WARWICK
THE UNIVERSITY OF WARWICK

Manuscript version: Author's Accepted Manuscript
The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/151368

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

warwick.ac.uk/lib-publications

Article Summary Line: This proposal of a human respiratory syncytial virus (HRSV) nomenclature below the species level is brought forward by a group of members of the International RSV Society, a special interest group of the International Society for Influenza and other Respiratory Virus Diseases (isirv), and by members of the WHO Global RSV Surveillance Project and will facilitate sample handling, database submission and analysis of HRSV sequence data.

Running Title: Universal HRSV nomenclature below species level

Keywords: Human respiratory syncytial virus, strain, isolate, specimen, nomenclature, species

PeerReview

Title: Proposal for Human Respiratory Syncytial Virus (HRSV) Nomenclature below the Species Level

Dedicated to José Antonio Melero

Authors: Vahid Salimi, Mariana Viegas[#], Alfonsina Trento[#], Charles N. Agoti, Larry J. Anderson, Vasanthi Avadhanula, Justin Bahl, Louis Bont, J. Rodney Brister, Patricia A. Cane, Mónica Galiano, Barney S. Graham, Eneida L. Hatcher, Orienka Hellferscee, David M. Henke, Siddhivinayak Hirve, Sandra Jackson, Els Keyaerts, Leyla Kragten-Tabatabaie, Stephen Lindstrom, Inne Nauwelaers, D. James Nokes, Peter J. Openshaw, Teresa C. Peret, Pedro A. Piedra, Kaat Ramaekers, Annabel Rector, Nídia Sequeira Trovão, Anne von Gottberg, Maria Zambon, Wengqing Zhang, Thomas C. Williams, Ian G. Barr[^*], Ursula J. Buchholz[^*]

Affiliations:

Tehran University of Medical Sciences, Tehran, Iran (V. Salimi)

22 Hospital de Niños Dr. Ricardo Gutiérrez, Buenos Aires, Argentina (M. Viegas)

23 National Center of Biotechnology, Madrid, Spain (A. Trento)

24 KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya (C. N. Agoti, D. J. Nokes)

25 Emory University, Atlanta, Georgia, USA (L. J. Anderson)

26 Baylor College of Medicine, Houston, Texas, USA (V. Avadhanula, D. M. Henke, P. A. Piedra)

27 University of Georgia, Athens, Georgia, USA (J. Bahl)

28 University Medical Center, Utrecht, The Netherlands (L. Bont)

29 National Center for Biotechnology Information, NIH, Bethesda, Maryland, USA (J. R. Brister, E.

30 L. Hatcher)

31 Public Health England, London, UK (P. A. Cane, M. Zambon)

32 Public Health England, London, UK, and Francis Crick Institute, London (M. Galiano)

33 National Institute of Allergy and Infectious Diseases, NIH, Bethesda (B. S. Graham, U. J.

34 Buchholz)

35 National Institute for Communicable Diseases of the National Health Laboratory Service,

36 Johannesburg, South Africa (O. Hellferscee, A. von Gottberg)

37 World Health Organization, Geneva, Switzerland (S. Hirve, S. Jackson, W. Zhang)

38 University Hospitals Leuven, Leuven, Belgium (E. Keyaerts)

39 ReSViNET Foundation, Zeist, The Netherlands and Julius Clinical, Zeist (L. Kragten-

40 Tabatabaie)

Peer Review

41 Centers for Disease Control and Prevention, Atlanta (S. Lindstrom, T. Peret)

42 Public Health England, London, UK and Imperial College London (I. Nauwelaers)

43 UK and Imperial College London (P. Openshaw)

44 Rega Institute, Leuven, Belgium (K. Ramaekers, A. Rector)

45 Fogarty International Center, NIH, Bethesda (N. Sequeira Trovão)

46 University of Edinburgh, Edinburgh, UK (T. C. Williams)

47 Peter Doherty Institute for Infection & Immunity, Melbourne, Australia (I. Barr)

48 [#]These authors contributed equally to this article

49 [^] These last authors contributed equally to this article.

50 * to whom correspondence should be addressed.

51 Ian Barr, Ian.Barr@influenzacentre.org; (613) 9342-9311

52 Ursula Buchholz, ubuchholz@niaid.nih.gov; (301) 594-1533

Peer Review

Abstract - Word Count: 149

Human respiratory syncytial virus (HRSV) is the leading viral cause of serious pediatric respiratory disease, and reinfections are common throughout life. HRSV has two major subgroups, A and B, which exhibit some antigenic variability, contributing to the ability of HRSV to circulate annually. The molecular epidemiology of HRSV is being studied worldwide, and increasing numbers of partial and full genomic sequences are becoming available. To facilitate molecular epidemiology analyses, we propose a uniform nomenclature for HRSV-positive samples, isolates and sequences as follows: HRSV/subgroup identifier/geographic identifier (ISO3166-1 alpha-3 letter country designation or a simple country name)/unique sequence identifier/calendar year of sampling. We also propose a basic template for submission of associated metadata. A universal nomenclature for HRSV isolates would facilitate data retrieval, analysis of sequence data, and would improve publications of figures such as sequence alignments and phylograms, which will overall aid our understanding of the evolution of this virus.

Peer Review

ScholarOne support: (434) 964-4100

68 **Text - Word Count: 2737**

69 **Introduction**

70 Human respiratory syncytial virus (HRSV) is the leading single cause of severe respiratory
71 illness in children less than 5 years of age, and is associated with substantial morbidity related to
72 lower respiratory tract infections in industrialized countries, and significant morbidity and
73 mortality in low- and/or middle-income countries (*1-5*). HRSV also causes severe disease in the
74 elderly and in high-risk adults (*6*). In 2016, HRSV (species *Human orthopneumovirus*) was
75 reclassified by the International Committee on Virus Taxonomy (ICTV) to the new family
76 *Pneumoviridae*, genus *Orthopneumovirus* (*7*). With the increasing availability of viral
77 sequencing technologies, the number of HRSV sequence submissions to databases is increasing
78 (Figure 1), and it is anticipated that this trend will continue. While the ICTV provides
79 nomenclature standards for virus taxa, no system currently exists for HRSV nomenclature below
80 the species level. Given the current interest in HRSV and in database submissions, a standard
81 nomenclature for HRSV strains and isolates below the species level is needed and would
82 facilitate sample handling and studies of the genomic diversity of HRSV strains and variants.
83 Following the taxonomical reassignment, it is timely to (i) propose a consistent and universal
84 naming convention for HRSV strains, sequences and isolates, and (ii) provide a framework for
85 the submission of associated basic metadata, facilitating database submissions that are rich in
86 context information.

87 Currently, several large laboratory HRSV surveillance and epidemiology studies are
88 ongoing. These include the World Health Organization's (WHO) Global Respiratory Syncytial
89 Virus (RSV) Surveillance Project, with large-scale testing for HRSV and extensive sequencing

ScholarOne support: (434) 964-4100

of HRSV-positive clinical specimens from more than 20 countries worldwide (https://www.who.int/influenza/rsv/en/). Focused molecular analyses have helped elucidate HRSV household transmission (*8*) as well as local transmission dynamics (*9*), and may guide the control of HRSV transmission. For example, molecular analysis showed that acquisition of HRSV in healthcare facilities occurs through both transmission within facilities and introduction from the community (*10, 11*). HRSV subgroups A and B usually dominate in alternate years but can also co-circulate in the same season, but it remains to be seen how future subgroup circulation patterns will be affected by the current COVID-19 pandemic. Early work showed that HRSV of subgroup A was associated with slightly greater clinical severity than subgroup B (*15*). HRSV disease severity has also been correlated with specific HRSV strains, genotypes or clades, but to date, there has been no consistent association between specific strains and virulence across studies (*12-14*), and no genotype or clade has been consistently associated with severe disease (*16-19*). Thus, a possible role of HRSV strain differences in disease severity remains to be elucidated. The lack of a standard nomenclature and the scarcity of rich metadata in databases currently limit and complicate such studies. In this context, a simple and standardized nomenclature would minimize data entry errors for handling, storing and shipping of HRSV-containing samples. Moreover, consistency in nomenclature will improve the ability of researchers to pool and analyze sequence data and associated information from a variety of different sources.

Reliable and concise nomenclature systems below the species level are available for measles virus, influenza virus, rotavirus, and filovirus isolates (*20-23*), and for many other human viral pathogens. A similar nomenclature system for HRSV is urgently needed and will provide a standard interface for submission, searching and display of sequencing data and

ScholarOne support: (434) 964-4100

113 associated metadata (i.e. an interface tailored to HRSV and its pathology), supporting the
114 requirements of HRSV research and the public health community. To fill this need, an
115 international group of researchers, in conjunction with the WHO Global RSV Surveillance
116 Project, proposes a concise HRSV nomenclature system.

117 **HRSV Genome Organization.** HRSV has a single-stranded non-segmented negative-
118 sense RNA genome of about 15,191 to 15,277 nucleotides in length (Figure 2A) (*7*). The HRSV
119 genome contains 10 genes, each encoding a separate mRNA with a single open reading frame
120 (ORF) (Figure 2A, Table 1), except for the M2 mRNA which contains two overlapping open
121 reading frames. The 11 HRSV proteins include two nonstructural proteins NS1 and NS2, the
122 nucleoprotein N, phosphoprotein P, matrix protein M, small hydrophobic envelope protein SH,
123 attachment glycoprotein G, fusion glycoprotein F, matrix protein M2-1 and RNA regulatory
124 factor M2-2, and the RNA polymerase protein L (Table 1) (*24, 25*). The fusion glycoprotein F is
125 the major viral neutralization and protective antigen, followed by the attachment glycoprotein G
126 (26).

127 **HRSV Subgroups and Genotype Designations: Status and Outlook**. HRSV has two
128 major genetic and antigenic subgroups, A and B, with genome-wide nucleotide and amino acid
129 divergence between these subgroups (Figure 2A) (*25, 27*). The reference sequences for the two
130 subgroups are derived from strains HRSV A2 [(*28, 29*); GenBank Accession number M74568.1,
131 RefSeq accession number NC_038235, curated by the National Center for Biotechnology
132 Information (NCBI), NIH (*30, 31)*], and HRSV B1 [(*32*), GenBank Accession number
133 AF013254.1, RefSeq accession number NC_001781.1] (Figure 2B). Between the two subgroups,
134 the F glycoprotein sequences are well conserved, with 89% amino acid identity, while the G
135 glycoproteins are the most divergent amongst the HRSV proteins (only 53% amino acid identity

ScholarOne support: (434) 964-4100

136 between the subgroups) (Figure 2A) and undergo continuous molecular evolution. The
137 ectodomain of the G glycoproteins of both subgroups contains a conserved central domain,
138 representing an important antigenic site, flanked by two hypervariable domains (*33*). Except for
139 the central conserved region, the antigenic cross reactivity between G glycoproteins of the two
140 subgroups is low (*26*).

141 Since the G ORF exhibits the greatest degree of genetic variability between isolates, it is
142 most commonly used for studies on the molecular evolution of HRSV. The genetic variability of
143 HRSV strains over time has been commonly determined by sequencing of the distal C-terminal
144 third of the G ORF, which includes the second hypervariable domain. The variability in the G
145 ORF is characterized by a high rate of non-synonymous nucleotide changes, suggesting that
146 evolution may be driven by immune pressure, even though this may be partially antibody
147 independent (*34*). It is likely that this variability in the G protein contributes to the capacity of
148 HRSV to cause yearly outbreaks in the community (*35-37*). In a parallel effort, several research
149 groups are working together on a genotyping proposal, providing a consensus on uniform
150 genotype designations (*38, 39*). With continuing virus evolution, the emergence of new
151 genotypes is expected, along with the extinction of older genotypes. HRSV genotyping
152 designations will capture the present status and will be adaptable to reflect future molecular
153 evolution, and will need to be re-evaluated periodically by a global consortium. While the
154 nomenclature proposal outlined herein does not extend to genotype designations, it should
155 facilitate the sequence analysis required to follow the molecular evolution of HRSV.

156 **Nomenclature Proposal for HRSV Strains and Isolates.** For molecular epidemiology
157 studies, it would be useful to have a concise standard for short identifiers of specific HRSV
158 sequences, suitable for the short "Definition" lines that give context to a sample and its derived

ScholarOne support: (434) 964-4100

159 sequence. Ideally, concise standardized names should convey key information about each
160 individual sequence in an alignment or phylogram such as the source, date and type (if known).
161 Here, we aim to define such a common naming convention for HRSV samples and isolates. We
162 also propose the use of standard names and appropriate annotations for HRSV genes, provide
163 examples to guide the annotation of sequence data during the sequence submission process, and
164 suggest the submission of metadata associated with the source materials of HRSV sequences. As
165 indicated above, this nomenclature proposal does not extend to genotype designations, but
166 proposes that genotype information be submitted with metadata to sequence and other relevant
167 databases.

168 **Definition Lines for HRSV Sequence Submissions**. GenBank records available through
169 NCBI are identified by two elements, a unique alphanumeric accession number and a definition
170 line. The definition line is the "title" that one commonly associates with GenBank records shown
171 in BLAST results and other searches. Definition lines are generated by the submitter during the
172 sequence submission process and include the species and isolate name; for example "*Human*
173 *orthopneumovirus* isolate HRSV/A/USA/001/2011, complete genome." (*40*). For each HRSV
174 positive clinical sample or isolate, and for the submission of HRSV sequences to sequence
175 databases, we propose to capture five sample-specific parameters in a standardized fashion,
176 compatible with these isolate name requirements. Specifically, we propose the following
177 convention for the naming of HRSV samples, which will be included in sequence definition lines
178 (Figure 3): <virus name abbreviation>/<HRSV subgroup>/<geographic identifier>/<unique
179 sequence identifier>/<year of sampling>

180 I. Organism name; virus name abbreviation: HRSV

ScholarOne support: (434) 964-4100

181 a. The ICTV proposed species name is "*Human orthopneumovirus*" (7), which will
182 be seen in the NCBI definition line. During submission, the organism name can
183 be provided as either "Human orthopneumovirus" or "Human respiratory
184 syncytial virus". The abbreviation "HRSV" should be included as part of the
185 isolate name regardless of which organism name is provided.
186 II. HRSV subgroup: A or B; X if unknown
187 III. Geographic Identifier of the location of sampling. As individual HRSV research networks
188 have pre-defined requirements, it is suggested that this field have some flexibility.
189 Specifically,
190 i. If not specified by a research network, the ISO 3166-1 alpha-3 letter designation
191 should be used to indicate the country of sampling (XXX if the country of sampling is
192 unknown). The ISO 3166-1 alpha-3 can be accessed through this link:
193 https://en.wikipedia.org/wiki/ISO_3166-1_alpha-3. It is strongly suggested to provide
194 any additional geographic information on the sampling location (e.g. city or state of
195 sampling) in metadata fields, rather than within the concise definition line.
196 ii. the WHO Global RSV Surveillance Project plans to use the simple name of the
197 country (in English).
198 iii. Individual national studies may require a state/province/city designation, as well as
199 the country name (which is mandatory). If state/province/city designation is required,
200 a period should be used to separate the country name ([state/province/city].[country
201 name])
202 IV. Unique isolate identifier: This field must be restricted to 8 alphanumeric characters.
203 Underscores are permitted, but hyphens, slashes, or any other types of special characters

Peer Review

204 (e.g., %, $, @, etc.) or spaces cannot be used. Controversial names/phrases, names of
205 prominent people, and trademarked names/phrases as unique identifiers cannot be used.
206 To prevent duplications of sample/isolate numbers by different groups, we recommend
207 inserting an acronym identifying a study or institute prior to the sample number. For
208 example, unique isolate identifiers for samples from the INFORM-RSV study might
209 include the acronym "INF" followed by a number
210 ("HRSV/A/COUNTRY/INF001/2019", unique isolate identifier underlined).
211 V. Year of sampling; (YYYY) or XXXX if unknown.

212 The proposed nomenclature options (HRSV/subgroup/geographic identifier/unique
213 identifier/year of collection) are demonstrated in the following examples:

214 - HRSV/A/USA/001/2011
215 - HRSV/B/Denver.USA/14617/1985
216 - HRSV/A/IRN/001/2017
217 - HRSV/A/Iran/001/2017
218 - HRSV/X/IRN/001/2017 (subgroup unknown)
219 - HRSV/B/New_Zealand/FR123/2020

220 This nomenclature proposal prioritizes a short and concise definition line that will be easy
221 to handle in the laboratory, be easily readable and consistent in public databases. Additional
222 information may be submitted in metadata fields. This would allow researchers, epidemiologists
223 and database users to apply specific metadata filters, as needed for data retrieval and specific
224 applications, analyses, or for displaying designations, such as in dendrograms.

Peer Review

ScholarOne support: (434) 964-4100

225 **Terminology for Annotations.** To support efficient data analysis, it is essential that
226 uniform designations be used at the stage of database submission. Table 1 shows commonly
227 accepted names for HRSV genes and proteins. An HRSV gene comprises a gene start signal
228 GGGGCAAAT(A/G), an open reading frame (ORF) with adjacent noncoding regions if present,
229 and the gene end signal through the last adenosine residue
230 [AGT(T/A)A(T/A/G)(A/T)(A/T)(A/T)A_n] [Figure 4, (*41*)]. Each HRSV gene contains a single
231 ORF, except for the M2 gene, which comprises two ORFs M2-1 and M2-2. Nucleotide
232 annotations of genes and ORFs for the HRSV A2 (Figure 4) and HRSV B1 reference sequences
233 are shown in Table 1.

234 **Metadata to Accompany HRSV Samples and Original Sequence Submissions.** The
235 most pertinent host data will depend on the interests of individual study groups, with conditions
236 such as prematurity being of interest in a pediatric setting. Groups studying HRSV in an adult
237 setting may have an interest in different metadata, such as if the individual was
238 immunocompromised. Suggested further information that may be included in metadata fields:

239 1. Isolation_source: sample type (upper or lower airways);
240 viral RNA can be extracted directly from a clinical sample, or from an isolate grown in cell
241 culture, or possibly from a cDNA derived recombinant virus. The sources of sequences from
242 isolated viruses can be identified by the following suffices:
243 a. wt: wild-type; sequences derived from RNA extracted directly from clinical specimens
244 b. tc: tissue culture; sequences derived from RNA extracted from HRSV isolates
245 propagated in tissue culture (a single round qualifies as "tc")
246 c. rec: recombinant; sequences of cDNA-derived recombinant virus (including vaccine
247 strain)

ScholarOne support: (434) 964-4100

2. Host: *Homo sapiens*; subject age. Indicate years and months if under 5 years of age; years of age if above 5 years of age. Subject sex (if known) should be spelled out.
3. Country, state, and (nearest) city of sampling. Metadata information must include the country (full name, not 3 letter abbreviation; for a list of accepted country designations, see https://www.ncbi.nlm.nih.gov/GenBank/collab/country/), can also include city or state/province. Names should be written based on the standard ASCII letters including spaces if required. City or coordinates of the location of sampling should be included if known.
4. Collection date (at least month and year). It is highly recommended that the exact date of specimen collection (DD-Mon-YYYY; example: 17-Feb-2002) be used; if not available the year and month should be indicated (Mon-YYYY).
5. Genotype according to the consensus in genotype classification by the RSV GeNom Working Group (consensus definition in progress).
6. Metadata on the host and the clinical disease should be included in the "Notes" field in a structured format. Protected health information will be excluded from metadata submissions.
    a. Host:
        I. If under 6 months of age, birthweight and gestational age at birth.
        II. Significant pediatric co-morbidities, including prematurity, congenital cardiac disease and broncho-pulmonary dysplasia (BPD).
        III. Twins (y/n)
        IV. Exposed to specific HRSV therapeutic, vaccine, antibody or antiviral (y/n)
        V. Viral or bacterial co-infections if known; pathogen species should be spelled out.

Peer Review

ScholarOne support: (434) 964-4100

VI. Adult co-morbidities such as chronic obstructive pulmonary disease (COPD) or asthma, or altered immune status (e.g. immunocompromised or a bone marrow transplant recipient).

b. Disease outcomes: Five grades are distinguished (no medical care; out-patient or emergency room; hospital admission; ICU admission; death)

For NCBI submissions, data can be entered through the web interface, or uploaded as tab-delimited text files (examples provided in Supplementary Information). Sequences can be uploaded in FASTA format (Supplementary file 1), with associated metadata provided in a plain text tab-delimited "source modifier" table (virtual examples provided in Supplementary file 2), and gene/protein annotations provided in a plain text tab-delimited "5 column feature" table (Supplementary file 3).

Peer Review

Outlook

Molecular surveillance has revealed that multiple HRSV genotypes circulate simultaneously in a community. Circulating genotypes often vary between communities, and circulation patterns can change from year to year. Extended monitoring of circulating viruses is necessary to better understand transmission and molecular evolution (*42*). As HRSV vaccine candidates and antivirals are being developed, molecular epidemiology studies may reveal potential effects of prevention strategies on viral evolution and possible antibody-escape variants. Timely HRSV data sharing worldwide through the use of public databases is essential. We propose that sequence data be uploaded to a publicly accessible database such as the NCBI [https://www.ncbi.nlm.nih.gov; (*31*)][1]. Whilst the most complete repository for HRSV sequence

ScholarOne support: (434) 964-4100

information is NCBI, studies may require that sequences be submitted in the first instance to other databases such as GISAID (www.gisaid.org).

Public access will provide useful for investigators to submit, query, and analyze HRSV sequence data, facilitating the evolutionary analysis of sequence diversity within or between HRSV genotypes. This has been clearly demonstrated with the emergence of SARS-CoV-2 and the critical role that genetic sequence analysis has provided. Notably there has been an immediate adoption of a nomenclature similar to what has been proposed here for HRSV, although some differences remain between databases (e.g. NCBI: SARS-CoV-2/human/USA/COVID20-1096/2020; GISAID: hCoV-19/Australia/VIC12/2020). When an HRSV vaccine becomes available, it will be important to monitor and trace possible evolutionary changes in response to vaccine-induced selective pressure (*43, 44*); this will require high-quality and geographically representative country-specific data on circulating strains, and rich datasets of well-curated, standardized, and parsable data.

This proposal will profit from strong support by members of the International RSV Society, a special interest group of the International Society for Influenza and other Respiratory Virus Diseases (isirv), members of the WHO Global RSV Surveillance Project, and the HRSV research community.

Dedication

This manuscript is in memory of Jose A. Melero, a superb scientist and leader in the HRSV field, and a wonderful and generous friend to us all.

Acknowledgments

ScholarOne support: (434) 964-4100

Support for this work was provided by the Intramural Research Programs of the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health. Maria Zambon was supported by Public Health England (PHE), and Inne Neuwelaers was supporte by PHE and by the National Institute of Health Research (NIHR) Respiratory Health Protection Unit. The work of J. Rodney Brister and Eneida L. Hatcher was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health. TC Williams is the recipient of a Wellcome Trust Award [204802/Z/16/Z]. The WHO's Global RSV Surveillance is supported by the Bill and Melinda Gates Foundation grant award no. OPP1127419.

