

1 **Tiled PCR amplification-based Whole Genome Sequencing and 2 Phylogenetic Classification Accelerate the Implementation of 3 Respiratory Syncytial Virus Genomic surveillance in Canada as a 4 Pilot Study**

5 Ruimin Gao^{a*}, Cody Buchanan^a, Kerry Dust^b, Paul Van Caeseele^b, Henry Wong^c, Calvin Sjaarda^c, Prameet M.
6 Sheth^c, Agatha N. Jassem^d, Jessica Minion^e, and Nathalie Bastien^{a,*}

7 ^aNational Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB, Canada

8 ^bVirus Detection, Cadham Provincial Laboratory, Winnipeg, MB, Canada

9 ^cKingston Health Sciences Centre and Queen's University, Kingston, ON, Canada

10 ^dBritish Columbia Centre for Disease Control, Provincial Health Services Authority, Vancouver, BC, Canada

11 ^eLaboratory Medicine, Saskatchewan Health Authority, Regina, SK, Canada

12 * Co-Correspondence:

13 Ruimin Gao: ruimin.gao@phac-aspc.gc.ca; Nathalie Bastien: nathalie.bastien@phac-aspc.gc.ca

14 **Running title:** Multiplex PCR WGS and comparative analysis of RSV

15 The authors declare no conflict of interest.

16 **IMPORTANCE**

17 We present assays to efficiently sequence genomes of the RSVA and RSVB. This enables
18 researchers and public health agencies to acquire high-quality genomic data using rapid and cost-
19 effective approaches. Genomic data based comparative analysis can be used to conduct
20 surveillance and monitor circulating isolates for efficacy of vaccines and antiviral therapeutics.

21 **ABSTRACT**

22 Whole genome sequencing (WGS) has emerged as a powerful tool to facilitate the study of
23 existing and emerging infectious diseases. WGS-based genomic surveillance provides
24 information on the genetic diversity and tracks the evolution of important viral pathogens
25 including Respiratory Syncytial Virus (RSV). Development and implementation of robust tiled
26 multiplex PCR amplification-based WGS assays will facilitate high-throughput RSV

27 surveillance initiatives. In this study, we developed multiplex PCR assays for targeted
28 enrichment of viral genomes using PrimalScheme (<http://primal.zibraproject.org>) to amplify over
29 97% of the genome in the majority of contemporaneous specimens tested. A pilot dataset
30 comprising 52 RSVA and 37 RSVB genomes derived from Canadian clinical specimens during
31 the 2022-2023 respiratory virus season were used to perform phylogenetic analyses using both
32 near complete genome and Glycoprotein (G) sequences. Overall, the RSV phylogenetic tree built
33 with whole genomes showed identical lineage clusters as that compared to the G gene, but
34 showed more confidence and discriminatory features within individual lineage. Moreover,
35 availability of whole genomes enabled the identification of a broader range of mutations, for
36 instance the identified S377N, K272M, S276N, S211N, S206I and S209Q in Canadian fusion
37 proteins that could be potentially associated with effectiveness of vaccines or antiviral-based
38 therapeutics. In conclusion, the tiled-PCR amplification assays described offer a more
39 streamlined approach to facilitate high-throughput, high sensitivity of RSV WGS, which is
40 capable of supporting enhanced genomic surveillance initiatives, as well as the more
41 comprehensive genomic analyses required to inform public health strategies for the development
42 and usage of vaccines and antiviral drugs.

43 **KEYWORDS**

44 RSV, multiplex PCR, Whole genome sequencing, genomic surveillance, phylogeny, vaccine and
45 antiviral drug

46 **INTRODUCTION**

47 Respiratory Syncytial Virus (RSV) also called human respiratory syncytial virus (hRSV) is a
48 common, contagious airborne viral infection that primarily affects the respiratory tract (1, 2).

49 RSV is a negative-sense enveloped single-stranded non-segmented RNA virus belonging to the
50 *Paramyxoviridae* family, genus *Orthopneumovirus*. This virus has a ~15.2 kb genome that
51 encodes 11 proteins, three surface glycoproteins Fusion protein (F), attachment glycoprotein (G)
52 and the small hydrophobic protein (SH)), RNA dependent RNA polymerase Large protein (L),
53 nucleocapsid (N), phosphoprotein (P), transcriptional regulators (M2-1, M2-2), matrix (M), ,
54 non-structural proteins (NS1, NS2) (3). The F and G protein promote the production of
55 protective immune response, and antigenic differences in the G, F and SH envelope proteins led
56 to the classification of RSV into two major antigenic group RSV A and B. The F protein is
57 responsible for the fusion of the viral envelope with the host cell membrane for the viral entry
58 into the cell, and is highly conserved and immunogenic which is attractive for vaccine
59 development. In contrast, the G protein, which is responsible for cellular attachment, is prone to
60 frequent mutations (3-5).

61 RSV diagnosis is usually based on clinical symptoms, but laboratory tests such as rapid antigen
62 tests (6, 7), polymerase chain reaction (PCR) (8) are commonly used diagnostic methods. In
63 recent years, whole genome sequencing (WGS) and its analysis have provided a more in-depth
64 understanding of RSV outbreaks (9-11). By identifying and characterizing viral isolates causing
65 outbreaks, public health authorities and institutions can implement appropriate control measures,
66 such as isolation, contact tracing, and targeted vaccination campaigns (11). WGS-based genomic
67 surveillance involves the systematic monitoring and analysis of viral genomes to understand their
68 genetic diversity, track transmission patterns, and inform public health interventions (12). It has
69 been instrumental in characterizing RSV isolates circulating globally, identifying emerging
70 clades or subclades, and assessing their impact on disease severity and vaccine efficacy. Initially
71 based on G gene sequences, five RSVA and four RSVB clades were identified, named GA1 to

72 GA5 and GB1 to GB5, respectively. Nextclade (<https://clades.nextstrain.org/>) has now included
73 both G_Clade (13), and recent standardized hRSV Genotyping Consensus Consortium (RGCC)
74 lineages for classifications (14, 15).

75 WGS based genomic surveillance provides complete genetic profiles on the pathogens and
76 enables researchers to identify specific genetic mutations including single nucleotide
77 polymorphisms (SNPs), insertions, deletions, and rearrangements throughout the viral genomes.
78 Especially since the RNA-dependent replication cycle of RSV is error prone with no
79 proofreading mechanism (16). Analysis of genetic variations and evolutionary patterns informs
80 our understanding of how the virus evolves over time and how new isolates or variants emerge.
81 Furthermore, genomic surveillance helps monitoring genetic changes that may affect the
82 antigenicity of RSV isolates (13). By examining specific genomic regions associated with viral
83 surface proteins (e.g., the G, F and SH), researchers can identify potential variations that might
84 affect the effectiveness of vaccines or antiviral-based therapeutics or prophylactic monoclonal
85 antibodies. This information informs vaccine development strategies and the selection of vaccine
86 isolates (17).

87 Although there are a number of RSV vaccine candidates in the pipeline (18-20), since late 2023,
88 only two have been approved for use in Canada for adults over 60 years of age including
89 RSVPreF3 (Arexvy) and RSVpreF (AbrysvoTM) (20, 21). Both comprise elements of the RSV F
90 stabilized in its pre-fusion conformation (22). Additionally, there are two prophylactic RSV
91 monoclonal antibodies (mAbs) approved for use in Canada including palivizumab (SYNAGIS[®]),
92 which is targeted against the F antigenic site II, and nirsevimab (BeyfourtusTM), which is targeted
93 against the F antigenic site Ø. In Canada, the use of palivizumab has been reserved for at risk
94 infants and children under two years of age largely due to cost, limited efficacy, and the

95 requirement of multi-dose regimen. Conversely, nirsevimab, approved for use in Canada in 2023,
96 requires only a single dose and is being recommended for all newborns (22). Nirsevimab makes
97 use of specifically engineered amino acid (aa) changes (M257Y/S259T/T261E) to the fragment
98 crystallizable region of the immunoglobulin G (IgG) antibody that results in an extended half-life
99 in serum capable of providing protection with one injection for the entire RSV season (23, 24).
100 Thus, with wider use of both vaccines and mAbs, it is critical to perform both serological and
101 genomics-based surveillance to identify any shift in antigenicity amongst circulating viruses that
102 could impact RSV vaccines and therapeutics used in Canada (20).

103 It is worth mentioning that the COVID-19 pandemic has significantly influenced research
104 activities across various fields, including RSV. Although several WGS based methods for RSV
105 have been reported previously, their strategies relied upon amplification of many individual PCR
106 products before pooling them for sequencing (25, 26), hybridization to capture probes, or
107 metagenomic sequencing (27), which are time-consuming and not ideal for high-throughput
108 routine analysis. In this study, we have designed multiplex PCR assays for targeted enrichment
109 to obtain both RSVA and RSVB genomes using Primalscheme (28). For validating our
110 developed multiplex PCR-based WGS methods, and exploring subsequent phylogenetic and
111 mutations analyses, we included a pilot of 89 representative clinical specimens from four
112 different provinces including Ontario (ON), Manitoba (MB), British Columbia (BC) and
113 Saskatchewan (SK) in Canada. Our results demonstrate that the multiplex PCR targeted
114 enrichment method and subsequent genomic analysis have great potential to accelerate large-
115 scale RSV sequencing, diagnosis, and genomic surveillance in Canada. Furthermore, this
116 multiplex PCR assay enables the simultaneous amplification of multiple genomic regions of this
117 virus, allowing for a more comprehensive analysis of its genetic diversity and evolution with

118 enhanced efficiency and sensitivity, which eventually enables researchers to better understand
119 the RSV and inform public health strategies regarding vaccine and antiviral development and
120 usage.

121 **MATERIALS AND METHODS**

122 **Reference viruses and clinical specimens**

123 Reference RSV isolates used in this study were purchased from Cedarlane and include VR-955
124 (Strain 9320; originally collected in 1977), VR-1803 (Strain ATCC-2012-11; originally collected
125 in 2012) and VR-1794 (Strain ATCC-2012-10; originally collected in 2012). RSV-positive
126 nasopharyngeal swabs were kindly provided by four provincial laboratories including the British
127 Columbia Centre for Disease Control (BC), Roy Romanow Provincial Laboratory (SK), Cadham
128 Provincial Laboratory (MB), and Kingston Health Sciences Centre (ON). All specimens were
129 collected during the 2022-2023 respiratory virus season with the exception of four specimens
130 from 2016 (SK), one from 2019 (MB) and two from 2021 (BC).

131 **Development of the Tiled-PCR Amplification Assays with PrimalScheme**

132 **RSVA**

133 The initial assay was designed against a dataset comprising all publicly available complete
134 RSVA genomes (n=869; accessed November 22nd, 2022) downloaded from the Bacterial and
135 Viral Bioinformatics Resource Center (BVBRC) (<https://www.bv-brc.org/>). The genomes were
136 clustered at 99% sequence identity using cd-hit-est v. 4.8.1 (29) resulting in 71 clusters, and the
137 corresponding genomes representing each cluster were aligned using MAFFT v 7.505 (30). The
138 alignment was re-ordered such that the first sequence corresponded to the largest cluster with
139 subsequent sequences representing smaller clusters. The re-ordered alignment was processed

140 using Primalscheme v. 1.3.2 (28) to develop a multiplex PCR using tiling amplification scheme
141 with default settings, except the amplicon size range was iterated in 100 nucleotides (nts)
142 increments from 400-500 nts through 1100-1200 nts to identify a primer set with the broadest
143 coverage across the genome. The final assay used an amplicon size range of 800-900 nts and the
144 RSVA isolate, SE01A-0167-V02 (Genbank accession: MZ515773.1), was selected as the
145 coordinate system for Primalscheme (28). An additional 2,527 RSVA genomes were acquired
146 from GISAID (accessed November 24th, 2022) and combined with those from the BVBRC for a
147 combined dataset comprising 3,396 genomes. After low quality (>10% ambiguous bases) and
148 mislabeled sequences were removed, a total of 3,356 sequences were aligned using MAFFT v
149 7.505 and the primers output by Primalscheme were mapped against this alignment in Geneious
150 Prime (V2023.2.1, Biomatters Ltd). The primers were manually modified (i.e., shifted
151 up/downstream, application of degenerate nucleotides, or the development of an
152 alternate/supplementary primers) as required to account for any genetic diversity or to correct
153 obvious flaws within the priming region.

