

# 1 **Discovery of influenza-like virus clades in invertebrates and the evolutionary history of** 2 **Orthomyxoviridae in metazoans**

3 Ricky Wai Tak Leung<sup>1,2,^</sup>, Ziwei Wu<sup>3,^</sup>, Ling Ming Tsang<sup>4</sup>, Ka Hou Chu<sup>4,5</sup>, Ka Wah Leung<sup>6</sup>, Jing  
4 Qin<sup>1,\*</sup>, Ka Yan Ma<sup>3,\*</sup>

5 <sup>1</sup> School of Pharmaceutical Sciences (Shenzhen), Shenzhen Campus of Sun Yat-sen University, Shenzhen,  
6 Guangdong 518107, China

7 <sup>2</sup> College of Professional and Continuing Education, The Hong Kong Polytechnic University, Kowloon, Hong  
8 Kong

9 <sup>3</sup> State Key Laboratory of Biocontrol, Southern Marine Science and Engineering Guangdong Laboratory  
10 (Zhuhai), School of Ecology, Sun Yat-sen University, Shenzhen, Guangdong 518107, China

11 <sup>4</sup> School of Life Sciences, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong.

12 <sup>5</sup> Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 510301, China

13 <sup>6</sup> Science Unit, Lingnan University, Tuen Mun, New Territories, Hong Kong.

14 <sup>^</sup>Equal contribution. <sup>\*</sup>Corresponding author.

## 16 **Abstract**

17 Epidemics are often initiated by emerging and re-emerging infectious diseases caused by viruses of  
18 animal origin. It is thus important to identify the reservoirs of potentially zoonotic viruses and  
19 understand the dynamics of their host shifts. The flu viruses belong to the virus family  
20 Orthomyxoviridae, which also contains “Isavirus”, “Quaranjavirus” and “Thogotovirus”. Many  
21 members from this virus family are known to be pathogenic to humans. To be well prepared for future  
22 epidemics caused by animal-originated or zoonotic Orthomyxoviridae, this study identified 96 novel  
23 Orthomyxoviridae members with the discovery of two new lineages of virus, one sister to Isavirus and  
24 the other sister to Influenza + Thogotovirus. Furthermore, our phylogenetic analysis did not recognize  
25 a monophyletic Quaranjavirus. By ancestral state and host association analysis, we have discovered at  
26 least four instances of host shifting from arthropods to chordates and one incident of host shifting event  
27 from Cnidaria+Ctenophora to Mollusca. Hence host shifts between Arthropoda and Chordata were the  
28 most frequent, and the Orthomyxoviridae virus is capable of invading diverse hosts from chordates.

## 30 **Background**

31 The current pandemic caused by SARS-CoV-2 virus has spread globally for three years with billions of  
32 cases and millions of deaths so far (Saeed, et al. 2021). The devastating effect has caused us to  
33 understand the terror of emerging infectious diseases. Most if not all emerging and re-emerging  
34 infectious diseases such as SARS, MERS and influenzas are associated with viruses of zoonotic origins,  
35 i.e., that they are transmitted from animals to humans (McArthur 2019; Zhu, et al. 2020). During the  
36 initial outbreak of SARS-CoV-2, the early detection, diagnosis, and treatment strategies were based on

37 the lessons learned from previously known coronaviruses (Guo, et al. 2020), while the origin of  
38 SARS-CoV-2 was linked and identified quickly thanks to the large scale coronavirus screening from  
39 animals over the past two decades (Zhou, et al. 2020). This indicates the importance of screening novel  
40 virus lineages from animals so as to brace ourselves for the next zoonotic-originated pandemic. Another  
41 virus group that is famous for repeatedly jumping from animals to human is the “Influenza”, which  
42 includes the swine flu and avian flu (Dhama, et al. 2012; Tanner, et al. 2015) and has already caused  
43 six pandemics in the last 140 years (Potter 2001). All influenza viruses belong to the virus family  
44 “Orthomyxoviridae” (Dowdle, et al. 1975) which consists of three genera other than the influenzas,  
45 namely “Isavirus”, “Quaranjavirus” and “Thogotovirus”. Both ‘Influenza’ and ‘Isavirus’ are being  
46 hosted only in vertebrates, while Isavirus is infamous for traumatizing the aquaculture industry by  
47 killing farmed fishes (Rimstad and Markussen 2020). For quaranjaviruses and thogotoviruses,  
48 invertebrates are their “reservoirs”, or “natural hosts”. However, both types of viruses are able to cross  
49 species barriers and cause diseases in humans (Taylor, et al. 1966; Kosoy, et al. 2015; Savage, et al.  
50 2018; Mourya, et al. 2019) as well as other vertebrates such as ducks (Kessell, et al. 2012) and sheep  
51 (Lledo, et al. 2020) as vector-borne diseases, capable of inducing pathological condition and cytokine  
52 responses same as the highly virulent influenza (Li, et al. 2008). A plausible explanation is that the viral  
53 proteins might undergo genetic recombination in the shared host due to antigenic drift, and thus evolve  
54 rapidly to evade immune mechanisms (Geoghegan, et al. 2017). Despite the capability of these  
55 Orthomyxoviridae members in causing multiple pandemics through various independent animal to  
56 human jumps, most studies have focused on influenzas only. There are limited studies focusing on the  
57 other three genera (Da Silva, et al. 2005; Briese, et al. 2014; Contreras-Gutierrez, et al. 2017; Cholleti,  
58 et al. 2018; Pettersson, et al. 2020). Moreover, there has been no study on the entire family in a holistic  
59 matter, with no literature review available. Therefore, most members of the Orthomyxoviridae family  
60 and their evolutionary histories are poorly documented.