Peer Review

Disclaimer

Siddhivinayak Hirve, Sandra Jackson and Wenqing Zhang work with the World Health Organization. The authors are responsible for the views expressed in this publication and they do not necessarily represent the decisions, policy or views of the World Health Organization. Names of specific vendors, manufacturers, or products are included for public health and informational purposes; inclusion does not imply endorsement of the vendors, manufacturers, or products by the World Health Organization. Larry Anderson has done paid consultancies on RSV vaccines for Bavarian Nordic, Novavax, ClearPath Vaccines Company, and Pfizer. Larry Anderson's laboratory is currently receiving funding through Emory University from Pfizer and Advac for laboratory studies for HRSV surveillance studies in adults, and he holds a subcontract on an NIH SBIR award to Sciogen on G protein HRSV vaccines. Larry Anderson is a co-inventor on several CDC patents on the HRSV G protein and its CX3C chemokine motif relative to immune therapy and vaccine development, and on a patent filing for use of HRSV platform VLPs with the F and G proteins for vaccines. Ursula Buchholz reports CRADA support to NIH

336 from Sanofi Pasteur, outside the submitted work; in addition, Ursula Buchholz has patents on
337 live-attenuated HRSV with royalties paid to NIH by Sanofi Pasteur. There are no additional
338 conflicts of interest by any of the authors.

339 **Author Bio**

340 Dr. Vahid Salimi is a virologist, associate professor and head of the virology department at the
341 School of Public Health, Tehran University of Medical Sciences. His primary research interests
342 include genetic characterization, immunopathogenesis and vaccine design of respiratory viruses.

343

344 **References:**

345 1. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, et al. Global,
346 regional, and national disease burden estimates of acute lower respiratory infections due to
347 respiratory syncytial virus in young children in 2015: a systematic review and modelling
348 study. Lancet. 2017 Sep 2;390(10098):946-58.
349 2. Nair H, Nokes DJ, Gessner BD, Dherani M, Madhi SA, Singleton RJ, et al. Global burden
350 of acute lower respiratory infections due to respiratory syncytial virus in young children: a
351 systematic review and meta-analysis. Lancet. 2010 May 1;375(9725):1545-55.
352 3. Salimi V, Tavakoli-Yaraki M, Yavarian J, Bont L, Mokhtari-Azad T. Prevalence of human
353 respiratory syncytial virus circulating in Iran. J Infect Public Health. 2016 Mar-
354 Apr;9(2):125-35.
355 4. Stein RT, Bont LJ, Zar H, Polack FP, Park C, Claxton A, et al. Respiratory syncytial virus
356 hospitalization and mortality: Systematic review and meta-analysis. Pediatr Pulmonol. 2017
357 Apr;52(4):556-69.

ScholarOne support: (434) 964-4100

Peer Review

358 5. Pneumonia Etiology Research for Child Health Study G. Causes of severe pneumonia
359 requiring hospital admission in children without HIV infection from Africa and Asia: the
360 PERCH multi-country case-control study. Lancet. 2019 Jun 27.
361 6. Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. Respiratory syncytial virus
362 infection in elderly and high-risk adults. N Engl J Med. 2005 Apr 28;352(17):1749-59.
363 7. Rima B, Collins P, Easton A, Fouchier R, Kurath G, Lamb RA, et al. ICTV Virus
364 Taxonomy Profile: Pneumoviridae. J Gen Virol. 2017 Dec;98(12):2912-3.
365 8. Agoti CN, Phan MVT, Munywoki PK, Githinji G, Medley GF, Cane PA, et al. Genomic
366 analysis of respiratory syncytial virus infections in households and utility in inferring who
367 infects the infant. Sci Rep. 2019 Jul 11;9(1):10076.
368 9. Trovao NS, Khuri-Bulos N, Tan Y, Puri V, Shilts MH, Halpin RA, et al. Molecular
369 characterization of respiratory syncytial viruses circulating in a paediatric cohort in Amman,
370 Jordan. Microb Genom. 2019 Sep 18.
371 10. Geis S, Prifert C, Weissbrich B, Lehners N, Egerer G, Eisenbach C, et al. Molecular
372 characterization of a respiratory syncytial virus outbreak in a hematology unit in Heidelberg,
373 Germany. J Clin Microbiol. 2013 Jan;51(1):155-62.
374 11. Nabeya D, Kinjo T, Parrott GL, Uehara A, Motooka D, Nakamura S, et al. The clinical and
375 phylogenetic investigation for a nosocomial outbreak of respiratory syncytial virus infection
376 in an adult hemato-oncology unit. J Med Virol. 2017 Aug;89(8):1364-72.
377 12. Malekshahi SS, Razaghipour S, Samieipoor Y, Hashemi FB, Manesh AAR, Izadi A, et al.
378 Molecular characterization of the glycoprotein and fusion protein in human respiratory
379 syncytial virus subgroup A: Emergence of ON-1 genotype in Iran. Infect Genet Evol. 2019
380 Jul;71:166-78.

Peer Review

ScholarOne support: (434) 964-4100

381 13. Vos LM, Oosterheert JJ, Kuil SD, Viveen M, Bont LJ, Hoepelman AIM, et al. High
382 epidemic burden of RSV disease coinciding with genetic alterations causing amino acid
383 substitutions in the RSV G-protein during the 2016/2017 season in The Netherlands. J Clin
384 Virol. 2019 Mar;112:20-6.
385 14. Pierangeli A, Viscido A, Bitossi C, Frasca F, Gentile M, Oliveto G, et al. Differential
386 interferon gene expression in bronchiolitis caused by respiratory syncytial virus-A genotype
387 ON1. Medical microbiology and immunology. 2020 Feb;209(1):23-8.
388 15. Hall CB, Walsh EE, Schnabel KC, Long CE, McConnochie KM, Hildreth SW, et al.
389 Occurrence of groups A and B of respiratory syncytial virus over 15 years: associated
390 epidemiologic and clinical characteristics in hospitalized and ambulatory children. J Infect
391 Dis. 1990 Dec;162(6):1283-90.
392 16. Vandini S, Biagi C, Lanari M. Respiratory Syncytial Virus: The Influence of Serotype and
393 Genotype Variability on Clinical Course of Infection. Int J Mol Sci. 2017 Aug 6;18(8).
394 17. Laham FR, Mansbach JM, Piedra PA, Hasegawa K, Sullivan AF, Espinola JA, et al.
395 Clinical Profiles of Respiratory Syncytial Virus Subtypes A AND B Among Children
396 Hospitalized with Bronchiolitis. Pediatr Infect Dis J. 2017 Aug;36(8):808-10.
397 18. Gilca R, De Serres G, Tremblay M, Vachon ML, Leblanc E, Bergeron MG, et al.
398 Distribution and clinical impact of human respiratory syncytial virus genotypes in
399 hospitalized children over 2 winter seasons. J Infect Dis. 2006 Jan 1;193(1):54-8.
400 19. Martinello RA, Chen MD, Weibel C, Kahn JS. Correlation between respiratory syncytial
401 virus genotype and severity of illness. J Infect Dis. 2002 Sep 15;186(6):839-42.
402 20. Measles virus nomenclature update: 2012. Wkly Epidemiol Rec. 2012 Mar 2;87(9):73-81.

Peer Review

ScholarOne support: (434) 964-4100

403 21. Kuhn JH, Bao Y, Bavari S, Becker S, Bradfute S, Brister JR, et al. Virus nomenclature
404 below the species level: a standardized nomenclature for laboratory animal-adapted strains
405 and variants of viruses assigned to the family Filoviridae. Arch Virol. 2013
406 Jun;158(6):1425-32.
407 22. Matthijnssens J, Ciarlet M, McDonald SM, Attoui H, Banyai K, Brister JR, et al. Uniformity
408 of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group
409 (RCWG). Arch Virol. 2011 Aug;156(8):1397-413.
410 23. Reconsideration of influenza A virus nomenclature: a WHO memorandum. Bull World
411 Health Organ. 1979;57(2):227-33.
412 24. Collins PL, Fearns R, Graham BS. Respiratory syncytial virus: virology, reverse genetics,
413 and pathogenesis of disease. Curr Top Microbiol Immunol. 2013;372:3-38.
414 25. Collins PL, Karron RA. Respiratory syncytial virus and metapneumovirus. In: Knipe DM,
415 Howley PM, Cohen JI, Griffin DE, Lamb RA, Martin MA, et al., editors. Fields Virology.
416 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2013. p. 1086-123.
417 26. Collins PL, Melero JA. Progress in understanding and controlling respiratory syncytial
418 virus: still crazy after all these years. Virus Res. 2011 Dec;162(1-2):80-99.
419 27. Mufson MA, Orvell C, Rafnar B, Norrby E. Two distinct subtypes of human respiratory
420 syncytial virus. J Gen Virol. 1985 Oct;66 (Pt 10):2111-24.
421 28. Collins PL, Olmsted RA, Spriggs MK, Johnson PR, Buckler-White AJ. Gene overlap and
422 site-specific attenuation of transcription of the viral polymerase L gene of human respiratory
423 syncytial virus. Proc Natl Acad Sci U S A. 1987 Aug;84(15):5134-8.
424 29. Mink MA, Stec DS, Collins PL. Nucleotide sequences of the 3' leader and 5' trailer regions
425 of human respiratory syncytial virus genomic RNA. Virology. 1991 Dec;185(2):615-24.

Peer Review

ScholarOne support: (434) 964-4100

426 30. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference
427 sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional
428 annotation. Nucleic Acids Res. 2016 Jan 4;44(D1):D733-45.
429 31. Sayers EW, Beck J, Brister JR, Bolton EE, Canese K, Comeau DC, et al. Database resources
430 of the National Center for Biotechnology Information. Nucleic Acids Res. 2020 Jan
431 8;48(D1):D9-D16.
432 32. Karron RA, Buonagurio DA, Georgiu AF, Whitehead SS, Adamus JE, Clements-Mann ML,
433 et al. Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral
434 replication in vitro: clinical evaluation and molecular characterization of a cold-passaged,
435 attenuated RSV subgroup B mutant. Proc Natl Acad Sci U S A. 1997 Dec 9;94(25):13961-6.
436 33. Martinez I, Dopazo J, Melero JA. Antigenic structure of the human respiratory syncytial
437 virus G glycoprotein and relevance of hypermutation events for the generation of antigenic
438 variants. J Gen Virol. 1997 Oct;78 (Pt 10):2419-29.
439 34. Trento A, Abrego L, Rodriguez-Fernandez R, Gonzalez-Sanchez MI, Gonzalez-Martinez F,
440 Delfraro A, et al. Conservation of G-Protein Epitopes in Respiratory Syncytial Virus (Group
441 A) Despite Broad Genetic Diversity: Is Antibody Selection Involved in Virus Evolution? J
442 Virol. 2015 Aug;89(15):7776-85.
443 35. Melero JA, Garcia-Barreno B, Martinez I, Pringle CR, Cane PA. Antigenic structure,
444 evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. J
445 Gen Virol. 1997 Oct;78 (Pt 10):2411-8.
446 36. Sullender WM, Mufson MA, Anderson LJ, Wertz GW. Genetic diversity of the attachment
447 protein of subgroup B respiratory syncytial viruses. J Virol. 1991 Oct;65(10):5425-34.

Peer Review

ScholarOne support: (434) 964-4100

448 37. Cane PA, Matthews DA, Pringle CR. Analysis of respiratory syncytial virus strain variation
449 in successive epidemics in one city. J Clin Microbiol. 1994 Jan;32(1):1-4.
450 38. Goya S, Galiano M, Nauwelaers I, Trento A, Openshaw PJ, Mistchenko AS, et al. Toward
451 unified molecular surveillance of RSV: A proposal for genotype definition. Influenza Other
452 Respir Viruses. 2020 Feb 5.
453 39. Ramaekers K, Rector A, Cuypers L, Lemey P, Keyaerts E, Van Ranst M. Towards a Unified
454 Classification for Human Respiratory Syncytial Virus Genotypes. Virus Evolution, accepted
455 for publication. 2020 Jul 24,veaa052.
456 40. The GenBank Submissions Handbook. 2011 ed. Bethesda (MD): National Center for
457 Biotechnology Informationn (US).
458 41. Collins PL, Dickens LE, Buckler-White A, Olmsted RA, Spriggs MK, Camargo E, et al.
459 Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal
460 distinctive features of intergenic structure and gene order. Proc Natl Acad Sci U S A. 1986
461 Jul;83(13):4594-8.
462 42. Trento A, Casas I, Calderon A, Garcia-Garcia ML, Calvo C, Perez-Brena P, et al. Ten years
463 of global evolution of the human respiratory syncytial virus BA genotype with a 60-
464 nucleotide duplication in the G protein gene. J Virol. 2010 Aug;84(15):7500-12.
465 43. Graham BS. Vaccine development for respiratory syncytial virus. Curr Opin Virol. 2017
466 Apr;23:107-12.
467 44. Mazur NI, Higgins D, Nunes MC, Melero JA, Langedijk AC, Horsley N, et al. The
468 respiratory syncytial virus vaccine landscape: lessons from the graveyard and promising
469 candidates. Lancet Infect Dis. 2018 Oct;18(10):e295-e311.

Peer Review

ScholarOne support: (434) 964-4100

470 45. Collins PL, Hill MG, Camargo E, Grosfeld H, Chanock RM, Murphy BR. Production of
471 infectious human respiratory syncytial virus from cloned cDNA confirms an essential role
472 for the transcription elongation factor from the 5' proximal open reading frame of the M2
473 mRNA in gene expression and provides a capability for vaccine development. Proc Natl
474 Acad Sci U S A. 1995 Dec 5;92(25):11563-7.
475 46. Crowe JE, Jr., Bui PT, Firestone CY, Connors M, Elkins WR, Chanock RM, et al. Live
476 subgroup B respiratory syncytial virus vaccines that are attenuated, genetically stable, and
477 immunogenic in rodents and nonhuman primates. J Infect Dis. 1996 Apr;173(4):829-39.

478

479 Address for correspondence:
480 Ian Barr, Peter Doherty Institute for Infection & Immunity, Melbourne, Australia; email:
481 Ian.Barr@influenzacentre.org; (613) 9342-9311

482 Ursula Buchholz, NIAID, NIH, 50 South Dr, 50/6503, Bethesda, MD 20892-8007, USA; email:
483 ubuchholz@niaid.nih.gov; (301) 594-1533

Peer Review

ScholarOne support: (434) 964-4100

484 **Table 1: Widely Accepted Nomenclature for HRSV Genes and Proteins; Gene Annotation**

485 **of Strains HRSV A2 and HRSV B1**

Gene or Genome Region			Protein		Open Reading Frame	
Annotation [a]	Genome Position (nt) [a]		Annotation	Abbre-viation	Genome Position (nt)	
	Strain A2[b]	Strain B1[c]			Strain A2	Strain B1
Leader region	1-44	1-44				
NS1	45-576	45-577	Nonstructural protein 1	NS1	99-518	99-518
NS2	596-1098	594-1098	Nonstructural protein 2	NS2	628-1002	626-1000
N	1126-2328	1125-2327	Nucleoprotein	N	1141-2316	1140-2315
P	2330-3243	2331-3244	Phosphoprotein	P	2347-3072	2348-3073
M	3253-4210	3254-4208	Matrix protein	M	3262-4032	3263-4033
SH	4220-4629	4218-4630	Small hydrophobic protein	SH	4304-4498	4303-4500
G	4674-5596	4675-5600	Attachment glycoprotein	G	4689-5585	4690-5589
F	5649-7551	5653-7552	Fusion glycoprotein	F	5662-7386	7666-7390
			Matrix protein 2			
M2	7598-8558	7609-8568	Matrix protein M2-1	M2-1	7607-8191	7618-8205
			Matrix protein M2-2	M2-2	8160-8432	8171-8443

Peer Review

ScholarOne support: (434) 964-4100

L	8491-15068	8501-15080	Polymerase protein	L	8499-14996	8509-15009
Trailer region	15069-15223	15081-15225				

[a] Nucleotide annotations for leader and trailer region, and for indicated HRSV genes from the first nucleotide (nt) of the gene start signal [GGGGCAAAT(A/G); GGAGCAAAAT in case of the L gene] through the last adenosine residue of the gene end signal [AGT(T/A)A(T/A/G)(A/T)(A/T)(A/T)A_n]

[b] HRSV/A/USA/A2/2015 (*45*); GenBank Accession Number KT992094

[c] HRSV/B/USA/B1/1985/B1 (*46*); GenBank Accession Number AF013254.1

Peer Review

ScholarOne support: (434) 964-4100

491 **Figure Legends**

492 **Figure 1.** Annual numbers of HRSV whole-genome sequences released in Genbank since
493 publication of the whole-genome sequence of HRSV A2 (M74568, released in 1993).

494 **Figure 2. (A)** Schematic overview of the HRSV gene order and comparison of the amino
495 acid identities of the reference strains of subgroup A (HRSV A2: GenBank Accession number
496 M74568/NC_038235, and HRSV B1: GenBank Accession number AF013254/NC_001781). **(B)**
497 International Committee on Virus Taxonomy (ICTV) species designation, virus name, and
498 associated GenBank reference sequences, curated by the National Center for Biotechnology
499 Information (NCBI), NIH.

500 **Figure 3.** Schematic representation of the five consensus nomenclature elements of HRSV
501 strains and isolates, with examples (top) and a legend for each element (bottom).

502 **Figure 4.** Schematic aid to gene annotations of HRSV whole-genome sequences. An
503 HRSV gene comprises the sequence from the first nucleotide of the conserved HRSV gene start
504 signal (GGGGCAAATa) to the last adenosine residue of the HRSV gene end signal
505 (AGTTAnnnnAAAA) (25, 41). Gene start signals are represented by black triangles, and gene
506 end signals are shown as black rectangles, separated by intergenic regions (underlined). Note the
507 M2/L gene overlap (annotations derived from HRSV A2; GenBank Accession number
508 M74568/NC_038235).

509

Peer Review

ScholarOne support: (434) 964-4100

Article Summary Line: This proposal of a human respiratory syncytial virus (HRSV) nomenclature below the species level is brought forward by a group of members of the International RSV Society, a special interest group of the International Society for Influenza and other Respiratory Virus Diseases (isirv), and by members of the WHO Global RSV Surveillance Project and will facilitate sample handling, database submission and analysis of HRSV sequence data.

Running Title: Universal HRSV nomenclature below species level

Keywords: Human respiratory syncytial virus, strain, isolate, specimen, nomenclature, species

PeerReview

Title: Proposal for Human Respiratory Syncytial Virus (HRSV) Nomenclature below the Species Level

Dedicated to José Antonio Melero

Authors: Vahid Salimi, Mariana Viegas[#], Alfonsina Trento[#], Charles N. Agoti, Larry J. Anderson, Vasanthi Avadhanula, Justin Bahl, Louis Bont, J. Rodney Brister, Patricia A. Cane, Mónica Galiano, Barney S. Graham, Eneida L. Hatcher, Orienka Hellferscee, David M. Henke, Siddhivinayak Hirve, Sandra Jackson, Els Keyaerts, Leyla Kragten-Tabatabaie, Stephen Lindstrom, Inne Nauwelaers, D. James Nokes, Peter J. Openshaw, Teresa C. Peret, Pedro A. Piedra, Kaat Ramaekers, Annabel Rector, Nídia Sequeira Trovão, Anne von Gottberg, Maria Zambon, Wengqing Zhang, Thomas C. Williams, Ian G. Barr[^*], Ursula J. Buchholz[^*]

Affiliations:

Tehran University of Medical Sciences, Tehran, Iran (V. Salimi)

Hospital de Niños Dr. Ricardo Gutiérrez, Buenos Aires, Argentina (M. Viegas)

National Center of Biotechnology, Madrid, Spain (A. Trento)

KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya (C. N. Agoti, D. J. Nokes)

Emory University, Atlanta, Georgia, USA (L. J. Anderson)

Baylor College of Medicine, Houston, Texas, USA (V. Avadhanula, D. M. Henke, P. A. Piedra)

University of Georgia, Athens, Georgia, USA (J. Bahl)

University Medical Center, Utrecht, The Netherlands (L. Bont)

National Center for Biotechnology Information, NIH, Bethesda, Maryland, USA (J. R. Brister, E. L. Hatcher)

Public Health England, London, UK (P. A. Cane, M. Zambon)

Public Health England, London, UK, and Francis Crick Institute, London (M. Galiano)

National Institute of Allergy and Infectious Diseases, NIH, Bethesda (B. S. Graham, U. J. Buchholz)

National Institute for Communicable Diseases~~,~~ of the National Health Laboratory Service, Johannesburg, South Africa (O. Hellfersce~~Johannesburg, South Africa (O. Helferscee~~, A. von Gottberg)

World Health Organization, Geneva, Switzerland (S. Hirve, S. Jackson, W. Zhang)

University Hospitals Leuven, Leuven, Belgium (E. Keyaerts)

ScholarOne support: (434) 964-4100

Peer Review

40 ReSViNET Foundation, Zeist, The Netherlands and Julius Clinical, Zeist (L. Kragten-
41 Tabatabaie)

42 Centers for Disease Control and Prevention, Atlanta (S. Lindstrom, T. Peret)

43 Public Health England, London, UK and Imperial College London (I. Nauwelaers)

44 UK and Imperial College London (P. Openshaw)

45 Rega Institute, Leuven, Belgium (K. Ramaekers, A. Rector)

46 Fogarty International Center, NIH, Bethesda (N. Sequeira Trovão)

47 University of Edinburgh, Edinburgh, UK (T. C. Williams)

48 Peter Doherty Institute for Infection & Immunity, Melbourne, Australia (I. Barr)

49 [#]These authors contributed equally to this article

50 [^] These last authors contributed equally to this article.

51 * to whom correspondence should be addressed.

52 Ian Barr, Ian.Barr@influenzacentre.org; (613) 9342-9311

53 Ursula Buchholz, ubuchholz@niaid.nih.gov; (301) 594-1533

Peer Review

Abstract - Word Count: 149

Human respiratory syncytial virus (HRSV) is the leading viral cause of serious pediatric respiratory disease, and reinfections are common throughout life. HRSV has two major subgroups, A and B, which exhibit some antigenic variability, contributing to the ability of HRSV to circulate annually. The molecular epidemiology of HRSV is being studied worldwide, and increasing numbers of partial and full genomic sequences are becoming available. To facilitate molecular epidemiology analyses, we propose a uniform nomenclature for HRSV-positive samples, isolates and sequences as follows: HRSV/subgroup identifier/geographic identifier (ISO3166-1 alpha-3 letter country designation or a simple country name)/unique sequence identifier/calendar year of sampling. We also propose a basic template for submission of associated metadata. A universal nomenclature for HRSV isolates would facilitate data retrieval, analysis of sequence data, and would improve publications of figures such as sequence alignments and phylograms, which will overall aid our understanding of the evolution of this virus.

Peer Review

ScholarOne support: (434) 964-4100

Text - Word Count: ~~2592~~2737

Introduction

Human respiratory syncytial virus (HRSV) is the leading single cause of severe respiratory illness in children less than 5 years of age, and is associated with substantial morbidity related to lower respiratory tract infections in industrialized countries, and significant morbidity and mortality in low- and/or middle-income countries (*1-5*). HRSV also causes severe disease in the elderly and in high-risk adults (*6*). In 2016, HRSV (species *Human orthopneumovirus*) was reclassified by the International Committee on Virus Taxonomy (ICTV) to the new family *Pneumoviridae*, genus *Orthopneumovirus* (*7*). With the increasing availability of viral sequencing technologies, the number of HRSV sequence submissions to databases is increasing (Figure 1), and it is anticipated that this trend will continue. While the ICTV provides nomenclature standards for virus taxa, no system currently exists for HRSV nomenclature below the species level. Given the current interest in HRSV and in database submissions, a standard nomenclature for HRSV strains and isolates below the species level is needed and would facilitate sample handling and studies of the genomic diversity of HRSV strains and variants. Following the taxonomical reassignment, it is timely to (i) propose a consistent and universal naming convention for HRSV strains, sequences and isolates, and (ii) provide a framework for the submission of associated basic metadata, facilitating database submissions that are rich in context information.