154 RSVB

155 The assay for RSVB was designed similarly as described above for RSVA, but with some
156 changes to simplify the primer design process. The RSVB design was targeted against more
157 contemporary isolates that have been isolated since 2018, which comprised a total of 2,126
158 partial (\geq 8000 bases) and complete genomes downloaded from BVBRC (n=437; accessed March
159 2nd, 2023) and GISAID (n=1,692; accessed March 2nd, 2023). After low quality (>10%
160 ambiguous bases) and mislabeled sequences were removed, a total of 2,097 sequences were
161 clustered at 99%, 98% and 97% sequence identity using cd-hit-est v. 4.8.1 (29). The set of
162 representative sequences output by cd-hit-est from each sequence identity level were aligned

163 using MAFFT v 7.505 (30), and a consensus sequence was generated from each dataset in
164 Geneious Prime (v2023.2.1, Biomatters Ltd) using the majority nts at each position. The
165 consensus sequences were used as input for Primalscheme v. 1.3.2, and the RSVB isolates,
166 HRSV/B/Bern/2019 (Genbank accession: MT107528.1), was used as the coordinate system to
167 develop a tiled PCR amplification assay using default settings except the amplicon size was set
168 to 500 nts. The primers output by Primalscheme were mapped against an alignment comprising
169 the initial dataset of 2,097 sequences and manually modified as required to account for any
170 genetic diversity or to correct obvious flaws within the priming region.

171 Prior to ordering, the primers from both assays were re-assessed *in silico* against their respective
172 datasets to identify any primers that required modification (i.e., shifted up/downstream,
173 application of degenerate nucleotides, or the development of an alternate/supplementary primers)
174 to correct for mutations in their extending ends, and/or to account for diversity not captured
175 during the design with PrimalScheme.

176 **RNA extraction, qRT-PCR and cDNA Synthesis**

177 Viral RNA was extracted from 265 µl of clinical specimens using the Magmax-96 Viral RNA
178 Isolation Kit (Cat No: AMB1836-5, Life Technologies - Invitrogen) as per the manufacturer's
179 protocol, and the samples were processed on the Thermo Scientific KingFisher Flex Purification
180 System. Viral RNA was eluted in 90 µL of Tris elution buffer and either used immediately or
181 stored at -80°C. The presence of RSV and subtypes was confirmed by quantitative real-time PCR
182 (qRT-PCR) using Invitrogen SuperScript III Platinum One-Step qRT-PCR System (Cat#
183 11732088) with RSV subtyping primers and probes based on *L* gene (31) and listed here in Table
184 1. The temperature cycles were one cycle of 50°C for 30 min, one cycle of 95°C for 2 min, 4
185 cycles of 98°C for 15 s, 63°C for 5 min. SuperScript IV Reverse Transcriptase (Cat#18090200,

186 Invitrogen) was used to synthesize cDNA from 5 µL of RNA in conjunction with 0.5 µL of 60
187 µM random hexamers [Cat# S1330S, New England Biolabs (NEB)] in a final reaction volume of
188 10 µL. The reverse transcription reaction was incubated at 42°C for 50 min followed by 70°C for
189 10 min.

190 **Preparation of primer pools and multiplex PCR amplification**

191 For each assay, multiplex primer pools were prepared for each set of primers (i.e., Pool 1 and
192 Pool 2) by combining equal volumes of the appropriate primers (LabReady, 100 uM IDTE, pH
193 8.0; IDT) then diluting them to 5 µM prior to use. Two separate 25 µL PCR reactions, one for
194 each primer pool, were prepared using the Q5 Hot Start High-Fidelity 2 × Master Mix (Cat#
195 M0494L, NEB) as follows: 12.5 µL of Q5 2 × mastermix, 7.5 µL of nuclease-free water, 2 µL of
196 cDNA and 3 µL of either 5 µM primer Pool 1 or Pool 2. PCR amplification was carried out on
197 MiniAmp™ Plus Thermal Cycler (Applied Biosystems™ by ThermoFisher Scientific) with an
198 initial denaturation stage at 98°C for 30 s, followed by 34 cycles of 98°C for 15 s, 63°C for 5
199 min, and a final hold at 4°C. The PCR products from each reaction were combined then purified
200 using an equal volume of Ampure XP SPRI Reagent (Cat# A63881, Beckman Coulter) as per the
201 manufacturer's protocol. The 1 × dsDNA High Sensitivity Kit (Cat# Q33230, Invitrogen) was
202 used to quantify 1 µL of the purified PCR product on the Qubit Flex fluorometer (Invitrogen) as
203 per the manufacturer's protocol, and the purified PCR products were normalized to 16 ng/µL
204 with 0.01 M Tris in preparation for sequencing.

205 Each assay was then optimized iteratively against specimens available in-house to identify and
206 correct poorly-amplifying or dropout regions either by modulation of the relative primer
207 concentrations or development of additional replacement primers. The finalized primer
208 sequences and corresponding ratios for both assays are listed in Tables 2 and 3.

209 **Nanopore library preparation and sequencing workflow**

210 A total of 12.5 µL of normalized PCR product (~200 ng) from each sample was used as input to
211 generate barcoded sequencing libraries using the Native Barcoding Kit (Cat# SQK-NBD114-96,
212 ONT). The protocol was followed exactly with the exception that the NEBNext FFPE DNA
213 Buffer and Repair Mix were not used, and replaced with 1.75 µL of Ultra II End-prep Reaction
214 Buffer. The final library, comprising 12 µL of the pooled libraries, 37.5 µL of Sequencing Buffer
215 and 25.5 µL of Library Beads, was loaded onto a FLO-MIN114 R10.4.1 flow cell with max of
216 96 specimens and sequenced for 72 h on the MinION Mk1C Sequencing Platform.

217 **Bioinformatics analysis of sequence data obtained from Nanopore platform**

218 The FAST5 files were basecalled and demultiplexed using Guppy v 6.5.7 with default settings
219 except specification of dna_r10.4.1_e8.2_400bps_5khz_hac as the basecaller model and SQK-
220 NBD114-96 as the barcode kit with the “--require_barcode_both_ends” parameter. The
221 basecalled FASTQ files were processed using the Nextflow-enabled viralassembly pipeline
222 (<https://github.com/phac-nml/viralassembly>), which is a genericized version of the ncov2019-
223 artic-nf pipeline (<https://github.com/connor-lab/ncov2019-artic-nf>) that automates the ARTIC
224 Network’s Field Bioinformatics Toolkit (<https://github.com/artic-network/fieldbioinformatics>).
225 The viralassembly pipeline is capable of processing any tiled-PCR amplification assay as long as
226 the reference sequence (fasta format) and primer coordinate file (bed format) used to create the
227 assay are provided. Briefly, the viralassembly pipeline automates read mapping, primer trimming,
228 variant calling and consensus sequence generation. Default settings were used, except Medaka
229 (<https://github.com/nanoporetech/medaka>) was selected as the variant caller in conjunction with
230 the r1041_e82_400bps_hac_v4.2.0 model, and the maximum read length to keep was set to 1500
231 nts for the RSVA assay and 1000 nts for the RSVB assay.

232 For the purposes of this study, a genome was considered to be complete if it encompassed the
233 first codon of the first NS1 and the last codon of L protein (15). The Open Reading Frames
234 (ORFs) encoding the RSV genes were identified using Geneious Prime (v2023.2.1, Biomatters
235 Ltd) and characterized as a means to detect and correct any sequencing errors, or to validate
236 legitimate biological mutations that would result in the formation of a truncated protein.

237 **RSVA and RSVB Datasets**

238 In order to contextualize our sequences within the broader population structure of RSV, 124
239 RSVA and 83 RSVB lineage exemplar reference sequences (<https://github.com/rsv-lineages>)
240 were downloaded from NCBI. They were combined with the 52 RSVA and 40 RSVB sequences
241 generated here (37 Canadian + 3 ATCC isolates) resulting in datasets comprising 176 RSVA and
242 123 RSVB sequences, respectively (Supplementary Table 1).

243 **Characterisation of RSVA and RSVB sequences using Nextclade**

244 The RSVA and RSVB datasets were processed using Nextclade's webportal on 9 Aug 2024
245 (<https://clades.nextstrain.org/>) and analyzed using the RSVA module against the reference isolate
246 hRSV/A/England/397/2017 (EPI_ISL_412866) and the RSVB module against the reference
247 isolate hRSV/B/Australia/VIC-RCH056/2019 (EPI_ISL_1653999), respectively. The RGCC
248 lineage assignments and G_Clade typing nomenclatures (15, 32), as well as quality metrics and
249 genomic features including the genome length, breadth of coverage, and the presence of variants
250 including nts and aa substitutions, insertions and deletions and frameshifts are described in
251 Supplementary Table 1. Amino acid changes identified in the F protein amongst the RSVA and
252 RSVB sequences were tabulated and used to generate heatmaps depicting their carriage across
253 the datasets. The heatmaps were generated using a custom R script implemented in RStudio

254 2023.12.1 Build 402 running R version 4.3.2 (2023-10-31) with the following packages:
255 ComplexHeatmap v2.10.0, tidyverse v2.0.0, magrittr v2.0.3, grid v4.1.2, and vegan v2.6.4.

256 **Whole genome phylogenetic analysis**

257 The RSVA and RSVB datasets were aligned separately with MAFFT (v7.520) using default
258 settings (30), then processed with FastTree to infer approximately-maximum-likelihood
259 phylogenetic trees with bootstrap using default settings and the Generalised time reversible
260 (GTR) model (33, 34). The phylogenetic trees were visualized using the Interactive Tree Of Life
261 (iTOL) tool (35) and annotated with the corresponding sequence metadata captured in
262 Supplementary Table 1.

263 **Glycoprotein G gene phylogenetic analysis**

264 The Glycoprotein (G) ORF for all RSVA and RSVB samples was identified using the “Find
265 ORFs” function in Geneious Prime (v2023.2.1, Biomatters Ltd), and the corresponding coding
266 sequences were extracted and subjected to BLAST analysis to confirm they were in-frame and
267 intact. The G sequences from RSVA and RSVB were each aligned separately using MAFFT
268 (v7.520) with default settings (30), then maximum-likelihood trees were inferred using FastTree
269 with default settings and the GTR model (33, 34). The phylogenetic trees were visualized using
270 the iTOL tool (35) and annotated with the corresponding RGCC lineage and G_Clade
271 assignments from Nextclade, as well as the sample collection year and location (Supplementary
272 Table 1).

273 **Co-phylogeny analysis**

274 For both RSVA and RSVB, a custom R script was used to construct a co-phylogeny comprising
275 the tree files generated from the whole genome and G gene coding sequences, and annotated

276 with the corresponding RGCC lineage for each sample. The script was implemented as described
277 above using the following packages managed with pacman v0.5.1: ape v5.7.1, biocmanager
278 v1.30.22, tidyverse v2.0.0, ggtree 3.10.1, phangorn v2.11.1, treeio v1.26.0, phytools 2.1.1, viridis,
279 here v1.0.1, and scico v1.5.0.

280 **Identification of mutations linked to phenotypic or epidemiologic traits using RSVsurver**

281 The RSVsurver, developed by Singapore's Agency for Science, Technology and Research
282 (A*STAR) Bioinformatics Institute (BII) and enabled by GISAID (<https://rsvsurver.bii.a-star.edu.sg/faq.html>), was used to screen the RSV sequences against their curated database for
283 the presence of mutations that may be linked to important phenotypic and epidemiological traits.
284 The RSVA and RSVB datasets were uploaded to the RSVsurver and each sequence was
285 compared against the reference isolates hRSV/A/England/397/2017 and hRSV/B/Australia/VIC-
286 RCH056/2019, respectively.