61 To prepare the world for future pandemics, it is of vital importance to screen and understand potential  
62 emerging infectious viruses and novel members from the Orthomyxoviridae of zoonotic origins. Here,  
63 we identified novel virus species from the emerging infectious virus family – Orthomyxoviridae from  
64 different animal hosts, uncover their evolutionary history and host shift dynamics by analyzing high-  
65 throughput RNA data.

66

## 67 **Results and discussions**

### 68 Phylogenetic relationships within Orthomyxoviridae

69 To assess how Orthomyxoviridae dispersed in metazoans, we analyzed 3813 transcriptomes covering  
70 2827 metazoan species from the Transcriptome Shotgun Assembly (TSA) Database (Sayers, et al. 2019).

71 Then, 114 transcribed RNA sequences from 96 species were identified to be related to  
72 Orthomyxoviridae PB1 genes (RNA-directed RNA polymerase catalytic subunit) which were not  
73 reported before. The known Orthomyxoviridae members were used to infer the classification of the  
74 newly identified viruses [Figure 1]. We found two new lineages of virus, one sister to Isavirus and the  
75 other sister to Influenza + Thogotovirus. These viruses were identified from various metazoan lineages,  
76 including some basal lineages such as Ctenophora (commonly known as comb jellies), Cnidaria  
77 (composed of jelly fishes, hydras and corals), Mollusca (shellfish), and Xenacoelomorpha (specifically,  
78 *Symsagittifera roscoffensis*, which is the only animal that is capable of photosynthesis). Cnidaria and  
79 Ctenophora are two of the most primitive metazoan lineages which have been proposed to be the sister  
80 group of all other animals (bilaterians) and are very important for studying the early evolution of  
81 animals and the origin of multicellular organisms (Rytkonen 2018; Daley and Antcliffe 2019; Nielsen  
82 2019).

83 In addition, we identified new viruses belonging to Thogotovirus and Quaranjavirus. However, our  
84 phylogenetic analysis did not recognize a monophyletic Quaranjavirus.

85 The occurrence of Orthomyxoviridae in multiple ancient metazoan lineages indicates ancestral  
86 Orthomyxoviridae might have been infecting and co-existing with animals since the dawn of Metazoa,  
87 or there might be frequent host shift events happened throughout the co-evolution of animals and  
88 Orthomyxoviridae. Further, apart from newly identified viruses belonged to the named genera of  
89 Orthomyxoviridae, several novel viruses were found to cluster together and formed completely new  
90 clades, suggesting the discovery of new lineages of Orthomyxoviridae. In short, Orthomyxoviridae are  
91 widely dispersed in a wide range of metazoans, including ancient animal lineages.

#### 92 Ancestral states of the novel viruses and host associations analysis

93 We used the maximum parsimony model in Mesquite (Maddison and Maddison, 2023) to infer the  
94 ancestral states (taxonomic group of the host). The ancestor of orthomyxovirids was found to infest  
95 arthropods (predominantly hexapods)[Figure 2]. Subsequently, host shift mainly occurred from  
96 arthropod hosts. We identified at least four incidents of shift from arthropods to chordates with one  
97 incident of host shifting event from Cnidaria+Ctenophora to Mollusca identified. Hence, host shifts  
98 between Arthropoda and Chordata were most frequent, and orthomyxovirid virus is capable of invading  
99 diverse hosts from chordates.

100 Clade A shows a higher frequency of host shift and invading a broader taxonomic spectrum of hosts  
101 than the basal lineages, which consist mainly of Quaranjavirus. The ancestor of Clade A was inferred  
102 to infest chordates. Hence, orthomyxovirids might have evolved the ability of frequent host shift after  
103 invading chordates, and the ability could be attributed to (1) protein evolution that enable to quickly adapt  
104 to new hosts, and (2) the mobility and wide distribution of chordates, when compared with hexapods,

105 facilitated physical dispersal of the virus across taxonomic groups. The two hypotheses are not mutually  
106 exclusive, and might have occurred simultaneously or sequentially with positive feedback. Such  
107 mechanisms, and especially the first hypothesis, warrant empirical tests. Nonetheless, the fast  
108 evolution antigen of influenza A (Morens and Taubenberger 2019), which belong to Clade A, might  
109 lend supports to this hypothesis.

110 Emerging infectious diseases are often characterised by viral host switching events (Geoghegan, et al.  
111 2017), which entails strong and stringent adaptive evolution on the virus as they colonize a new niche  
112 (Simmonds, et al. 2019). Understanding how viruses overcome ecological and genetic barriers in host  
113 shifts is of paramount importance in disease control.