Currently, several large laboratory HRSV surveillance and epidemiology studies are ongoing. These include the World Health Organization's (WHO) Global Respiratory Syncytial Virus (RSV) Surveillance Project, with large-scale testing for HRSV and extensive sequencing

ScholarOne support: (434) 964-4100

of HRSV-positive clinical specimens from more than 20 countries worldwide (https://www.who.int/influenza/rsv/en/). Focused molecular analyses have helped elucidate HRSV household transmission (*8*) as well as local transmission dynamics (*9*), and may guide the control of HRSV transmission. For example, molecular analysis showed that acquisition of HRSV in healthcare facilities occurs through both transmission within facilities and introduction from the community (*10, 11*)~~. Molecular analyses also have suggested that strain-associated differences in virulence may exist (*12-14*), but to date, there is no consistent association between specific strains and virulence across studies~~. HRSV subgroups A and B usually dominate in alternate years but can also co-circulate in the same season, but it remains to be seen how future subgroup circulation patterns will be affected by the current COVID-19 pandemic. Early work showed that HRSV of subgroup A ~~is~~ was associated with slightly greater clinical severity than subgroup B (*15*). HRSV disease severity has also been correlated with specific HRSV strains, genotypes or clades, but to date, there has been no consistent association between specific strains and virulence across studies (*12-14*), and no genotype or clade ~~is~~ has been consistently associated with severe disease (*16-19*). Thus, a possible role of HRSV strain differences in disease severity remains to be elucidated. The lack of a standard nomenclature and the scarcity of rich metadata in databases currently limit and complicate such studies. In this context, a simple and standardized nomenclature would minimize data entry errors for handling, storing and shipping of HRSV-containing samples. Moreover, consistency in nomenclature will improve the ability of researchers to pool and analyze sequence data and associated information from a variety of different sources.

Reliable and concise nomenclature systems below the species level are available for measles virus, influenza virus, rotavirus, and filovirus isolates (*20-23*), and for many other

ScholarOne support: (434) 964-4100

114 human viral pathogens. A similar nomenclature system for HRSV is urgently needed and will
115 provide a standard interface for submission, searching and display of sequencing data and
116 associated metadata (i.e. an interface tailored to HRSV and its pathology), supporting the
117 requirements of HRSV research and the public health community. To fill this need, an
118 international group of researchers, in conjunction with the WHO Global RSV Surveillance
119 Project, proposes a concise HRSV nomenclature system.

120 **HRSV Genome Organization.** HRSV has a single-stranded non-segmented negative-
121 sense RNA genome of about 15,191 to 15,277 nucleotides in length (Figure 2A) (*7*). The HRSV
122 genome contains 10 genes, each encoding a separate mRNA with a single open reading frame
123 (ORF) (Figure 2A, Table 1), except for the M2 mRNA which contains two overlapping open
124 reading frames. The 11 HRSV proteins include two nonstructural proteins NS1 and NS2, the
125 nucleoprotein N, phosphoprotein P, matrix protein M, small hydrophobic envelope protein SH,
126 attachment glycoprotein G, fusion glycoprotein F, matrix protein M2-1 and RNA regulatory
127 factor M2-2, and the RNA polymerase protein L (Table 1) (*24, 25*). The fusion glycoprotein F is
128 the major viral neutralization and protective antigen, followed by the attachment glycoprotein G
129 (26).

130 **HRSV Subgroups and Genotype Designations: Status and Outlook**. HRSV has two
131 major genetic and antigenic subgroups, A and B, with genome-wide nucleotide and amino acid
132 divergence between these subgroups (Figure 2A) (*25, 27*). The reference sequences for the two
133 subgroups are derived from strains HRSV A2 [(*28, 29*); GenBank Accession number M74568.1,
134 RefSeq accession number NC_038235, curated by the National Center for Biotechnology
135 Information (NCBI), NIH (*30, 31)*], and HRSV B1 [(*32*), GenBank Accession number
136 AF013254.1, RefSeq accession number NC_001781.1] (Figure 2B). Between the two subgroups,

ScholarOne support: (434) 964-4100

137 the F glycoprotein sequences are well conserved, with 89% amino acid identity, while the G
138 glycoproteins are the most divergent amongst the HRSV proteins (only 53% amino acid identity
139 between the subgroups) (Figure 2A) and undergo continuous molecular evolution. The
140 ectodomain of the G glycoproteins of both subgroups contains a conserved central domain,
141 representing an important antigenic site, flanked by two hypervariable domains (*33*). Except for
142 the central conserved region, the antigenic cross reactivity between G glycoproteins of the two
143 subgroups is low (*26*).

144 Since the G ORF exhibits the greatest degree of genetic variability between isolates, it is
145 most commonly used for studies on the molecular evolution of HRSV. The genetic variability of
146 HRSV strains over time has been commonly determined by sequencing of the distal C-terminal
147 third of the G ORF, which includes the second hypervariable domain. The variability in the G
148 ORF is characterized by a high rate of non-synonymous nucleotide changes, suggesting that
149 evolution may be driven by immune pressure, even though this may be partially antibody
150 independent (*34*). It is likely that this variability in the G protein contributes to the capacity of
151 HRSV to cause yearly outbreaks in the community (*35-37*). In a parallel effort, several research
152 groups are working together on a genotyping proposal, providing a consensus on uniform
153 genotype designations (*38, 39*). With continuing virus evolution, the emergence of new
154 genotypes is expected, along with the extinction of older genotypes. HRSV genotyping
155 designations will capture the present status and will be adaptable to reflect future molecular
156 evolution, and will need to be re-evaluated periodically by a global consortium. While the
157 nomenclature proposal outlined herein does not extend to genotype designations, it should
158 facilitate the sequence analysis required to follow the molecular evolution of HRSV.

ScholarOne support: (434) 964-4100

159 **Nomenclature Proposal for HRSV Strains and Isolates.** For molecular epidemiology
160 studies, it would be useful to have a concise standard for short identifiers of specific HRSV
161 sequences, suitable for the short "Definition" lines that give context to a sample and its derived
162 sequence. Ideally, concise standardized names should convey key information about each
163 individual sequence in an alignment or phylogram such as the source, date and type (if known).
164 Here, we aim to define such a common naming convention for HRSV samples and isolates. We
165 also propose the use of standard names and appropriate annotations for HRSV genes, provide
166 examples to guide the annotation of sequence data during the sequence submission process, and
167 suggest the submission of metadata associated with the source materials of HRSV sequences. As
168 indicated above, this nomenclature proposal does not extend to genotype designations, but
169 proposes that genotype information be submitted with metadata to sequence and other relevant
170 databases.

171 **Definition Lines for HRSV Sequence Submissions**. GenBank records available through
172 NCBI are identified by two elements, a unique alphanumeric accession number and a definition
173 line. The definition line is the "title" that one commonly associates with GenBank records shown
174 in BLAST results and other searches. Definition lines are generated by the submitter during the
175 sequence submission process and include the species and isolate name; for example "*Human*
176 *orthopneumovirus* isolate HRSV/A/USA/001/2011, complete genome." (*40*). For each HRSV
177 positive clinical sample or isolate, and for the submission of HRSV sequences to sequence
178 databases, we propose to capture five sample-specific parameters in a standardized fashion,
179 compatible with these isolate name requirements. Specifically, we propose the following
180 convention for the naming of HRSV samples, which will be included in sequence definition lines

ScholarOne support: (434) 964-4100

(Figure 3): <virus name abbreviation>/<HRSV subgroup>/<geographic identifier>/<unique sequence identifier>/<year of sampling>

I. Organism name; virus name abbreviation: HRSV
   a. The ICTV proposed species name is "*Human orthopneumovirus*" (7), which will be seen in the NCBI definition line. During submission, the organism name can be provided as either "Human orthopneumovirus" or "Human respiratory syncytial virus". The abbreviation "HRSV" should be included as part of the isolate name regardless of which organism name is provided.

II. HRSV subgroup: A or B; X if unknown

III. Geographic Identifier of the location of sampling. As individual HRSV research networks have pre-defined requirements, it is suggested that this field have some flexibility. Specifically,
   i. If not specified by a research network, the ISO 3166-1 alpha-3 letter designation should be used to indicate the country of sampling (XXX if the country of sampling is unknown). The ISO 3166-1 alpha-3 can be accessed through this link: https://en.wikipedia.org/wiki/ISO_3166-1_alpha-3. It is strongly suggested to provide any additional geographic information on the sampling location (e.g. city or state of sampling) in metadata fields, rather than within the concise definition line.
   ii. the WHO Global RSV Surveillance Project plans to use the simple name of the country (in English).
   iii. Individual national studies may require a state/province/city designation, as well as the country name (which is mandatory). If state/province/city designation is required,

ScholarOne support: (434) 964-4100

203 a period should be used to separate the country name ([state/province/city].[country
204 name])
205 IV. Unique isolate identifier: This field must be restricted to 8 alphanumeric characters.
206 Underscores are permitted, but hyphens, slashes, or any other types of special characters
207 (e.g., %, $, @, etc.) or spaces cannot be used. Controversial names/phrases, names of
208 prominent people, and trademarked names/phrases as unique identifiers cannot be used.
209 To prevent duplications of sample/isolate numbers by different groups, we recommend
210 inserting an acronym identifying a study or institute prior to the sample number. For
211 example, unique isolate identifiers for samples from the INFORM-RSV study might
212 include the acronym "INF" followed by a number
213 ("HRSV/A/COUNTRY/INF001/2019", unique isolate identifier underlined).
214 V. Year of sampling; (YYYY) or XXXX if unknown.

215 The proposed nomenclature options (HRSV/subgroup/geographic identifier/unique
216 identifier/year of collection) are demonstrated in the following examples:

217 • HRSV/A/USA/001/2011
218 • HRSV/B/Denver.USA/14617/1985
219 • HRSV/A/IRN/001/2017
220 • HRSV/A/Iran/001/2017
221 • HRSV/X/IRN/001/2017 (subgroup unknown)
222 • HRSV/B/New_Zealand/FR123/2020

223 This nomenclature proposal prioritizes a short and concise definition line that will be easy
224 to handle in the laboratory, be easily readable and consistent in public databases. Additional

ScholarOne support: (434) 964-4100

Free Review

information may be submitted in metadata fields. This would allow researchers, epidemiologists and database users to apply specific metadata filters, as needed for data retrieval and specific applications, analyses, or for displaying designations, such as in dendrograms.

Terminology for Annotations. To support efficient data analysis, it is essential that uniform designations be used at the stage of database submission. Table 1 shows commonly accepted names for HRSV genes and proteins. An HRSV gene comprises a gene start signal GGGGCAAAT(A/G), an open reading frame (ORF) with adjacent noncoding regions if present, and the gene end signal through the last adenosine residue [AGT(T/A)A(T/A/G)(A/T)(A/T)(A/T)A_n] [Figure 4, (*41*)]. Each HRSV gene contains a single ORF, except for the M2 gene, which comprises two ORFs M2-1 and M2-2. Nucleotide annotations of genes and ORFs for the HRSV A2 (Figure 4) and HRSV B1 reference sequences are shown in Table 1.

Metadata to Accompany HRSV Samples and Original Sequence Submissions. The most pertinent host data will depend on the interests of individual study groups, with conditions such as prematurity being of interest in a pediatric setting. Groups studying HRSV in an adult setting may have an interest in different metadata, such as if the individual was immunocompromised. Suggested further information that ~~Further information~~ may be included in metadata fields:

1. Isolation_source: sample type (upper or lower airways);
   viral RNA can be extracted directly from a clinical sample, or from an isolate grown in cell culture, or possibly from a cDNA derived recombinant virus. The sources of sequences from isolated viruses can be identified by the following suffices:

ScholarOne support: (434) 964-4100

247 a. wt: wild-type; sequences derived from RNA extracted directly from clinical specimens
248 b. tc: tissue culture; sequences derived from RNA extracted from HRSV isolates
249 propagated in tissue culture (a single round qualifies as "tc")
250 c. rec: recombinant; sequences of cDNA-derived recombinant virus (including vaccine
251 strain)
252 2. Host: *Homo sapiens*; subject age. Indicate years and months if under 5 years of age; years of
253 age if above 5 years of age. Subject sex (if known) should be spelled out.
254 3. Country, state, and (nearest) city of sampling. Metadata information must include the country
255 (full name, not 3 letter abbreviation; for a list of accepted country designations, see
256 https://www.ncbi.nlm.nih.gov/GenBank/collab/country/), can also include city or
257 state/province. Names should be written based on the standard ASCII letters including spaces
258 if required. City or coordinates of the location of sampling should be included if known.
259 4. Collection date (at least month and year). It is highly recommended that the exact date of
260 specimen collection (DD-Mon-YYYY; example: 17-Feb-2002) be used; if not available the
261 year and month should be indicated (Mon-YYYY).
262 5. Genotype according to the consensus in genotype classification by the RSV GeNom Working
263 Group (consensus definition in progress).
264 6. Metadata on the host and the clinical disease should be included in the "Notes" field in a
265 structured format. Protected health information will be excluded from metadata submissions.
266 a. Host:
267 I. If under 6 months of age, birthweight and gestational age at birth.
268 II. Significant pediatric co-morbidities, including prematurity, congenital cardiac
269 disease and broncho-pulmonary dysplasia (BPD).

ScholarOne support: (434) 964-4100

Peer Review

III. Twins (y/n)

IV. Exposed to specific HRSV therapeutic, vaccine, antibody or antiviral (y/n)

V. Viral or bacterial co-infections if known; pathogen species should be spelled out.

~~V.~~VI. Adult co-morbidities such as chronic obstructive pulmonary disease (COPD) or asthma, or altered immune status (e.g. immunocompromised or a bone marrow transplant recipient).

b. Disease ~~severity~~outcomes: Five grades are distinguished (no medical care; out-patient or emergency room; hospital admission; ICU admission; death)

For NCBI submissions, data can be entered through the web interface, or uploaded as tab-delimited text files (examples provided in Supplementary Information). Sequences can be uploaded in FASTA format (Supplementary file 1), with associated metadata provided in a plain text tab-delimited "source modifier" table (virtual examples provided in Supplementary file 2), and gene/protein annotations provided in a plain text tab-delimited "5 column feature" table (Supplementary file 3).

Outlook

Molecular surveillance has revealed that multiple HRSV genotypes circulate simultaneously in a community. Circulating genotypes often vary between communities, and circulation patterns can change from year to year. Extended monitoring of circulating viruses is necessary to better understand transmission and molecular evolution (*42*). As HRSV vaccine candidates and antivirals are being developed, molecular epidemiology studies may reveal potential effects of

ScholarOne support: (434) 964-4100

292 prevention strategies on viral evolution and possible antibody-escape variants. Timely HRSV
293 data sharing worldwide through the use of public databases is essential. We propose that
294 sequence data be uploaded to a publicly accessible database such as the NCBI
295 [https://www.ncbi.nlm.nih.gov; (*31*)][1]. Whilst the most complete repository for HRSV sequence
296 information is NCBI, studies may require that sequences be submitted in the first instance to
297 other databases such as GISAID (www.gisaid.org).

298 Public access will provide useful for investigators to submit, query, and analyze HRSV
299 sequence data, facilitating the evolutionary analysis of sequence diversity within or between
300 HRSV genotypes. This has been clearly demonstrated with the emergence of SARS-CoV-2 and
301 the critical role that genetic sequence analysis has provided. Notably there has been an
302 immediate adoption of a nomenclature similar to what has been proposed here for HRSV,
303 although some differences remain between databases (e.g. NCBI: SARS-CoV-
304 2/human/USA/COVID20-1096/2020; GISAID: hCoV-19/Australia/VIC12/2020). When an
305 HRSV vaccine becomes available, it will be important to monitor and trace possible evolutionary
306 changes in response to vaccine-induced selective pressure (*43, 44*); this will require high-quality
307 and geographically representative country-specific data on circulating strains, and rich datasets
308 of well-curated, standardized, and parsable data.

309 This proposal will profit from strong support by members of the International RSV Society,
310 a special interest group of the International Society for Influenza and other Respiratory Virus
311 Diseases (isirv), members of the WHO Global RSV Surveillance Project, and the HRSV research
312 community.

313 **Dedication**

ScholarOne support: (434) 964-4100

This manuscript is in memory of Jose A. Melero, a superb scientist and leader in the HRSV field, and a wonderful and generous friend to us all.

Acknowledgments

Support for this work was provided by the Intramural Research Programs of the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health. Maria Zambon was supported by Public Health England (PHE), and Inne Neuwelaers was supporte by PHE and by the National Institute of Health Research (NIHR) Respiratory Health Protection Unit. The work of J. Rodney Brister and Eneida L. Hatcher was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health. TC Williams is the recipient of a Wellcome Trust Award [204802/Z/16/Z]. The WHO's Global RSV Surveillance is supported by the Bill and Melinda Gates Foundation grant award no. OPP1127419.

Disclaimer

Siddhivinayak Hirve, Sandra Jackson and Wenqing Zhang work with the World Health Organization. The authors are responsible for the views expressed in this publication and they do not necessarily represent the decisions, policy or views of the World Health Organization. Names of specific vendors, manufacturers, or products are included for public health and informational purposes; inclusion does not imply endorsement of the vendors, manufacturers, or products by the World Health Organization. Larry Anderson has done paid consultancies on RSV vaccines for Bavarian Nordic, Novavax, ClearPath Vaccines Company, and Pfizer. Larr~~r~~y Anderson's laboratory is currently receiving funding through Emory University from Pfizer and Advac for laboratory studies for HRSV surveillance studies in adults, and he holds a subcontract

ScholarOne support: (434) 964-4100

on an NIH SBIR award to Sciogen on G protein HRSV vaccines. Larry Anderson is a co-inventor on several CDC patents on the HRSV G protein and its CX3C chemokine motif relative to immune therapy and vaccine development, and on a patent filing for use of HRSV platform VLPs with the F and G proteins for vaccines. Ursula Buchholz reports CRADA support to NIH from Sanofi Pasteur, outside the submitted work; in addition, Ursula Buchholz has patents on live-attenuated HRSV with royalties paid to NIH by Sanofi Pasteur. There are no additional conflicts of interest by any of the authors.

Author Bio

Dr. Vahid Salimi is a virologist, associate professor and head of the virology department at the School of Public Health, Tehran University of Medical Sciences. His primary research interests include genetic characterization, immunopathogenesis and vaccine design of respiratory viruses.

~~**Footnote**~~

~~[1]Whilst the most complete repository for HRSV sequence information is NCBI, we acknowledge that local study restrictions may require that sequences may be submitted in the first instance to other databases such as GISAID (www.gisaid.org).~~

References:

1. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet. 2017 Sep 2;390(10098):946-58.

ScholarOne support: (434) 964-4100

Peer Review

2. Nair H, Nokes DJ, Gessner BD, Dherani M, Madhi SA, Singleton RJ, et al. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. Lancet. 2010 May 1;375(9725):1545-55.
3. Salimi V, Tavakoli-Yaraki M, Yavarian J, Bont L, Mokhtari-Azad T. Prevalence of human respiratory syncytial virus circulating in Iran. J Infect Public Health. 2016 Mar-Apr;9(2):125-35.
4. Stein RT, Bont LJ, Zar H, Polack FP, Park C, Claxton A, et al. Respiratory syncytial virus hospitalization and mortality: Systematic review and meta-analysis. Pediatr Pulmonol. 2017 Apr;52(4):556-69.
5. Pneumonia Etiology Research for Child Health Study G. Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. Lancet. 2019 Jun 27.
6. Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. Respiratory syncytial virus infection in elderly and high-risk adults. N Engl J Med. 2005 Apr 28;352(17):1749-59.
7. Rima B, Collins P, Easton A, Fouchier R, Kurath G, Lamb RA, et al. ICTV Virus Taxonomy Profile: Pneumoviridae. J Gen Virol. 2017 Dec;98(12):2912-3.
8. Agoti CN, Phan MVT, Munywoki PK, Githinji G, Medley GF, Cane PA, et al. Genomic analysis of respiratory syncytial virus infections in households and utility in inferring who infects the infant. Sci Rep. 2019 Jul 11;9(1):10076.
9. Trovao NS, Khuri-Bulos N, Tan Y, Puri V, Shilts MH, Halpin RA, et al. Molecular characterization of respiratory syncytial viruses circulating in a paediatric cohort in Amman, Jordan. Microb Genom. 2019 Sep 18.

Peer Review

ScholarOne support: (434) 964-4100

379 10. Geis S, Prifert C, Weissbrich B, Lehners N, Egerer G, Eisenbach C, et al. Molecular
380 characterization of a respiratory syncytial virus outbreak in a hematology unit in Heidelberg,
381 Germany. J Clin Microbiol. 2013 Jan;51(1):155-62.
382 11. Nabeya D, Kinjo T, Parrott GL, Uehara A, Motooka D, Nakamura S, et al. The clinical and
383 phylogenetic investigation for a nosocomial outbreak of respiratory syncytial virus infection
384 in an adult hemato-oncology unit. J Med Virol. 2017 Aug;89(8):1364-72.
385 12. Malekshahi SS, Razaghipour S, Samieipoor Y, Hashemi FB, Manesh AAR, Izadi A, et al.
386 Molecular characterization of the glycoprotein and fusion protein in human respiratory
387 syncytial virus subgroup A: Emergence of ON-1 genotype in Iran. Infect Genet Evol. 2019
388 Jul;71:166-78.
389 13. Vos LM, Oosterheert JJ, Kuil SD, Viveen M, Bont LJ, Hoepelman AIM, et al. High
390 epidemic burden of RSV disease coinciding with genetic alterations causing amino acid
391 substitutions in the RSV G-protein during the 2016/2017 season in The Netherlands. J Clin
392 Virol. 2019 Mar;112:20-6.
393 14. Pierangeli A, Viscido A, Bitossi C, Frasca F, Gentile M, Oliveto G, et al. Differential
394 interferon gene expression in bronchiolitis caused by respiratory syncytial virus-A genotype
395 ON1. Medical microbiology and immunology. 2020 Feb;209(1):23-8.
396 15. Hall CB, Walsh EE, Schnabel KC, Long CE, McConnochie KM, Hildreth SW, et al.
397 Occurrence of groups A and B of respiratory syncytial virus over 15 years: associated
398 epidemiologic and clinical characteristics in hospitalized and ambulatory children. J Infect
399 Dis. 1990 Dec;162(6):1283-90.
400 16. Vandini S, Biagi C, Lanari M. Respiratory Syncytial Virus: The Influence of Serotype and
401 Genotype Variability on Clinical Course of Infection. Int J Mol Sci. 2017 Aug 6;18(8).

Peer Review

ScholarOne support: (434) 964-4100

402 17. Laham FR, Mansbach JM, Piedra PA, Hasegawa K, Sullivan AF, Espinola JA, et al.
403 Clinical Profiles of Respiratory Syncytial Virus Subtypes A AND B Among Children
404 Hospitalized with Bronchiolitis. Pediatr Infect Dis J. 2017 Aug;36(8):808-10.
405 18. Gilca R, De Serres G, Tremblay M, Vachon ML, Leblanc E, Bergeron MG, et al.
406 Distribution and clinical impact of human respiratory syncytial virus genotypes in
407 hospitalized children over 2 winter seasons. J Infect Dis. 2006 Jan 1;193(1):54-8.
408 19. Martinello RA, Chen MD, Weibel C, Kahn JS. Correlation between respiratory syncytial
409 virus genotype and severity of illness. J Infect Dis. 2002 Sep 15;186(6):839-42.
410 20. Measles virus nomenclature update: 2012. Wkly Epidemiol Rec. 2012 Mar 2;87(9):73-81.
411 21. Kuhn JH, Bao Y, Bavari S, Becker S, Bradfute S, Brister JR, et al. Virus nomenclature
412 below the species level: a standardized nomenclature for laboratory animal-adapted strains
413 and variants of viruses assigned to the family Filoviridae. Arch Virol. 2013
414 Jun;158(6):1425-32.
415 22. Matthijnssens J, Ciarlet M, McDonald SM, Attoui H, Banyai K, Brister JR, et al. Uniformity
416 of rotavirus strain nomenclature proposed by the Rotavirus Classification Working Group
417 (RCWG). Arch Virol. 2011 Aug;156(8):1397-413.
418 23. Reconsideration of influenza A virus nomenclature: a WHO memorandum. Bull World
419 Health Organ. 1979;57(2):227-33.
420 24. Collins PL, Fearns R, Graham BS. Respiratory syncytial virus: virology, reverse genetics,
421 and pathogenesis of disease. Curr Top Microbiol Immunol. 2013;372:3-38.
422 25. Collins PL, Karron RA. Respiratory syncytial virus and metapneumovirus. In: Knipe DM,
423 Howley PM, Cohen JI, Griffin DE, Lamb RA, Martin MA, et al., editors. Fields Virology.
424 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2013. p. 1086-123.