288 **RESULTS**

289 **Multiplex-PCR based amplification using tiling scheme of RSV and complete genome
290 sequences collection**

291 For RSVA, PrimalScheme generated an assay comprising 24 overlapping amplicons with 0 gaps
292 accounting for 95.5% of the genome. Additional primer sets were manually designed to generate
293 amplicons capturing regions closer to the 5' and 3' ends of the genome extending coverage of the
294 assay to >99% of the genome. For RSVB, Primalscheme generated an assay comprising 38
295 overlapping regions with 0 gaps accounting for 99.6% of the genome. The finalized primer
296 sequences and corresponding ratios for both assays are listed in Tables 2 and 3. Among the in-
297 house 52 RSVA, and 37 RSVB clinical specimens, three ATCC RSVB stains, and Nextclade

298 curated NCBI 124 RSVA and 83 RSVB reference genomes, the antigenic group, subtypes,
299 genome length, identified mutations and GISAID IDs were summarized in the Supplementary
300 Table 1. The near complete genome length ranged from 14,994 nts to 15,225 nts.

301 **Whole genome phylogenetic analysis of 52 Canadian clinical RSVA isolates**

302 The RSVA dataset (n=176; 52 in-house + 124 reference) was demarcated into nine clades using
303 the G_Clade scheme developed by Goya *et al.*, 2020 (Supplementary Table 1 and Fig. 1). The
304 majority of sequences were classified as GA.2.3.5, and largely represent more contemporaneous
305 isolates collected since 2007 including all Canadian isolates sequenced in this study (Fig.1). In
306 contrast, the new scheme proposed by the hRSV RGCC and implemented in Nextclade was more
307 discriminatory and demarcated the same dataset into 25 lineages that more closely reflects the
308 structure of the phylogenetic tree. (Fig. 1; represented by coloured range with the tree nodes).
309 Overall, lineage assignment was well supported by the phylogenetic tree, though not all
310 sequences from lineages A.D and A.D.5 clustered cohesively. However, this was also apparent in
311 the RSVA maximum likelihood tree described in Goya *et al.*, 2024 (15) and the publicly
312 available Nextclade RSVA phylogeny.

313 Isolates derived from clinical specimens characterized in this study, collected between 2016 and
314 2023, were assigned to nine lineages all descending from A.D with the plurality of sequences
315 (25/52) assigned to lineage A.D.5.2 (Fig. 1). The oldest two isolates collected from SK in 2016
316 were assigned to lineages A.D and AD.2.2, respectively and clustered with reference sequences
317 collected between 2012-2016 (blue stars, Fig.1). Based on available data, lineage A.D.2.2 was
318 prevalent amongst isolates characterized in 2016, but detections have declined since then (15).
319 Interestingly, the remaining Canadian isolates, collected during the 2022-2023 respiratory virus
320 season and one from 2021, were assigned to multiple lineages indicating that a diverse array of

321 viruses can be co-circulating simultaneously. With the exception A.D.5.2 and A.D.1, which
322 represent the most common (n=25) and second most common (n=8) lineages detected amongst
323 the Canadian isolates, a lineage tended to be dominated by sequences from a single province
324 (excluding reference sequences) (Fig.1).

325 **RSVA glycoprotein G gene based phylogenetic analyses and its co-phylogeny analysis with**
326 **whole genome-based tree**

327 The structure of the phylogenetic tree derived from the complete coding sequence of the G gene
328 (Fig.2A) was largely congruent with the phylogeny derived from the WGS (Fig.1). Overall, the
329 lineage designations corresponded to well-defined groups in both phylogenies, though use of the
330 WGS showed higher bootstrap values and were better able to resolve the phylogenetic structure
331 within certain sublineages, such as A.D.5.2, and differentiate closely-related sequences that were
332 indistinguishable based on the glycoprotein gene sequence (Fig.1 & Fig.2A). For instance, within
333 the A.D.5.2 lineage, four isolates, RV00279 (CA-SK 2023), RV00288 (CA-SK 2023), RV00290
334 (CA-SK 2023) and RV00292 (CA-SK 2023) (green stars, Fig. 2A), were indistinguishable using
335 the G gene sequence, but at the WGS level, could be resolved into two pairs of identical
336 sequences: RV00279 (CA-SK 2023) and RV00288 (CA-SK 2023), as well as RV00290 (CA-SK
337 2023) and RV00292 (CA-SK 2023) (green stars, Fig. 1).

338 **Whole genome phylogenetic analysis of 37 Canadian clinical RSVB isolates**

339 Using the new RSVB scheme developed by the RGCC, the RSVB dataset (n=123; 37 Canadian
340 + 86 references) were demarcated into 15 lineages that largely reflects the structure of the
341 phylogenetic tree (Fig.3). Comparatively, using the G_Clade scheme developed by Goya *et al.*,
342 2020, the dataset was characterized across seven clades with the majority of sequences classified

343 as GB5.05a, encompassing isolates collected since 2013 including all Canadian isolates
344 sequenced in this study (Fig.3, Supplementary Table 1). One sequence from each of the B
345 (KP856965) and B.D (MH594451) lineages did not cluster cohesively with other sequences from
346 their respective lineages, but this is also reflected in the RSVB maximum likelihood tree
347 described in Goya *et al.*, 2024 (15) and the publicly-available Nextclade RSVB phylogeny.
348 Isolates derived from the Canadian clinical specimens were demarcated into 3 lineages with the
349 majority (30/37) assigned to B.D.E.1, which included those from the 2022-2023 respiratory virus
350 season from all provinces, as well as one collected in 2021 from BC. The two isolates collected
351 in 2016 from SK were assigned to B.D.4.1 and clustered with reference sequences collected
352 between 2013-2015, while the remaining three isolates, two collected from MB in 2019 and 2023,
353 and one collected from BC in 2022, were assigned to B.D.4.1.1, which showed a bit further
354 distance as compared to the rest of NCBI reference isolates (Fig.3). The reference viruses VR-
355 955 (collected in 1977), VR-1794 (collected in 2012), and VR-1803 (collected in 2012) were
356 assigned to lineages B, B.D and B.D.4, respectively, and clustered with appropriately
357 contemporaneous reference sequences.

358 **RSVB glycoprotein G gene based phylogenetic analyses and its co-phylogeny analysis with
359 whole genome-based tree**

360 Similar to RSVA, the structure of the phylogenetic trees derived for RSVB from the complete
361 coding sequences of the G gene (Fig.4A) and WGS (Fig.3) were largely congruent (Fig.4B).
362 Overall, the lineage designations corresponded to well-defined groups in both phylogenies,
363 though use of the complete genome sequences was better able to resolve the phylogenetic
364 structure within certain sublineages such as B.D.E.1, and differentiate closely-related sequences
365 that were indistinguishable based on the glycoprotein gene sequence (Fig.3 & Fig.4A). For

366 instance, in the B.D.E.1 lineage, sequences from four isolates including RV01401 (CA-ON
367 2022), RV01402 (CA-ON 2022), RV01409 (CA-ON 2022) and RV01419 (CA-ON 2022) (green
368 stars, Fig. 4A) were indistinguishable, but at the whole genome sequence level, could be further
369 resolved into separate clusters (green stars, Fig.3& 4A).

370

371 **Characterization of aa substitutions identified amongst the Canadian RSVA and RSVB**
372 **Sequences**

373 Amongst the 52 Canadian RSVA isolates sequenced in this study, a total of 271 unique aa
374 substitutions were observed across all proteins (NS1=4, NS2=4, N=9, P=11, M=6, SH=4, G=100,
375 F=20, M2-1=14, M2-2=13, and L=86) ranging in frequency from 1.9% to 100% relative to
376 hRSV/A/England/397/2017. Notable mutations flagged by RSVsurver include F:T122A (n=29)
377 and F:T122N (n=2), which negates potential glycosylation of residue 120 (magenta mutations
378 Supplementary Fig. 1), as well as F:K272M (n=1, RV00295, CA-SK 2016, A.D.) and F:S276N
379 (n=5, RV00301-00302, RV00305, RV00316 and RV00326)(Fig. 5A, Supplementary Fig.1).
380 Likewise, for the 37 Canadian RSVB isolates sequenced here, a total of 228 unique aa
381 substitutions were observed across all proteins (NS1=3, NS2=8, N=8, P=6, M=22, SH=4, G=62,
382 F=30, M2-1=4, M2-2=10, and L=71) ranging in frequency from 2.7% to 100% relative to
383 hRSV/B/Australia/VIC-RCH056/2019. RSVsurver flagged the presence of F:R191K in two
384 isolates, RV00297 and RV00298 (CA-SK 2016, B.D.4.1) (orange mutation, Supplementary
385 Fig.1), which was shown to be amongst residues capable of modulating RSV fusion activity *in*
386 *vitro* (36).

387 For the Canadian RSVA isolates, no variability was observed within the aa residues
388 corresponding to the nirsevimab targeted antigenic site Ø (62-96, 195-227), while variability was
389 observed at two residues corresponding to the palivizumab targeted antigenic site II (254-277)
390 including F:K272M, as flagged by RSVsurver, as well as F:S276N (n=5); variability was
391 observed at corresponding to the RSVPreF3 targeting antigenic site III (45-54, 301-311, 345-352
392 and 367-378) including F:S377N (n=3). For RSVB, three residues within antigenic site Ø
393 showed variability including F:M206I (n=3), F:R209Q (n=3), F:S211N (n=32) while F:S276N
394 (n=1) represented the only variability observed within antigenic site II, and F:S377N (n=1)
395 observed within antigenic site III (Fig. 5, Supplementary Figs. 1&2).

396 DISCUSSION

397 The recent licensing of two new RSV vaccines and an additional monoclonal antibody for use in
398 Canada will necessitate implementation of a more robust surveillance system to monitor
399 evolution of this pathogen. To that end, we sought to develop tiled-PCR amplification-based
400 assays for both RSVA and RSVB that can be used to quickly generate near-complete genome
401 sequence data from clinical specimens. This method uses highly-multiplexed PCR amplification
402 to generate overlapping amplicons spanning the targeted genome, and has been used to facilitate
403 sequencing of a variety of different viruses (28, 37, 38). Enrichment by PCR is less expensive
404 and complex, faster, and more portable relative to other methods such as capture probe-based
405 assays, or using techniques like ultracentrifugation and host RNA depletion in combination with
406 brute-force metagenomic sequencing, which is attractive for outbreak investigations, surveillance
407 programs and use in clinical settings. However, the primer design and subsequent multiplex PCR
408 optimization can be challenging, especially for more diverse viral species, and primers may need
409 to be updated over time as the targeted virus evolves (28). With the continued democratization of

410 next-generation sequencing and development of more accessible computational tools for assay
411 design (i.e., Primalscheme), tiled-PCR amplification-based assays are increasingly being
412 developed and deployed to support outbreak investigation and response (Zika), as well as for
413 large-scale genomic epidemiology initiatives (i.e., SARS-CoV-2) (28, 39).

414 Here, we conducted a small pilot study to test our assays against RSV-positive clinical
415 specimens collected from four different provinces in Canada between 2016-2023. A total of 52
416 RSVA and 37 RSVB near-complete genomes were recovered and subjected to downstream
417 phylogenetic and comparative genomics analyses. To help contextualize the Canadian sequences
418 within the overall population structure of RSV, we downloaded a collection of lineage exemplar
419 reference sequences and generated phylogenetic trees based on the whole genome and the G
420 gene coding sequences (Supplementary Table 1). Nextclade was used to assign both the new
421 RGCC lineage and G_clade designations for each sequence, and these data were overlaid onto
422 the phylogenetic trees (Figs. 1-4). We observed that the overall structures of the phylogenetic
423 trees derived from the WGS-based and G sequences were largely congruent, though the former
424 was able to distinguish more closely related isolates given the presence of additional sequence
425 data available for interrogation (Figs. 1, 2A, 3 & 4A). This highlights the sensitivity and utility of
426 the whole genome approach to support high resolution outbreak and trace back investigations.
427 Given the contemporaneous nature of the specimens used in this study, all of the RSVA and
428 RSVB sequences contained the A.D and B.D lineage-defining G-gene sequence duplications,
429 respectively. Lineage A.D contains a 72 nts G-gene duplication that emerged in 2011 and by
430 2017, its descendants had replaced all other lineages (15). Similarly, lineage B.D, first detected
431 in 1999, contains a 60-nt G gene duplication and its descendants replaced all other lineages by

432 2009 (Supplementary Fig. 3). The evolutionary impact of the duplications in these lineages are
433 not well understood.

