 2.15×10^{-3} - 4.54×10^{-3} substitutions/site/year, Figure 7), coinciding with inter-
199 genotype evolutionary rates previously published [24,25]. The time of the most recent common
200 ancestor (tMRCA) varied from 1982 to 2007. Ancestry of the RV circulating in Puget Sound in
201 2021 and 2022 showed that most of the sequences were closely associated into a few large
202 clades, though some sequences were associated with clades that diverged decades ago. For
203 example, sequences from the A101 genotype detected in 2021 were grouped into different
204 clades sharing a MRCA from 1989, highlighting the evolutionary diversity within this genotype
205 (Supplementary Figure 4). We wondered how the high evolutionary rate and divergence time
206 might affect intra-genotypic variability at the amino acid level with possible antigenic
207 consequences. Intra-genotype selection pressure estimation across the RV polyprotein unveiled
208 overall strong negative selection spanning the entire open reading frame. However, positively
209 selected sites for three genotypes were seen within the VP1 capsid protein, at positions 274
210 (genotype A1B), 43 (genotype A39) and 86 (genotype C15) (Appendix). Positively selected sites
211 were also found within 3D (253 in genotype B6, 66 in genotype C3 and 133 in genotype C11),
212 3B (position 5, genotype C17) and 2A (position 58, genotype C3). In addition, the Shannon
213 entropy calculation in the polyprotein revealed a single position (residue 86) in VP1 for the C15
214 genotype with moderate entropy (Supplementary Figure 5). Notably, all residues in all the other
215 genotypes exhibited entropy values consistently below 1, indicating that all genotypes have a
216 highly conserved polyprotein sequence.

217

218 **Discussion**

219 In this study we analyzed the largest collection of newly sequenced rhinovirus (RV) genomes to
220 date, comprising more than half of publicly available RV genomes in NCBI GenBank as of
221 December 2023. We observed a high RV incidence during the studied period, even during
222 November and December 2022 when the “triple-demic” of SARS-CoV-2, RSV, and influenza
223 virus caused significant public health concern [26].

224
225 Our study also supports previous findings indicating that symptomatology is associated with
226 higher viral loads than asymptomatic individuals. [27–29]. We observed a high incidence of RV-
227 A in spring and summer, with RV-C predominating during the winter, consistent with prior
228 reports [30]. However, unlike previous reports, we did not find age groups or symptomatology to
229 be associated with specific RV species [3]. Our findings were limited to mild outpatient cases
230 taken from community testing sites, and broader surveillance in severe infections could yield
231 different results. Our study showed that the containment measures against the COVID-19
232 pandemic had no effect on the geographically spread of RV species, co-circulation, and
233 replacement in the Puget Sound region, consistent with the high positivity rate seen for RV in
234 hospital-based surveillance data. Further investigations globally are needed to assess if the
235 seasonality we observed reflects a local or global circulation pattern, helping to determine if RV
236 spread extends beyond strict border containment measures.

237
238 We detected 100 of 173 described RV genotypes, including 2 newly described genotypes and 3
239 other genotypes described in published manuscripts but not yet confirmed by the ICTV or
240 *Picornaviridae* Study Group (Supplementary Figure 6). Overall, our work more than doubled the
241 number of genomes available in the NCBI database for 65 RV genotypes. We detected the
242 cocirculation of 30 RV-A genotypes in June 2021, that to the best of our knowledge is the
243 highest number of genotypes observed in a single month for one RV species in a small
244 geographic region. However, our sequencing efforts covered an estimated >70% of the

245 genotypic richness, indicating there still were unseen genotypes. Although we have studied a
246 distinctive epidemiological period that may have facilitated an increase in RV viral diversity in
247 2021, these results demonstrate the complexity of designing optimal RV molecular surveillance.

248
249 During our study period, 11 genotypes comprised >60% of studied RV cases. Interestingly,
250 genotypes that predominated in 2021 were replaced in 2022. However, we did not perform
251 serological analysis to determine whether population-level immunity to prevalent genotypes in
252 2021 defined the genotypic incidence the following year, or if it is related to another seasonality
253 pattern. Prior work has shown that RV serotype-specific IgA in serum can last for at least one
254 year, indicating that immunity is likely a contributing factor for shifting genotypic incidence [31].

255
256 Our genomic analysis detected a recombinant monophyletic clade between A105 and A21
257 genotypes with a breakpoint in the 3C protein region. A similar 3C recombination breakpoint has
258 been reported between A76 and A56 genotypes [24,25]. This region of the genome is a hot
259 spot for recombination events in poliovirus [32]. However, our A105-A21 recombination
260 breakpoint was located upstream of the corresponding RNase L inhibitory RNA secondary
261 structure element, which was shown to be associated with higher rates of recombination in
262 poliovirus [33,34]. Further investigation of the RNA secondary structure of RV may help to
263 understand our findings.

264
265 We observed that RV co-circulating in the Puget Sound area from the same genotype diverged
266 over a decade ago, though the polyprotein sequence exhibited a significant constraint on amino
267 acid variability. Only three genotypes exhibited positive selection in single different sites of the
268 VP1 protein, the major architectural component of the viral capsid and host-cell receptor
269 attachment, while four other genotypes showed positively-selected sites in non-structural
270 proteins. Further studies are needed to evaluate the biological impact of amino acid changes in

271 these sites, but overall results suggest that the intra-genotypic amino acid variability is low, a
272 promising fact for the development of an RV vaccine. However, this study emphasizes the
273 importance of investigating correlations between genotypes and serotypes to understand long-
274 term immunity and cross-protection. Recent research by Bochkov et al revealed cross-
275 protection among some genetically similar genotypes in RV-A and RV-C species [35]. This fact
276 was originally mentioned in 1982 when Cooney et al demonstrated cross-protection between 50
277 RV types, categorized them into 16 antigenic groups [36]. In our study, hierarchical clustering
278 based on polyprotein genetic distances helped to associate some genotypes into a small
279 number of clusters. We found statistical association of some clusters with the year of detection
280 which might indicate a possible relationship with the long-term immunity interference. While this
281 clustering does not necessarily define antigenic and serotypic relatedness, it did cluster together
282 RV-A12, RV-20, and RV-A78 which have previously shown cross-neutralizing protection [35].