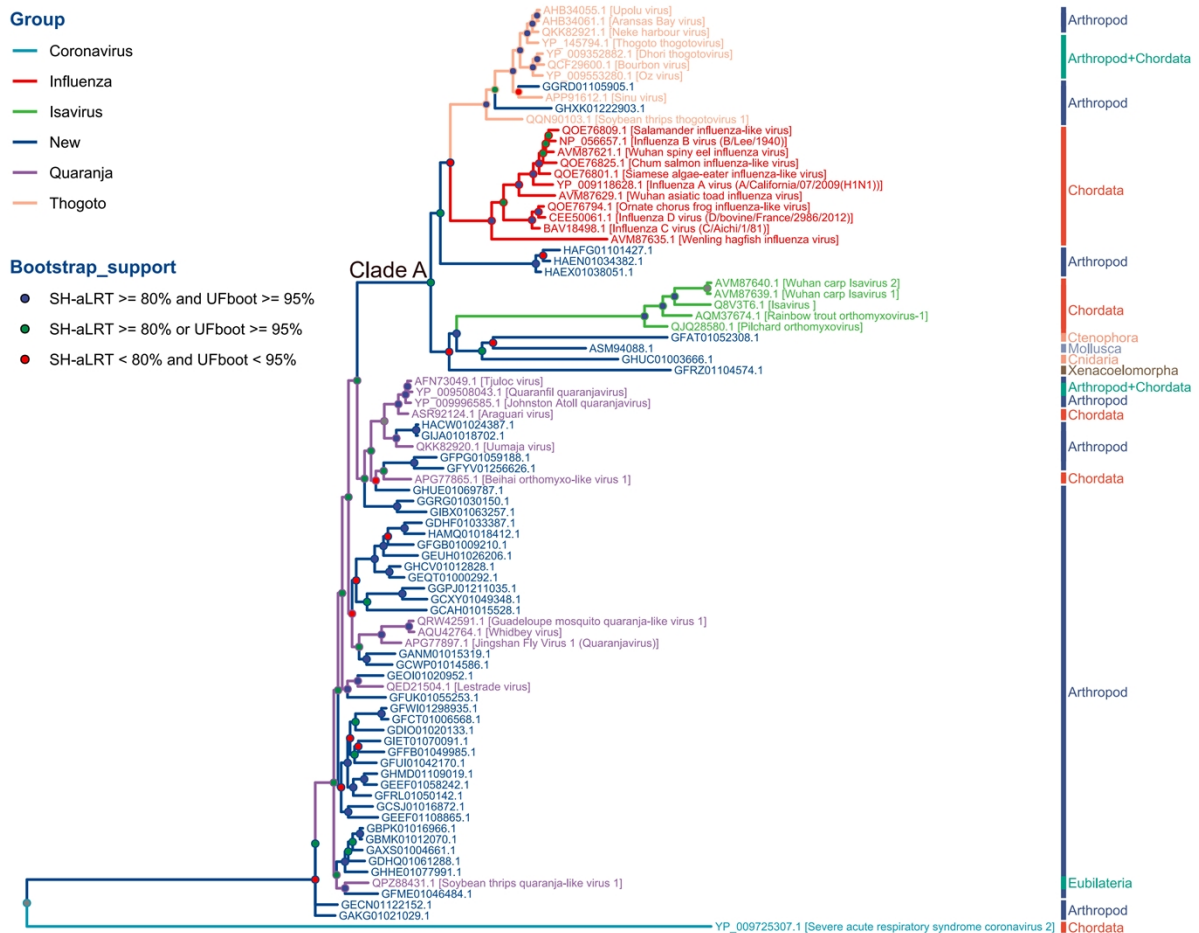
#### 114 Conclusion

115 This study identified 96 novel Orthomyxoviridae members with the discovery of two new lineages of  
116 virus, one sister to Isavirus and the other sister to Influenza + Thogotovirus. The other 89 new viruses  
117 were identified belonging to Thogotovirus and Quaranjavirus. However, our phylogenetic analysis did  
118 not recognize a monophyletic Quaranjavirus [Figure 1]. By ancestral state and host association  
119 analysis, we have discovered at least four instances of host shifting from arthropods to chordates and  
120 one incident of host shifting event from Cnidaria+Ctenophora to Mollusca. Hence host shifts between  
121 Arthropoda and Chordata were the most frequent, and the Orthomyxoviridae virus can infect a variety  
122 of hosts from chordates, while Clade A shows a higher frequency of host shift and invaded a broad  
123 taxonomic spectrum of hosts. [Figure 2].

124

125

126



128

129 Figure 1. Maximum likelihood phylogeny of Orthomyxoviridae. Branch colors and colors of the tip  
 130 labels denote the four orthomyxovirid genera. Viruses that were not named are labelled as “new” and  
 131 in blue color except for ASM94088.1 being named as ‘Barns Ness dog whelk orthomyxo-like virus 1’  
 132 (Waldron, et al. 2018). The GenBank IDs and virus names (where applicable) are shown at the tips.  
 133 Bars on the right represent the hosts infected by the corresponding viruses. Ultrafast bootstrap supports  
 134 of the nodes are indicated by the color of the circle at each node.