434 From a public health perspective, RSV genomes from different regions and time points provide
435 important information on genetic changes that may affect viral pathogenicity, antigenicity, and
436 vaccine efficacy. This knowledge can ultimately aid in the development of more effective
437 vaccines and antiviral therapies. WGS provides a more complete understanding of the genetic
438 diversity of RSV and permits a comprehensive genomic analysis of specific genomic regions
439 associated with virulence, antigenicity and drug resistance, which is important given the fact that
440 RSV does not employ a proofreading mechanism during replication (16). The assays described
441 here generate near complete genomes that can be readily characterized using both internal and
442 publicly available tools, such as Nextclade and RSVSurver, to facilitate identification of
443 emerging lineages, as well as the presence of biologically important or novel mutations
444 (Supplementary Table 1). RSVsurver was used to visualize the aa changes identified for RSVA
445 and RSVB (colored-balls) among all available antigenic sites within the three-dimensional
446 prefusion and postfusion structure of the RSV F glycoprotein in complex with AM22 (magenta)
447 and Infant Antibody AD-19425 (green) respectively (Supplementary Figs. 1 & 2). The prefusion
448 conformation possesses all six major antigenic sites (\emptyset , I, II, III, IV, V), of which only I, II, III,
449 and IV are present in the postfusion conformation (40). Analysis using RSVSurver revealed that
450 amongst the Canadian RSVA isolates, one isolate from SK collected in 2016 and three isolates
451 collected from MB in 2023 possess the F:S377N mutation located in antigenic site III targeted by
452 the RSVPreF3 vaccine, and this mutation might be immunodominant sites for both RSVA and
453 RSVB isolates (41). Moreover, one RSVA isolate collected from SK in 2016 contained the
454 F:K272M mutation located in antigenic site II, which has been shown to impact the efficacy of

455 palivizumab (42). Similarly, amongst the Canadian RSVB isolates, three mutations were
456 identified including F:S211N (n=32), F:M206I (n=3) and F:R209Q (n=3) that are located in
457 antigenic site Ø targeted by nirsevimab (Fig. 5, Supplementary Table 1). Thus, these isolates
458 present an opportunity to conduct downstream antigenicity and antiviral testing to study the
459 effect of both the well-characterized and rarer mutations observed across the dataset.

460 In conclusion, our tiled PCR amplification-based assays represent a convenient and inexpensive,
461 method for the rapid generation of near complete RSV genomes from clinical specimens. These
462 enhanced WGS methods can ultimately contribute to the advancement of RSV research,
463 particularly in the areas of sequencing, diagnosis, and genomic surveillance. We have also
464 demonstrated that the sequence data generated using our assays can be readily analysed using
465 downstream tools including Nextclade for lineage assignment and RSVSurver for screening of
466 important mutations, and can also be used to support epidemiological investigations within a
467 genomic framework. These assays and subsequent genomic analyses offer potential for serving
468 large-scale RSV genomic surveillance with enhanced efficiency and sensitivity, which will allow
469 researchers to better monitor genomic variability in RSV and inform public health strategies for
470 the development and usage of vaccines and antivirals.

471 **SUPPLEMENTAL MATERIAL**

472 Supplemental material is available online only.

473 **SUPPLEMENTAL TABLE 1** Profiles and identified mutations among tested 176 RSVA and
474 123 RSVB genomes

475 **SUPPLEMENTAL FIG.1** Presence of all amino acid mutations as identified by RSVsurver
476 from the 176 RSVA sequences used in this study contextualized within the 3D structure of the

477 RSV F glycoprotein. (A) prefusion RSVA F glycoprotein (PDB: 6apd, X-ray 4.1 Angstrom) in
478 complex with AM22 (magenta ribbon) & Infant Antibody ADI-19425 (green ribbon). (B)
479 postfusion RSVA F glycoprotein (PDB: 6apb, X-ray 3.0 Angstrom) in complex with Infant
480 Antibody ADI-14359 (green ribbon). The mutations are color-coded by RSVsurver according to
481 their known or predicted biological effect significance. When there are no known effects for the
482 mutation, the mutation will appear in black colored font and assigned interestlevel 0 (least
483 significant). Mutations occurring at a site of interaction will appear in blue colored font and
484 assigned interestlevel 1 (moderately significant). If the mutation occurs at a site known to
485 involved in drug-binding or alters host-cell specificity, it will appear in orange and assigned
486 interestlevel 2 (significant). Mutations will also appear in orange and assigned interest level 2
487 when its equivalent site is known to result in antigenic shifts or causes mild drug resistance.
488 Mutations that create or remove a potential glycosylation site are colored magenta with assigned
489 interestlevel 2.

490 **SUPPLEMENTAL FIG. 2** Identified mutations distributed in the different format of 123 RSVB
491 F glycoprotein in complex with antibody. (A) prefusion RSVB F glycoprotein (PDB: 6apd, X-
492 ray 4.1 Angstrom) in complex with AM22 (magenta ribbon) & Infant Antibody ADI-19425
493 (green ribbon). (B) postfusion RSVB F glycoprotein (PDB: 6apb, X-ray 3.0 Angstrom) in
494 complex with Infant Antibody ADI-14359 (green ribbon). The mutations are color-coded by
495 RSVsurver according to their known or predicted biological effect significance. When there are
496 no known effects for the mutation, the mutation will appear in black colored font and assigned
497 interestlevel 0 (least significant). Mutations occurring at a site of interaction will appear in blue
498 colored font and assigned interestlevel 1 (moderately significant). If the mutation occurs at a site
499 known to involved in drug-binding or alters host-cell specificity, it will appear in orange and

500 assigned interestlevel 2 (significant). Mutations will also appear in orange and assigned interest
501 level 2 when its equivalent site is known to result in antigenic shifts or causes mild drug
502 resistance. Mutations that create or remove a potential glycosylation site are colored magenta
503 with assigned interestlevel 2.

504 **SUPPLEMENTAL FIG. 3** Duplication to define RSV lineage. (A) Representative sequences
505 demonstration for demonstrating A.D lineage defined by the 72-nt duplication. (B)
506 Representative sequences demonstration for demonstrating A.D lineage defined by the 60-nt
507 duplication.

508

509

510 **DATA AVAILABILITY**

511 All the complete RSV genome sequences were deposited in the GISAID database, with the
512 accession number being listed in Supplementary Table 1.

513 **ACKNOWLEDGMENTS**

514 This study was funded by Public Health Agency of Canada. We thank Cole Slater from IRVC
515 NMLB for performing wetlab work in this study, Darian Hole from Computational Operational
516 Genomics Section at NMLB, PHAC to develop the viralassembly pipeline
517 (<https://github.com/phac-nml/viralassembly>), and Ben Hetman for his assistance adapting the R
518 code used to generate the co-phylogeny.

519 **REFERNECES**

- 520 1. Tan L, Coenjaerts FE, Houspie L, Viveen MC, van Bleek GM, Wiertz EJ, Martin DP, Lemey P. 2013.
521 The comparative genomics of human respiratory syncytial virus subgroups A and B: genetic
522 variability and molecular evolutionary dynamics. *J Virol* 87:8213-26.
523 2. Hall CB, Simoes EA, Anderson LJ. 2013. Clinical and epidemiologic features of respiratory
524 syncytial virus. *Curr Top Microbiol Immunol* 372:39-57.
525 3. Battles MB, McLellan JS. 2019. Respiratory syncytial virus entry and how to block it. *Nat Rev
526 Microbiol* 17:233-245.
527 4. Tan L, Lemey P, Houspie L, Viveen MC, Jansen NJ, van Loon AM, Wiertz E, van Bleek GM, Martin
528 DP, Coenjaerts FE. 2012. Genetic variability among complete human respiratory syncytial virus
529 subgroup A genomes: bridging molecular evolutionary dynamics and epidemiology. *PLoS One*
530 7:e51439.
531 5. Rebuffo-Scheer C, Bose M, He J, Khaja S, Ulatowski M, Beck ET, Fan J, Kumar S, Nelson MI,
532 Henrickson KJ. 2011. Whole genome sequencing and evolutionary analysis of human respiratory
533 syncytial virus A and B from Milwaukee, WI 1998-2010. *PLoS One* 6:e25468.
534 6. Zhang N, Wang L, Deng X, Liang R, Su M, He C, Hu L, Su Y, Ren J, Yu F, Du L, Jiang S. 2020. Recent
535 advances in the detection of respiratory virus infection in humans. *J Med Virol* 92:408-417.
536 7. Jha A, Jarvis H, Fraser C, Openshaw PJM. 2016. Respiratory Syncytial Virus. In Hui DS RG, Johnston
537 SL (ed), SARS, MERS and other Viral Lung Infections
538 doi:<https://pubmed.ncbi.nlm.nih.gov/28742304>. European Respiratory Society, Sheffield, UK.
539 8. Henrickson KJ, Hall CB. 2007. Diagnostic assays for respiratory syncytial virus disease. *Pediatr
540 Infect Dis J* 26:S36-40.
541 9. Zhu Y, Zembower TR, Metzger KE, Lei Z, Green SJ, Qi C. 2017. Investigation of respiratory
542 syncytial virus outbreak on an adult stem cell transplant unit by use of whole-genome
543 sequencing. *J Clin Microbiol* 55:2956-2963.
544 10. Johnson PCD, Hagglund S, Naslund K, Meyer G, Taylor G, Orton RJ, Zohari S, Haydon DT,
545 Valarcher JF. 2022. Evaluating the potential of whole-genome sequencing for tracing
546 transmission routes in experimental infections and natural outbreaks of bovine respiratory
547 syncytial virus. *Vet Res* 53:107.
548 11. Lin GL, Golubchik T, Drysdale S, O'Connor D, Jefferies K, Brown A, de Cesare M, Bonsall D, Ansari
549 MA, Aerssens J, Bont L, Openshaw P, Martinon-Torres F, Bowden R, Pollard AJ, Investigators R.
550 2020. Simultaneous viral whole-genome sequencing and differential expression profiling in
551 respiratory syncytial virus infection of infants. *J Infect Dis* 222:S666-S671.
552 12. Ferdinand AS, Kelaher M, Lane CR, da Silva AG, Sherry NL, Ballard SA, Andersson P, Hoang T,
553 Denholm JT, Easton M, Howden BP, Williamson DA. 2021. An implementation science approach
554 to evaluating pathogen whole genome sequencing in public health. *Genome Med* 13:121.
555 13. Panatto D, Domnich A, Lai PL, Ogliastro M, Bruzzone B, Galli C, Stefanelli F, Pariani E, Orsi A,
556 Icardi G. 2023. Epidemiology and molecular characteristics of respiratory syncytial virus (RSV)
557 among Italian community-dwelling adults, 2021/22 season. *BMC Infect Dis* 23:134.
558 14. Aksamentov I, Roemer C, Hodcroft EB, Neher RA. 2021. Nextclade: clade assignment, mutation
559 calling and quality control for viral genomes. *Journal of Open Source Software* 6:3773.
560 15. Goya S, Ruis C, Neher RA, Meijer A, Ammar Aziz A, Hinrichs AS, Gottberg A, Roemer C, Amoako
561 DG, Acuña D, McBroome J, Otieno JR, Bhiman JN, Everatt J, Muñoz-Escalante JC, K. R, Duggan K,
562 Presser LD, Urbanska L, Venter M, Wolter N, Peret TCT, Salimi V, Varsha Potdar V, Borges V,
563 Viegas M. 2024. Standardized phylogenetic classification of human respiratory syncytial virus
564 below the subgroup Level. *Emerg Infect Dis* 30:1631-1641.
565 16. Griffiths C, Drews SJ, Marchant DJ. 2017. Respiratory syncytial virus: infection, detection, and
566 new options for prevention and treatment. *Clin Microbiol Rev* 30:277-319.