Peer Review

ScholarOne support: (434) 964-4100

425 26. Collins PL, Melero JA. Progress in understanding and controlling respiratory syncytial
426 virus: still crazy after all these years. Virus Res. 2011 Dec;162(1-2):80-99.
427 27. Mufson MA, Orvell C, Rafnar B, Norrby E. Two distinct subtypes of human respiratory
428 syncytial virus. J Gen Virol. 1985 Oct;66 (Pt 10):2111-24.
429 28. Collins PL, Olmsted RA, Spriggs MK, Johnson PR, Buckler-White AJ. Gene overlap and
430 site-specific attenuation of transcription of the viral polymerase L gene of human respiratory
431 syncytial virus. Proc Natl Acad Sci U S A. 1987 Aug;84(15):5134-8.
432 29. Mink MA, Stec DS, Collins PL. Nucleotide sequences of the 3' leader and 5' trailer regions
433 of human respiratory syncytial virus genomic RNA. Virology. 1991 Dec;185(2):615-24.
434 30. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference
435 sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional
436 annotation. Nucleic Acids Res. 2016 Jan 4;44(D1):D733-45.
437 31. Sayers EW, Beck J, Brister JR, Bolton EE, Canese K, Comeau DC, et al. Database resources
438 of the National Center for Biotechnology Information. Nucleic Acids Res. 2020 Jan
439 8;48(D1):D9-D16.
440 32. Karron RA, Buonagurio DA, Georgiu AF, Whitehead SS, Adamus JE, Clements-Mann ML,
441 et al. Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral
442 replication in vitro: clinical evaluation and molecular characterization of a cold-passaged,
443 attenuated RSV subgroup B mutant. Proc Natl Acad Sci U S A. 1997 Dec 9;94(25):13961-6.
444 33. Martinez I, Dopazo J, Melero JA. Antigenic structure of the human respiratory syncytial
445 virus G glycoprotein and relevance of hypermutation events for the generation of antigenic
446 variants. J Gen Virol. 1997 Oct;78 (Pt 10):2419-29.

Peer Review

ScholarOne support: (434) 964-4100

447 34. Trento A, Abrego L, Rodriguez-Fernandez R, Gonzalez-Sanchez MI, Gonzalez-Martinez F,
448 Delfraro A, et al. Conservation of G-Protein Epitopes in Respiratory Syncytial Virus (Group
449 A) Despite Broad Genetic Diversity: Is Antibody Selection Involved in Virus Evolution? J
450 Virol. 2015 Aug;89(15):7776-85.
451 35. Melero JA, Garcia-Barreno B, Martinez I, Pringle CR, Cane PA. Antigenic structure,
452 evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. J
453 Gen Virol. 1997 Oct;78 (Pt 10):2411-8.
454 36. Sullender WM, Mufson MA, Anderson LJ, Wertz GW. Genetic diversity of the attachment
455 protein of subgroup B respiratory syncytial viruses. J Virol. 1991 Oct;65(10):5425-34.
456 37. Cane PA, Matthews DA, Pringle CR. Analysis of respiratory syncytial virus strain variation
457 in successive epidemics in one city. J Clin Microbiol. 1994 Jan;32(1):1-4.
458 38. Goya S, Galiano M, Nauwelaers I, Trento A, Openshaw PJ, Mistchenko AS, et al. Toward
459 unified molecular surveillance of RSV: A proposal for genotype definition. Influenza Other
460 Respir Viruses. 2020 Feb 5.
461 39. Ramaekers K, Rector A, Cuypers L, Lemey P, Keyaerts E, Van Ranst M. Towards a Unified
462 Classification for Human Respiratory Syncytial Virus Genotypes. Virus Evolution, accepted
463 for publication. 2020 Jul 24,veaa052.
464 40. The GenBank Submissions Handbook. 2011 ed. Bethesda (MD): National Center for
465 Biotechnology Informationn (US).
466 41. Collins PL, Dickens LE, Buckler-White A, Olmsted RA, Spriggs MK, Camargo E, et al.
467 Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal
468 distinctive features of intergenic structure and gene order. Proc Natl Acad Sci U S A. 1986
469 Jul;83(13):4594-8.

Peer Review

ScholarOne support: (434) 964-4100

470 42. Trento A, Casas I, Calderon A, Garcia-Garcia ML, Calvo C, Perez-Brena P, et al. Ten years
471 of global evolution of the human respiratory syncytial virus BA genotype with a 60-
472 nucleotide duplication in the G protein gene. J Virol. 2010 Aug;84(15):7500-12.
473 43. Graham BS. Vaccine development for respiratory syncytial virus. Curr Opin Virol. 2017
474 Apr;23:107-12.
475 44. Mazur NI, Higgins D, Nunes MC, Melero JA, Langedijk AC, Horsley N, et al. The
476 respiratory syncytial virus vaccine landscape: lessons from the graveyard and promising
477 candidates. Lancet Infect Dis. 2018 Oct;18(10):e295-e311.
478 45. Collins PL, Hill MG, Camargo E, Grosfeld H, Chanock RM, Murphy BR. Production of
479 infectious human respiratory syncytial virus from cloned cDNA confirms an essential role
480 for the transcription elongation factor from the 5' proximal open reading frame of the M2
481 mRNA in gene expression and provides a capability for vaccine development. Proc Natl
482 Acad Sci U S A. 1995 Dec 5;92(25):11563-7.
483 46. Crowe JE, Jr., Bui PT, Firestone CY, Connors M, Elkins WR, Chanock RM, et al. Live
484 subgroup B respiratory syncytial virus vaccines that are attenuated, genetically stable, and
485 immunogenic in rodents and nonhuman primates. J Infect Dis. 1996 Apr;173(4):829-39.

486

487 Address for correspondence:
488 Ian Barr, Peter Doherty Institute for Infection & Immunity, Melbourne, Australia; email:
489 Ian.Barr@influenzacentre.org; (613) 9342-9311

490 Ursula Buchholz, NIAID, NIH, 50 South Dr, 50/6503, Bethesda, MD 20892-8007, USA; email:
491 ubuchholz@niaid.nih.gov; (301) 594-1533

ScholarOne support: (434) 964-4100

Peer Review

492 **Table 1: Widely Accepted Nomenclature for HRSV Genes and Proteins; Gene Annotation**

493 **of Strains HRSV A2 and HRSV B1**

Gene or Genome Region			Protein		Open Reading Frame	
Annotation [a]	Genome Position (nt) [a]		Annotation	Abbre-viation	Genome Position (nt)	
	Strain A2[b]	Strain B1[c]			Strain A2	Strain B1
Leader region	1-44	1-44				
NS1	45-576	45-577	Nonstructural protein 1	NS1	99-518	99-518
NS2	596-1098	594-1098	Nonstructural protein 2	NS2	628-1002	626-1000
N	1126-2328	1125-2327	Nucleoprotein	N	1141-2316	1140-2315
P	2330-3243	2331-3244	Phosphoprotein	P	2347-3072	2348-3073
M	3253-4210	3254-4208	Matrix protein	M	3262-4032	3263-4033
SH	4220-4629	4218-4630	Small hydrophobic protein	SH	4304-4498	4303-4500
G	4674-5596	4675-5600	Attachment glycoprotein	G	4689-5585	4690-5589
F	5649-7551	5653-7552	Fusion glycoprotein	F	5662-7386	7666-7390
			Matrix protein 2			
M2	7598-8558	7609-8568	Matrix protein M2-1	M2-1	7607-8191	7618-8205
			Matrix protein M2-2	M2-2	8160-8432	8171-8443

Peer Review

ScholarOne support: (434) 964-4100

L	8491-15068	8501-15080	Polymerase protein	L	8499-14996	8509-15009
Trailer region	15069-15223	15081-15225				

494 [a] Nucleotide annotations for leader and trailer region, and for indicated HRSV genes from the first nucleotide (nt) of
495 the gene start signal [GGGGCAAAT(A/G); GGAGCAAAAT in case of the L gene] through the last adenosine
496 residue of the gene end signal [AGT(T/A)A(T/A/G)(A/T)(A/T)(A/T)A_n]
497 [b] HRSV/A/USA/A2/2015 (*45*); GenBank Accession Number KT992094
498 [c] HRSV/B/USA/B1/1985/B1 (*46*); GenBank Accession Number AF013254.1

Peer Review

ScholarOne support: (434) 964-4100

Figure Legends

Figure 1. Annual numbers of HRSV whole-genome sequences released in Genbank since publication of the whole-genome sequence of HRSV A2 (M74568, released in 1993).

Figure 2. (A) Schematic overview of the HRSV gene order and comparison of the amino acid identities of the reference strains of subgroup A (HRSV A2: GenBank Accession number M74568/NC_038235, and HRSV B1: GenBank Accession number AF013254/NC_001781). **(B)** International Committee on Virus Taxonomy (ICTV)~~ICTV~~ species designation, virus name, and associated GenBank reference sequences, curated by the National Center for Biotechnology Information (NCBI), NIH.

Figure 3. Schematic representation of the five consensus nomenclature elements of HRSV strains and isolates, with examples (top) and a legend for each element (bottom).

Figure 4. Schematic aid to gene annotations of HRSV whole-genome sequences. An HRSV gene comprises the sequence from the first nucleotide of the conserved HRSV gene start signal (GGGGCAAATa) to the last adenosine residue of the HRSV gene end signal (AGTTAnnnnAAAA) (25, 41). Gene start signals are represented by black triangles, and gene end signals are shown as black rectangles, separated by~~.~~ intergenic regions (underlined). Note the M2/L gene overlap (annotations derived from HRSV A2; GenBank Accession number M74568/NC_038235).

Peer Review

ScholarOne support: (434) 964-4100

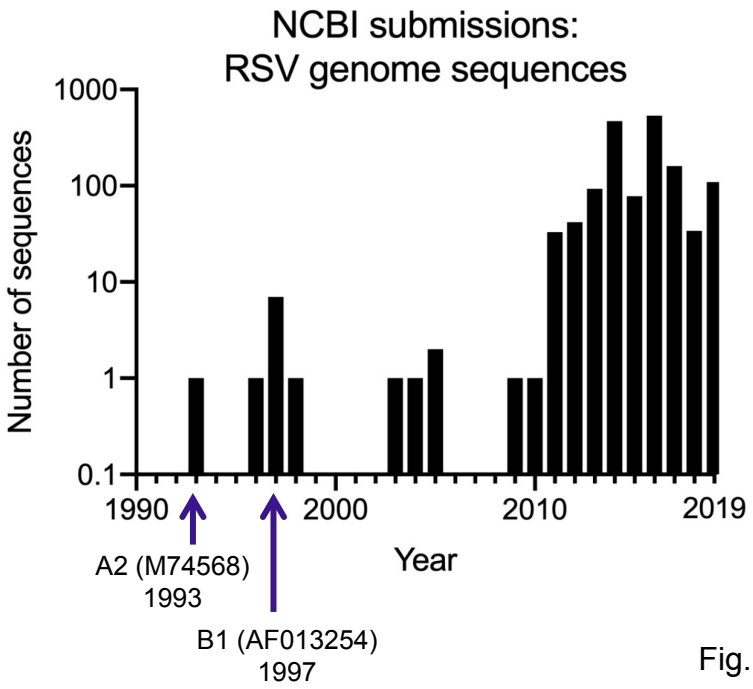
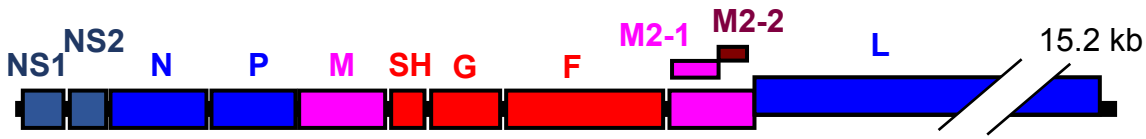

NCBI submissions:
RSV genome sequences
Number of sequences
1000
100
10
1
0.1
1990
2000
2010
2019
Year
A2 (M74568)
1993
B1 (AF013254)
1997

Fig. 1

A



HRSV Isolate A2 versus B1:
% amino acid sequence identity for the indicated protein

NS1	NS2	N	P	M	SH	G	F	M2-1	M2-2	L
87	92	96	91	91	76	**53**	**89**	92	72	93

B

NCBI-curated reference sequences for human respiratory syncytial virus subgroups A, B

Species	Virus name(s)	Virus abbrev	Isolate	GenBank accession	RefSeq* accession	Other information
Human orthopneumo-virus	human orthopneumovirus; human respiratory syncytial virus	HRSV	A2	M74568	NC_038235	ICTV exemplar
			B1	AF013254	NC_001781	

* RefSeq: NCBI-curated

Fig. 2

- HRSV/A/USA/001/2011
- HRSV/B/USA/14617/1985
- HRSV/A/IRN/002/2017

1. Virus Name: human respiratory syncytial virus
   Abbreviation: HRSV
2. HRSV subgroup identifier (A or B; X if unknown)
3. Country of sampling: ISO 3166-1 alpha-3 letter designation; XXX = unknown
4. Unique isolate identifier, restricted to 8 alphanumeric characters
5. Calendar year of sampling (YYYY) or XXXX if unknown

Fig. 3

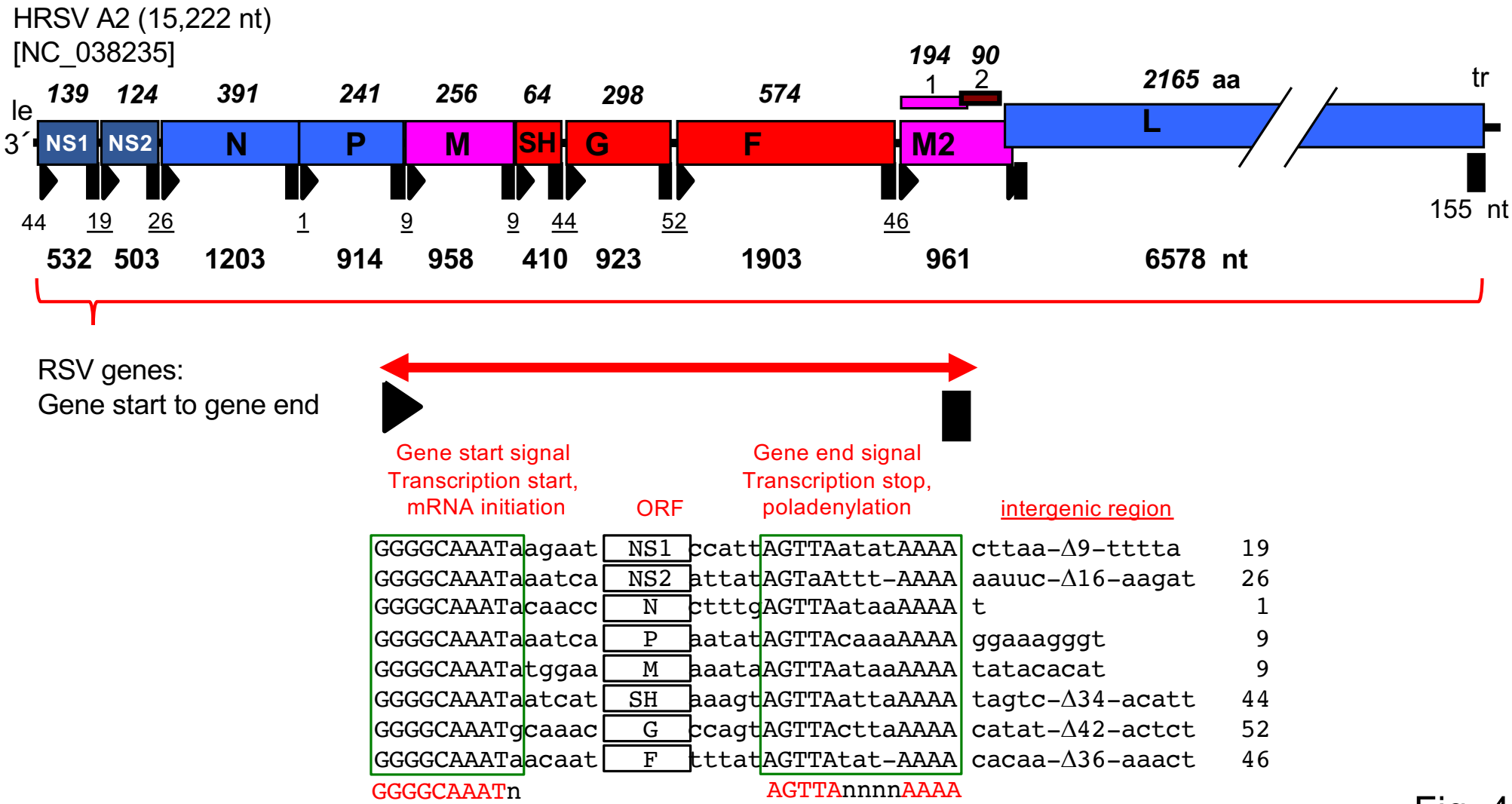

HRSV A2 (15,222 nt)
[NC_038235]
le
3´
139
124
391
241
256
64
298
574
194
90
1
2
2165 aa
tr
NS1
NS2
N
P
M
SH
G
F
M2
L
44
19
26
1
9
9
44
52
46
155 nt
532
503
1203
914
958
410
923
1903
961
6578 nt
RSV genes:
Gene start to gene end
Gene start signal
Transcription start,
mRNA initiation
ORF
Gene end signal
Transcription stop,
poladenylation
intergenic region
GGGGCAAATaagaat NS1 ccattAGTTAatatAAAA cttaa-Δ9-tttta 19
GGGGCAAATaaatca NS2 attatAGTaAttt-AAAA aauuc-Δ16-aagat 26
GGGGCAAATacaacc N ctttgAGTTAataAAAA t 1
GGGGCAAATaaatca P aatatAGTTAcaaaAAAA ggaaagggt 9
GGGGCAAATatggaa M aaataAGTTAataAAAA tatacacat 9
GGGGCAAATaatcat SH aaagtAGTTAattaAAAA tagtc-Δ34-acatt 44
GGGGCAAATgcaaac G ccagtAGTTActtaAAAA catat-Δ42-actct 52
GGGGCAAATaacaat F tttatAGTTAtat-AAAA cacaa-Δ36-aaact 46
GGGGCAAATn
AGTTAnnnnAAAA

Fig. 4

>rsv011
ACGCGAAAAAATGCGTACAACAAACTTGCGTAAACCAAAAAAATGGGGCAAATAA
GAATTTGATAAGTACCACTTAAATT
TAACTCCTTTGGTTAGAGATGGGCAGCAACTCATTGAGTATGATAAAAGTTAGATTG
CAAAATCTGTTTGACAATGATGA
AGTAGCATTGTTAAAAATAACATGCTATACTGACAAATTAATACAGTTAACTAATGC
TTTGGCTAAGGCAGTTATACATA
CAATCAAATTGAATGGCATTGTATTTGTGCATGTTATTACAAGTAGTGATATTTGCCC
TAATAATAATATTGTAGTGAAA
TCCAATTTCACAACAATGCCAGTATTACAAAATGGAGGTTATATATGGGAAATGATG
GAATTAACACACTGCTCTCAACC
TAATGGCCTAATAGATGACAATTGTGAAATTAAATTCTCCAAAAAACTAAGTGATTC
AACAATGACCAATTATATGAATC
AATTATCTGAATTACTTGGATTTGACCTCAATCCATAAATCATAATAAATATCAACT
AGCAAATCAATGTCACTAACACC
ATTAGTTAATATAAAACTTGACAGAAGATAAAAATGGGGCAAATAAATCAATTCAG
CCGACCCAACCATGGACACAACAC
ACAATGATACCACACCACAAAGACTGATGATCACAGACATGAGACCATTATCGCTT
GAGACTATAATAACATCTCTAACC
AGAGATATCATAACACATAAATTTATATACTTGATAAATCATGAATGCATAGTGAGA
AAACTTGATGAAAGACAGGCCAC
ATTTACATTTCTGGTCAACTATGAAATGAAACTATTGCACAAAGTGGGAAGCACTAA
ATATAAAAAATATACTGAATACA
ACACAAAATATGGCACTTTCCCTATGCCAATATTTATCAATCATGATGGGTTCTTAG
AATGCATTGGCATTAAGCCTACC
AAGCACACACCCATAATATACAAGTATGATCTCAATCCATGAATATCAAACCAAGA
TTCAAACAATCCGAAATAACAACT
TTATGCATAATCACACTCCATAGTCCAAATGGAGCCTGAAAATTATAGTTATTTAAA
ATTAAGGAGAGACATAAGATGAA
AGATGGGGCAAATACAAAAATGGCTCTTAGCAAAGTCAAGTTGAATGATACACTCA
ACAAAGATCAACTTCTATCATCCA
GCAAATATACCATCCAACGGAGCACAGGAGACAGCATTGACACTCCTAATTATGAT
GTGCAGAAACACATTAATAAGTTA
TGTGGCATGTTATTAATCACAGAAGATGCTAATCATAAATTCACTGGGTTAATAGGT
ATGTTATATGCTATGTCTAGATT
AGGAAGAGAAGACACCATAAAAATACTTAAAGATGCGGGATATCATGTTAAGGCAA
ATGGAGTGGATGTAACAACACATC
GTCAAGACATTAATGGGAAAGAAATGAAATTTGAAGTGTTAACATTAGCAAGCTTA
ACAACTGAAATTCAAATCAACATT
GAGATAGAATCTAGAAAATCCTACAAAAAAATGCTAAAAGAAATGGGAGAGGTGG
CTCCAGAATACAGGCATGACTCTCC
AGATTGTGGGATGATAATATTATGTATAGCAGCATTAGTAATAACCAAATTAGCAGC
AGGAGATAGATCAGGTCTTACAG
CTGTGATTAGGAGAGCTAATAATGTCCTAAAAAATGAAATGAAACGTTACAAAGGT
TTATTACCCAAGGATATAGCCAAC