135

136

137

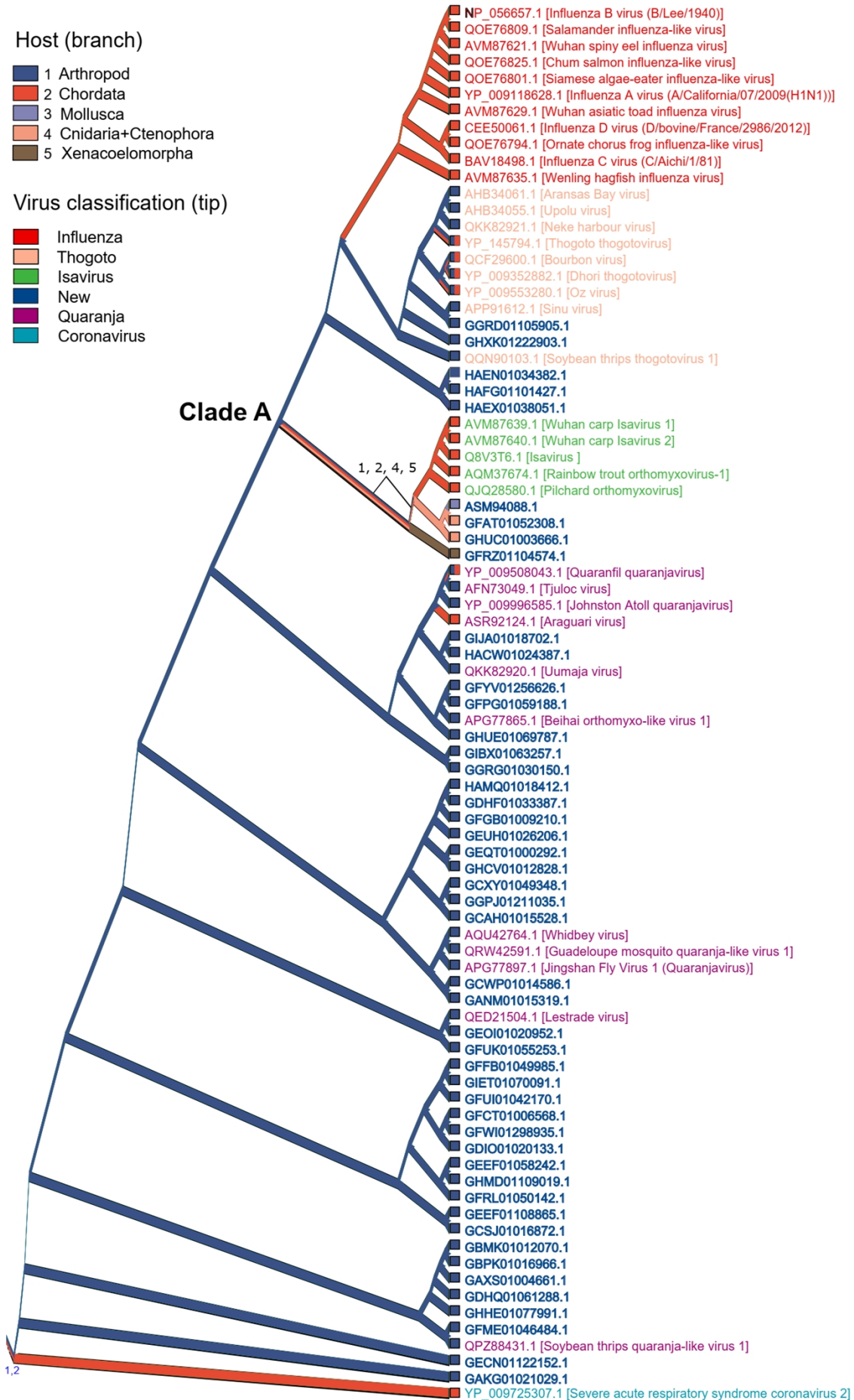
138

Host (branch)

- 1 Arthropod
- 2 Chordata
- 3 Mollusca
- 4 Cnidaria+Ctenophora
- 5 Xenacoelomorpha

Virus classification (tip)

- Influenza
- Thogoto
- Isavirus
- New
- Quaranja
- Coronavirus



140 Figure 2. **Ancestral host reconstruction of orthomyxovirid viruses.** The color of branches  
141 represents host taxonomic group, while the color of tips represents classification of viruses.

142

## 143 **Methods**

### 144 Data collection

145 The PB1 protein, which exhibits RNA-dependent polymerase activity, is the most conserved protein in  
146 Orthomyxoviridae (Chu, et al. 2012), and is thus suitable for evolutionary analyses of viruses, which  
147 generally have high substitution rates. All 3813 transcriptomes covering 2827 metazoan species from  
148 the Transcriptome Shotgun Assembly (TSA) Database (Sayers, et al. 2019) were used for analyzes in  
149 July 2021. These RNA sequences were matched with PB1 protein sequences of known  
150 Orthomyxoviridae downloaded from NCBI virus database in July 2021 (Sayers, et al. 2019) by  
151 TBLASTN (Gertz, et al. 2006) with the expected threshold of 0.05 using BLOSUM62. Then, 114  
152 transcribed RNA sequences from 96 species were identified for the first time to be related to  
153 Orthomyxoviridae PB1 genes (RNA-directed RNA polymerase catalytic subunit) which were not  
154 reported before as belonging to known virus after cross checking them with virus database  
155 [Supplementary table 1]. The taxonomy of the 96 host species was identified by NCBI Taxonomy tool  
156 (Schoch, et al. 2020) [Supplementary figure 1]. These RNA sequences were then translated to protein  
157 sequences according to the matched frame by Clustal X (Larkin, et al. 2007) and aligned by MAFFT  
158 (Kato and Standley 2013) together with other Orthomyxoviridae PB1 proteins downloaded from NCBI  
159 virus database. After comparing with known complete PB1 protein sequences (~700 AA long), 25  
160 newly identified Orthomyxoviridae PB1 proteins with relatively high protein completeness (at least 650  
161 AA) were retained and a neighbor-joining phylogenetic tree with gaps ignored was constructed with  
162 other known Orthomyxoviridae PB1 proteins and SARS-CoV-2 RNA-dependent RNA polymerase as  
163 outgroup [Figure 1]. The known Orthomyxoviridae members were used to position the lineage of the  
164 new viruses.