283
284 Our study has several limitations which may limit generalizability of findings, including the
285 convenience sampling of remnant SARS-CoV-2 negative specimens, restriction of sampling to
286 outpatient community testing sites, limited clinical metadata associated with high-throughput
287 drive-thru testing, and sampling only during 2021-22 period within Puget Sound area.
288 Nonetheless, the high volume of outpatient respiratory testing performed during the SARS-CoV-
289 2 pandemic and omnipresent high co-circulation of RV made this study possible. Genomic
290 recovery also requires specimens with higher viral loads, which may bias genomic
291 epidemiological analyses such as to individuals with a symptomatic infection.

292
293 Overall, our deep look at RV genomic epidemiology during the 2021-2022 period in Puget
294 Sound significantly increased the availability of genomic data for RV. We highlighted the
295 seemingly unrestricted ability of RV species to co-circulate, including up to 30 RV-A genotypes
296 in one month, despite public health containment measures that significantly limited co-

297 circulation of other respiratory viruses. In addition, we showed the importance of update the RV
298 surveillance to genomic approach to monitor recombination, and the continued need to examine
299 antigenic and serotypic relatedness of RV genotypes that may reveal relationships with viral
300 seasonality and vaccine development.

301

302 **Materials and methods**

303 *Clinical samples and data collection*

304 SARS-CoV-2-negative nasal swabs collected from symptomatic and asymptomatic individuals
305 at University of Washington (UW) Medicine COVID-19 community testing sites were screened
306 for RV by RT-qPCR (Appendix). Individuals declaring at least one respiratory symptom at time
307 of collection were defined as symptomatic. This study was approved the UW Medicine
308 Institutional Review Board with a consent waiver (STUDY00000408).

309

310 *NGS and assembly of viral consensus genomes*

311 RNA from nasal swabs with RV Ct value below 33 were sequenced by metagenomic next-
312 generation sequencing as described previously [37]. Twelve samples with Ct values 33-37 were
313 also sequenced. Briefly, viral RNA was extracted with Quick-RNA Viral Kit (Zymo Research),
314 double stranded cDNA synthesis was performed with random hexamers with SuperScript IV
315 Reverse Transcriptase and Sequenase Version 2.0 DNA Polymerase (ThermoFisher Scientific)
316 and purified with AMPure XP Magnetic Beads (Beckman Coulter). Library preparation was
317 performed with Illumina DNA Prep, (S) Tagmentation kit (Illumina) and sequenced as 2 × 100-
318 bp or 2 × 150-bp runs on a NovaSeq sequencer (Illumina). Consensus genomes were called
319 using custom pipelines of mapping against references of all existing RV genomes from the
320 International Committee on Taxonomy of Viruses (ICTV) as of March 2022 (Supplementary
321 Material).

322

323 *Sequence analysis*

324 RV complete genomes longer than 6,000 nt from human hosts were downloaded from in NCBI
325 Virus (May 25, 2023) using the taxonomic IDs 147711 (RV-A), 147712 (RV-B), 463676 (RV-C).
326 Partial genomes were downloaded from the *Picornaviridae* Study Group database to represent
327 genotypes without complete genome available (<https://www.picornastudygroup.com/>). MEGA11
328 software was used to calculate the pairwise genetic p-distance for RV genotypes with 1,000
329 replicates bootstrapping [38]. Alignment-based recombination analysis was performed with
330 RDP4 software using the RDP, GENECONV, Bootscan, Maximum Chi Square, Chimaera,
331 SiScan, and 3SEQ methods with default settings [39]. Comparative co-phylogeny of VP1 and
332 3D regions was performed using ‘cophylo’ tool from the ‘phytools’ package in R [40].

333

334 *Statistical analyses*

335 Rarefaction curves and the coverage-based extrapolation curves were calculated to estimate
336 the genotypic diversity using the packages ‘vegan’ and ‘iNEXT’ in RStudio 2022.07.2 [41,42].
337 Association among categorical and continuous variables were evaluated with ANOVA, Kruskal-
338 Wallis test, Wilcoxon test, G-test, or Wald odds ratio using AMR or epitools libraries in RStudio
339 [43]. Principal component analysis (PCA) and hierarchical clustering (HC) based on pairwise
340 genetic distances (p-distance) were performed using the packages ‘FactoMineR’, ‘factoextra’
341 and ‘eclust’ in RStudio (Supplementary Material). R markdown code is available at
342 https://github.com/greninger-lab/HRV_epidemiology.

343

344 *Phylogenetic analyses*

345 Sequences were aligned with MAFFT v7.490 and visualized with Aliview v1.28 [44,45].
346 Maximum likelihood trees were inferred with IQ-TREE v2.1 using ModelFinder to select the
347 substitution model, and SH-aLRT test (1,000 replicates) and UFBoot2 method (1,000 replicates)
348 to evaluate reliability of phylogenetic clades [46,47]. Comparative analysis of the genotyping

349 procedure with maximum likelihood instead of neighbor joining inferences are detailed in
350 Appendix.

351 Temporal signal was assessed with TempEst v1.5.3 [48]. The evolutionary rate was inferred
352 with BEAST2 package v2.7.5 using an optimized relaxed clock and the tree priors selected
353 according to the Nested Sampling test [49,50]. The convergence of the BEAST inference was
354 assessed from the estimations of the Effective Sampling Size (ESS) and the highest posterior
355 density interval (95% HPD) after a 10% burn-in using Tracer v1.7. TreeAnnotator was used to
356 summarize the information from the sampled trees onto a single tree (the maximum clade
357 credibility tree, MCCCT).

358

359 **Sequence data availability.**

360 FASTQ and RV consensus genome data are available in NCBI BioProject PRJNA1066815
361 (SRA and GenBank accession numbers in Supplementary Table 1).

362

363 **Acknowledgments.**

364 We thank the excellent technical support of Nathan Breit and Carlos Avendaño. This research
365 received no specific funding and was supported by departmental funds.

366

367 **Conflict of interest.**

368 ALG reports contract testing from Abbott, Cepheid, Novavax, Pfizer, Janssen and Hologic and
369 research support from Gilead, outside of the described work.