- 567 17. Schobel SA, Stucker KM, Moore ML, Anderson LJ, Larkin EK, Shankar J, Bera J, Puri V, Shilts MH,
568 Rosas-Salazar C, Halpin RA, Fedorova N, Shrivastava S, Stockwell TB, Peebles RS, Hartert TV, Das
569 SR. 2016. Respiratory syncytial virus whole-genome sequencing identifies convergent evolution
570 of sequence duplication in the C-terminus of the G gene. *Sci Rep* 6:26311.
- 571 18. Schaerlaekens S, Jacobs L, Stobbelaar K, Cos P, Delputte P. 2024. All eyes on the prefusion-
572 stabilized F construct, but are we missing the potential of alternative targets for respiratory
573 syncytial virus vaccine design? *Vaccines (Basel)* 12.
- 574 19. Mazur NI, Terstappen J, Baral R, Bardaji A, Beutels P, Buchholz UJ, Cohen C, Crowe JE, Jr.,
575 Cutland CL, Eckert L, Feikin D, Fitzpatrick T, Fong Y, Graham BS, Heikkinen T, Higgins D, Hirve S,
576 Klugman KP, Kragten-Tabatabaie L, Lemey P, Libster R, Lowenstein Y, Mejias A, Munoz FM,
577 Munywoki PK, Mwananyanda L, Nair H, Nunes MC, Ramilo O, Richmond P, Ruckwardt TJ, Sande
578 C, Srikanthi P, Thacker N, Waldstein KA, Weinberger D, Wildenbeest J, Wiseman D, Zar HJ,
579 Zambon M, Bont L. 2023. Respiratory syncytial virus prevention within reach: the vaccine and
580 monoclonal antibody landscape. *Lancet Infect Dis* 23:e2-e21.
- 581 20. Langedijk AC, Bont LJ. 2023. Respiratory syncytial virus infection and novel interventions. *Nat Rev Microbiol* 21:734-749.
- 582 21. Canada H. 2023. Summary basis of decision (SBD) for Arexvy. Office of Regulatory Affairs
583 BaRDDO-B, Health Canada, Drug and Health Product Portal. <https://dhpp.hppfb-dgpsa.ca/review-documents/resource/SBD1701180125532>.
- 584 22. Canada Go. 2024. Respiratory syncytial virus (RSV) vaccines: Canadian Immunization Guide.
- 585 23. Ko S, Park S, Sohn MH, Jo M, Ko BJ, Na JH, Yoo H, Jeong AL, Ha K, Woo JR, Lim C, Shin JH, Lee D,
586 Choi SY, Jung ST. 2022. An Fc variant with two mutations confers prolonged serum half-life and
587 enhanced effector functions on IgG antibodies. *Exp Mol Med* 54:1850-1861.
- 588 24. Domachowske JB, Khan AA, Esser MT, Jensen K, Takas T, Villafana T, Dubovsky F, Griffin MP.
589 2018. Safety, tolerability and pharmacokinetics of MEDI8897, an extended half-life single-dose
590 respiratory syncytial virus prefusion F-targeting monoclonal antibody administered as a single
591 dose to healthy preterm infants. *Pediatr Infect Dis J* 37:886-892.
- 592 25. Agoti CN, Otieno JR, Munywoki PK, Mwihuri AG, Cane PA, Nokes DJ, Kellam P, Cotten M. 2015.
593 Local evolutionary patterns of human respiratory syncytial virus derived from whole-genome
594 sequencing. *J Virol* 89:3444-54.
- 595 26. Pangesti KNA, Ansari HR, Bayoumi A, Kesson AM, Hill-Cawthorne GA, Abd El Ghany M. 2023.
596 Genomic characterization of respiratory syncytial virus genotypes circulating in the paediatric
597 population of Sydney, NSW, Australia. *Microb Genom* 9.
- 598 27. Goya S, Sereewit J, Pfalmer D, Nguyen TV, Bakhsh S, Sobolik EB, Greninger AL. 2023. Genomic
599 characterization of respiratory syncytial virus during 2022-23 outbreak, Washington, USA. *Emerg
600 Infect Dis* 29:865-868.
- 601 28. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, Oliveira G, Robles-Sikisaka
602 R, Rogers TF, Beutler NA, Burton DR, Lewis-Ximenez LL, de Jesus JG, Giovanetti M, Hill SC, Black
603 A, Bedford T, Carroll MW, Nunes M, Alcantara LC, Jr., Sabino EC, Baylis SA, Faria NR, Loose M,
604 Simpson JT, Pybus OG, Andersen KG, Loman NJ. 2017. Multiplex PCR method for MinION and
605 Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc*
606 12:1261-1276.
- 607 29. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or
608 nucleotide sequences. *Bioinformatics* 22:1658-9.
- 609 30. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence
610 alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059-66.
- 611 31. Todd AK, Costa AM, Waller G, Daley AJ, Barr IG, Deng YM. 2021. Rapid detection of human
612 respiratory syncytial virus A and B by duplex real-time RT-PCR. *J Virol Methods* 294:114171.
- 613 614

- 615 32. Goya S, Galiano M, Nauwelaers I, Trento A, Openshaw PJ, Mistchenko AS, Zambon M, Viegas M.
616 2020. Toward unified molecular surveillance of RSV: a proposal for genotype definition.
617 Influenza Other Respir Viruses 14:274-285.
618 33. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with
619 profiles instead of a distance matrix. Mol Biol Evol 26:1641-50.
620 34. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2--approximately maximum-likelihood trees for
621 large alignments. PLoS One 5:e9490.
622 35. Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree
623 display and annotation. Nucleic Acids Res 49:W293-W296.
624 36. Hotard AL, Lee S, Currier MG, Crowe JE, Jr., Sakamoto K, Newcomb DC, Peebles RS, Jr., Plempert
625 RK, Moore ML. 2015. Identification of residues in the human respiratory syncytial virus fusion
626 protein that modulate fusion activity and pathogenesis. J Virol 89:512-22.
627 37. Li K, Shrivastava S, Brownley A, Katzel D, Bera J, Nguyen AT, Thovarai V, Halpin R, Stockwell TB.
628 2012. Automated degenerate PCR primer design for high-throughput sequencing improves
629 efficiency of viral sequencing. Virol J 9:261.
630 38. Isabel S, Eshaghi A, Duvvuri VR, Gubbay JB, Cronin K, Li A, Hasso M, Clark ST, Hopkins JP, Patel
631 SN, Braukmann TWA. 2024. Targeted amplification-based whole genome sequencing of
632 Monkeypox virus in clinical specimens. Microbiol Spectr 12:e0297923.
633 39. Lambisia AW, Mohammed KS, Makori TO, Ndwiga L, Mburu MW, Morobe JM, Mora EO,
634 Musyoki J, Murunga N, Mwangi JN, Nokes DJ, Agoti CN, Ochola-Oyier LI, Githinji G. 2022.
635 Optimization of the SARS-CoV-2 ARTIC network V4 primers and whole genome sequencing
636 protocol. Front Med (Lausanne) 9:836728.
637 40. Tang A, Chen Z, Cox KS, Su HP, Callahan C, Fridman A, Zhang L, Patel SB, Cejas PJ, Swoyer R,
638 Touch S, Citron MP, Govindarajan D, Luo B, Eddins M, Reid JC, Soisson SM, Galli J, Wang D, Wen
639 Z, Heidecker GJ, Casimiro DR, DiStefano DJ, Vora KA. 2019. A potent broadly neutralizing human
640 RSV antibody targets conserved site IV of the fusion glycoprotein. Nat Commun 10:4153.
641 41. Bin L, Liu H, Tabor DE, Tovchigrechko A, Qi Y, Ruzin A, Esser MT, Jin H. 2019. Emergence of new
642 antigenic epitopes in the glycoproteins of human respiratory syncytial virus collected from a US
643 surveillance study, 2015-17. Sci Rep 9:3898.
644 42. Zhu Q, McAuliffe JM, Patel NK, Palmer-Hill FJ, Yang CF, Liang B, Su L, Zhu W, Wachter L, Wilson S,
645 MacGill RS, Krishnan S, McCarthy MP, Losonsky GA, Suzich JA. 2011. Analysis of respiratory
646 syncytial virus preclinical and clinical variants resistant to neutralization by monoclonal
647 antibodies palivizumab and/or motavizumab. J Infect Dis 203:674-82.
648

649 **Figure legends**

650 **FIG 1** Phylogenetic trees demonstrating 176 RSVA isolates including 52 Canadian isolates and
651 124 Nextclade references based on whole-genome sequences. hRSV Genotyping Consensus
652 Consortium (RGCC) lineages were shown in different color range and G_Clade was overlaid as
653 a color strip in the phylogenetic tree. The year of collection of 2022 and 2023 were shown in
654 light red and red respectively, and the rest were shown in black. Representative Canadian isolates

655 primarily collected in the year of 2022–2023 and from four provinces, namely British Columbia
656 (BC), Manitoba (MB), Ontario (ON) and Saskatchewan (SK), with the text color of blue, red,
657 purple and green respectively. Bootstrap was annotated in each branch of the tree with light
658 purple dots.

659 **FIG 2** Phylogenetic tree of glycoprotein (G) gene and its comparison with the whole genome-
660 based tree based on 176 RSVA isolates including 52 Canadian isolates and 124 Nextclade
661 references. **(A)** Phylogenetic trees based on G gene sequences. RSV RGCC lineages were shown
662 in different color range and G_Clade was overlaid as a color strip in the phylogenetic tree. The
663 year of collection of 2022 and 2023 were shown in light red and red respectively, and the rest
664 were shown in black. Representative Canadian isolates primarily collected in the year of 2022–
665 2023 and from four provinces, namely BC, MB, ON and SK, with the text color of blue, red,
666 purple and green respectively. Bootstrap was annotated in each branch of the tree with light red
667 dots. **(B)** Co-phylogenetic tree comparison between whole genome and G gene sequences.
668 RGCC lineages were shown in different colors.

669 **FIG 3** Phylogenetic trees demonstrating 123 RSVB isolates including 37 Canadian and three
670 ATCC isolates, and 86 references based on whole-genome sequences. RGCC lineages were
671 shown in different color range and G_Clade was overlaid as a color strip in the phylogenetic tree.
672 The year of collection of 2022 and 2023 were shown in light red and red respectively, and the
673 rest were shown in black. Representative Canadian isolates primarily collected in the year of
674 2022–2023 and from four provinces, namely BC, MB, ON and SK, with the text color of blue,
675 red, purple and green respectively. Bootstrap was annotated in each branch of the tree with light
676 cyan dots.

677 **FIG 4** Phylogenetic tree of glycoprotein (G) gene and its comparison with the whole genome-
678 based tree based on 123 RSVB isolates including 37 Canadian isolates and 86 references. (A)
679 Phylogenetic trees based on G gene sequences. RGCC lineages were shown in different color
680 range and G_Clade was overlaid as a color strip in the phylogenetic tree. The year of collection
681 of 2022 and 2023 were shown in light red and red respectively, and the rest were shown in black.
682 Representative Canadian isolates primarily collected in the year of 2022–2023 and from four
683 provinces, namely BC, MB, ON and SK, with the text color of blue, red, purple and green
684 respectively. Bootstrap was annotated in each branch of the tree with light green dots. (B) Co-
685 phylogenetic tree comparison between whole genome and G gene sequences. RGCC lineages
686 were shown in different colors.

687 **FIG 5** Heatmap distribution of amino acids substitutions in (A) 176 RSVA and (B) 123 RSVB
688 isolates. Red and blue colors represent gene presence and absence, respectively. The x axis
689 contains the list of respective 176 RSVA and 123 RSVB isolates with their lineages indicated
690 and the y axis shows the identified aas substitutions. For 52 Canadian RSVA and 37 RSVB
691 isolates, blue font mutations are assigned interestlevel 1 (moderately significant); orange font
692 mutations are assigned interestlevel 2 (significant), which is known to involved in drug-binding.
693 Vaccines RSVpreF/F3 is in red with red ball pointing to its detected antigenic site mutation. Two
694 mAbs of palivizumab Nirsevimab are shown in purple and green, respectively.