### 165 Phylogenetic analysis

166 Sequences of PB1 proteins with more than 600 AA identified from NGS data, together with those  
167 available from NCBI virus (Hatcher, et al. 2017) [Supplementary table 2] , were aligned and then  
168 trimmed by ALICUT/ALISCORE (Misof and Misof 2009). The best substitution model was searched  
169 by ModelTest implemented in IQ-Tree 2 (Minh, et al. 2020), which was used for maximum likelihood  
170 phylogenetic reconstruction with 1000 ultrafast bootstrap replicates using the same software .

### 171 Ancestral states of the novel viruses and host associations analysis

172 To trace the evolutionary history of host association of Orthomyxoviridae, we conducted ancestral  
173 character reconstruction in Mesquite 3.8.1 (Maddison and Maddison, 2023) using the maximum  
174 likelihood tree constructed by IQtree. Host association was coded in five categories, namely, Arthropod,  
175 Chordata, Mollusca, Cnidaria+Ctenophora, and Xenacoelomorpha. We used the “trace ancestral  
176 character” function to estimate the ancestral state of the nodes using maximum parsimony model  
177 (Appendices S1).

178

## 179 **Acknowledgement**

180 Ancestral host associations analysis was performed on the Secevo HPC cluster of the School of  
181 Ecology, Sun Yat-sen University (SYSU).

182

## 183 **Supplementary materials**

184 Appendices S1 are available at [https://github.com/ziweiwuzw/Trait-dependent-analyses-](https://github.com/ziweiwuzw/Trait-dependent-analyses-Orthomyxoviridae)  
185 [Orthomyxoviridae](https://github.com/ziweiwuzw/Trait-dependent-analyses-Orthomyxoviridae).

186

## 187 **Funding**

188 The work was supported by grants from Shenzhen Science and Technology Program (project nos.  
189 202206193000001 and 20220817122906001)

190

## 191 **References**

- 192 Briese T, Chowdhary R, Travassos da Rosa A, Hutchison SK, Popov V, Street C, Tesh RB, Lipkin  
193 WI. 2014. Upolu virus and Aransas Bay virus, two presumptive bunyaviruses, are novel members  
194 of the family Orthomyxoviridae. *J Virol* 88:5298-5309.
- 195 Cholleti H, Hayer J, Mulandane FC, Falk K, Fafetine J, Berg M, Blomstrom AL. 2018. Viral  
196 metagenomics reveals the presence of highly divergent quaranjavirus in Rhipicephalus ticks from  
197 Mozambique. *Infect Ecol Epidemiol* 8:1478585.
- 198 Chu C, Fan S, Li C, Macken C, Kim JH, Hatta M, Neumann G, Kawaoka Y. 2012. Functional  
199 analysis of conserved motifs in influenza virus PB1 protein. *PLoS One* 7:e36113.
- 200 Contreras-Gutierrez MA, Nunes MRT, Guzman H, Uribe S, Suaza Vasco JD, Cardoso JF, Popov VL,  
201 Widen SG, Wood TG, Vasilakis N, et al. 2017. Sinu virus, a novel and divergent orthomyxovirus



202 related to members of the genus *Thogotovirus* isolated from mosquitoes in Colombia. *Virology*  
203 501:166-175.

204 Da Silva EV, Da Rosa AP, Nunes MR, Diniz JA, Tesh RB, Cruz AC, Vieira CM, Vasconcelos PF.  
205 2005. Araguari virus, a new member of the family *Orthomyxoviridae*: serologic, ultrastructural,  
206 and molecular characterization. *Am J Trop Med Hyg* 73:1050-1058.

207 Daley AC, Antcliffe JB. 2019. Evolution: The battle of the first animals. *Curr Biol* 29:R257-R259.

208 Dhama K, Verma AK, Rajagunalan S, Deb R, Karthik K, Kapoor S, Mahima, Tiwari R, Panwar PK,  
209 Chakraborty S. 2012. Swine flu is back again: a review. *Pak J Biol Sci* 15:1001-1009.

210 Dowdle WR, Davenport FM, Fukumi H, Schild GC, Tumova B, Webster RG, Zakstelskaja LY. 1975.  
211 *Orthomyxoviridae*. *Intervirology* 5:245-251.

212 Geoghegan JL, Duchêne S, Holmes EC. 2017. Comparative analysis estimates the relative frequencies  
213 of co-divergence and cross-species transmission within viral families. *PLoS Pathog* 13:e1006215.

214 Gertz EM, Yu YK, Agarwala R, Schaffer AA, Altschul SF. 2006. Composition-based statistics and  
215 translated nucleotide searches: improving the TBLASTN module of BLAST. *BMC Biol* 4:41.

216 Guo YR, Cao QD, Hong ZS, Tan YY, Chen SD, Jin HJ, Tan KS, Wang DY, Yan Y. 2020. The origin,  
217 transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update  
218 on the status. *Mil Med Res* 7:11.