370

371

372 **References**

- 373 1. Hansen C, Perofsky AC, Burstein R, et al. Trends in Risk Factors and Symptoms
374 Associated With SARS-CoV-2 and Rhinovirus Test Positivity in King County, Washington,
375 June 2020 to July 2022. *JAMA Netw Open*. **2022**; 5(12):e2245861.
- 376 2. Chow EJ, Casto AM, Roychoudhury P, et al. The Clinical and Genomic Epidemiology of
377 Rhinovirus in Homeless Shelters-King County, Washington. *J Infect Dis*. **2022**; 226(Suppl
378 3):S304–S314.
- 379 3. Zlateva KT, Rijn AL van, Simmonds P, et al. Molecular epidemiology and clinical impact of
380 rhinovirus infections in adults during three epidemic seasons in 11 European countries
381 (2007-2010). *Thorax*. **2020**; 75(10):882–890.
- 382 4. Giardina FAM, Piralla A, Ferrari G, Zavaglio F, Cassaniti I, Baldanti F. Molecular
383 Epidemiology of Rhinovirus/Enterovirus and Their Role on Cause Severe and Prolonged
384 Infection in Hospitalized Patients. *Microorganisms*. **2022**; 10(4):755.
- 385 5. Makhous N, Goya S, Avendaño CC, et al. Within-host rhinovirus evolution in upper and
386 lower respiratory tract highlights capsid variability and mutation-independent
387 compartmentalization. *J Infect Dis*. **2023**; :jiad284.
- 388 6. Boonyaratanakornkit J, Vivek M, Xie H, et al. Predictive Value of Respiratory Viral
389 Detection in the Upper Respiratory Tract for Infection of the Lower Respiratory Tract With
390 Hematopoietic Stem Cell Transplantation. *The Journal of Infectious Diseases*. **2020**;
391 221(3):379–388.
- 392 7. Grissell TV, Powell H, Shafren DR, et al. Interleukin-10 gene expression in acute virus-
393 induced asthma. *Am J Respir Crit Care Med*. **2005**; 172(4):433–439.
- 394 8. Atmar RL, Guy E, Guntupalli KK, et al. Respiratory tract viral infections in inner-city
395 asthmatic adults. *Arch Intern Med*. **1998**; 158(22):2453–2459.
- 396 9. Castillo JR, Peters SP, Busse WW. Asthma Exacerbations: Pathogenesis, Prevention, and
397 Treatment. *J Allergy Clin Immunol Pract*. **2017**; 5(4):918–927.
- 398 10. Wat D, Gelder C, Hibbitts S, et al. The role of respiratory viruses in cystic fibrosis. *J Cyst*
399 *Fibros*. **2008**; 7(4):320–328.
- 400 11. McManus TE, Marley A-M, Baxter N, et al. Respiratory viral infection in exacerbations of
401 COPD. *Respir Med*. **2008**; 102(11):1575–1580.
- 402 12. Sim YS, Lee JH, Lee EG, et al. COPD Exacerbation-Related Pathogens and Previous
403 COPD Treatment. *J Clin Med*. **2022**; 12(1):111.
- 404 13. Hamparian VV, Colonno RJ, Cooney MK, et al. A collaborative report: rhinoviruses--
405 extension of the numbering system from 89 to 100. *Virology*. **1987**; 159(1):191–192.
- 406 14. Stanway G, Hughes PJ, Mountford RC, Minor PD, Almond JW. The complete nucleotide
407 sequence of a common cold virus: human rhinovirus 14. *Nucleic Acids Res*. **1984**;
408 12(20):7859–7875.

- 409 15. Esneau C, Bartlett N, Bochkov YA. Chapter 1 - Rhinovirus structure, replication, and
410 classification. In: Bartlett N, Wark P, Knight D, editors. Rhinovirus Infections. Academic
411 Press; 2019. p. 1–23.
- 412 16. McIntyre CL, Knowles NJ, Simmonds P. Proposals for the classification of human
413 rhinovirus species A, B and C into genotypically assigned types. *J Gen Virol.* **2013**; 94(Pt
414 8):1791–1806.
- 415 17. Esneau C, Duff AC, Bartlett NW. Understanding Rhinovirus Circulation and Impact on
416 Illness. *Viruses.* **2022**; 14(1):141.
- 417 18. Goya S, Sereewit J, Pfallmer D, et al. Genomic Characterization of Respiratory Syncytial
418 Virus during 2022-23 Outbreak, Washington, USA. *Emerg Infect Dis.* **2023**; 29(4):865–868.
- 419 19. Varela FH, Sartor ITS, Polese-Bonatto M, et al. Rhinovirus as the main co-circulating virus
420 during the COVID-19 pandemic in children. *J Pediatr (Rio J).* **2022**; 98(6):579–586.
- 421 20. Champredon D, Bancej C, Lee L, Buckrell S. Implications of the unexpected persistence of
422 human rhinovirus/enterovirus during the COVID-19 pandemic in Canada. *Influenza Other
423 Respir Viruses.* **2022**; 16(2):190–192.
- 424 21. Georgieva I, Stoyanova A, Angelova S, Korsun N, Stoitsova S, Nikolaeva-Glomb L.
425 Rhinovirus Genotypes Circulating in Bulgaria, 2018-2021. *Viruses.* **2023**; 15(7):1608.
- 426 22. Pendrey CG, Strachan J, Peck H, et al. The re-emergence of influenza following the
427 COVID-19 pandemic in Victoria, Australia, 2021 to 2022. *Eurosurveillance.* European
428 Centre for Disease Prevention and Control; **2023**; 28(37):2300118.
- 429 23. Morobe JM, Nyiro JU, Brand S, et al. Human rhinovirus spatial-temporal epidemiology in
430 rural coastal Kenya, 2015-2016, observed through outpatient surveillance. *Wellcome Open
431 Res.* **2019**; 3:128.
- 432 24. McIntyre CL, Savolainen-Kopra C, Hovi T, Simmonds P. Recombination in the evolution of
433 human rhinovirus genomes. *Arch Virol.* **2013**; 158(7):1497–1515.
- 434 25. Zhao P, Shao N, Dong J, et al. Genetic diversity and characterization of rhinoviruses from
435 Chinese clinical samples with a global perspective. *Microbiology Spectrum.* American
436 Society for Microbiology; **2023**; 0(0):e00840-23.
- 437 26. Nirappil F. Covid, flu, RSV declining in hospitals as ‘triple-demic’ threat fades. *Washington
438 Post [Internet].* **2023** [cited 2024 Jan 8]; . Available from:
439 <https://www.washingtonpost.com/health/2023/01/22/covid-declining-flu-rsv-triple-demic/>
- 440 27. Sanchez-Codez MI, Moyer K, Benavente-Fernández I, Leber AL, Ramilo O, Mejias A. Viral
441 Loads and Disease Severity in Children with Rhinovirus-Associated Illnesses. *Viruses.*
442 Multidisciplinary Digital Publishing Institute; **2021**; 13(2):295.
- 443 28. Gerna G, Piralla A, Rovida F, et al. Correlation of rhinovirus load in the respiratory tract and
444 clinical symptoms in hospitalized immunocompetent and immunocompromised patients. *J
445 Med Virol.* **2009**; 81(8):1498–1507.