Peer Review

AGCTTCTATGAAGTGTTTGAAAAATATCCTCACTTTATAGATGTTTTTGTTCATTTTG
GTATAGCACAATCTTCTACCAG
AGGTGGCAGTAGAGTTGAAGGGATTTTTGCAGGATTGTTTATGAATGCCTATGGTGC
AGGGCAAGTGATGTTACGGTGGG
GGGTCTTAGCAAAATCAGTTAAAAACATTATGTTAGGACACGCTAGTGTACAAGCA
GAAATGGAACAAGTTGTGGAGGTG
TATGAGTATGCTCAGAAATTGGGTGGAGAAGCAGGATTCTACCATATATTGAACAA
CCCAAAAGCATCACTATTATCTTT
GACTCAATTTCCTCACTTCTCTAGTGTAGTATTGGGCAATGCTGCTGGCCTAGGCATA
ATGGGAGAATACAGAGGTACAC
CAAGGAATCAAGATTTATATGATGCTGCAAAAGCATATGCTGAACAACTCAAAGAA
AATGGTGTGATTAACTACAGTGTA
TTAGATTTGACAGCAGAAGAACTAGAGGCTATCAAACATCAGCTTAATCCAAAAGA
TAATGATGTAGAGCTTTGAGTTAA
TAAAAAGGTGGGGCAAATAAATCATCATGGAAAAGTTTGCTCCTGAATTCCATGGA
GAAGATGCAAACAACAGAGCCACC
AAATTCCTAGAATCAATAAAGGGCAAATTCACATCACCCAAAGATCCCAAGAAAAA
AGATAGTATCATATCTGTCAACTC
AATAGATATAGAAGTAACCAAAGAAAGCCCTATAACATCAAATTCAACCATTATAA
ACCCAATAAATGAGACAGATGATA
CTGTAGGGAACAAGCCCAATTATCAAAGAAAGCCTCTAGTAAGTTTCAAAGAAGAC
CCTATGCCAAGTGATAATCCTTTT
TCAAAACTATACAAAGAAACCATAGAAACATTTGATAACAATGAAGAAGAATCTAG
CTATTCATATGAAGAAATAAATGA
TCAGACAAACGATAATATAACAGCAAGATTAGATAGGATTGATGAGAAATTAAGTG
AAATACTAGGAATGCTTCACACAT
TAGTAGTAGCGAGTGCAGGACCCACATCTGCTCGGGATGGTATAAGAGATGCCATG
GTTGGTTTAAGAGAAGAAATGATA
GAAAAAATCAGAACTGAAGCATTAATGACCAATGACAGACTAGAAGCTATGGCAAG
ACTCAGGAATGAAGAAAGTGAAAA
GATGGCAAAAGACACATCAGATGAAGTGTCTCTCAATCCAACATCAGAGAAACTGA
ACAACCTGTTGGAAGGGAATGATA
GTGACAATGATCTATCACTTGAAGATTTCTGATTAGCTACCAAACTGTACATCAAAA
CACAACACCAATAGAAAACCAAC
AAACAAACCAACTCACCCATCCAACCAAACATCTATCTGCTGATTAGCCAACCAGCC
AAAAAACAACCAGCCAATCTAAA
ACTAGCCACCCGGAAAAAATCGATACTATAGTTACAAAAAAAGATGGGGCAAATAT
GGAAACATACGTGAATAAACTTCA
CGAGGGCTCCACATACACAGCTGCTGTTCAATACAATGTCCTAGAAAAAGACGATG
ATCCTGCATCACTTACAATATGGG
TGCCCATGTTCCAATCATCCATGCCAGCAGATCTACTCATAAAAGAACTAGCCAATG
TCAATATACTAGTGAAACAAATA
TCCACACCCAAGGGACCCTCATTAAGAGTCATGATAAACTCAAGAAGTGCAGTGCT
AGCACAAATGCCCAGCAAATTTAC
CATATGTGCCAATGTGTCCTTGGATGAAAGAAGCAAGCTAGCATATGATGTAACCAC
ACCCTGTGAAATTAAGGCATGCA

For Review

GTCTAACATGCCTAAAATCAAAAAATATGTTAACTACAGTTAAAGATCTCACTATGA
AAACACTCAACCCAACACATGAC
ATCATTGCTTTATGTGAATTTGAAAATATAGTAACATCAAAAAAAGTCATAATACCA
ACATACCTAAGATCTATCAGCGT
CAGAAATAAAGATCTGAACACACTTGAAAATATAACAACCACTGAATTCAAAAATG
CCATTACAAATGCAAAAATCATCC
CTTACTCAGGATTACTGTTAGTCATCACAGTGACTGACAACAAAGGAGCATTCAAAT
ACATAAAGCCACAAAGTCAATTC
ATAGTAGATCTTGGAGCCTACCTAGAAAAAGAAAGTATATATTATGTTACAACAAAT
TGGAAGCACACAGCTACACGATT
TGCAATCAAACCCATGGAAGATTAACCTTTTTCCTATACATCAATGAGTAGATTCAT
ACAAACTTTCTAACTACATTCTT
CACTTCACAATCATAATCACCAACCCTCTGTGGTTCAATCAATCAAACAAAACTCAT
CAGGAGTTCCAGATCATCCCAAG
TCATTGTTCATCAGATCCAGTACTCAAATAAGTTAATAAAAAATCCACATGGGGCAA
ATAATCATTGAGGGAAATCCAAC
TAATCACAACATCTGTCAACATAGACAAGTCAACACGCTAGATAAAATCAACCAAT
GGAAAATACATCCATAACTATAGA
ATTCTCAAGCAAATTCTGGCCTTACTTTACACTAATACACATGATAACAACAATAAT
CTCTTTGATAATCATAATCTCCA
TCATGATTGCAATACTAAACAAACTCTGCGAATATAATGTATTCCATAACAAAACCT
TTGAGCTACCAAGAGCTCGAGTC
AATACATAGCATTCACCAATCTGATAGCTCAAAACAGTAACCTTGCATTTGTAAATG
AACTACCCTCACTTCTTCACAAA
ACCACATCAACATCTCACCATGCAAGCCATCATCTATACCATAAAGTAGTTAATTAA
AAAATAGTCATAACAATGAACTA
GGATATTAAGACCAAAAACAACGCTGGGGCAAATGCAAACATGTCCAAAACCAAGG
ACCAACGCACCGCCAAGACACTAG
AAAGGACCTGGGACACTCTCAATCATCTGTTATTCATATCATCGTGCTTATACAAGT
TAAATCTTAAATCTATAGCACAA
ATCACATTATCTATTTTGGCAATGATAATCTCAACCTCACTTATAATTGCAGCCATCA
TATTCATAGCCTCGGCAAACCA
CAAAGTCACACTAACAACTGCAATCATACAAGATGCAACGAACCAGATCAAGAACA
CAACCCCAACATACCTCACCCAGA
ATCCCCAGCTTGGAATCAGCTTCTCCAATCTGTCCGGAACTACATCACAATCCACCA
CCATACTAGCTTCAACAACACCA
AGTGCTGAGTCAACCCCACAATCCACAACAGTCAAGATCAAAAACACAACAACAAC
CCAAATATTACCTAGCAAACCCAC
CACAAAACAACGCCCAAATAAACCACAAAACAAACCCAACAATGATTTTCACTTTG
AAGTGTTCAATTTTGTACCCTGCA
GCATATGCAGCAACAATCCAACCTGCTGGGCCATCTGCAAGAGAATACCAAACAAA
AAACCTGGAAAGAAAACCACCACC
AAGCCCACAAAAAAACCAACCCTCAAGACAACCAAAAAAGATCCCAAACCTCAAA
CCACAAAACCAAAGGAAGTACTCAC
TACCAAGCCTACAGGAAAGCCAACCATCAACACCACCAAAACAAACATCAGAACTA
CACTGCTCACCTCCAACACCAAAG

For Peer Review

GAAATCCAGAACACACAAGTCAAGAGGAAACCCTCCACTCAACCACCTCCGAAGGC
TATCTAAGCCCATCACAAGTCTAT
ACAACATCCGGTCAAGAGGAAACCCTCCACTCAACCACCTCCGAAGGCTATCTAAG
CCCATCACAAGTCTATACAACATC
CGAGTACCTATCACAATCTCTATCTTCATCCAACACAACAAAATGATAGTCATTAAA
AAGCGTATTGTTGCAAAAAGCCA
TGACCAAATCAAACAGAATCAAAATCAACTCTGGGGCAAATAACAATGGAGTTGCC
AATCCTCAAAACAAATGCTATTAC
CACAATCCTTGCTGCAGTCACACTCTGTTTCGCTTCCAGTCAAAACATCACTGAAGA
ATTTTATCAATCAACATGCAGTG
CAGTTAGCAAAGGCTATCTTAGTGCTCTAAGAACTGGTTGGTATACTAGTGTTATAA
CTATAGAATTAAGTAATATCAAG
GAAAATAAGTGTAATGGTACAGACGCTAAGGTAAAATTAATAAAACAAGAATTAGA
TAAATATAAAAATGCTGTAACAGA
ATTGCAGTTGCTCATGCAAAGCACACCAGCAGCCAACAGTCGAGCCAGAAGAGAAC
TACCAAGATTTATGAATTATACAC
TCAACAATACCAAAAACACCAATGTAACATTAAGTAAGAAAAGGAAAAGAAGATTT
CTTGGATTTTTGTTAGGTGTTGGA
TCTGCAATCGCCAGTGGCATTGCCGTATCCAAGGTCCTGCACCTAGAAGGGGAAGTG
AACAAAATCAAAAGTGCTCTACT
ATCCACAAACAAGGCTGTAGTCAGCTTATCTAATGGAGTCAGTGTCTTAACCAGCAA
GGTGTTAGACCTCAAAAACTATA
TAGATAAACAGTTGTTACCTATTGTTAACAAGCAAAGCTGCAGCATATCAAACATTG
AAACTGTGATAGAGTTCCAACAA
AAGAACAACAGACTACTAGAGATTACCAGAGAATTTAGTGTTAATGCAGGTGTAAC
TACACCTGTAAGCACTTATATGTT
AACTAATAGTGAGTTATTATCATTAATCAATGATATGCCTATAACAAATGATCAGAA
AAAGTTAATGTCCAGCAATGTTC
AAATAGTTAGACAGCAAAGTTACTCTATCATGTCAATAATAAAAGAGGAAGTCTTA
GCATATGTAGTACAATTACCACTA
TATGGTGTAATAGATACTCCTTGTTGGAAACTACACACATCCCCTCTATGTACAACC
AACACAAAGGAAGGATCCAACAT
CTGCTTAACAAGAACCGACAGAGGATGGTACTGTGACAATGCAGGATCAGTATCCT
TTTTCCCACAAGCTGAAACATGTA
AAGTTCAATCGAATCGGGTGTTTTGTGACACAATGAACAGTTTAACATTACCAAGTG
AGGTAAATCTCTGCAACATTGAC
ATATTCAACCCCAAATATGATTGCAAAATTATGACTTCAAAAACAGATGTAAGCAGC
TCCGTTATCACATCTCTAGGAGC
CATTGTGTCATGCTATGGCAAAACCAAATGTACAGCATCCAATAAAAATCGTGGGAT
CATAAAGACATTCTCTAACGGGT
GTGATTATGTATCAAATAAGGGGGTGGATACTGTGTCTGTAGGTAATACATTATATT
ATGTAAATAAGCAAGAAGGCAAA
AGTCTCTATGTAAAAGGTGAACCAATAATAAATTTCTATGATCCATTAGTGTTCCCC
TCTGATGAATTTGATGCATCAAT
ATCTCAAGTCAATGAGAAAATTAATCAGAGTCTAGCATTTATCCGTAAATCAGATGA
ATTATTACATAATGTAAATGCTG

For Peer Review

GTAAATCCACCACAAATATCATGATAACTACCATAATTATAGTAATTATAGTAATAT
TGTTAGCATTAATTGCAGTTGGA
CTGCTTCTATACTGCAAGGCCAGAAGCACACCAGTCACATTAAGTAAGGATCAACTG
AGTGGTATAAATAATATTGCATT
TAGTAACTGAATAAAAATAGCACCTAATCATATTCTTACAATGGTTCGCTATTTGAC
CATAGATAACCCATCTATCATTA
GATTATCCTAAAATTTGAACTTCATCACAACTTTCATCTATAAACCATCTCACTTACA
CTTTTTAAGTGGATTCCTATTT
TATAGTTATATAAAACAATTGAATACCAAATTAACTTACTATTTGTAAAAATGAGAA
CTGGGGCAAATATGTCACGAAGG
AATCCTTGCAAATTCGAAATTCGAGGTCATTGCTTGAATGGTAAAAGGTGTCATTTT
AGTCATAATTATTTTGAATGGCC
ACCCCATGCACTGCTTGTAAGACAAAACTTTATGTTAAACAGAATACTTAAGTCTAT
GGATAAAAGCATAGATACTTTGT
CAGAAATAAGTGGAGCTGCAGAGTTGGACAGAACAGAAGAGTATGCCCTCGGTGTA
GTTGGAGTGCTAGAGAGTTATATA
GGATCAATAAATAATATAACTAAACAATCAGCATGTGTTGCCATGAGCAAACTCCTT
ACTGAACTCAACAGCGATGACAT
AAAAAAACTAAGGGACAATGAAGAGCCAAACTCACCCAAAGTAAGAGTGTACAAT
ACTGTCATATCATATATTGAAAGCA
ACAGGAGGAACAATAAACAAACTATCCATCTGTTAAAAAGATTGCCAGCAGACGTA
TTGAAGAAAACCATCAAAAACACA
TTGGATATCCACAAGAGCATAACCATCAATAACCCAAAAGAATCAACTGTTAGTGA
TACGAACGACCATGCCAAAAATAA
TGATACTACCTGACAAATATCCTTGTAGTATAAATTCCATACTAATAACAAGTAATT
GTAGAGTCACTATGTATAATCAA
AAAAACACACTATATATCAATCAAAACAACCAAAATAACCATATATACCCACCGGA
TCAACCATTCAATGAAATCCATTG
GACCTCTCAAGACTTGATTGATGCAACTCAAAATTTTCTACAACATCTAGGTATTAC
TGATGATATATACACAATATATA
TATTAGTGTCATAATACTCAATCCTAATACTTACCACATCATCAAATTATTAACTCAA
ACAATTCAAGCTATGGGACAAA
ATGGATCCCATTATTAGTGGAAATTCTGCTAATGTTTATCTAACTGATAGTTATTTAA
AAGGTGTTATTTCTTTCTCAGA
ATGTAACGCTTTAGGAAGTTACATATTCAATGGTCCTTATCTCAAAAATGATTATAC
CAACTTAATTAGTAGACAAAATC
CATTAATAGAACACATAAATCTAAAGAAACTAAATATAACACAGTCCTTAATATCTA
AGTATCATAAAGGTGAAATAAAA
ATAGAAGAACCTACTTACTTTCAGTCATTACTTATGACATACAAGAGTATGACCTCG
TCAGAACAGACTACTACTACTAA
TTTACTTAAAAAGATAATAAGAAGAGCTATAGAAATCAGTGATGTCAAAGTCTATG
CTATATTGAATAAACTGGGGCTCA
AAGAAAAAGACAAGATTAAATCCAATAATGGACAAGATGAAGACAACTCAGTCATT
ACTACCATAATCAAAGATGATATA
CTTTTAGCTGTCAAGGATAATCAATCTCATCTTAAAGCAGACAAAAATCAATCCACA
AAACAAAAAGATACAATCAAAAC

For Review

AACACTTTTGAAGAAATTAATGTGTTCGATGCAACATCCTCCATCATGGTTAATACA
TTGGTTTAATTTATACACACAAT
TAAACAGCATATTAACACAATATCGATCTAGTGAGGTAAAAAACCATGGTTTTATAT
TGATAGATAATCATACTCTTAGT
GGATTCCAATTTATTTTGAATCAATATGGTTGTATAGTTTATCATAAGGAACTCAAA
AGAATTACTGTGACAACTTATAA
TCAATTCTTGACATGGAAAGATATTAGCCTTAGTAGATTAAATGTTTGTTTGATTACA
TGGATTAGTAACTGTTTGAACA
CATTAAACAAAAGCTTAGGCTTAAGATGTGGATTCAATAATGTTATCTTGACACAAT
TATTCCTTTATGGAGATTGTATA
CTAAAACTATTCCACAATGAGGGGTTCTACATAATAAAAGAGGTAGAGGGATTTATT
ATGTCTCTAATTTTAAATATAAC
AGAAGAAGATCAATTCAGAAAACGGTTTTATAATAGTATGCTCAACAACATCACAG
ATGCCGCCAACAAAGCTCAAAAAA
ATCTGCTATCAAGAGTATGTCATACATTATTAGATAAGACAATATCAGATAATATAA
TAAATGGCAGATGGATAATTCTA
TTGAGTAAGTTCCTAAAATTAATTAAGCTTGCAGGTGACAATAACCTCAACAATCTG
AGTGAATTATATTTTTTGTTCAG
AATATTTGGACACCCAATGGTAGATGAAAGACAAGCCATGGATGCTGTTAAAGTTA
ATTGCAACGAGACCAAATTTTACT
TGTTAAGTAGTTTGAGTATGTTAAGAGGAGCTTTTATATATAGAATTATAAAAGGGT
TTGTAAATAATTACAACAGATGG
CCTACTTTAAGAAATGCCATTGTCTTACCCTTAAGATGGTTAACTTACTATAAACTAA
ACACTTATCCTTCCTTGTTGGA
ACTTACAGAAAGAGATTTGATTGTTCTATCAGGACTACGTTTCTATCGAGAGTTTCG
GTTGCCTAAAAAAGTGGATCTTG
AAATGATCATAAATGATAAGGCTATATCACCTCCTAAAAATTTAATATGGACTAGTT
TCCCTAGAAATTATATGCCGTCA
CACATACAAAATTATATAGAACATGAAAAATTAAAATTCTCTGATAGTGATAAATCA
AGAAGAGTATTAGAGTATTATTT
AAGAGATAACAAATTCAATGAATGTGATTTACACAACTGTGTAGTTAATCAAAGTTA
TCTTAACAACCCGAATCATGTGG
TATCATTGACAGGCAAAGAAAGAGAACTCAGTGTAGGTAGAATGTTTGCAATGCAA
CCAGGAATGTTCAGACAAGTTCAA
ATATTAGCAGAGAAAATGATAGCAGAAAACATATTACAATTTTTCCCTGAAAGTCTT
ACAAGATATGGTGATCTAGAACT
ACAAAAAATATTAGAATTGAAAGCAGGAATAAGTAACAAATCAAATCGTTACAATG
ATAATTACAACAATTACATTAGTA
AGTGCTCTATCATCACAGATCTCAGCAAATTCAATCAAGCATTTCGATATGAAACAT
CATGTATTTGTAGTGATGTACTG
GATGAACTGCATGGTGTACAATCTCTATTTTCCTGGGGACATTTAACTATTCCTCCTG
TCACAATAATATGCACATATAG
GCATGCACCCCCCTATATAAAGGATCATATTGTAGATCTTAACAATGTAGATGAGCA
AAGTGGACTATATAGATATCATA
TGGGTGGTATTGAAGGGTGGTGTCAAAAACTATGGACCATAGAAGCTATATCACTAT
TAGATCTAATATCTCTCAAAGGG

For Peer Review

AAATTCTCAATTACTGCTTTAATTAATGGTGACAATCAATCAATAGATATAAGTAAA
CCAGTCAGACTCATGGAAGGTCA
AACTCATGCTCAAGCAGATTATATGCTAGCATTAAATAGTCTCAAATTACTGTATAA
AGAGTATGCAGGAATAGGCCACA
AATTAAAAGGAACTGAGACTTATATATCGAGAGATATGCAATTTATGAGTAAAACG
ATCCAACATAACGGTGTATATTAC
CCAGCTAGTATAAAGAAAGTCCTAAGAGTGGGACCGTGGATAAACACTATACTTGA
TGACTTCAAAGTGAGTCTAGAATC
TATAGGTAGTTTGACACAAGAATTAGAATATAGAGGTGAAAGTCTATTATGCAGTTT
AATATTTAGAAATGTATGGTTAT
ATAATCAAATTGCATTACAACTTAAAAATCATGCATTATGTAACAACAAATTATATT
TGGATATATTAAAAGTTCTAAAA
CACTTAAAAACCTTTTTTAATCTTGATAACATTGATACAGCATTAACATTGTATATGA
ATTTGCCCATGTTATTTGGTGG
TGGTGATCCCAACTTGTTATATCGAAGTTTCTATAGAAGAACTCCTGATTTCCTCACA
GAGGCTATAGTTCACTCTGTGT
TCATACTTAGTTATTATACAAACCATGATTTAAAAGATAAACTTCAAGATCTGTCAG
ATGATAGATTGAATAAGTTCTTA
ACATGCATAATCACGTTTGACAAAAACCCCAATGCTGAATTCGTTACATTGATGAGA
GATCCTCAAGCTTTAGGATCTGA
GAGGCAAGCTAAAATTACTAGCGAAATCAATAGACTGGCAGTTACCGAGGTTTTGA
GCACAGCTCCAAACAAAATATTTT
CCAAAAGTGCACAACACTATACCACTACAGAGATAGATCTTAATGATATTATGCAA
AATATAGAACCTACATATCCTCAC
GGGCTAAGAGTTGTTTATGAAAGTTTACCCTTTTATAAAGCAGAGAAAATAGTAAAT
CTTATATCCGGTACAAAATCTAT
AACTAACATACTGGAAAAGACTTCTGCAATAGACTTAACAGATATTGATAGAGCCA
CTGAGATGATGAGGAAAAACATAA
CTTTGCTTATAAGGATATTACCATTAGATTGTAACAGAGATAAAAGAGAAATATTGA
GTATGGAAAACCTAAGTATTACT
GAATTAAGCAAATACGTTAGAGAAAGATCTTGGTCTTTATCCAATATAGTTGGTGTT
ACATCACCCAGTATCATGTATAC
AATGGACATAAAATATACAACAAGCACTATAGCTAGTGGCATAATCATAGAGAAAT
ATAATGTCAACAGTTTAACACGTG
GTGAGAGAGGACCCACTAAACCATGGGTTGGTTCATCTACACAAGAGAAAAAGACA
ATGCCAGTTTATAATAGACAAGTT
TTAACCAAAAAACAGAGAGATCAAATAGATCTATTAGCAAAATTGGATTGGGTGTA
TGCATCTATAGATAACAAGGATGA
ATTTATGGAGGAACTTAGCATAGGAACTCTTGGGTTAACATATGAGAAGGCCAAAA
AATTATTCCCACAATATTTAAGTG
TTAACTATTTGCATCGTCTTACAGTCAGTAGTAGACCATGTGAATTCCCTGCATCTAT
ACCAGCTTATAGAACTACAAAT
TATCACTTTGATACTAGCCCTATTAATCGCATATTAACAAAAAAGTATGGTGATGAA
GATATTGATATAGTATTCCAAAA
CTGTATAAGCTTTGGCCTTAGCTTAATGTCTGTAGTAGAACAATTTACTAATGTATGT
CCTAACAGAATTATTCTCATAC