219 Hatcher EL, Zhdanov SA, Bao Y, Blinkova O, Nawrocki EP, Ostapchuck Y, Schaffer AA, Brister JR.  
220 2017. Virus Variation Resource - improved response to emergent viral outbreaks. *Nucleic Acids*  
221 *Res* 45:D482-D490.

222 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7:  
223 improvements in performance and usability. *Mol Biol Evol* 30:772-780.

224 Kessell A, Hyatt A, Lehmann D, Shan S, Crameri S, Holmes C, Marsh G, Williams C, Tachedjian M,  
225 Yu M, et al. 2012. Cygnet River virus, a novel orthomyxovirus from ducks, Australia. *Emerg*  
226 *Infect Dis* 18:2044-2046.

227 Kosoy OI, Lambert AJ, Hawkinson DJ, Pastula DM, Goldsmith CS, Hunt DC, Staples JE. 2015.  
228 Novel thogotovirus associated with febrile illness and death, United States, 2014. *Emerg Infect Dis*  
229 21:760-764.

230 Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F,  
231 Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*  
232 23:2947-2948.

233 Li G, Wang N, Guzman H, Sbrana E, Yoshikawa T, Tseng CT, Tesh RB, Xiao SY. 2008. Dhori virus  
234 (*Orthomyxoviridae*: *Thogotovirus*) infection of mice produces a disease and cytokine response  
235 pattern similar to that of highly virulent influenza A (H5N1) virus infection in humans. *Am J Trop*  
236 *Med Hyg* 78:675-680.

237 Lledo L, Gimenez-Pardo C, Gegundez MI. 2020. Epidemiological study of Thogoto and Dhori virus  
238 infection in people bitten by ticks, and in sheep, in an area of Northern Spain. *Int J Environ Res*  
239 *Public Health* 17(7), 2254.

240 McArthur DB. 2019. Emerging infectious diseases. *Nurs Clin North Am* 54:297-311.

241 Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R.  
242 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic  
243 era. *Mol Biol Evol* 37:1530-1534.

244 Misof B, Misof K. 2009. A Monte Carlo approach successfully identifies randomness in multiple  
245 sequence alignments: a more objective means of data exclusion. *Syst Biol* 58:21-34.

246 Morens DM, Taubenberger JK. 2019. Making Universal Influenza Vaccines: Lessons From the 1918  
247 Pandemic. *The Journal of Infectious Diseases* 219:S5-S13.

248 Mourya DT, Yadav PD, Nyayanit DA, Majumdar TD, Jain S, Sarkale P, Shete A. 2019.  
249 Characterization of a strain of quaranfil virus isolated from soft ticks in India. Is quaranfil virus an  
250 unrecognized cause of disease in human and animals?". *Heliyon* 5:e01368.

251 Nielsen C. 2019. Early animal evolution: a morphologist's view. *R Soc Open Sci* 6:190638.

252 Pettersson JH, Ellstrom P, Ling J, Nilsson I, Bergstrom S, Gonzalez-Acuna D, Olsen B, Holmes EC.  
253 2020. Circumpolar diversification of the *Ixodes uriae* tick virome. *PLoS Pathog* 16:e1008759.

254 Potter CW. 2001. A history of influenza. *J Appl Microbiol* 91:572-579.

255 Rimstad E, Markussen T. 2020. Infectious salmon anaemia virus-molecular biology and pathogenesis  
256 of the infection. *J Appl Microbiol* 129:85-97.

257 Rytkonen KT. 2018. Evolution: Oxygen and early animals. *Elife* 7: e34756.

258 Saeed H, Osama H, Madney YM, Harb HS, Abdelrahman MA, Ehrhardt C, Abdelrahim MEA. 2021.  
259 COVID-19; current situation and recommended interventions. *Int J Clin Pract* 75:e13886.

260 Savage HM, Godsey MS, Jr., Panella NA, Burkhalter KL, Manford J, Trevino-Garrison IC, Straily A,  
261 Wilson S, Bowen J, Raghavan RK. 2018. Surveillance for Tick-Borne Viruses Near the Location  
262 of a Fatal Human Case of Bourbon Virus (Family Orthomyxoviridae: Genus Thogotovirus) in  
263 Eastern Kansas, 2015. *J Med Entomol* 55:701-705.

264 Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I. 2019. GenBank. *Nucleic*  
265 *Acids Res* 47:D94-D99.

266 Schoch CL, Ciufu S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, McVeigh R,  
267 O'Neill K, Robbertse B, et al. 2020. NCBI Taxonomy: a comprehensive update on curation,  
268 resources and tools. *Database (Oxford)* 2020.

269 Simmonds P, Aiewsakun P, Katzourakis A. 2019. Prisoners of war — host adaptation and its  
270 constraints on virus evolution. *Nature Reviews Microbiology* 17:321-328.

271 Tanner WD, Toth DJ, Gundlapalli AV. 2015. The pandemic potential of avian influenza A(H7N9)  
272 virus: a review. *Epidemiol Infect* 143:3359-3374.