695 **Table 1: Primer/probe sequences based on L gene used for identifying RSV subtypes**

Primer/probe names	Sequences (5'—3')
RSV_Forward	AATACAGCCAAATCTAACCAACTTACA
RSV_Reverse	GCCAAGGAAGCATGCAATAAA
RSVA_probe	FAM-TGCTATTGTGCACTAAAG-MGBNFQ
RSVB_probe	FAM-CACTATTCTTACTAAAGATGTC-MGBNFQ

696

697

698

699

700

701

Table 2: RSVA multiplex PCR primer pools

	Name	Sequences	Ratio
Pool 1	hRSVA_99_800-900_v3_1_LEFT	TGTTATTACAAGTAGTGATATTGCCCY	5×
	hRSVA_99_800-900_v3_1_RIGHT	TTGCCCATCTTCATCTTATRT	5×
	hRSVA_99_800-900_v3_3_LEFT	TGAAATGAAACGTTATAAAGGYTTAYTACC	5×
	hRSVA_99_800-900_v3_3_RIGHT	AATTGGGCTTGTCCCTRCAGT	5×
	hRSVA_99_800-900_v3_5_LEFT	GAAGCTATGGCAAGACTCAGGA	1×
	hRSVA_99_800-900_v3_5_RIGHT	TCAAGTGTGTTAGATCTTATTCTGA	1×
	hRSVA_99_800-900_v3_7_LEFT	AAGCAAATTCTGGCCTTACTTTAC	1×
	hRSVA_99_800-900_v3_7_RIGHT	GTTGGATTGTTGCTGCATATGCT	1×
	hRSVA_99_800-900_v3_9_LEFT	CAAACCTCAAACCACAAAACCAAAR	2×
	hRSVA_99_800-900_v3_9_RIGHT	GCRATTGCAGATCCAACACCTA	2×
	hRSVA_99_800-900_v3_11_LEFT	TGTAACACACCTGTAAGCACTTATATGT	1×
	hRSVA_99_800-900_v3_11_RIGHT	GTAGTTATCATGATATTGTGGTGGATTAC	1×
	hRSVA_99_800-900_v3_13_LEFT	ATAAGTGGAGCTGCAGAGTTGG	1×
	hRSVA_99_800-900_v3_13_RIGHT	AGGACCATTGAATATGTAACCTCCTAARG	1×
	hRSVA_99_800-900_v3_15_LEFT	CTATGCTATATTGAATAAACTGGGGCT	3×
	hRSVA_99_800-900_v3_15_RIGHT	AGGTTRTTGTCACCTGCAAGYT	3×
	hRSVA_99_800-900_v3_17_LEFT	AGATGGTTAACCTACTATAAACTAAACAC	2×
	hRSVA_99_800-900_v3_17_RIGHT	TCCATAGTTTGACACCACCCCTT	2×
	hRSVA_99_800-900_v3_19_LEFT	AGTATAAAGAAAGTCCTAACAGAGTGGGAC	1×
	hRSVA_99_800-900_v3_19_LEFT_alt1	TAGTATAAAGAAAGTCCTAACAGAGTRGG	1×
	hRSVA_99_800-900_v3_19_RIGHT	GCAGAAGTCTTCCAGTATGTTAGT	1×
	hRSVA_99_800-900_v3_21_LEFT	ACTATAGCTAGTGGCATAATCATAGARAA	2×
	hRSVA_99_800-900_v3_21_RIGHT	TCCCTCTCCCCAATCTTTCAA	2×
	hRSVA_99_800-900_v3_23_LEFT	AATTACAACAAATTATCATCCYACACC	3×
	hRSVA_99_800-900_v3_23_RIGHT	CATCTTGAGCATGATATTACTATTAAYGTC	3×
	hRSVA_99_800-900_v3_25_LEFT	AGAGTGTGTTAGTGGAGATATACTATC	5×
	hRSVA_99_800-900_v3_25_RIGHT	ACGAGAAAAAAAGTGTCAAAACTAATRTC	5×
	hRSVA_99_800-900_v3_25_RIGHT_alt2	AATATACATATAAACCAATTAGATTGGATTAA	5×
Pool 2	hRSVA_99_800-900_v3_0_LEFT	CGAAAAAAATGCGTACAACAAACTT	2×
	hRSVA_99_800-900_v3_0_LEFT_alt1	GGGCAAATAAGAATTGATAAGTACCA	2×
	hRSVA_99_800-900_v3_0_RIGHT	ACTTTGTGCAATAGTTCATTTCATAGTT	2×
	hRSVA_99_800-900_v3_2_LEFT	CCCATAATACAAAGTATGATCTCAATCCRT	1×
	hRSVA_99_800-900_v3_2_RIGHT	CCACCTCTGGTAGAAGATTGTGC	1×
	hRSVA_99_800-900_v3_4_LEFT	TGACAGCAGAAGAACTAGAGGC	1×

hRSVA_99_800-900_v3_4_RIGHT	TCAGAAATCTTCAAGTGATAGATCATTRTC	1×
hRSVA_99_800-900_v3_6_LEFT	GATCTCACTATGAAAACACTCAAYCC	1×
hRSVA_99_800-900_v3_6_RIGHT	TTGACTCGAGCTCTGGTARY	1×
hRSVA_99_800-900_v3_8_LEFT	CCAGATCAAGAACACAACCCCCAA	1×
hRSVA_99_800-900_v3_8_RIGHT	GCAACTCCATTGTTATTCGCCCC	1×
hRSVA_99_800-900_v3_10_LEFT	CAAAGCACACCAGCAGCMAA	2×
hRSVA_99_800-900_v3_10_RIGHT	TTGAATATGTCAATGTTGCAGAGATTAC	2×
hRSVA_99_800-900_v3_12_LEFT	TCCCYTCTGATGAATTGATGCA	1×
hRSVA_99_800-900_v3_12_RIGHT	TGAGTTCAGTRAGGAGTTGCTCA	1×
hRSVA_99_800-900_v3_14_LEFT	TGATACTACCTGACAAATATCCTTAG	1×
hRSVA_99_800-900_v3_14_RIGHT	TTGTTTGWGGATTGATTTGYCTG	1×
hRSVA_99_800-900_v3_14_RIGHT_alt1	TATTAACCATGATGGAGGATGTTGC	1×
hRSVA_99_800-900_v3_16_LEFT	TGCTCAACAACATCACAGATGC	1×
hRSVA_99_800-900_v3_16_RIGHT	TCTGCTAATATTGAACCTGTCTGAACA	1×
hRSVA_99_800-900_v3_18_LEFT	TCCTGGTTACATTAACTATTCCCATG	1×
hRSVA_99_800-900_v3_18_RIGHT	AGGAAATCAGGAGTTCTCTATAGAAACT	1×
hRSVA_99_800-900_v3_20_LEFT	AACCTACATATCCTCAYGGGCT	2×
hRSVA_99_800-900_v3_20_RIGHT	AGCTAAGGCCAARCTTATACAGT	2×
hRSVA_99_800-900_v3_22_LEFT	TCTMATGTTAATTCTAATTAAATTGGCRCA	5×
hRSVA_99_800-900_v3_22_RIGHT	TCTTGTYTGCTGTAATTGGTTCTAAC	5×
hRSVA_99_800-900_v3_24_LEFT	GTTCTACAGGTTGAAAATTAGTATAGAGT	3×
hRSVA_99_800-900_v3_24_RIGHT	CTGTTGATCTGAAATTAAAACATGRTTGAACC	3×

703

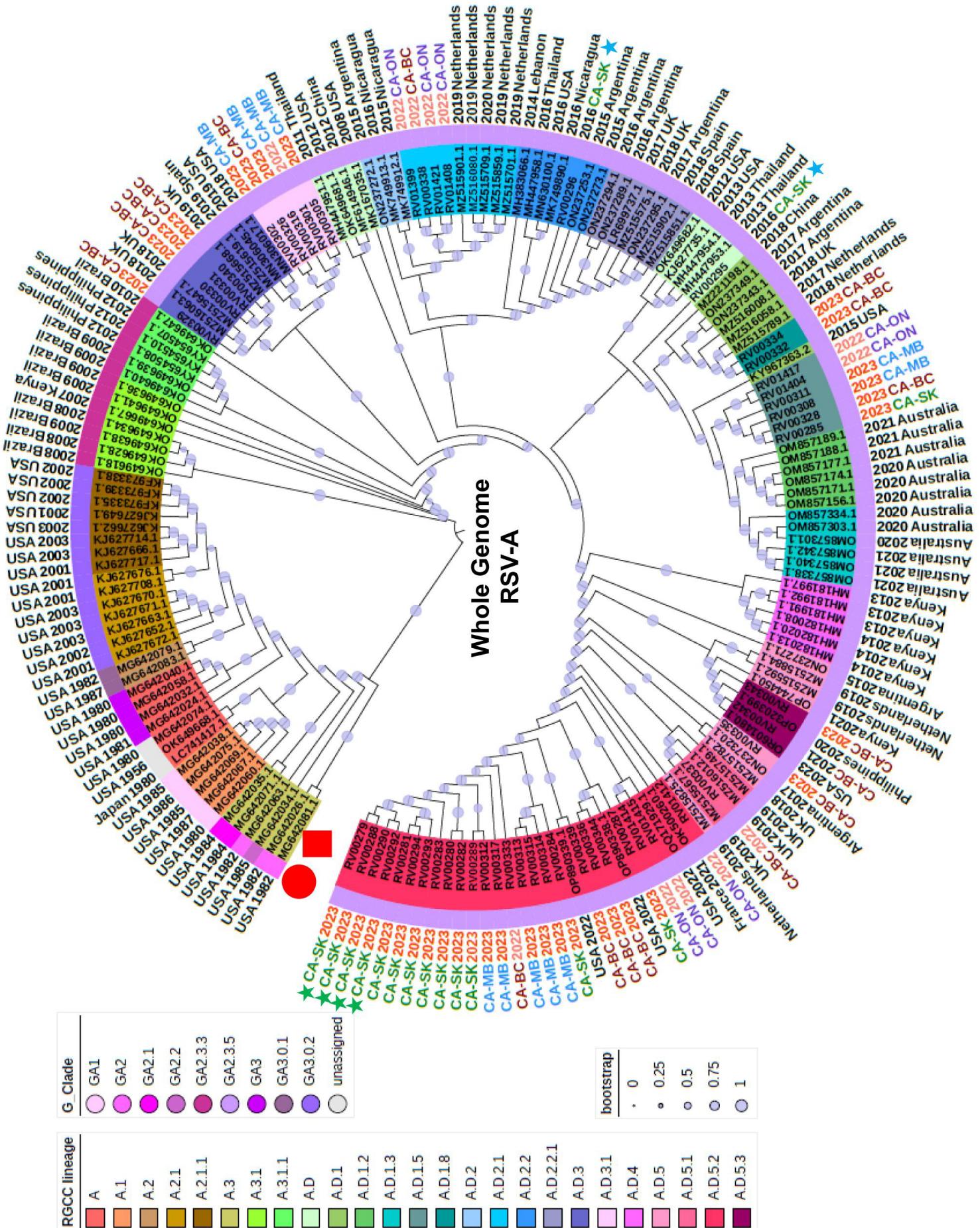
704 **Table 3: RSVB multiplex PCR primer pools**