Peer Review

CCAAGCTTAATGAGATACATTTGATGAAACCTCCCATATTCACAGGTGATGTTGATA
TTCACAAGTTAAAACAAGTGATA
CAAAAACAACATATGTTTTTACCAGACAAAATAAGTTTGACTCAATATGTGGAATTA
TTCTTAAGTAATAAAACACTCAA
ATCCGGATCTAATGTTAATTCTAATTTAATATTGGCGCATAAGATATCTGACTATTTT
CATAATACTTACATTTTAAGTA
CTAATTTAGCTGGACATTGGATTCTTATTATACAACTTATGAAAGATTCTAAGGGTA
TTTTTGAAAAAGATTGGGGAGAG
GGATATATAACTGATCATATGTTCATTAATTTGAAAGTTTTCTTCAATGCTTATAAGA
CATATCTCTTGTGTTTTCATAA
AGGTTACGGCAGAGCAAAGCTGGAGTGTGATATGAATACTTCAGATCTCCTATGTGT
ATTGGAATTAATAGACAGTAGTT
ATTGGAAGTCTATGTCTAAGGTGTTTTTAGAACAAAAAGTTATCAAATACATTCTTA
GCCAAGATGCAAGTTTACATAGA
GTAAAAGGATGTCATAGCTTCAAACTATGGTTTCTTAAACGTCTTAATGTAGCAGAA
TTCACAGTTTGCCCTTGGGTTGT
TAACATAGATTATCATCCAACACATATGAAAGCAATATTAACTTATATTGATCTTGT
TAGAATGGGATTGATAAATATAG
ATAGAATATACATTAAAAATAAACACAAGTTCAATGATGAGTTTTATACTTCTAATC
TGTTTTACATTAATTATAACTTC
TCAGATAATACTCATCTATTAACTAAACATATAAGGATTGCTAATTCCGAATTAGAA
AGTAATTACAACAAATTATATCA
TCCCACACCAGAAACCCTAGAAAATATACTAACCAATCCGGTTAAAAGTAATGGAA
AAAAGACACTGAGTGACTATTGTA
TAGGTAAAAATGTTGACTCAATAATGTTACCATCGTTATCTAATAAGAAGCTTATTA
AATCGTCTACAATGATTAGAACC
AATTACAGCAGACAAGATTTGTATAATTTATTTCCTACGGTTGTGATTGATAAAATT
ATAGATCATTCAGGTAATACAGC
CAAATCTAACCAACTTTACACTACTACTTCTCATCAAATATCCTTAGTGCACAATAG
CACATCACTTTATTGCATGCTTC
CTTGGCATCATATTAATAGATTCAATTTTGTATTTAGTTCTACAGGTTGTAAAATTAG
TATAGAGTATATTTTAAAAGAT
CTTAAAATTAAGGATCCTAATTGTATAGCATTCATAGGTGAAGGAGCAGGGAATTTA
TTATTGCGTACAGTAGTGGAACT
TCATCCTGATATAAGATATATTTACAGAAGTCTGAAAGATTGCAATGATCATAGTTT
ACCAATTGAGTTTTTAAGGCTGT
ACAATGGACATATCAACATTGATTATGGTGAAAATTTGACCATTCCTGCTACAGATG
CAACCAACAACATTCATTGGTCT
TATTTACATATAAAGTTTGCTGAACCTATCAGTCTTTTTGTCTGTGATGCTGAATTGC
CTGTAACAGTCAACTGGAGTAA
GATTATAATAGAGTGGAGCAAGCATGTAAGAAAATGCAAGTACTGTTCTTCAGTTA
ATAAATGTACATTAATAGTAAAAT
ATCATGCTCAAGATGATATCGATTTCAAATTAGACAACATAACTATATTAAAAACTT
ATGTATGCTTAGGCAGTAAGTTA
AAGGGATCTGAAGTTTACTTAGTCCTTACAATAGGTCCTGCAAATGTGTTCCCAGTA
TTTAATGTAGTACAAAATGCTAA

For Peer Review

ATTGATACTATCAAGAACCAAAAATTTCATCATGCCTAAAAAAGCTGATAAAGAGT
CTATTGATGCAAATATTAAGAGTT
TGATACCCTTTCTTTGTTACCCTATAACAAAAAAAGGAATTAATACTGCATTGTCTA
AATTAAAGAGTGTTGTTAGTGGA
GATATACTATCATATTCTATAGCTGGACGTAATGAAGTTTTCAGCAATAAACTTATA
AATCATAAGCATATGAACATCTT
AAAGTGGTTCAATCATGTTTTAAATTTCAGATCAACAGAATTAAACTATAATCATTT
ATATATGGTAGAATCTACTTATC
CTCATCTAAGTGAATTGTTAAACAGCTTGACAACCAATGAACTTAAAAAACTGATTA
AAATCACAGGTAGTTTGTTATAC
AACTTTTATAATGAATAATGAGCAAAAATCTTATAACAAAAATAGCTACACACTAAC
ATTGTATTCAATTATAGTTATTT
AAAATTAATAATTATATAATTTTTAATAACTTCTAGTGAACTAATCCTAAAATTATCA
TTTTGATCTAGGAAGAATAAGT
TTAAATCCAAATCTAATTGGTTTATATGTATATTAACTAAATTACGAGATATTAGTTT
TTGACACTTTTTTTCTCGT
>rsv012
ACGCGAAAAAATGCGTACTACAAACTTGCACATTCGGAAAAAATGGGGCAAATAAG
AATTTGATAAGTGCTATTTAAGTC
TAACCTTTTTAATCAGAAATGGGGTGCAATTCACTGAGCATGATAAAGGTTAGATTA
CAAAATTTGTTTGATAATGACGA
AGTAGCATTGTTAAAAATAACATGTTATACTGACAAATTAATTCTTCTGACTAATGC
ATTAGCCAAAGCAACAATACATA
CAATTAAATTAAACGGCATAGTTTTTATACATGTTATAACAAGCAGTGAAGTGTGCC
CTGATAACAATATTGTAGTGAAA
TCTAACTTTACAACAATGCCAATATTACAAAATGGAGGATACATATGGGAATTGATT
GAATTGACACACTGCTCTCAATT
AAATGGTCTAATAGATGATAATTGTGAAATCAAATTTTCTAAAAGACTAAGTGACTC
AGTAATGACTGATTATATGAATC
AAATATCTGATTTACTTGGGCTTGATCTCCATTCATGAATTATGTTTAGTCTAATTCA
ATAGACATGTGTTTATTACCAT
TTTAGTTAATATAAAACCTCATCAAAGGGAAATGGGGCAAATAAACTCACCCAATC
AATCAAACCATGAGCACTACAAAC
GACAACACCACCATGCAAAGATTGATGATCACAGACATGAGACCCCTGTCGATGGA
TTCAATAATAACATCTCTCACCAA
AGAAATCATCACACACAAATTCATATACTTGATAAACAATGAATGTATTGTAAGAA
AACTCGATGAAAGACAAGCTACAT
TTACATTCCTAGTCAATTATGAGATGAAGCTATTGCACAAAGTAGGGAGTACCAAAT
ACAAGAAATACACCGAATATAAT
ACAAAATATGGCACATTCCCTATGCCTATATTTATCAATCATGGCGGGTTTCTAGAA
TGTATTGGCATCAAGCCTACAAA
ACATACTCCTATAATATACAAATATGACCTCAACCCGTAACTTCCAACAAAAAAACC
AACTCATCCAAACCAAACCATTC
TCCAAACAACAATGCTCAACAGTTAAGAAGGAGCTAATCCATTTTAGTAATTAAAA
ATAAGGGTGAAACCAGTAACATAA

Peer Review

ATTGGGGCAAATACAAAGATGGCTCTTAGCAAAGTCAAGTTGAATGATACATTAAA
TAAGGATCAGCTGCTGTCATCCAG
CAAATACACTATTCAACGTAGTACAGGAGATAATATTGACACTCCCAATTATGATGT
GCAAAAACACCTAAACAAACTAT
GTGGTATGCTATTAATCACTGAAGATGCAAATCATAAATTCACAGGATTAGTAGGTA
TGCTATATGCTATGTCCAGATTA
GGAAGGGAAGACACTATAAAGATACTTAAAGATGCTGGATATCATGTTAAAGCTAA
TGGAGTAGATATAACAACATATCG
TCAAGATATAAATGGAAAGGAAATGAAATTCGAAGTATTAACATTATCAAGCTTGA
CATCAGAAATACAAGTCAATATTG
AGATAGAATCTAGAAAGTCCTACAAAAAAATGCTAAAAGAGATGGGAGAAGTGGCT
CCAGAATATAGGCATGATTCTCCA
GACTGTGGGATGATAATACTGTGTATAGCTGCCCTTGTAATAACCAAATTAGCAGCA
GGAGATAGATCAGGTCTTACAGC
AGTAATTAGGAGGGCAAACAATGTCTTAAAAAACGAAATAAAACGCTACAAGGGCC
TAATACCAAAAGACATAGCCAACA
GTTTTTATGAAGTGTTTGAAAAATACCCTCATCTTATAGATGTTTTTGTGCATTTTGG
CATAGCACAATCATCCACAAGA
GGGGGCAGTAGAGTTGAAGGAATCTTTGCAGGATTGTTTATGAATGCCTATGGTTCA
GGACAAGTAATGCTAAGATGGGG
AGTTTTAGCCAAATCTGTAAAAAATATCATGCTAGGACATGCTAGTGTCCAAGCAGA
AATGGAGCAAGTTGTGGAAGTCT
ATGAGTATGCACAGAAGTTGGGAGGAGAAGCTGGTTTCTACCATATATTGAACAAT
CCAAAAGCATCATTGCTGTCATTA
ACTCAATTTCCTAACTTCTCAAGTGTGGTCCTAGGCAATGCAGCAGGTCTAGGCATA
ATGGGAGAGTATAGAGGTACACC
AAGAAACCAAGATCTCTATGATGCAGCCAAAGCATATGCAGAGCAACTCAAAGAAA
ATGGAGTAATAAACTACAGTGTAT
TAGACTTAACAACAGAAGAATTGGAAGCCATAAAGCATCAACTCAACCCCAAAGAA
GATGACGTAGAGCTTTAAGTTAAC
AAAAAATACGGGGCAAATAAGTCAACATGGAGAAGTTTGCACCTGAATTTCATGGA
GAAGATGCAAATAACAAAGCTACC
AAATTCCTAGAATCAATAAAAGGCAAGTTTGCATCATCCAAAGATCCTAAGAAGAA
AGATAGCATAATATCTGTCAACTC
AATAGACATAGAAGTCACTAAAGAGAGCCCGATAACATCTGGCACCAACATTATCA
ATCCAACAAGTGACGCCGACAGTA
CCCCAGAAACTAAAGCCAACTACCCAAGAAAACCCCTAGTAAGCTTCAAAGAAGAT
CTCACCCCAAGTGATAACCCTTTC
TCTAAGTTGTACAAAGAAACCATAGAAACATTTGATAACAATGAAGAAGAATCTAG
CTACTCATATGAGGAGATCAATGA
CCAAACAAATGACAACATTACAGCAAGACTAGATAGAATTGATGAAAAATTAAGTG
AAATATTAGGAATGCTCCATACAT
TAGTAGTTGCAAGTGCAGGACCTACTTCGGCTCGTGACGGAATAAGAGATGCTATG
GTTGGTCTAAGAGAAGAAATGATA
GAAAAAATAAGAGCAGAAGCATTAATGACCAATGATAGGTTAGAGGCTATGGCAAG
ACTTAGGAATGAGGAAAGCGAAAA

For Peer Review

AATGGCAAAAGACACCTCAGATGAAGTGTCTCTCAATCCAACCTCCAAAAAATTGA
GTGACTTGTTGGAAGACAACGATA
GCGACAATGATCTATCACTTGATGATTTTTGATCAGTGATCAACTCACTCAGCAATC
AACAACATCAATAAAACAGACAC
CAATCCATTGAATCAATTGCCAGACTGAAAAAACAAACATCCATCAGCAGAACCAC
CAGCCAATCAATCAACCAATTGAT
CAATCAGCACCCTGACAAAATTAACAATATAGTAACAAAAAAAGAACAAGATGGGG
CAAATATGGAAACATACGTGAACA
AGCTTCACGAAGGCTCCACATACACAGCAGCTGTCCAGTACAATGTTCTAGAAAAA
GATGATGATCCCGCATCACTAACA
ATATGGGTGCCTATGTTCCAGTCATCTGTGCCAGCAGACTTGCTCATAAAAGAACTT
GCAAGCATCAACATACTAGTAAA
GCAGATCTCTACGCCCAAAGGACCTTCACTACGAGTCACGATCAACTCAAGAAGTG
CTGTGCTGGCTCAAATGCCTAGTA
ATTTCACCATAAGTGCAAATGTATCATTAGATGAAAGAAGCAAATTAGCATATGATG
TAACTACACCTTGTGAAATCAAA
GCATGCAGTTTAACATGCTTAAAAGTAAAAAGTATGTTAACTACAGTCAAAGATCTA
ACCATGAAGACATTCAACCCCAC
TCATGAGATCATTGCTCTATGTGAATTTGAAAATATTATGACATCAAAAAGAGTAAT
AATACCAACCTATCTAAGATCAA
TTAGTGTCAAAAACAAGGATCTGAACTCACTGGAAAATATAGCAACCACCGAATTC
AAAAATGCTATCACCAATGCTAAA
ATTATTCCCTATGCAGGATTAGTGTTAGTTATCACGGTTACTGATAATAAAGGAGCA
TTCAAGTATATCAAGCCACAGAG
TCAATTTATAGTGGATCTTGGTGCCTACCTAGAAAAAGAGAGCATATATTATGTGAC
TACTAATTGGAAGCATACAGCTA
CACGTTTTTCAATCAAACCACTAGAGGATTAAACTCAATTATCAACATTGAATGACA
GGTTCACATATGTCCTCAACTGC
ACACTATATCTAAACATCATAAACATCTACACTACACACTTCATCACACAAACCAAT
CCCACTCAAAATCTAAAATCACT
TCCAGCCATTGTCTGCCAGACCTAGAGTGCGAATAGGTAAATAAAACAAGAATATG
GGGTAAATAGATATCAGTTAGAGT
TCAACCAATCTCAACAACCATCTATACCGCCAATCCAATACATACATTGCAAATCTT
AAAATGGGAAACACATCCATCAC
AATAGAATTCACAAGCAAATTTTGGCCCTATTTTACACTAATACATATGATCTTAAC
TCTAATCTCTTTACTAATTATAA
TCACTATTATGATTGCAATACTAAATAAGCTAAGTGAACATAAAATATTCTGCAACA
AAACTCTTGAACAAGGACAGATG
TATCAAATCGACACATAGTGTTCTCCCATCATGCTGTATCAAATTATAATCTTGTGTA
TGTAAATAAACAAATCCAATCT
TCTCACAGAGTCATGGCATCACAAAACCATGCCAACCATCATGGTAGCATAGAGTA
GTTANTTAAAAATTAACATAATGA
TGAATTATTAGTATGGGATCAAAAACAACATTGGGGCAAATGCAACAATGTCCAAA
AACAAGAATCAACGCACTGCCAGG
ACTCTAGAAAAGACCTGGGATACTCTTAATCATCTAATTGTAATATCCTCTTGTTTAT
ACAAATTAAATTTAAAATCTAT

For Peer Review

AGCACAAATAGCACTATCAGTTTTGGCAATGATAATCTCAACCTCTCTTATAATTGC
AGCCATAATATTCATCATCTCTG
CCAATCACAAAGTTACACTAACAACTGTCACAGTTCAAACAATAAAAAACCACACT
GAGAAAAACATAACCACTTACCTT
ACTCAAGTCTCACCAGAAAGGGTTAGCCCATCCAAACAACCCACAGCTACACCACC
AATCCACACAAACTCAGCCACAAT
ATCACCCAATACAAAATCAGAAACACACCATACAACAGCACAAACCAAAGGCACA
ACCTCTACTCCAACACAGAACAACA
AGCCAAGCACAAAACCACGTCCAAAAAATCCACCAAAAAAAGATGATTACCATTTT
GAAGTGTTCAACTTTGTTCCCTGT
AGTATATGTGGCAACAATCAGCTCTGCAAATCCATTTGCAAAACAATACCAAGCAAT
AAACCAAAGAAAAAACCAACTAC
AAAACCTACAAACAAACCACCTACCAAAACCACAAACAAAAGAGACCCCAAAACA
CTAGCCAAACCACCGAAAAAAGAAA
CCACCATTAACCCAACAAAAAAACCAACCCCCAAGACTACAGAAAGAGACACCAAC
ACCCCACAATCCACTGTGCTCGAC
ATAACCACATCAAAACACACAGAAAGGGACACCAGCACCTCACAATCCATTGCGCT
TGACACAACCACATCAAAACACAC
AACCCAACAGCAATCTCTCTACTCAACCACCCCCGAAAACACACCCAACTCCACAC
AAACACCCACAGCATCCGAGCCCT
CCACATCAAATTCTACCTAAAGACTCCAGTCATATGCTTAGTTATTTAAAAACTACA
TCTTAGCAGAGAACCGTGATCCC
TCAAGCAAGAACGAAATTAAATCTGGGGCAAATAACCATGGAGTTGCTGATCCATA
GATCAAGTGCAATCTTCCTAACTC
TTGCTATTAATGCATTGTACCTTACCTCAAGTCAGAACATAACTGAGGAGTTTTACC
AATCAACATGTAGTGCAGTTAGC
AGAGGTTACTTGAGTGCTTTAAGAACAGGTTGGTATACCAGTGTCATAACAATAGAA
TTAAGTAATATAAAAGAAACCAA
ATGCAATGGAACTGACACTAAAGTAAAACTTATAAAACAAGAATTAGATAAGTATA
AGAATGCAGTAACTGAATTACAGT
TACTTATGCAAAACACACCAGCTGTCAACAACCGGGCCAGAAGAGAAGCACCACAG
TATATGAACTACACAATCAATACC
ACTAAAAACCTAAATGTATCAATAAGCAAGAAGAGGAAACGAAGATTTCTGGGCTT
CTTGTTAGGTGTAGGATCTGCAAT
AGCAAGTGGTATAGCTGTATCCAAAGTTCTACACCTTGAAGGAGAAGTGAACAAGA
TCAAAAATGCTTTGCAGCTTACAA
ACAAAGCTGTAGTCAGTCTATCAAATGGGGTCAGTGTTTTAACCAGCAAAGTGTTAG
ATCTCAAGAATTATATAAACAAC
CAATTATTACCTATAGTAAATCAACAGAGTTGTCGCATATCCAACATTGAAACAGTT
ATAGAATTCCAGCAGAAGAACAG
CAGATTGTTGGAAATCACCAGAGAATTTAGTGTCAATGCAGGTGTAACGACACCTTT
AAGCACTTACATGTTAACAAACA
GTGAGTTACTATCATTAATCAATGATATGCCTATAACAAATGATCAGAAAAAATTAA
TGTCAAGCAATGTTCAGATAGTA
AGGCAACAAAGTTATTCTATCATGTCTATAATAAAGGAAGAAGTCCTTGCATATGTT
GTACAGCTACCTATCTATGGTGT

For Peer Review

AATTGATACACCTTGCTGGAAATTACACACATCACCTCTGTGCACCACCAACATCAA
AGAAGGATCAAATATTTGTTTAA
CAAGGACTGATAGAGGATGGTATTGTGATAATGCAGGATCAGTATCCTTCTTTCCAC
AAGCTGACACTTGTAAAGTACAG
TCCAATCGAGTATTTTGTGACACTATGAACAGTTTGACATTACCAAGTGAAGTCAGC
CTTTGTAACACTGACATATTCAA
TTCCAAGTATGACTGCAAAATTATGACATCAAAAACAGACATAAGCAGCTCAGTAA
TTACTTCTCTAGGAGCTATAGTGT
CATGCTATGGTAAGACTAAATGTACTGCATCCAACAAAAATCGTGGAATTATAAAG
ACATTTTCTAATGGTTGTGATTAT
GTGTCAAACAAAGGAGTAGATACTGTATCAGTGGGCAACACTTTATATTATGTCAAC
AAGCTGGAAGGCAAAAACCTTTA
TGTAAAAGGGGAACCTATAATAAATTACTATGACCCTCTAGTGTTTCCTTCTGATGA
GTTTGATGCATCAATATCTCAAG
TCAATGAAAAAATTAATCAAAGTTTAGCTTTTATTCGTAGATCCGATGAATTATTAC
ATAATGTAAATACTGGAAAATCT
ACTACAAATATTATGATAACTGCAATTATTATAGTAATCATTGTAGTATTGTTATCAT
TAATAGCTATTGGTTTACTGTT
GTATTGCAAAGCCAAAAACACACCAGTTACATTAAGCAAAGACCAACTAAGTGGAA
TCAATAATATTGCATTCAGCAAAT
AGACAAAAAACCACCTGATCATGTTCCAACAACAATCTGCTGACCATCAATCCCAA
ATCAACTTACAACAGATACTTCAA
CATCACAGCACAGGCTGAATCATTTCCTCGCATCATGCTACCCACACAACTAAGCTA
GATCCTTAATTCATAGTTACATA
AAAGCCTAAAATATCGCAATCAACACTAAATCAACACATCATTCACAAAACTAACA
GCTGGGGCAAATATGTCGCGAAGA
AATCCCTGCAAATTTGAGATTAGAGGTCATTGCTTGAATGGTAGAAGATGCCACTAC
AGTCATAATTACTTTGAATGGCC
TCCTCATGCATTGCTAGTGAGGCAAAACTTCATGTTAAACAAGATACTCAAGTCAAT
GGACAAGAGCATAGACACTTTGT
CTGAAATAAGTGGAGCTGCTGAACTTGATAGAACAGAAGAATATGCTCTTGGTATA
GTTGGAGTGCTAGAGAGTTACATA
GGATCTATAAACAACATAACAAAACAATCAGCATGTGTTGCTATGAGTAAACTTCTT
ATTGAGATCAATAGTGATGACAT
TAAAAAGCTGAGAGATAATGAAGAACCCAATTCACCTAAGATAAGAGTGTACAATA
CTGTTATATCATACATCGAGAGCA
ATAGAAAAAACAGCAAGCAAACCATCCATCTGCTTAAACGATTACCAGCAGACGTG
CTGAAGAAGACAATAAAGAACACA
TTGGATATCCACAAAAGCATAACCATAAGCAACCCAAAAGAGTCAACCATAAGTGA
TCAAAATGACCAAACCAAAAATAA
TGATATTACCGGATAAATATCCTTGTAGTATATCATCCATATTGATCTCAAGTGAAA
GCATGATTGCTACATTCAATCAT
AAAGACATATTACAATTTAACCACAACCATTTAGATAACCACCAGTGTTTATTAAAT
CATATATTTGATGAAATTCATTG
GACACCTAAAAACTTATTAGATGCCACTCAACAATTTCTCCAACATCTTAACATCCC
TGAAGATATGTATACAGTATATA

For Peer Review

TATTAGTGTCATAATGCTTGATCATAACGATTCTATATCATCCTACCATAAAACTGTC
TTAATAAAGTTATGGGACAAAA
TGGATCCCATTATTAATGGAAGCTCTGCTAATGTATATCTAACTGATAGTTATCTAA
AAGGTGTTATCTCTTTTTCAGAA
TGTAATGCTTTAGGGAGTTACCTTTTTAACGGCCCTTATCTTAAAAATGATTATACCA
ACTTAATTAGTAGACAAAGCCC
ACTAATAGAGCATATGAATCTAAAAAAACTAACTATAACACAATCATTAATATCTAG
ATACCATAAAGGTGAACTGAAAT
TAGAAGAACCAACTTATTTCCAGTCATTACTTATGACATATAAAAGCATGTCCTCGT
CTGAACAAATTGCTACAACTAAC
TTACTTAAAAAAATAATACGAAGAGCTATAGAAATAAGTGATGTAAAGGTGTACGC
CATCTTGAATAAACTAGGACTAAA
GGAAAAGGACAGAGTTAAGCCCAACAACAATTCAGGTGATGAAAACTCAGTACTTA
CAACCATAATTAAAGATGATATAC
TCTCAGCTGTGGAAAACAATCAATCATATACAAATTCAGATAAAAATTACTCAGTAA
ATCAAAATATCAATATCAAAACA
ACACTCTTAAAAAAGTTGATGTGTTCAATGCAACATCCTCCATCATGGTTAATACAC
TGGTTCAATTTATATACAAAATT
AAATAACATATTAACACAATATCGATCAAATGAGGTAAAAAGTCATGGGTTTATATT
AATAGATAATCAAACTTTGAGTG
GTTTTCAGTTTATTTTAAATCAATATGGTTGCATTGTTTATCATAAAGGGCTCAAAAA
AATTACAACTACTACATACAAT
CAATTTTTGACATGGAAAGACATCAGCCTCAGCAGATTAAATGTTTGCTTAATTACT
TGGATAAGTAATTGTTTAAATAC
ATTAAATAAAAGCTTAGGGTTGAGATGCGGATTCAATAATGTTGTGCTATCACAATT
ATTCCTTTACGGAGATTGTATAC
TGAAATTATTTCATAATGAAGGCTTTTACATAATAAAAGAAGTAGAAGGATTTATTA
TGTCTTTAATTCTAAACATAACA
GAAGAAGATCAATTTAGGAAACGATTTTATAATAGCATGCTAAATAACATCACAGA
TGCAGCTATTAAGGCTCAAAAGGA
TCTACTATCAAGAGTATGTCACACTTTATTAGACAAGACAGTGTCTGATAATATCAT
AAATGGTAAATGGATAATTCTAT
TAAGTAAATTTCTTAAATTGATTAAGCTTGCAGGTGATAATAATCTCAATAACTTGA
GTGAGCTATATTTTCTCTTCAGA
ATCTTTGGACATCCAATGGTTGATGAAAGACAAGCAATGGATGCTGTAAGAATTAA
CTGCAATGAAACTAAGTTCTATTT
ATTAAGTAGCCTAAGTACGTTGAGAGGTGCTTTCATTTATAGAATCATAAAAGGGTT
TGTAAATACCTACAACAGATGGC
CCACTTTAAGGAATGCTATTGTCCTACCTCTAAGATGGTTAAACTATTATAAACTTA
ATACTTATCCATCTCTACTTGAA
ATCACAGAAAATGATTTGATTATTTTATCAGGATTGCGGTTCTATCGTGAGTTTCATC
TGCCTAAAAAAGTGGATCTTGA
AATGATAATAAATGACAAAGCTATTTCTCCTCCAAAAGATCTAATATGGACTAGTTT
TCCTAGAAATTACATGCCATCAC
ATATACAAAATTATATAGAACATGAAAAGTTGAAGTTCTCTGAAAGCGACAGATCA
AGAAGAGTACTAGAGTATTACTTG