273 Taylor RM, Hurlbut HS, Work TH, Kingston JR, Hoogstraal H. 1966. Arboviruses isolated from  
 274 ARGAS TICKS IN Egypt: Quarafil, Chenuda, and Nyamanini. Am J Trop Med Hyg 15:76-86.  
 275 Waldron FM, Stone GN, Obbard DJ. 2018. Metagenomic sequencing suggests a diversity of RNA  
 276 interference-like responses to viruses across multicellular eukaryotes. PLoS Genet 14:e1007533.  
 277 Zhou H, Chen X, Hu T, Li J, Song H, Liu Y, Wang P, Liu D, Yang J, Holmes EC, et al. 2020. A  
 278 Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2  
 279 Cleavage Site of the Spike Protein. Curr Biol 30:2196-2203 e2193.  
 280 Zhu Z, Lian X, Su X, Wu W, Marraro GA, Zeng Y. 2020. From SARS and MERS to COVID-19: a  
 281 brief summary and comparison of severe acute respiratory infections caused by three highly  
 282 pathogenic human coronaviruses. Respir Res 21:224.

283

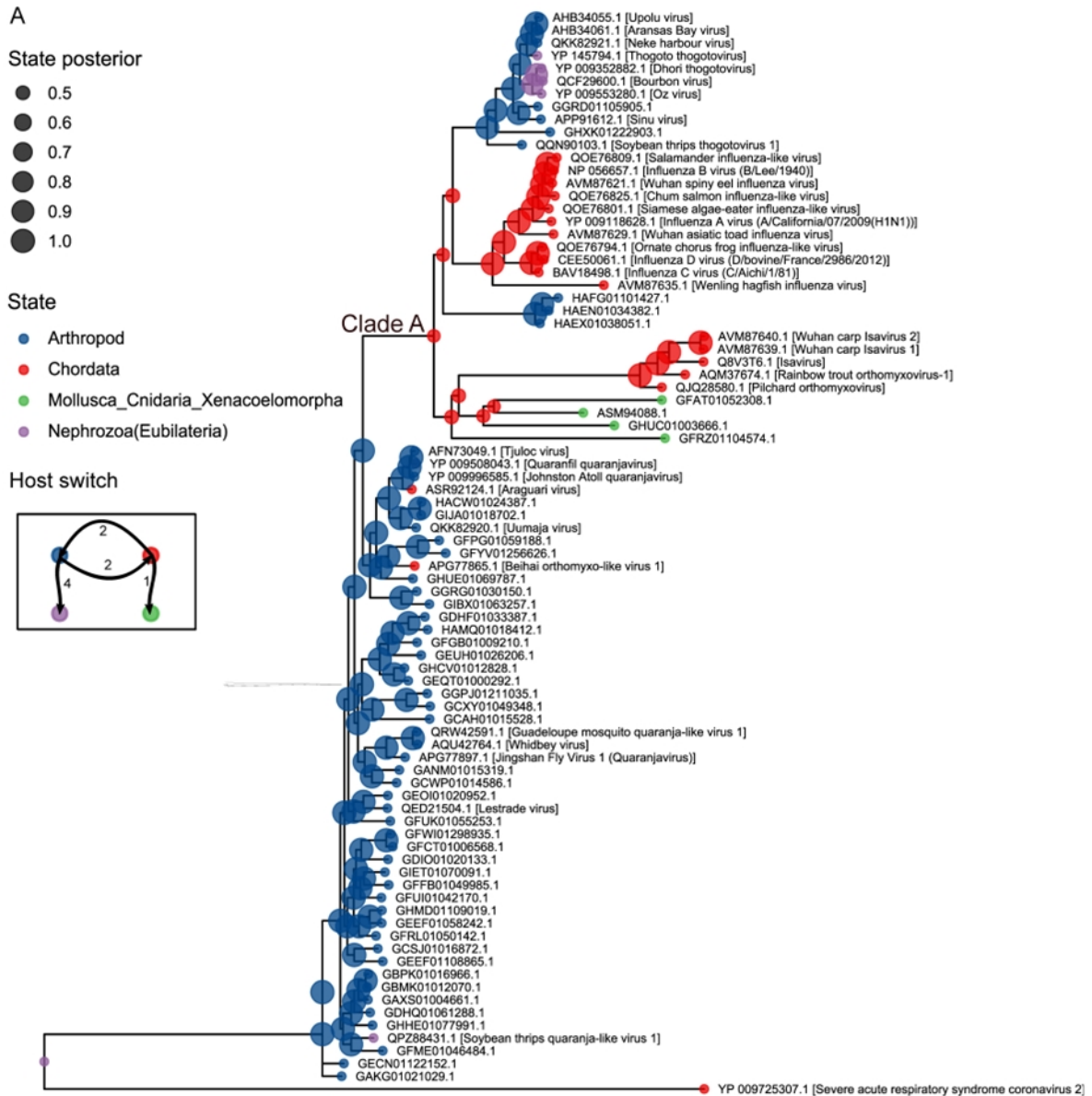
284 Supplementary table 1. **GenBank IDs of Orthomyxoviridae PB1 related RNA sequences.** These  
 285 transcribed RNAs contain 114 hypothetical Orthomyxoviridae PB1 genes that have not been reported  
 286 or identified before.