- 446 29. Granados A, Peci A, McGeer A, Gubbay JB. Influenza and rhinovirus viral load and
447 disease severity in upper respiratory tract infections. *Journal of Clinical Virology*. **2017**;
448 86:14–19.
- 449 30. Wildenbeest JG, Schee MP van der, Hashimoto S, et al. Prevalence of rhinoviruses in
450 young children of an unselected birth cohort from the Netherlands. *Clin Microbiol Infect*.
451 **2016**; 22(8):736.e9–736.e15.
- 452 31. Barclay WS, Nakib W al-, Higgins PG, Tyrrell DA. The time course of the humoral immune
453 response to rhinovirus infection. *Epidemiol Infect*. **1989**; 103(3):659–669.
- 454 32. Runckel C, Westesson O, Andino R, DeRisi JL. Identification and Manipulation of the
455 Molecular Determinants Influencing Poliovirus Recombination. *PLOS Pathogens*. Public
456 Library of Science; **2013**; 9(2):e1003164.
- 457 33. Dutkiewicz M, Stachowiak A, Swiatkowska A, Ciesiołka J. Structure and function of RNA
458 elements present in enteroviral genomes. *Acta Biochimica Polonica*. **2016**; 63(4):623–630.
- 459 34. Burrill CP, Westesson O, Schulte MB, Strings VR, Segal M, Andino R. Global RNA
460 Structure Analysis of Poliovirus Identifies a Conserved RNA Structure Involved in Viral
461 Replication and Infectivity. *J Virol*. **2013**; 87(21):11670–11683.
- 462 35. Bochkov YA, Devries M, Tetreault K, et al. Rhinoviruses A and C elicit long-lasting antibody
463 responses with limited cross-neutralization. *J Med Virol*. **2023**; 95(8):e29058.
- 464 36. Cooney MK, Fox JP, Kenny GE. Antigenic groupings of 90 rhinovirus serotypes. *Infect*
465 *Immun*. **1982**; 37(2):642–647.
- 466 37. Greninger AL, Waghmare A, Adler A, et al. Rule-Out Outbreak: 24-Hour Metagenomic
467 Next-Generation Sequencing for Characterizing Respiratory Virus Source for Infection
468 Prevention. *J Pediatric Infect Dis Soc*. **2017**; 6(2):168–172.
- 469 38. Stecher G, Tamura K, Kumar S. Molecular Evolutionary Genetics Analysis (MEGA) for
470 macOS. *Molecular Biology and Evolution*. **2020**; 37(4):1237–1239.
- 471 39. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of
472 recombination patterns in virus genomes. *Virus Evol*. **2015**; 1(1):vev003.
- 473 40. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things).
474 *Methods in Ecology and Evolution*. **2012**; 3(2):217–223.
- 475 41. Oksanen J, Simpson GL, Blanchet FG, et al. vegan: Community Ecology Package version
476 2.6-4 from CRAN [Internet]. 2022 [cited 2023 Dec 26]. Available from:
477 <https://rdrr.io/cran/vegan/>
- 478 42. Chao A, Gotelli NJ, Hsieh TC, et al. Rarefaction and extrapolation with Hill numbers: a
479 framework for sampling and estimation in species diversity studies. *Ecological Monographs*.
480 **2014**; 84(1):45–67.
- 481 43. Sullivan KM, Dean A, Soe MM. OpenEpi: A Web-based Epidemiologic and Statistical
482 Calculator for Public Health. *Public Health Rep*. **2009**; 124(3):471–474.

- 483 44. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7:
484 Improvements in Performance and Usability. *Mol Biol Evol.* **2013**; 30(4):772–780.
- 485 45. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets.
486 *Bioinformatics.* **2014**; 30(22):3276–3278.
- 487 46. Minh BQ, Schmidt HA, Chernomor O, et al. IQ-TREE 2: New Models and Efficient Methods
488 for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution.* **2020**;
489 37(5).
- 490 47. Hoang DT, Chernomor O, Haeseler A von, Minh BQ, Vinh LS. UFBoot2: Improving the
491 Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution.* **2018**; 35(2):518–522.
- 492 48. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. Exploring the temporal structure of
493 heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution.* **2016**;
494 2(1):vew007.
- 495 49. Bouckaert R, Vaughan TG, Barido-Sottani J, et al. BEAST 2.5: An advanced software
496 platform for Bayesian evolutionary analysis. *PLOS Computational Biology. Public Library of
497 Science*; **2019**; 15(4):e1006650.
- 498 50. Kalyaanamoorthy S, Minh BQ, Wong TKF, Haeseler A von, Jermini LS. ModelFinder: fast
499 model selection for accurate phylogenetic estimates. *Nat Methods. Nature Publishing
500 Group*; **2017**; 14(6):587–589.

501

502 **Table 1. Demographic and virological characteristic of RV-positive individuals.**

	Total	Feb-Jul 2021 (period 1)	Nov-Dec 2022 (period 2)
Individuals tested RV positives	1903	1083	820
Mean RV Ct values	28.8	22.4	29.5
Individuals with clinical data	992	406	586
Male	488 (49%)	210 (51%)	278 (47%)
<5 years old	190 (19%)	68 (17%)	122 (21%)
5 - 17 years old	196 (20%)	126 (31%)	70 (12%)
18 - 65 years old	569 (57%)	203 (50%)	366 (62%)
>65 years old	40 (4%)	9 (2%)	31 (5%)
Symptomatic individuals	655 (66%)	286 (70%)	369 (63%)
<5 years old	121 (18%)	47 (69%)	74 (61%)
5 - 17	129 (20%)	84 (67%)	45 (64%)
18 - 65	383 (58%)	148 (78%)	235 (64%)
>65y	22 (3%)	7 (78%)	15 (48%)

503 Percentages are shown regarding to the number of individuals with available clinical data,
504 which are a subset of the individuals that tested positives for RV.