	Name	Sequences	Ratio
Pool 1	hRSVB_2018-2023_500_1_LEFT	CTAGACTCCGTACCGCGAAAAA	2×
	hRSVB_2018-2023_500_1_LEFT_alt1	AAATGCGTACTACAAACTTGCACACTC	2×
	hRSVB_2018-2023_500_1_RIGHT	GGAGATCAAGGCCAAGTAAATCAGA	2×
	hRSVB_2018-2023_500_3_LEFT	TCAATTATGAGATGAAGCTATTGCACA	1×
	hRSVB_2018-2023_500_3_RIGHT	TGCATCTTCAGTGATTAATAGCATACC	1×
	hRSVB_2018-2023_500_5_LEFT_alt1	CCAGACTGTGGGATGATAAACTGTG	1×
	hRSVB_2018-2023_500_5_RIGHT	TTGCCTAGGACCACACTTGAGA	1×
	hRSVB_2018-2023_500_7_LEFT_alt1	CAAGTTGCATCATCCAAAGATCCTAA	1×
	hRSVB_2018-2023_500_7_RIGHT_alt1	TCTTGCCTAGGCCCTAACCTAT	1×

	hRSVB_2018-2023_500_7_RIGHT_alt2	GCCTCTAACCTATCATTGGTCAT	1×
	hRSVB_2018-2023_500_9_LEFT	GGAAACATACGTGAACAAGCTTCA	1×
	hRSVB_2018-2023_500_9_RIGHT_alt1	ACACTGATTGATCTTAGATAGGTTGGTATT	1×
	hRSVB_2018-2023_500_11_LEFT_alt1	TCCTCAACTGCACACTATATCTAACATC	1×
	hRSVB_2018-2023_500_11_RIGHT	TGTGATGCCATGACTCTGTGAG	1×
	hRSVB_2018-2023_500_13_LEFT	CACAAAGTTACACTAACAACTGTACACA	10×
	hRSVB_2018-2023_500_13_RIGHT	TCTGTAGTCTGGGGGTTGGTTT	10×
	hRSVB_2018-2023_500_15_LEFT_alt1	CTGGGGCAAATAACCATGGAGY	1×
	hRSVB_2018-2023_500_15_RIGHT_alt1	TGTTCACTTCTCCTCAAGGTG	1×
	hRSVB_2018-2023_500_17_LEFT	TCAATGATATGCCTATAACAAATGATCAGA	5×
	hRSVB_2018-2023_500_17_RIGHT_alt1	TCCACAGATTTTGTGGATGC	5×
	hRSVB_2018-2023_500_19_LEFT	TCGTAGATCCGATGAATTATTACATAATGT	1×
	hRSVB_2018-2023_500_19_RIGHT	TGACTGTAGTGGCATCTTCTACC	1×
	hRSVB_2018-2023_500_21_LEFT	CAAACTGTTATATCATAACATCGAGAGCA	1×
	hRSVB_2018-2023_500_21_RIGHT_alt1	GGATCCATTGTCCCATAACTTATTRAG	1×
	hRSVB_2018-2023_500_23_LEFT	ATAAAAGCATGTCCTCGTCTGAAC	2×
	hRSVB_2018-2023_500_23_RIGHT	AGCCCTTATGATAAACAAATGCAACC	2×
	hRSVB_2018-2023_500_25_LEFT_alt1	TCACAGATGCAGCTATTAAGGC	1×
	hRSVB_2018-2023_500_25_RIGHT	ACTCACGATAGAACCGCAATCC	1×
	hRSVB_2018-2023_500_27_LEFT	TGCAACCAGGTATGTTAGGCAA	1×
	hRSVB_2018-2023_500_27_RIGHT	AATGATATGGCTTCAATGGTCCAC	1×
	hRSVB_2018-2023_500_29_LEFT_alt1	AGGTCCATGGATAAATACAATACTTGATGA	2×
	hRSVB_2018-2023_500_29_RIGHT	GCCTGTGGATCCCTCATCAATG	2×
	hRSVB_2018-2023_500_31_LEFT	AAAAACATCAGCGATAGATAACAACGTGA	1×
	hRSVB_2018-2023_500_31_RIGHT_alt1	ATGACAGTCCAAGTGTCCAGTAC	1×
	hRSVB_2018-2023_500_31_RIGHT_alt2	CATATGACAGTCCAAGTGTCC	1×
	hRSVB_2018-2023_500_33_LEFT	ACATTGATGAAAACCTCTATATTACAGG	2×
	hRSVB_2018-2023_500_33_RIGHT	TGACTTTGTCTAGGAAAACTTAGACA	2×
	hRSVB_2018-2023_500_35_LEFT_alt1	CTGCTAACAAAACAAATAAGGATTGCTAA	1×
	hRSVB_2018-2023_500_35_RIGHT_alt1	ACCAGCTCCTCACCTATGAATG	1×
	hRSVB_2018-2023_500_37_LEFT	TGGAGTAAGCATGTAAGAAAGTGCA	1×
	hRSVB_2018-2023_500_37_RIGHT_alt1	TCTAAAGTTAAAACATGATCCAGCCAT	1×
Pool 2	hRSVB_2018-2023_500_2_LEFT_alt1	TGCTCTAACCTATGGCTAATAGATGATAA	1×
	hRSVB_2018-2023_500_2_RIGHT_alt1	GCATAGGAAATGTGCCATATTTGTA	1×
	hRSVB_2018-2023_500_4_LEFT_alt1	ACACTATTCAACGTAGTACAGGAGA	1×
	hRSVB_2018-2023_500_4_RIGHT	CATTGTTGCCCTCCTAATTACTGC	1×
	hRSVB_2018-2023_500_6_LEFT_alt1	CAGAAAGTTGGGAGGAGAACG	1×
	hRSVB_2018-2023_500_6_RIGHT	TGTTGGTGCCAGATGTTATCGG	1×

hRSVB_2018-2023_500_8_LEFT_alt1	CTCGTGACGGAATAAGAGATGC	1×
hRSVB_2018-2023_500_8_RIGHT	TGATGCGGGATCATCATCTTTTC	1×
hRSVB_2018-2023_500_10_LEFT	ACCCCACACTCATGAGATCATTGC	1×
hRSVB_2018-2023_500_10_RIGHT	GCAGACAATGGCTGGAAGTGAT	1×
hRSVB_2018-2023_500_12_LEFT_alt1	TCGACACATAGTGTCTCCCATTAT	10×
hRSVB_2018-2023_500_12_RIGHT_alt1	GGCTAACCCTTCTGGTGAGACT	10×
hRSVB_2018-2023_500_14_LEFT_alt1	ACCACAAACAAAAGAGACYCYA	10×
hRSVB_2018-2023_500_14_RIGHT_alt1	CCTCAGTTATGTTCTGACTTGAGGY	5×
hRSVB_2018-2023_500_16_LEFT_alt1	AGGAAACGAAGATTCTGGCTT	1×
hRSVB_2018-2023_500_16_RIGHT_alt1	AGCTGTACAACATATGCAAGGAC	1×
hRSVB_2018-2023_500_18_LEFT_alt1	CAAAAACAGACATAAGCAGCTCAG	1×
hRSVB_2018-2023_500_18_RIGHT	TGATTCACCTAGTTGGTCTTGCT	1×
hRSVB_2018-2023_500_20_LEFT	TCACAAAACTAACAGCTGGGC	1×
hRSVB_2018-2023_500_20_RIGHT	TTGTCTTCTTCAGCACGCTCTGC	1×
hRSVB_2018-2023_500_22_LEFT	ACATCTAACATCCCTGAAGATATATACAGT	2×
hRSVB_2018-2023_500_22_RIGHT_alt1	AGTTTATTCAAGATGGCGTACACYT	2×
hRSVB_2018-2023_500_24_LEFT_alt1	TCAAATGAGGTAAAAAGTCATGGGT	2×
hRSVB_2018-2023_500_24_RIGHT	ACCATTATGATATTATCAGACACTGTCTT	2×
hRSVB_2018-2023_500_26_LEFT	GCTATTGTCCTACCTCTAAGATGGT	1×
hRSVB_2018-2023_500_26_RIGHT_alt1	TGTCAAACCTCAGGGAGAATTG	1×
hRSVB_2018-2023_500_28_LEFT_alt1	TGAAGTTGATGAACAAAGTGGTTA	2×
hRSVB_2018-2023_500_28_RIGHT_alt1	ACTGCATAATAAGCTTCTCCTCTGTA	2×
hRSVB_2018-2023_500_30_LEFT_alt1	AGCTCCAGGATTTCCAGAYGAT	2×
hRSVB_2018-2023_500_30_RIGHT	TCTCTTGTCTTGTACAATCTAGTGG	2×
hRSVB_2018-2023_500_32_LEFT_alt1	CCAAATAGATTATTAGCAAAATTAGACTGG	2×
hRSVB_2018-2023_500_32_RIGHT_alt1	TTGGGTTAAACTTATTATCTGGTAGGAAC	2×
hRSVB_2018-2023_500_34_LEFT_alt1	GCAAAATTAGAATGTGATATGAACACTTCAGA	2×
hRSVB_2018-2023_500_34_RIGHT	ACAGGGATTAATGATACTTTCTAAAGCT	2×
hRSVB_2018-2023_500_36_LEFT_alt1	CCTTGGCATCATGTCAATAGATTTAACTT	1×
hRSVB_2018-2023_500_36_RIGHT	TGAAATCAATATCATCTTGAGCATGGT	1×
hRSVB_2018-2023_500_38_LEFT	ACTTCATTGTCAAAATTGAAGAGTGTAGT	1×
hRSVB_2018-2023_500_38_RIGHT	CATGCCGCCACGAGAAAAA	2×
hRSVB_2018-2023_500_38_RIGHT_alt1	AATGTCTCGTTGTGTTGAAATGCACATR	2×