GenBank ID				
GHYI01006715.1	GEPO01050000.1	HAMG01033019.1	GDNI01055862.1	GFCT01006568.1
GCWF01042188.1	GIBX01063257.1	GFME01046484.1	GEIL01040298.1	GCWP01016768.1
GFUY01045714.1	GAPE01012446.1	GEEF01108865.1	GEIL01040297.1	GCWP01014586.1
GGRD01105905.1	GEQT01000292.1	GEXY01504401.1	GHCV01012828.1	GEOI01020952.1
GDES01014887.1	GHUE01031179.1	GFYV01256626.1	GCAH01015528.1	GCVZ01022955.1
GIWR01317530.1	GDHQ01061288.1	GHUE01031182.1	GFAT01052308.1	GGIX01219007.1
GIWR01068549.1	GFUB01185180.1	GEQP01111649.1	GHCT01108755.1	GCZE01054765.1
GFDB01016754.1	GFUI01042170.1	GFWI01298935.1	IADR01000155.1	GDOD01006372.1
GCXY01049348.1	GIET01070091.1	GEUH01026206.1	GIQZ01347520.1	HAMG01022083.1
GFUY01038943.1	GHZM01177108.1	GFJU01146619.1	GFUK01055253.1	GFFB01049985.1
GCBC01024963.1	GHCU01086526.1	GCSJ01016872.1	GDHF01033387.1	GCYG01007971.1
JL173372.1	GAXS01004661.1	GDIO01020133.1	GGJM01068797.1	GHZM01148481.1
GCXB01020719.1	GDOP01009886.1	GAKG01021029.1	GFRZ01104574.1	GEZB01015213.1
GIJA01018702.1	GHFE01002498.1	GBES01001676.1	GFAT01014782.1	GCUD01022660.1
GHUE01069787.1	GAXD01017262.1	GDPN01021634.1	GHXK01222901.1	GEQP01035960.1
GIQZ01054476.1	GGKE01166827.1	GIOV01193485.1	GHXK01222902.1	GIXZ01027188.1
GIPB01005140.1	GFPG01059188.1	GAZA02096166.1	GGPJ01211035.1	HAEN01034382.1
GACR01007602.1	GGRG01030150.1	GHHE01077991.1	GEEF01058242.1	HAEX01038051.1
GBMK01012070.1	GEBI01000117.1	GGBB01006510.1	GCTA01003798.1	GHUC01003666.1
HACW01024387.1	GHMD01109019.1	HAMQ01018412.1	GCWL01021633.1	JP305384.1
GFRL01050142.1	GEZY01007301.1	GECN01122152.1	GBPK01016966.1	GFRZ01104573.1
GFGB01009210.1	HAMF01039940.1	GCYS01005111.1	GEZY01033642.1	GANM01015319.1
GGEW01122366.1	HAEP01035468.1	HAET01043341.1	HAFI01055399.1	

287

<b>Virus name</b>	<b>Gene name</b>	<b>Sequence ID</b>
Araguari virus	polymerase PB1	ASR92124.1
Aransas Bay virus	PB1 protein	AHB34061.1
Barns Ness dog whelk orthomyxo-like virus 1	putative polymerase PB1 subunit	ASM94088.1
Beihai orthomyxo-like virus 1	polymerase PB1	APG77865.1
Bourbon virus	polymerase subunit PB1	QCF29600.1
Chum salmon influenza-like virus	polymerase basic 1	QOE76825.1
Dhori thogotovirus	polymerase subunit PB1	YP 009352882.1
Guadeloupe mosquito quaranja-like virus 1	polymerase PB1	QRW42591.1
Influenza A virus (A/California/07/2009(H1N1))	polymerase PB1	YP 009118628.1
Influenza B virus (B/Lee/1940)	polymerase PB1	NP 056657.1
Influenza C virus (C/Aichi/1/81)	polymerase 1	BAV18498.1
Influenza D virus (D/bovine/France/2986/2012)	PB1	CEE50061.1
Jingshan Fly Virus 1 (Quaranjavirus)	polymerase PB1	APG77897.1
Johnston Atoll quaranjavirus	polymerase basic 1 protein	YP 009996585.1
Lestrade virus	PB1	QED21504.1
Neke harbour virus	PB1	QKK82921.1
Ornate chorus frog influenza-like virus	polymerase basic 1	QOE76794.1
Oz virus	polymerase basic subunit 1	YP 009553280.1
Pilchard orthomyxovirus	putative basic polymerase 1	QJQ28580.1
Quaranfil quaranjavirus	polymerase PB1	YP 009508043.1
Rainbow trout orthomyxovirus-1	putative polymerase basic 1	AQM37674.1
Salamander influenza-like virus	polymerase basic 1	QOE76809.1
Severe acute respiratory syndrome coronavirus 2	RNA-dependent RNA polymerase	YP 009725307.1
Siamese algae-eater influenza-like virus	polymerase basic 1	QOE76801.1
Sinu virus	PB1 protein	APP91612.1
Soybean thrips quaranja-like virus 1	putative polymerase PB1, partial	QPZ88431.1
Soybean thrips thogotovirus 1	RNA-dependent RNA polymerase subunit PB1	QQN90103.1
Thogoto thogotovirus	PB1 protein	YP 145794.1
Tjuloc virus	polymerase PB1	AFN73049.1
Upolu virus	polymerase basic subunit 1	AHB34055.1
Uumaja virus	PB1	QKK82920.1
Wenling hagfish influenza virus	PB1	AVM87635.1
Whidbey virus	PB1	AQU42764.1
Wuhan asiatic toad influenza virus	PB1	AVM87629.1
Wuhan carp Isavirus 1	PB1	AVM87639.1
Wuhan carp Isavirus 2	PB1	AVM87640.1





296

297 **Supplementary Figure 2. (A) The posterior probabilities of the identification of ancestral states**  
 298 **of the novel viruses in Orthomyxoviridae. The size of the branch's circle represents the branch's**  
 299 **confidence level, that is, the posterior probability.**

300