505

506

507 **Figure 1. Seasonality of respiratory viruses in symptomatic patients in Washington State**
508 **from 2019 to 2023.** The positivity rate per month is detailed for different respiratory viruses.
509 Rhinovirus is highlighted with red line including triangles. Grey shadow at the background of the
510 plot informs the number of screened samples per month. The two studied periods of this work
511 are denoted with dotted squares (P1= period 1, P2=period 2).

512
513 **Figure 2. RV infections and relationship with the individuals' age group and the**
514 **symptoms presence.**

515 The association between the age of the individuals, presence of respiratory symptoms and the
516 Ct value was evaluated. Violin plots illustrating the distribution of cases for Ct value and
517 stratified age group (A), Ct value and symptoms presence (B) and non-stratified age and
518 symptoms presence (C). Black dots represent individuals. Within the violin plot, a boxplot shows
519 the median, interquartile range and 95% confidence interval. Kruskal-Wallis tests was used to
520 assess the association between variables, but also Wilcoxon test for the age groups pairwise
521 analysis. Supported (p value <0.001) differences are denoted with *** and includes details on
522 the Kruskal-Wallis statistical test results. (D) Multivariate regression model including the Ct
523 value, individuals' age, and symptoms presence. Dots indicate individuals, colored by the
524 symptom's presence, same as the linear regression and the 95% confidence interval. Details on
525 the multivariate prediction model are detail at the bottom of the plot.

526
527 **Figure 3. Abundance of RV species and genotypes within the sequenced genomes.**

528 The stacked bar-plot showed the proportion of RV species (A), genotypes within RV-A (B),
529 genotypes within RV-B (C) and genotypes within RV-C (D) by studied month. For genotypes bar
530 plots the total number of genomes per month is detailed above each column.

531
532 **Figure 4. Geographic spread of RV species in Puget sound region.**

533 For each studied month the proportion of RV species (RV-A green, RV-B red and RV-C yellow)
534 is mapped according to the University of Washington (UW) Medicine COVID-19 community
535 testing sites where the individuals attended. Below each pie plot the total number of detected is
536 informed. February 2021 is not shown due to the low number of cases available. At the bottom
537 right a map highlighting the geographical location of Puget Sound is shown.

538

539 **Figure 5. Genotyping of RV detected in Washington State in 2021 and 2022.**

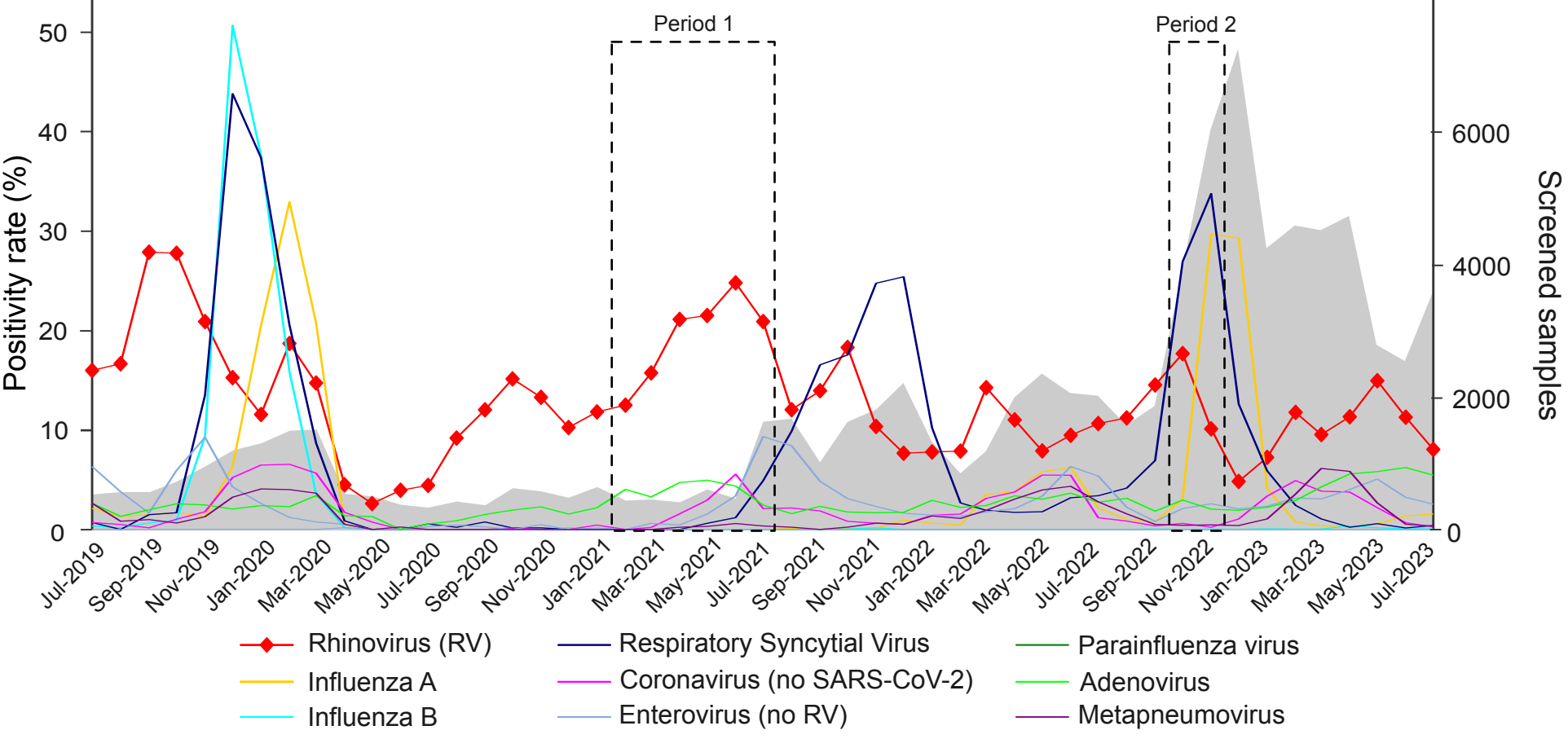
540 A) Seasonality of RV genotypes detected within RV-A, RV-B and RV-C. The bubble chart
541 details the number of cases (by the size of the circle) for each genotype per month of the
542 studied period. Color in the circles denotes the RV species. B) Maximum likelihood tree with the
543 VP1 nucleotide sequence of the cases studied in this work and reference sequences. Tree tips
544 of sequences from this study are denoted with a red or light blue circle if were collected in
545 period 1 or 2, respectively. The total number of sequences in the tree is informed below the
546 rhinovirus specie. Phylogenetic clades of genotypes comprising more than 1% of the cases
547 analyzed in this work are highlighted with yellow, including the name of the genotype in italics.
548 Scale bar indicates substitution per site.