For Peer Review

AGAGATAATAAATTCAATGAATGTGATCTATACAATTGTGTAGTTAATCAAAGCTAT
CTCAACAACTCTAATCATGTGGT
ATCACTAACTGGTAAAGAAAGAGAGCTCAGTGTGGGTAGAATGTTTGCTATGCAAC
CAGGTATGTTTAGGCAAATCCAAA
TCTTAGCAGAAAAAATGATAGCCGAAAATATTTTACAATTCTTCCCTGAGAGTTTGA
CAAGATATGGTGATCTAGAGCTT
CAAAAGATATTAGAATTAAAAGCAGGAATAAGCAACAAGTCAAATCGTTATAATGA
TAACTACAACAATTATATCAGTAA
ATGTTCTATAATAACAGATCTTAGCAAATTTAATCAAGCATTTAGATATGAAACATC
ATGTGTCTGCAGTGATGTATTAG
ATGAACTGCATGGGGTACAATCTCTATTCTCTTGGTTGCATTTAACAATACCTCTTGT
CACAATAATATGTACATATAGA
CATGCACCTCCTTTTATAAAGGATCATGTTGTCAATCTTAATGAAGTTGATGAACAA
AGTGGGTTATACAGATATCATAT
GGGTGGTATTGAGGGCTGGTGTCAAAAACTGTGGACCATTGAAGCCATATCATTATT
AGATCTAATATCTCTTAAAGGTA
AATTCTCCATCACAGCTCTGATAAATGGTGATAATCAGTCAATTGATATAAGTAAAC
CAGTTAGACTTATAGAGGGTCAG
ACCCATGCTCAAGCAGATTATTTGTTAGCATTAAATAGCCTTAAATTGCTATATAAA
GAGTATGCAGGCATAGGCCATAA
GCTTAAGGGAACTGAGACCTATATATCTCGAGATATGCAGTTCATGAGCAAAACAA
TCCAGCACAATGGAGTGTACTATC
CAGCCAGTATCAAAAAAGTCCTGAGAGTAGGTCCATGGATAAATACAATACTTGAT
GATTTTAAAGTTAGTTTAGAATCT
ATAGGTAGCTTAACACAGGAGTTAGAATACAGAGGAGAAAGCTTATTATGCAGTTT
AATATTTAGGAATATTTGGTTATA
CAATCAAATTGCTCTGCAACTCCGAAATCATGCTTTATGTAATAATAAGCTATATTT
AGATATATTGAAAGTATTAAAAC
ACTTAAAAACCTTTTTTAATCTTGATAGTATCGATACGGCGTTATCATTGTATATGAA
CTTGCCTATGCTGTTTGGTGGT
GGTGATCCTAATTTGTTATATCGAAGCTTTTATAGGAGAACTCCAGACTTCCTTACA
GAAGCTATAGTACATTCAGTGTT
TGTGTTGAGCTATTATACTGGTCACGATCTACAAGATAAGCTCCAGGATCTTCCAGA
TGATAGACTGAACAAATTCTTGA
CTTGTGTCATCACATTTGATAAAAATCCAAATGCCGAGTTTGTAACATTGATGAGGG
ATCCACAGGCTTTAGGGTCTGAA
AGGCAAGCTAAAATTACTAGTGAGATTAATAGATTAGCAGTAACAGAAGTCTTAAG
TATAGCTCCAAACAAAATATTTTC
TAAAAGTGCGCAACACTATACTACCACTGAGATTGATCTAAATGACATTATGCAAAA
TATAGAACCAACTTACCCTCATG
GATTAAGAGTTGTTTATGAAAGTTTACCTTTTTATAAAGCAGAAAAAATAGTTAATC
TTATATCAGGAACAAAATCCATA
ACTAATATACTTGAAAAAACATCAGCGATAGATACAACTGATATTAATAGGGCTACT
GATATGATGAGGAAAAATATAAC
CTTACTTATAAGGATACTTCCACTAGATTGTAACAAAGACAAAAGAGAGTTATTAAG
TTTAGAAAATCTTAGCATAACTG

For Peer Review

AATTAAGCAAGTATGTGAGAGAAAGATCTTGGTCATTATCCAATATAGTAGGAGTA
ACATCACCAAGTATTATGTTCACA
ATGGACATTAAATATACAACTAGCACTATAGCCAGTGGTATAATTATAGAAAAATAT
AATGTTAATGGTTTAACTCGTGG
TGAAAGAGGACCTACTAAGCCATGGGTAGGTTCATCTACACAGGAGAAAAAAACAA
TGCCAGTGTACAATAGACAAGTTT
TAACCAAAAAGCAAAGAGACCAAATAGATTTATTAGCAAAATTAGACTGGGTATAT
GCATCCATAGACAACAAAGATGAA
TTCATGGAAGAACTGAGTACTGGAACACTTGGACTGTCATATGAAAAGGCCAAAAA
GTTGTTTCCACAATATCTAAGTGT
CAATTATTTACACCGGTTAACAGTCAGTAGTAGACCATGCGAATTCCCTGCCTCAAT
ACCAGCTTATAGAACAACAAATT
ATCATTTTGATACTAGTCCTATCAATCATGTATTAACAGAAAAGTATGGAGATGAAG
ATATCGACATTGTGTTTCAAAAT
TGCATAAGTTTTGGTCTTAGCTTGATGTCAGTTGTGGAACAATTCACAAACATATGT
CCTAATAGAATTATTCTCATACC
GAAGCTGAATGAGATACATTTGATGAAACCTCCTATATTTACAGGAGATGTTGATAT
CATCAAGTTGAAGCAAGTGATAC
AAAAACAGCATATGTTCCTACCAGATAAAATAAGTTTAACCCAATATGTAGAATTAT
TCTTAAGTAACAAAGCACTTAAA
TCTGGATCTCACATCAACTCTAATTTAATATTAGTACATAAAATGTCTGATTATTTTC
ATAATGCGTATATTTTAAGTAC
TAATTTAGCTGGACATTGGATTCTGATTATTCAACTTATGAAAGATTCAAAAGGTAT
TTTTGAAAAAGATTGGGGAGAGG
GGTATATAACTGATCATATGTTCATTAATTTAAATGTTTTCTTTAATGCTTATAAGAC
TTATTTGCTATGTTTTCATAGA
GGTTATGGTAAAGCAAAATTAGAATGTGATATGAACACTTCAGATCTTCTTTGTGTT
TTGGAGTTAATAGACAGTAGCTA
CTGGAAATCTATGTCTAAAGTTTTCCTAGAACAAAAAGTCATAAAATACATAGTCAA
TCAAGACACAAGTTTGCATAGAA
TAAAAGGTTGTCACAGTTTTAAGTTGTGGTTTTTAAAACGCCTTAATAATGCTAAATT
TACCGTATGCCCTTGGGTTGTT
AACATAGATTATCACCCAACACACATGAAAGCTATATTATCTTACATAGATTTAGTT
AGAATGGGGTTAATAAATGTAGA
TAAATTAACCATTAAAAATAAAAACAAATTCAATGATGAATTTTACACATCAAATCT
CTTTTACATTAGTTACAACTTTT
CAGACAACACTCATCTGCTAACAAAACAAATAAGGATTGCTAATTCAGAATTAGAA
GATAATTATAACAAACTATATCAC
CCAACCCCAGAAGCTTTAGAAAATGTATCATTAATCCCTGTTAAAAGTAATAATAGA
ACCAAACCTAAATTTTGTATAAG
TGGAAGTACTGAATCTATGATGACGTCAACATTCTCTAATAAAATGCATATTAAATC
TTCCACTGTTACCACAAGATTCA
ATTATAGCAGACAAGACTTGTACAATTTATTTCCAATTGTTGTGATAGACAGGATTA
TAGATCATTCAGGTAATACAGAA
AAATCTAACCAACTTTATACCACCACTTCACATCAGACATCTTTAGTAAGGAATAGT
GCATCACTTTATTGCATGCTTCC

For Peer Review

TTGGCATCATGTCAATAGATTTAACTTTGTATTTAGTTCCACAGGATGCAAGATCAG
TATAGAATATATTTTAAAAGATC
TTAAGATTAAAGATCCCAGTTGTATAGCATTCATAGGTGAAGGAGCTGGTAACTTAT
TATTACGTACGGTAGTAGAACTT
CATCCTGACATAAGATATATTTACAGAAGTTTAAAAGATTGCAATGATCATAGTTTA
CCTATTGAATTTCTAAGGTTATA
CAACGGGCATATAAACATAGATTATGGTGAGAATTTAACCATTCCTGCTACAGATGC
AACTAACAACATACATTGGTCTT
ATTTACATATAAAATTTGCAGAACCTATTAGTATCTTTGTCTGTGATGCTGAATTACC
TGTCACAGCCAATTGGAGTAAA
ATCATAATTGAATGGAGTAAGCATGTAAGAAAGTGCAAATACTGTTCTTCTGTAAAT
AGATGCATTTTAATTGCAAAATA
CCATGCTCAAGATGATATTGATTTAAAATTAGATAACATTACTATATTAAAAACTTA
TGTGTGCCTAGGTAGCAAGTTAA
AAGGATCTGAAGTTTACTTAGTCCTTACAATAGGCCCTGCAAATATACTTCCTGTTTT
TGATGTTGTGCAAAATGCTAAA
TTGATTCTTTCAAGAACTAAAAATTTCATTATGCCTAAAAAGATTGACAAGGAATCT
ATCGATGCAAATATTAAAAGCTT
AATACCTTTCCTTTGTTACCCTATAACAAAAAATGGAATTAAGACTTCATTGTCAAA
ATTGAAGAGTGTAGTTAATGGAG
ATATATTATCATATTCTATAGCTGGACGTAATGAAGTATTCAGCAACAAGCTTATAA
ACCACAAGCATATGAATATCTTA
AAATGGCTGGATCATGTTTTAAACTTTAGATCAGCTGAACTTAATTACAATCATTTAT
ACATGATAGAGTCCACATATCC
TTACTTGAGTGAATTATTAAATAGTTTAACAACCAATGAGCTCAAGAAGCTGATTAA
AATAACAGGTAGTGTACTATACA
ACCTTCCTAATGAACAGTAACTTAAAATATCATTAACAAGTTTGGTCAAATTTAGAT
GCTAACACATTATTATATTATAG
TTATTAAAAAATATGCAAACTTTTCAATAATTTAGCATATTGATTCCAAAATTATCAT
TTTAGTCTTAAGGGATTAAATA
AAAGTCTAAAACTAACAATCACACATGTGCATTTACAACACAACGAGACATTAGTTT
TTGACACTTTTTTTCTCGT
>rsv013
TCAAATTAAATGGCATTGTATTTGTGCATGTTATTACAAGTAGTGATATTTGCCCTAA
TAATAATATTGTAGTGAAATCC
AATTTCACAACAATGCCAGTATTACAAAATGGAGGTTATATATGGGAAATGATGGA
ATTAACACACTGCTCTCAACCTAA
TGGCCTAATAGATGACAATTGTGAAATTAAATTCTCCAAAAAACTAAGTGATTCAAC
AATGACCAATTATATGAATCAAT
TATCTGAATTACTTGGATTTGACCTCAATCCATAAATCATAATAAATATCAACTAGC
AAATCAATGTCACTAACACCATT
AGTTAATATAAAACTTGACAGAAGATAAAAATGGGGCAAATAAATCAATTCAGCCG
ACCCAACCATGGACACAACACACA
ATGATACCACACCACAAAGACTGATGATCACAGACATGAGACCATTATCACTAGAA
ACTATAATAACATCTCTGACCAGA

For Peer Review

GATATCATAACACATAAATTTATATACTTGATAAATCATGAATGCATAGTGAGAAAA
CTTGATGAAAGACAGGCCACATT
TACATTTCTGGTCAACTATGAAATGAAACTATTGCACAAAGTGGGAAGCACTAAATA
TAAAAAATATACTGAATACAACA
CAAAATATGGCACTTTCCCTATGCCAATATTTATCAATCATGATGGGTTCTTAGAAT
GCATTGGCATTAAGCCTACCAAG
CACACACCCATAATATACAAGTATGATCTCAATCCATGAATATCAAAACAAGATTCA
AACAATCCGAAATAACAACTTTA
TGCATAATCACACTCCATAGTCCAAATGGAGCCTGAAAATTATAGTTATTTAAAATT
AAGGAGAGACATAAGATGAAAGA
TGGGGCAAATACAAAAATGGCTCTTAGCAAAGTCAAGTTGAATGATACACTCAACA
AAGATCAACTTCTATCATCCAGCA
AATATACCATCCAACGGAGCACAGGAGACAGCATTGACACTCCTAATTATGATGTG
CAGAAACACATTAATAAGTTATGT
GGCATGTTATTAATCACAGAAGATGCTAATCATAAATTCACTGGGTTAATAGGTATG
TTATATGCTATGTCTAGATTAGG
AAGAGAAGACACCATAAAAATACTCAAAGATGCGGGATATCATGTTAAGGCAAATG
GAGTGGATGTAACAACACATCGTC
AAGACATTAATGGGAAAGAAATGAAATTTGAAGTGTTAACATTAGCAAGCTTAACA
ACTGAAATTCAAATCAACATTGAG
ATAGAATCTAGAAAATCCTACAAAAAAATGCTAAAAGAAATGGGAGAGGTGGCTCC
AGAATACAGGCATGACTCTCCTGA
TTGTGGGATGATAATATTATGTATAGCAGCATTAGTAATAACCAAATTAGCAGCAGG
AGATAGATCTGGTCTTACAGCTG
TAATTAGGAGAGCTAATAATGTCCTAAAAAATGAAATGAAACGTTATAAAGGTTTA
TTACCCAAGGATATAGCCAACAGC
TTCTATGAAGTGTTTGAAAAATATCCTCACTTTATAGATGTTTTTGTTCATTTTGGTA
TAGCACAATCTTCTACCAGAGG
TGGCAGTAGAGTTGAAGGGATTTTTGCAGGATTGTTTATGAATGCCTATGGTGCAGG
GCAAGTGATGTTACGGTGGGGGG
TCTTAGCAAAATCAGTTAAAAACATTATGTTAGGACACGCTAGTGTACAAGCAGAA
ATGGAACAAGTTGTGGAGGTGTAT
GAGTATGCTCAGAAATTGGGTGGAGAAGCAGGATTCTACCATATATTGAACAACCC
AAAAGCATCACTATTATCTTTGAC
TCAATTTCCTCACTTCTCTAGTGTAGTATTGGGCAATGCTGCTGGCCTAGGCATAATG
GGAGAATACAGAGGTACACCAA
GGAATCAAGATTTATATGATGCTGCAAAAGCATATGCTGAACAACTCAAAGAAAAT
GGTGTGATTAACTACAGTGTATTA
GACTTGACAGCAGAAGAACTAGAGGCTATCAAACATCAGCTTAATCCAAAAGATAA
TGATGTAGAGCTTTGAGTTAATAA
AAAAATAAAAAAATGGGGCAAATAAATCATCATGGAAAAGTTTGCTCCTGAATTCC
ATGGAGAAGATGCAAATAACAGAG
CCACCAAATTCCTAGAATCAATAAAGGGCAAATTCACATCACCCAAAGATCCCAAG
AAAAAAGATAGTATCATATCTGTC
AACTCAATAGATATAGAAGTAACCAAAGAAAGCCCTATAACATCAAATTCAACCAT
TATAAATCCAATAAATGAGACAGA

For Peer Review

TGATACTATAGGGAACAAGCCCAATTATCAAAGAAAGCCTCTAGTAAGTTTCAAAG
AAGACCCTACGCCAAGTGATAATC
CTTTTTCAAAACTATACAAAGAAACCATAGAAACATTTGATAACAATGAAGAAGAA
TCTAGCTATTCATATGAAGAAATA
AATGATCAGACAAACGATAATATAACAGCAAGATTAGATAGGATTGATGAGAAATT
AAGTGAAATACTAGGAATGCTTCA
CACATTAGTAGTAGCGAGTGCAGGACCTACATCTGCTCGGGATGGTATAAGAGATG
CCATGGTTGGTTTAAGAGAAGAAA
TGATAGAAAAAATCAGAACTGAAGCATTAATGACTAATGACAGACTAGAAGCTATG
GCAAGACTCAGGAATGAAGAAAGT
GAAAAGATGGCAAAAGACACATCAGATGAAGTGTCTCTCAATCCAACATCAGAGAA
ACTGAACAACCTGTTGGAAGGGAA
TGACAGTGACAATGATCTATCACTTGAAGATTTCTGATTAGCTACCAAACTGTACAT
CAAAACACAACACCAATAGAAAA
CCAACAAACAAACCAACTCACCCATCCAACCAAACATCTATCTGCCGATTAGCCAA
CCAACCAAAAAACAACCAGCCAAT
CCAAAACTAGCCACCCGGAAAAAATCGATACTATAGTTACAAAAAAAGATGGGGCA
AATATGGAAACATACGTGAACAAA
CTTCACGAGGGCTCCACATACACAGCTGCTGTTCAATACAATGTCCTAGAAAAAGAC
GATGATCCTGCATCACTTACAAT
ATGGGTGCCCATGTTCCAATCATCCATGCCAGCAGATCTACTCATAAAAGAACTAGC
CAATGTCAATATACTAGTGAAAC
AAATATCCACACCCAAGGGACCCTCATTAAGAGTCATGATAAACTCAAGAAGTGCA
GTGCTAGCACAAATGCCCAGCAAA
TTTACCATATGTGCCAATGTGTCCTTGGATGAAAGAAGCAAGCTGGCATATGATGTA
ACCACACCCTGTGAAATTAAGGC
ATGCAGTCTAACATGCCTAAAATCAAAAAATATGTTAACTACAGTTAAAGATCTCAC
TATGAAAACACTCAACCCAACAC
ATGACATCATTGCTTTATGTGAATTTGAAAATATAGTAACATCAAAAAAAGTCATAA
TACCAACATACCTAAGATCTATC
AGCGTCAGAAATAAAGATCTGAACACACTTGAAAATATAACAACCACTGAATTCAA
AAATGCCATTACAAATGCAAAAAT
CATCCCTTACTCAGGATTACTGTTAGTCATCACAGTGACTGACAACAAAGGAGCATT
CAAATATATAAAGCCACAAAGTC
AATTCATAGTAGATCTTGGAGCTTACCTTGAAAAAGAAAGTATATATTATGTTACAA
CAAATTGGAAGCACACAGCTACA
CGATTTGCAATCAAACCCATGGAAGATTAACCTTTTTCCTCTACATCAATGAGTAGA
TTCATACAAACTTTCTAACTACA
TTCTTCACTTCACAATCATAATCACCAACCCTCTGTGGTTCAATCAATCAAACAAAA
CTCATCAGGAGTTCCAGATCATC
CCAAGTCATTATTCATCAGATCCAGTACTCAAATAAGTTAATAAAAAATCCACATGG
GGCAAATAATCATTGAGGGAAAT
CCAACTAATCACAACATCTGTCAACATAGACAAGTCAACACGCTAGATAAAATCAA
CCAATGGAAAATACATCCATAACT
ATAGAATTCTCAAGCAAATTCTGGCCTTACTTTACACTAATACACATGATAACAACA
ATAATCTCTTTGATAATCATAAT

For Peer Review

CTCCATCATGATTGCAATACTAAACAAACTCTGCGAATATAATGTATTCCATAACAA
AACCTTTGAGCTACCAAGAGCTC
GAGTCAATACATAGCATTCACCAATCTGATAGCTCAAAACAGTAACCTTGCATTTGT
AAGTGAACTACCCTCACCTCTTC
ACAAAACCACATCAACATCTCACCATGCAAGCCATCATCTATACCATAAAGTAGTTA
ATTAAAAATAGTCATAACAATGA
ACTAGGATATTAAGACCAAAAACAACGCTGGGGCAAATGCAAACATGTCCAAAACC
AAGGACCAACGCACCGCCAAGACA
CTAGAAAGGACCTGGGACACTCTCAATCATCTATTATTCATATCATCGTGCTTATAC
AAGTTAAATCTTAAATCTATAGC
ACAAATCACATTATCTATTTTGGCAATGATAATCTCAACCTCACTTATAATTGCAGCC
ATCATATTCATAGCTTCAGCAA
ACCACAAAGTCACACTAACAACTGCAATCATACAAGATGCAACGAACCAGATCAAG
AACACAACCCCAACATACCTCACC
CAGAATCCCCAGCTTGGAATCAGCTTCTCCAATCTGTCCGGAACTACATCACAATCC
ACCACCATACTAGCTTCAACAAC
ACCAAGTGCTGAGTCAACCCCACAATCCACAACAGTCAAGATCAAAAACACAACAA
CAACCCAAATACTACCTAGCAAAC
CCACCACAAAACAACGCCAAAATAAACCACAAAACAAACCCAACAATGATTTTCAC
TTTGAAGTGTTCAATTTTGTACCC
TGCAGCATATGCAGCAACAATCCAACCTGCTGGGCCATCTGCAAGAGAATACCAAA
CAAAAAACCTGGAAAGAAAACCAC
CACCAAGCCCACAAAAAAACCAACCATCAAGACAACCAAAAAAGATCCCAAACCTC
AAACCACAAAACCAAAGGAAGTAC
TCACCACCAAGCCTACAGAAAAGCCAACCATCGACACCACCAAAACAAACATCAGA
ACTACACTGCTCACCTCCAACACC
ACAGGAAATCCAGAACACACAAGTCAAGAGGAAACCCTCCACTCAACCACCTCCGA
AGGCAATCTAAGCCCATCACAAGT
CTATACAACATCCGAGTACCTATCACAATCTCCATCTTCATCCAACACAACAAAATG
ATAGTCATTAAAAAGTGTATTGT
TGCAAAAAGCCATGACCAAATCAAACAGAATCAAAATCAACTCTGGGGCAAATAAC
AATGGAGTTGCCAATCCTCAAAAC
AAATGCTATTATCACAATCCTTGCTGCAGTCACACTCTGTTTCGCTTCCAGTCAAAAC
ATCACTGAAGAATTTTATCAAT
CAACATGCAGTGCAGTTAGCAAAGGCTATCTTAGTGCTCTAAGAACTGGTTGGTATA
CTAGTGTTATAACTATAGAATTA
AGTAATATCAAGGAAAATAAGTGTAATGGTACAGACGCTAAGGTAAAATTAATAAA
ACAAGAATTAGATAAATATAAAAA
TGCTGTAACAGAATTGCAGTTGCTCATGCAAAGCACACCAGCAGCCAACAGTCGAG
CCAGAAGAGAACTACCAAGATTTA
TGAATTATACACTCAACAATACCAAAAACACCAATGTAACATTAAGCAAGAAAAGG
AAAAGAAGATTTCTTGGATTTTTG
TTAGGTGTTGGATCTGCAATCGCCAGTGGCATTGCCGTATCTAAGGTCCTGCACCTA
GAAGGGGAAGTGAACAAAATCAA
AAGTGCTCTACTATCCACAAACAAGGCTGTAGTCAGCTTATCTAATGGAGTCAGTGT
CTTAACCAGCAAAGTGTTAGACC