FIG 1



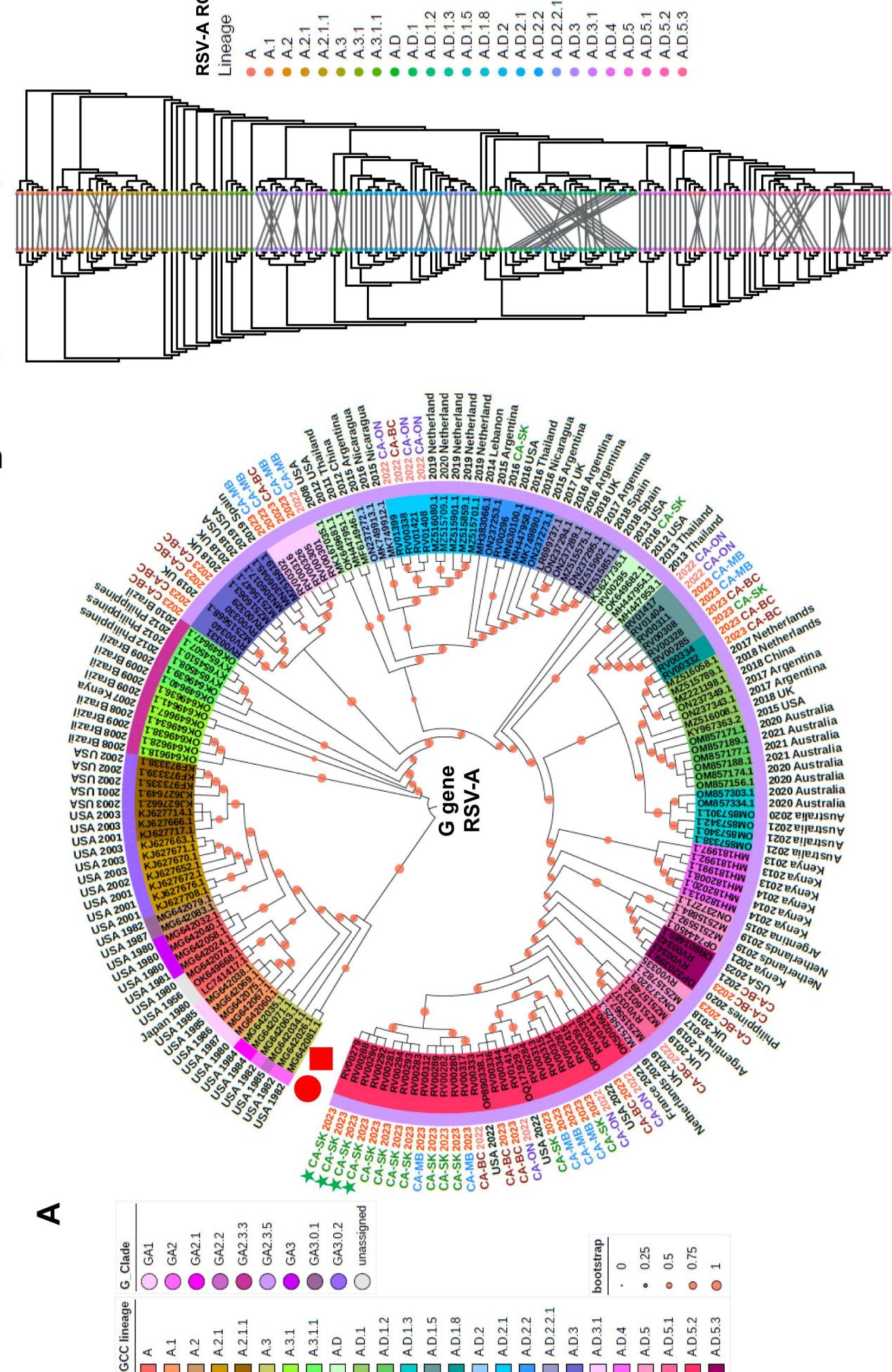


FIG 2

FIG 3

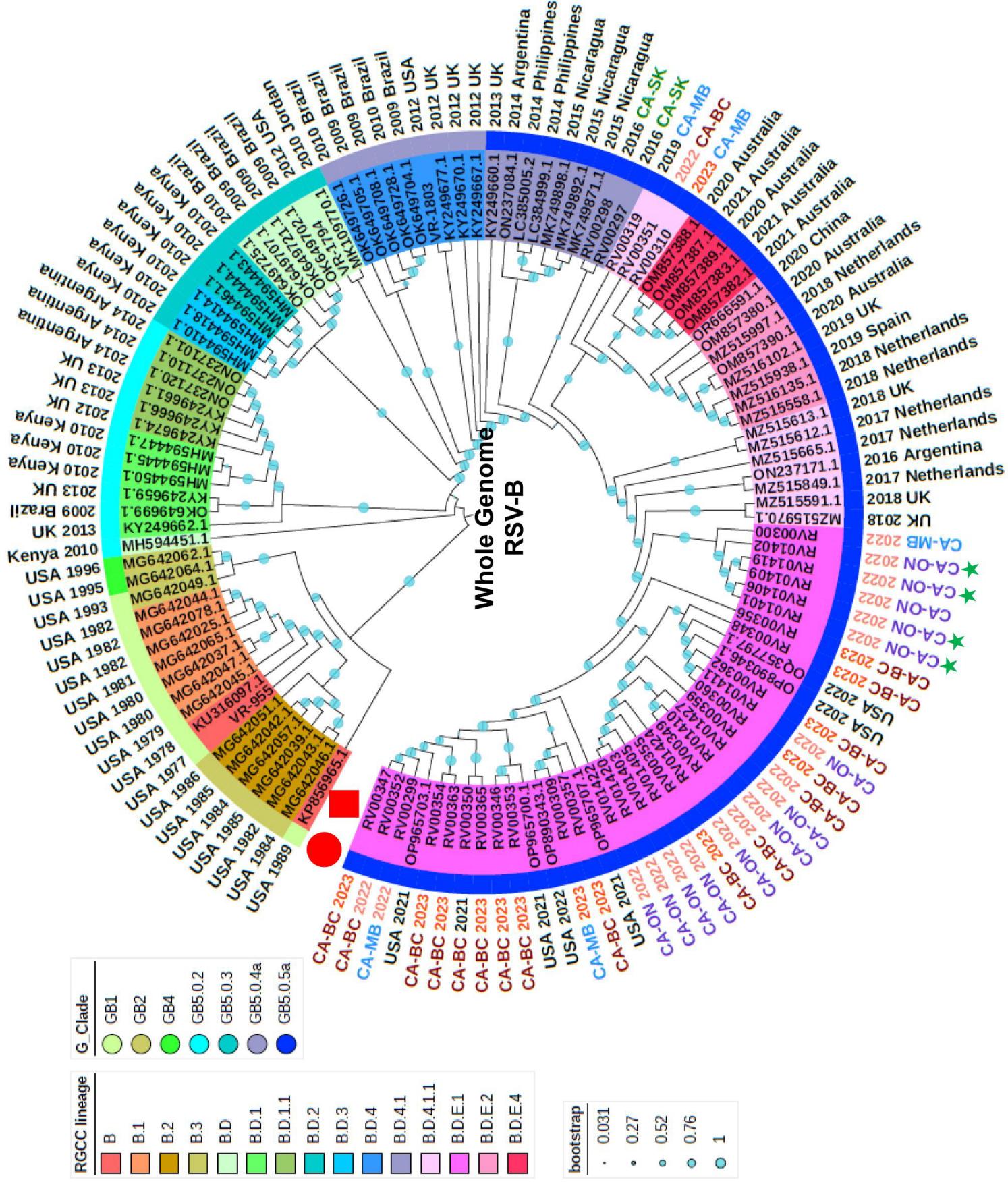


FIG 4

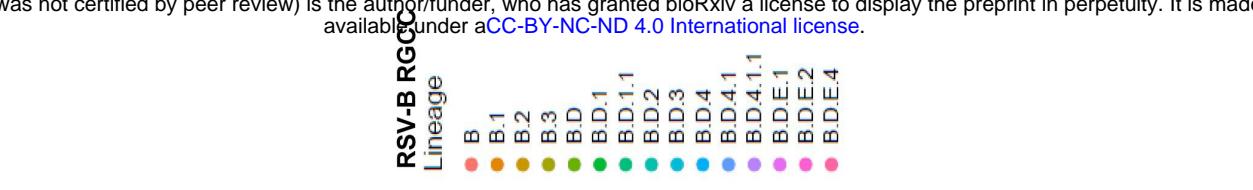
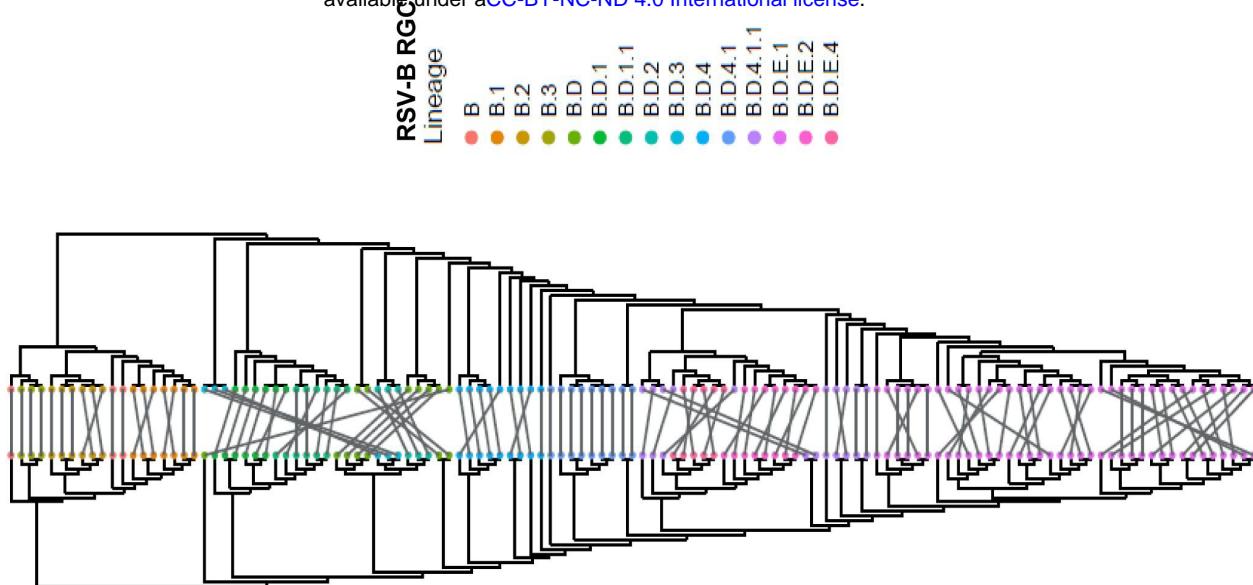
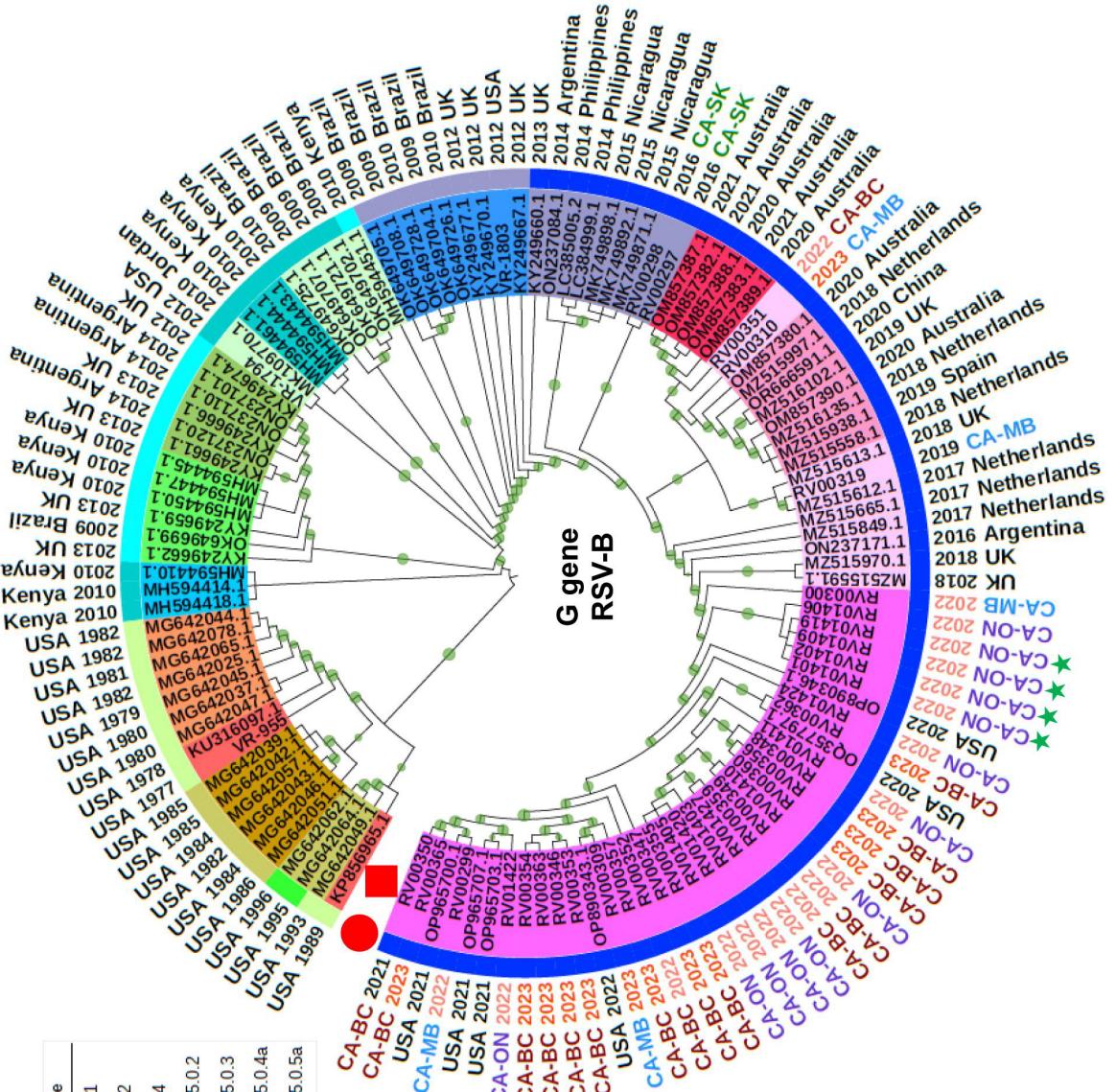
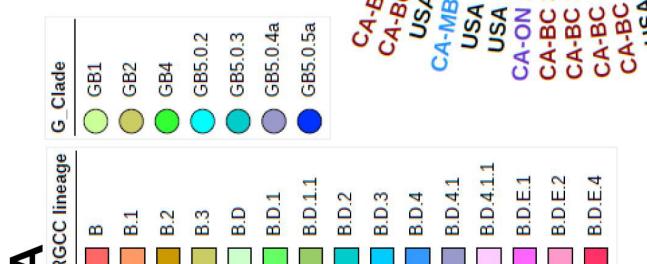


FIG 5