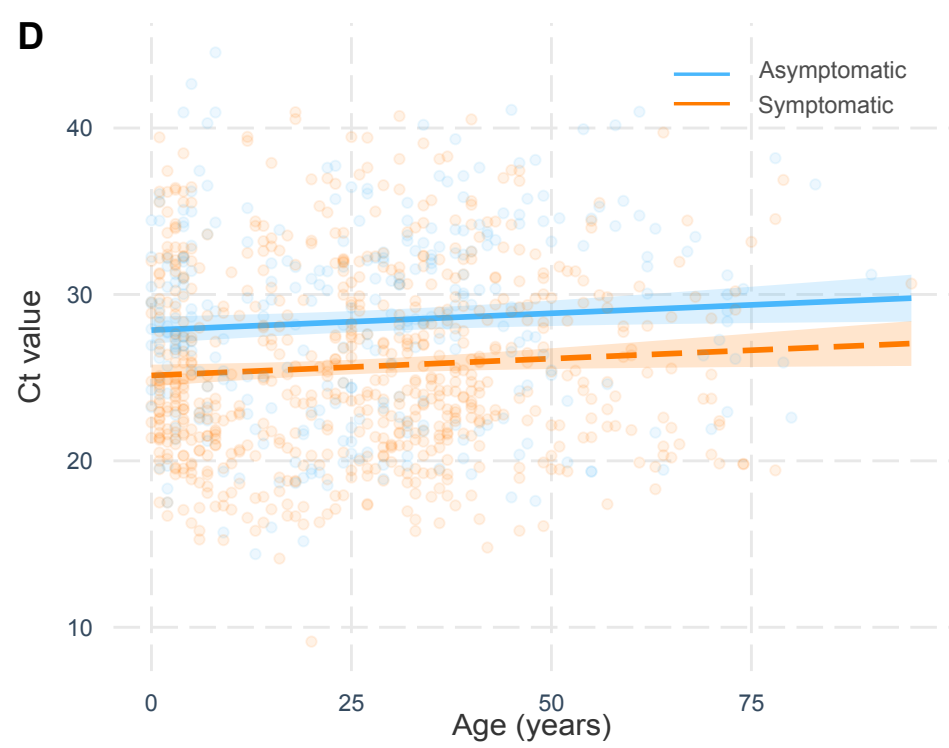
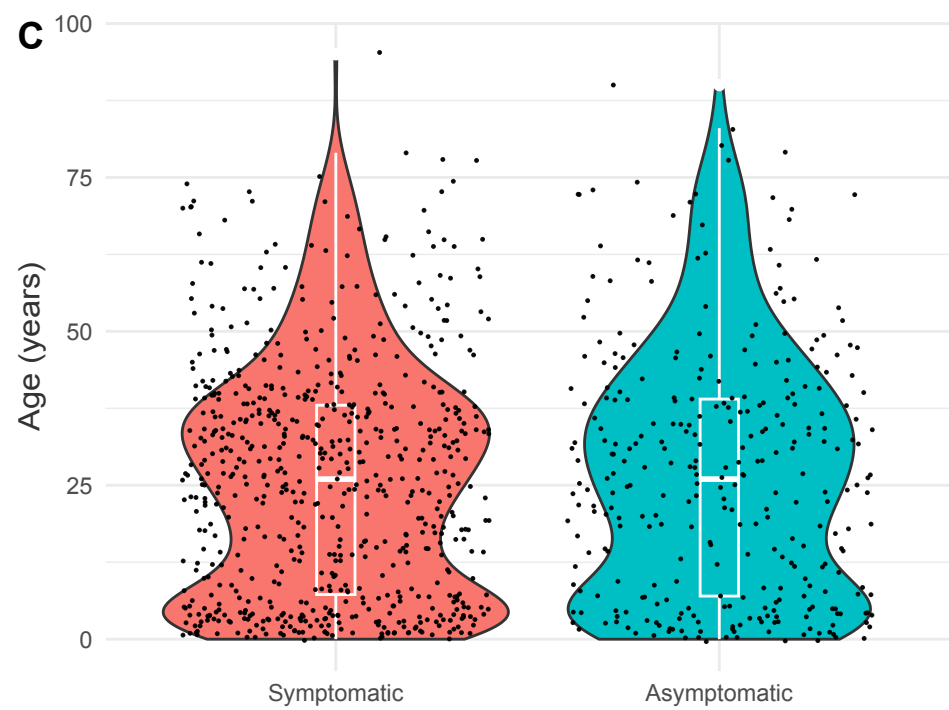
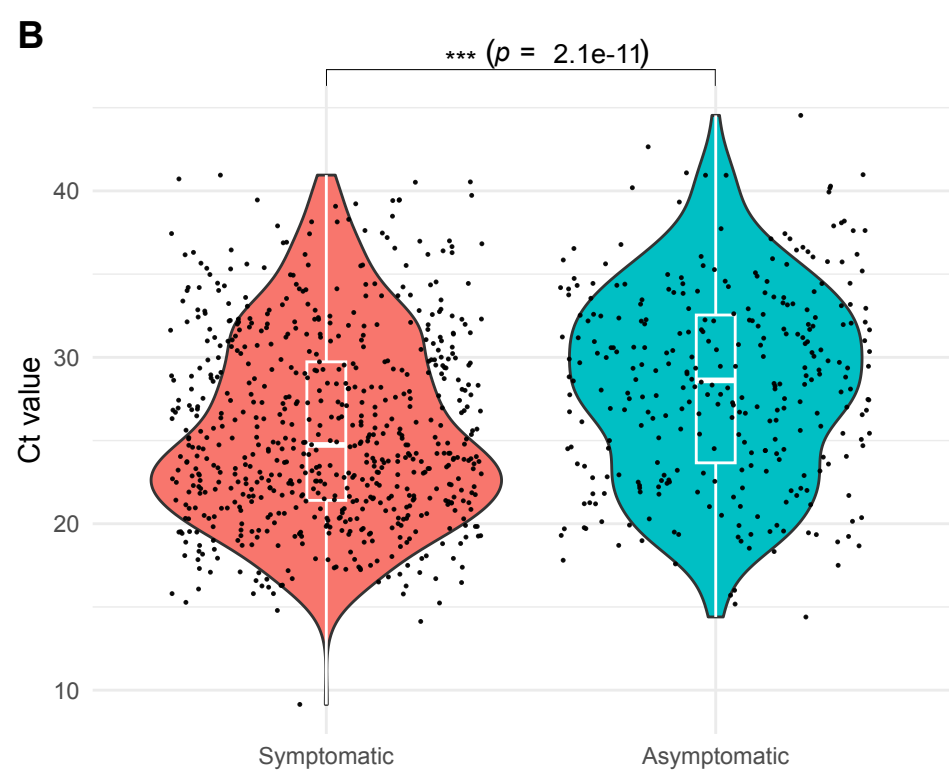
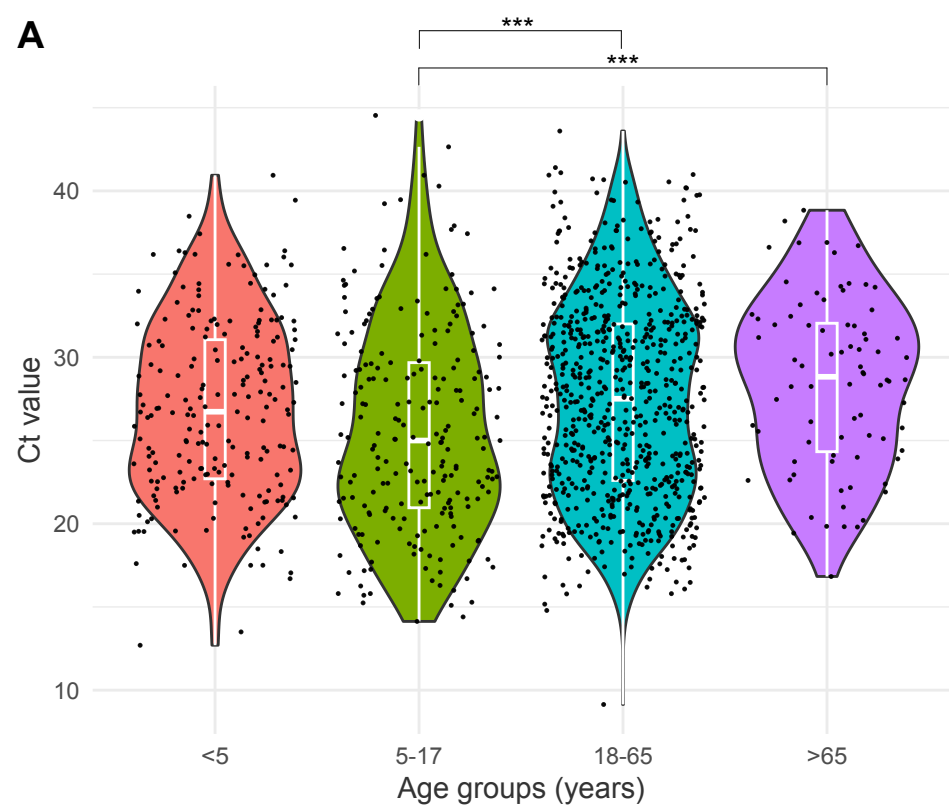
549