For Peer Review

TCAAAAACTATATAGATAAACAGTTGTTACCTATTGTTAACAAGCAAAGCTGCAGCA
TATCAAACATTGAAACTGTGATA
GAGTTCCAACAAAAGAACAACAGACTACTAGAGATTACCAGAGAATTTAGTGTTAA
TGCAGGTGTAACTACACCTGTAAG
CACTTATATGTTAACTAATAGTGAGTTATTATCATTAATCAATGATATGCCTATAACA
AATGATCAGAAAAAGTTAATGT
CCAGCAATGTTCAAATAGTTAGACAGCAAAGTTACTCTATCATGTCAATAATAAAAG
AGGAAGTCTTAGCATATGTAGTA
CAATTACCACTATATGGTGTAATAGATACTCCTTGTTGGAAACTACACACATCCCCT
CTATGTACAACCAACACAAAGGA
AGGATCCAACATCTGCTTAACAAGAACCGACAGAGGATGGTACTGTGACAATGCAG
GATCAGTATCCTTTTTCCCACAAG
CTGAAACATGTAAAGTTCAATCGAATCGGGTATTTTGTGACACAATGAACAGTTTAA
CATTACCAAGTGAGGTAAATCTC
TGCAATATTGATATATTCAACCCCAAATATGATTGCAAAATTATGACTTCAAAGACA
GATGTAAGCAGCTCCGTTATCAC
ATCTCTAGGAGCCATTGTGTCATGCTATGGCAAAACCAAATGTACAGCATCCAATAA
AAATCGTGGGATCATAAAGACAT
TCTCTAACGGGTGTGATTATGTATCAAATAAGGGGGTGGATACTGTGTCTGTAGGTA
ATACATTATATTATGTAAATAAG
CAAGAAGGCAAAAGTCTCTATGTAAAAGGTGAACCAATAATAAATTTCTATGATCC
ATTAGTGTTCCCCTCTGATGAATT
TGATGCATCAATATCTCAAGTCAATGAGAAAATTAATCAGAGTCTAGCATTTATCCG
TAAATCAGATGAATTATTACATA
ATGTAAATGCTGGTAAATCCACCACAAATATCATGATAACTACTATAATTATAGTAA
TTATAGTAATATTGTTAGCATTA
ATTGCAGTTGGACTGCTTCTATACTGCAAGGCCAGAAGCACACCAGTCACACTAAGT
AAGGATCAACTGAGTGGTATAAA
TAATATTGCATTTAGTAACTGAATAAAAATAGCACCTAATCATATTCTTACAATGGT
TCACTATTTGACCATAGATAACC
CATCTATCATTGGATTATCCTAAAATTTGAACTTCATCACAACTTTCATCTATAAACC
ATCTCACTTACACTATTTAAGT
AGATTCCTATTTTATAGTTATATAAAACAATTGAATACCAAATTAACTTACTATTTGT
AAAAATGAGAACTGGGGCAAAT
ATGTCACGAAGGAATCCTTGCAAATTCGAAATTCGAGGTCATTGCTTGAATGGTAAA
AGGTGTCATTTTAGTCATAATTA
TTTTGAATGGCCACCCCATGCACTACTTGTAAGACAAAACTTTATGTTAAACAGAAT
ACTTAAGTCTATGGATAAAAGCA
TAGATACTTTGTCAGAAATAAGTGGAGCTGCAGAGTTGGACAGAACAGAAGAGTAT
GCCCTCGGTGTAGTTGGAGTGCTA
GAGAGTTATATAGGATCAATAAATAATATAACTAAACAATCAGCATGTGTTGCCATG
AGCAAACTCCTCACTGAACTCAA
CAGCGATGACATCAAAAAACTAAGGGACAATGAAGAGCCAAACTCACCCAAAGTA
AGAGTGTACAATACTGTCATATCAT
ATATTGAAAGCAACAGGAAGAACAATAAACAAACTATCCATCTGTTAAAAAGATTG
CCAGCAGACGTATTGAAGAAAACC

For Peer Review

ATCAAAAACACATTGGATATCCACAAGAGCATAACCATCAATAACCCAAAAGAATC
AACTGTTAGTGATACGAATGACCA
TGCCAAAAATAATGATACTACCTGACAAATATCCTTGTAGTATAAATTCCATACTAA
TAACAAGTAATTGTAGAGTCACT
ATGTATAATCAAAAAAACACACTATATATCAATCAAAACAACCAAAATAACCATAT
ATACCCACCGGATCAACCATTCAA
TGAAATCCATTGGACCTCTCAAGACTTGATTGATGCAACTCAAAATTTTCTACAACA
TCTAGGTATTACTGATGATATAT
ACACAATATATATATTAGTGTCATAATACTCAATCCTAATACTTACCACATCATCAA
ATTATTAACTCAAACAATTCAAG
CTATGGGACAAAATGGATCCCATTATTAGTGGAAATTCTGCTAATGTTTATCTAACT
GATAGTTATTTAAAAGGTGTTAT
TTCTTTCTCAGAATGTAACGCCTTAGGTAGTTACATATTCAATGGTCCTTATCTCAAA
AATGATTATACCAACTTAATTA
GTAGACAAAATCCATTAATAGAACACATAAATCTAAAGAAACTAAATATAACACAG
TCCTTAATATCTAAGTATCATAAA
GGTGAAATAAAAATAGAAGAACCTACTTACTTTCAGTCATTACTTATGACATACAAG
AGTATGACCTCGTCAGAACAGAC
TACTACTACTAATTTACTTAAAAAGATAATAAGAAGAGCTATAGAAATCAGTGATGT
TAAAGTCTATGCTATATTGAATA
AACTGGGGCTCAAAGAAAAAGACAAGATTAAATCCAATAATGGACAAGATGAAGA
CAACTCAGTTATTACTACCATAATC
AAAGATGATATACTTTTAGCTGTCAAGGATAATCAATCTCATCTTAAAGCAGACAAA
AATCAATCCACAAAACAAAAAGA
TACAATCAAAACAACACTTTTGAAGAAATTAATGTGTTCGATGCAACATCCTCCATC
ATGGTTAATACATTGGTTTAACT
TATACGCAAAATTAAACAGCATATTAACACAGTATCGATCTAGTGAGGTAAAAAAC
CATGGTTTTATATTGATAGATAAT
CATACTCTTAGTGGATTCCAATTTATTTTGAATCAATATGGTTGTATAGTTTATCATA
AGGAACTCAAAAGAATTACTGT
GACAACTTATAATCAATTCTTGACATGGAAAGATATTAGCCTTAGTAGATTAAATGT
TTGTTTGATTACATGGATTAGTA
ACTGTTTGAACACATTAAATAAAAGCTTAGGCTTAAGATGTGGATTCAATAATGTTA
TCTTGACACAATTATTCCTTTAT
GGAGATTGTATACTAAAACTATTCCACAATGAGGGGTTCTACATAATAAAAGAGGT
AGAGGGGTTTATTATGTCTCTAAT
TTTAAATATAACAGAAGAAGATCAATTCAGAAAACGGTTTTATAATAGTATGCTCAA
CAACATCACAGATGCCGCTAATA
AAGCTCAGAAAAATCTGCTATCAAGAGTATGTCATACATTATTAGATAAGACAATAT
CAGATAATATAATAAATGGCAGA
TGGATAATTCTATTGAGTAAGTTCCTTAAATTAATTAAGCTTGCAGGTGACAATAAC
CTCAACAATCTGAGTGAATTATA
TTTTTTGTTCAGAATATTTGGACACCCAATGGTAGATGAAAGACAAGCCATGGATGC
TGTTAAAGTTAATTGCAACGAGA
CCAAATTTTACTTGTTAAGTAGTTTGAGTATGTTAAGAGGAGCTTTTATATATAGAAT
TATAAAAGGGTTTGTAAATAAT

For Review

TACAACAGATGGCCTACTTTAAGAAATGCCATTGTTTTACCCTTAAGATGGTTAACT
TACTATAAACTAAACACTTATCC
TTCCTTGTTGGAACTTACAGAAAGAGATTTGATTGTTCTATCAGGACTACGTTTCTAT
CGAGAGTTTCGGTTGCCTAAAA
AAGTGGATCTTGAAATGATCATAAATGATAAGGCTATATCACCTCCTAAAAATTTAA
TATGGACTAGTTTCCCTAGAAAT
TATATGCCGTCACACATACAAAATTATATAGAACATGAAAAATTAAAATTCTCTGAT
AGTGATAAATCAAGAAGAGTATT
AGAGTATTATTTAAGAGATAACAAATTCAATGAATGTGATTTATACAACTGTGTAGT
TAATCAAAGTTATCTTAACAACC
CGAATCATGTGGTATCATTGACAGGCAAAGAAAGAGAACTCAGTGTAGGTAGAATG
TTTGCAATGCAACCAGGAATGTTC
AGACAAGTTCAAATATTAGCAGAGAAAATGATAGCTGAAAACATATTACAATTTTTC
CCTGAAAGTCTTACAAGATATGG
TGATCTAGAACTACAGAAAATATTAGAATTGAAAGCAGGAATAAGTAACAAATCAA
ATCGTTACAATGATAATTACAACA
ATTACATTAGTAAGTGCTCTATCATCACAGATCTCAGCAAATTCAATCAAGCATTTC
GGTATGAAACATCATGTATTTGT
AGCGATGTACTGGATGAACTGCATGGTGTACAATCTCTATTTTCCTGGTTACATTTAA
CTATTCCTCATGTCACAATAAT
ATGCACATATAGGCATGCACCCCCCTATATAAAGGATCATATTGTAGATCTTAACAA
TGTAGATGAGCAAAGTGGATTAT
ATAGATATCATATGGGTGGTATCGAAGGGTGGTGTCAAAAACTATGGACCATAGAA
GCTATATCACTATTAGATCTAATA
TCTCTCAAAGGGAAATTCTCAATTACTGCTTTAATTAATGGTGACAATCAATCAATA
GATATAAGTAAACCAGTCAGACT
CATGGAAGGTCAAACTCATGCTCAAGCAGATTATTTGCTAGCATTAAATAGTCTTAA
ATTACTGTATAAAGAGTATGCAG
GAATAGGCCACAAATTAAAAGGAACTGAGACTTATATATCGAGAGATATGCAATTT
ATGAGTAAAACGATCCAACATAAC
GGTGTATATTACCCAGCTAGTATAAAGAAAGTCCTGAGAGTGGGACCGTGGATAAA
CACTATACTTGATGACTTCAAAGT
GAGTCTAGAATCTATAGGTAGTTTGACACAAGAATTAGAATATAGAGGTGAAAGTC
TATTATGCAGTTTAATATTTAGAA
ATGTATGGTTATATAATCAAATTGCATTACAACTTAAAAATCATGCATTATGTAACA
ACAAATTATATTTGGACATATTA
AAAGTTCTAAAACACTTAAAAACCTTTTTTAATCTTGATAACATTGATACAGCATTA
ACATTATATATGAATTTGCCCAT
GTTATTTGGTGGTGGTGATCCCAACTTGTTATATCGAAGTTTCTATAGAAGAACTCCT
GATTTCCTCACAGAGGCTATAG
TTCACTCTGTGTTCATACTTAGTTATTATACAAACCATGATTTAAAAGATAAACTTCA
AGATCTGTCAGATGATAGATTG
AATAAGTTCTTAACATGCATAATCACGTTTGACAAAAACCCCAATGCTGAATTCGTT
ACATTGATGAGAGATCCTCAAGC
TTTAGGATCTGAGAGGCAAGCTAAAATTACTAGCGAAATCAATAGACTGGCAGTTA
CCGAGGTTTTGAGCACAGCTCCAA

For Peer Review

ACAAAATATTCTCCAAAAGTGCACAACACTATACCACTACAGAGATAGATCTTAATG
ATATTATGCAAAATATAGAACCT
ACATATCCTCACGGGCTAAGAGTTGTTTATGAAAGTTTACCCTTTTATAAAGCAGAG
AAAATAGTAAATCTTATATCCGG
TACAAAATCTATAACTAACATACTGGAAAAGACTTCTGCCATAGACTTAACAGATAT
TGATAGAGCCACTGAGATGATGA
GGAAAAACATAACTTTGCTTATAAGGATATTACCATTAGATTGTAACAGAGATAAA
AGAGAAATATTGAGTATGGAAAAC
CTAAGTATTACTGAATTAAGCAAATACGTTAGAGAAAGATCTTGGTCTTTATCCAAT
ATAGTTGGTGTTACATCACCCAG
TATCATGTATACAATGGACATAAAATATACAACAAGCACTATAGCTAGTGGCATAAT
CATAGAGAAATATAATGTCAACA
GTTTAACACGTGGTGAGAGAGGACCCACTAAACCATGGGTTGGTTCATCTACACAA
GAGAAAAAGACAATGCCAGTTTAT
AATAGACAAGTTTTAACCAAAAAACAGAGAGATCAAATAGATCTATTAGCAAAATT
GGATTGGGTGTATGCATCTATAGA
TAACAAGGATGAATTTATGGAGGAACTTAGCATAGGAACTCTTGGGTTAACATATG
AGAAGGCCAAAAAATTATTTCCAC
AATATTTAAGTGTTAACTATTTGCATCGTCTTACAGTCAGTAGTAGACCATGTGAATT
CCCTGCATCTATACCAGCTTAT
AGAACTACAAATTATCACTTTGATACTAGCCCTATTAATCGCATATTAACAGAAAAG
TATGGTGATGAAGATATTGATAT
AGTATTCCAAAACTGTATAAGCTTTGGCCTTAGCTTAATGTCAGTAGTAGAACAATT
TACTAATGTATGTCCTAACAGAA
TTATTCTCATACCCAAGCTTAATGAGATACATTTGATGAAACCTCCCATATTCACAG
GTGATGTTGATATTCACAAGTTA
AAACAAGTGATACAAAAACAACATATGTTTTTACCAGACAAAATAAGTTTGACTCA
ATATGTGGAATTATTTTTAAGTAA
TAAAACACTCAAATCTGGATCTAATGTTAATTCTAATTTAATATTGGCGCATAAGAT
ATCTGACTATTTTCATAATACTT
ACATTTTAAGTACTAATTTAGCTGGACATTGGATTCTTATTATACAACTTATGAAAG
ATTCTAAGGGTATTTTTGAAAAA
GATTGGGGAGAGGGATATATAACTGATCATATGTTCATTAATTTGAAAGTTTTCTTC
AATGCTTATAAGACATATCTCTT
GTGTTTTCATAAAGGTTACGGCAGAGCAAAGCTGGAGTGTGATATGAATACTTCAGA
TCTCCTATGTGTATTGGAATTAA
TAGACAGTAGTTATTGGAAGTCTATGTCTAAGGTGTTTTTAGAACAAAAAGTTATCA
AATACATTCTTAGCCAGGATGCA
AGTTTACATAGAGTAAAAGGATGTCATAGCTTCAAACTATGGTTTCTTAAACGTCTT
AATGTAGCAGAATTCACAGTTTG
CCCTTGGGTTGTTAACATAGATTATCATCCAACACATATGAAAGCAATATTAACTTA
TATTGATCTTGTTAGAATGGGAT
TAATAAATATAGATAGAATATACATTAAAAATAAACACAAGTTCAATGATGAATTTT
ATACTTCTAATCTTTTTTACATT
AATTATAACTTCTCAGATAATACTCATCTATTAACTAAACATATAAGGATTGCTAAT
TCTGAATTAGAAAGTAATTACAA

For Peer Review

CAAATTATATCATCCTACACCAGAAACCCTAGAAAATATACTAACCAATCCGGTTAA
AAGTAATGATAAAAAGACACTGA
GTGACTATTGTATAGGTAAAAATGTTGACTCAATAATGTTACCATCGTTATCTAATA
AGAAGCTTATTAAATCGTCTACA
ATGATTAGAACCAATTACAGCAGACAAGATTTGTATAATTTATTTCCTACGGTTGTG
ATTGATAAAATTATAGATCATTC
AGGTAATACAGCCAAATCTAACCAACTTTACACTACTACTTCTCATCAAATATCCTT
AGTGCACAATAGCACATCACTTT
ATTGCATGCTTCCTTGGCATCATATTAATAGATTCAATTTTGTATTTAGTTCTACAGG
TTGTAAAATTAGTATAGAGTAT
ATTTTAAAAGATCTTAAAATTAAGGATCCTAATTGTATAGCATTCATAGGTGAAGGA
GCAGGGAATTTATTATTGCGTAC
AGTAGTGGAACTTCATCCTGATATAAGATATATTTACAGAAGTCTGAAAGATTGCAA
TGATCATAGTTTACCAATTGAGT
TTTTAAGGCTGTACAATGGGCATATCAACATTGATTATGGTGAAAATTTGACCATTC
CTGCTACAGATGCAACCAACAAC
ATTCATTGGTCTTATTTACATATAAAGTTTGCTGAACCTATCAGTCTTTTTGTCTGTG
ATGCTGAATTGCCTGTAACAGT
CAACTGGAGTAAGATTATAATAGAGTGGAGCAAGCATGTAAGAAAATGCAAGTACT
GTTCTTCAGTTAATAAATGTACAT
TAATAGTAAAATATCATGCTCAAGATGATATCGATTTCAAATTAGACAACATAACTA
TATTAAAAACTTATGTATGCTTA
GGCAGTAAGTTAAAGGGATCTGAAGTTTACTTAGTCCTTACAATAGGTCCTGCAAAT
GTGTTCCCAGTATTTAATGTAGT
ACAAAATGCTAAATTGATACTATCAAGAACCAAAAATTTCATCATGCCTAAAAAAG
CTGATAAAGAGTCTATTGATGCAA
ATATTAAGAGTTTGATACCCTTTCTTTGTTACCCTATAACAAAAAAAGGAATTAATA
CTGCATTGTCTAAATTAAAGAGT
GTTGTTAGTGGAGATATACTATCATATTCTATAGCTGGACGTAATGAAGTTTTCAGC
AATAAACTTATAAATCATAAGCA
TATGAACATCCTAAAGTGGTTCAATCATGTTTTAAATTTCAGATCAACAGAATTA

Peer Review

Sequence_ID	isolate*	isolation-source*	specific-host*	lab-host*	country	collection-date*	genotype	note*	subtype
rsv011	HRSV/A/PHL/pim16223/2016	wt; sputum	Homo sapiens; 3 months; male		Philippines	20-Dec-2016	ON1	Birthweight and gestational age at birth: 7 lbs, 38 weeks; twin: no; HRSV therapeutic: no; co-infection: group A streptococcus; disease severity: emergency room	A
rsv012	HRSV/B/ARG/352162/2019	wt; nasopharyngeal aspirate	Homo sapiens; 2 years 7 months; female		Argentina: Buenos Aires	May-2019	tbd	Co-morbidities: broncho-pulmonary dysplasia; disease severity: hospital admission	B
rsv013	HRSV/A/Piura.PER/PIU048/2020	tc; nasal swab	Homo sapiens; 87 years; female	HEp-2 cells x2	Peru: Piura	21-Apr-2020	tbd	Co-infection: influenza A virus; disease severity: no medical care	A

* Virtual examples of typical sequence-associated data, not based on existing primary data.

Peer Review

```
>Feature rsv011
45      576     gene
                        gene    NS1
99      518     CDS
                        product         nonstructural protein 1
596     1098    gene
                        gene    NS2
628     1002    CDS
                        product         nonstructural protein 2
1125    2326    Gene
                        gene    N
1140    2315    CDS
                        product         nucleoprotein
2330    3242    Gene
                        gene    P
2347    3072    CDS
                        product         phosphoprotein
3246    4203    Gene
                        gene    M
3255    4025    CDS
                        product         matrix protein
4211    4621    Gene
                        gene    SH
4295    4489    CDS
                        product         small hydrophobic protein
4666    5659    Gene
                        gene    G
4681    5646    CDS
                        product         attachment glycoprotein
5713    7615    Gene
                        gene    F
5726    7450    CDS
                        product         fusion glycoprotein
7660    8620    Gene
                        gene    M2
7669    8253    CDS
                        product         M2-1 protein
8222    8494    CDS
                        product         M2-2 protein
8552    15124   Gene
                        gene    L
8561    15058   CDS
                        product         polymerase protein

>Feature rsv012
```

Peer Review

45	576	gene		
			gene	NS1
99	518	CDS		
			product	nonstructural protein 1
594	1097	gene		
			gene	NS2
626	1000	CDS		
			product	nonstructural protein 2
1124	2326	Gene		
			gene	N
1139	2314	CDS		
			product	nucleoprotein
2330	3243	Gene		
			gene	P
2347	3072	CDS		
			product	phosphoprotein
3253	4206	Gene		
			gene	M
3262	4032	CDS		
			product	matrix protein
4216	4628	Gene		
			gene	SH
4301	4498	CDS		
			product	small hydrophobic protein
4673	5652	Gene		
			gene	G
4688	5620	CDS		
			product	attachment glycoprotein
5705	7603	Gene		
			gene	F
5718	7442	CDS		
			product	fusion glycoprotein
7660	8619	Gene		
			gene	M2
7669	8256	CDS		
			product	M2-1 protein
8222	8494	CDS		
			product	M2-2 protein
8552	15131	Gene		
			gene	L
8560	15060	CDS		
			product	polymerase protein

>Feature rsv013

Peer Review

<1	333	gene		
			gene	NS1 (partial)
<1	275	CDS		
			product	nonstructural protein 1
			codon_start	3
353	855	gene		
			gene	NS2
385	759	CDS		
			product	nonstructural protein 2
882	2085	Gene		
			gene	N
897	2072	CDS		
			product	nucleoprotein
2095	3007	Gene		
			gene	P
2112	2837	CDS		
			product	phosphoprotein
3011	3968	Gene		
			gene	M
3020	3790	CDS		
			product	matrix protein
3976	4385	Gene		
			gene	SH
4060	4254	CDS		
			product	small hydrophobic protein
4430	5351	Gene		
			gene	G
4445	5338	CDS		
			product	attachment glycoprotein
5405	7307	Gene		
			gene	F
5418	7142	CDS		
			product	fusion glycoprotein
7352	8312	Gene		
			gene	M2
7361	7945	CDS		
			product	M2-1 protein
7914	8186	CDS		
			product	M2-2 protein
8245	>14615	Gene		
			gene	L
8253	>14615	CDS		
			product	polymerase protein

Peer Review