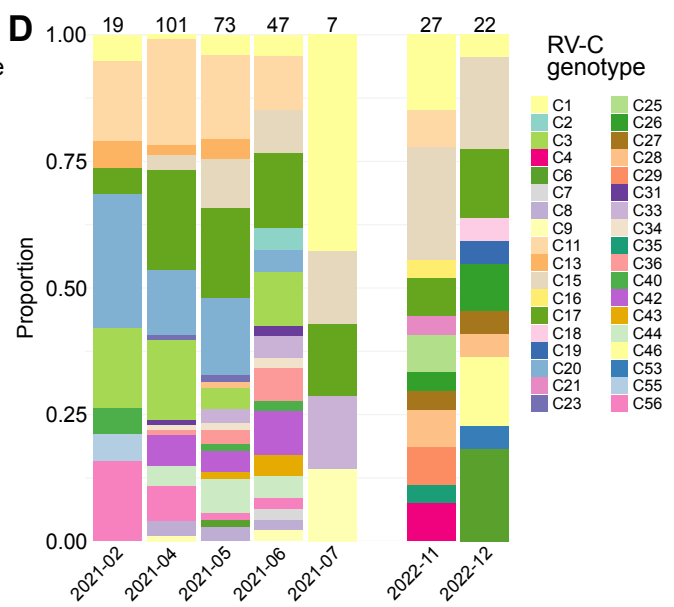
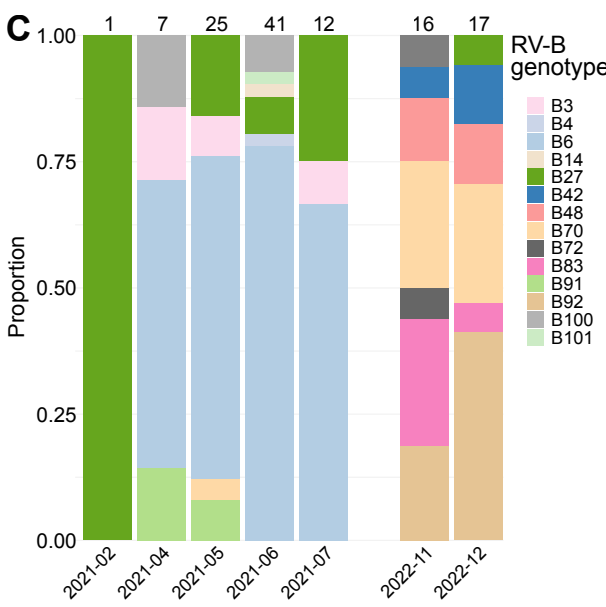
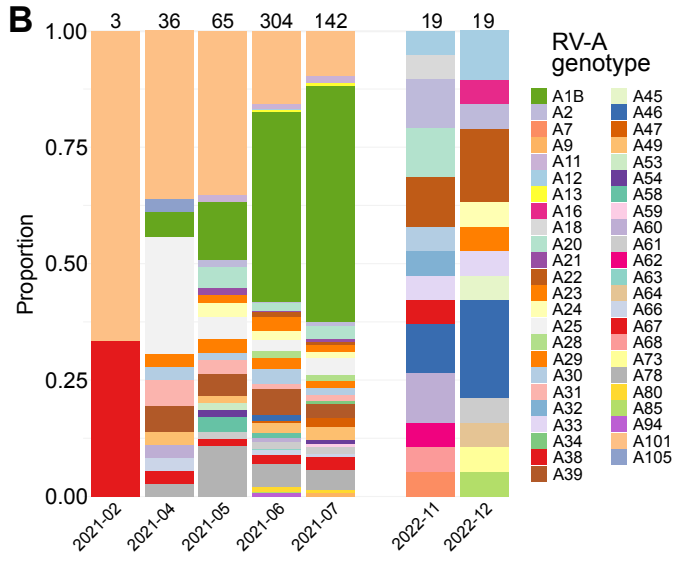
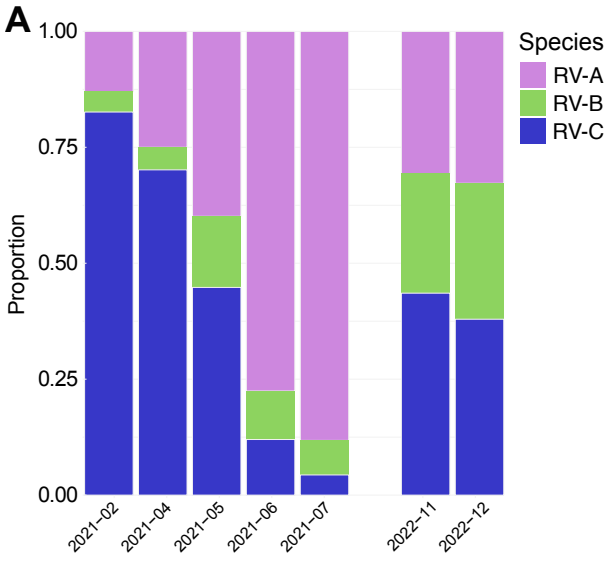
550 **Figure 6. Recombination between RV-A105 and RV-A21 genotypes.**

551 A) Phylogenetic clades from the maximum likelihood trees with VP-1 and 3D nucleotide
552 sequences. Branches are colored highlighting the inferred mayor (dark green) and minor (violet)
553 parents for the recombinant clade (highlighted in grey). Bootstrap values of relevant nodes are
554 indicated in the nodes. Scale bar indicates substitution per site. B) RDP pairwise identity plot
555 between the sequences MZ542285 (A105, major parent), JN837693 (A21, minor parent) and
556 MZ268661 (recombinant). The gray background denotes the 95% recombination breakpoint
557 confidence interval and the pink background highlight the region with recombinant origin. At the

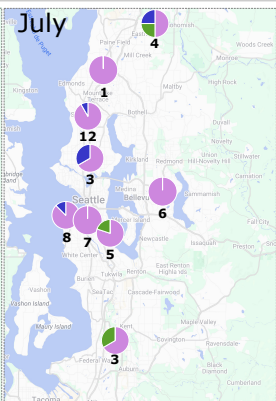
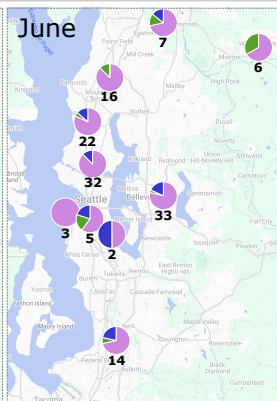
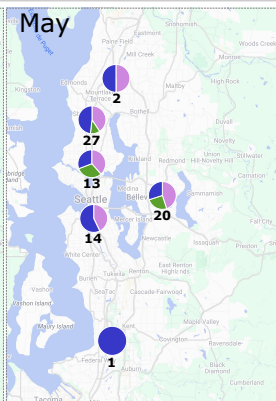
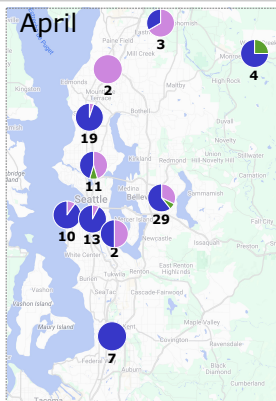
558 bottom of the plot the representation of the parent combination and the RV genome organization
559 following the same scale of nucleotide positions.



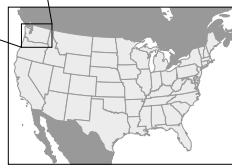
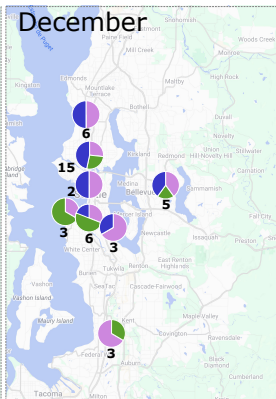
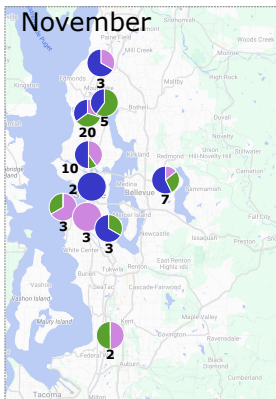




2021



2022



Species

- RV-A
- RV-B
- RV-C

