

1 **Title:** Computational and molecular analysis of conserved influenza A virus RNA
2 secondary structures involved in infectious virion production
3

4 **Authors:** ^{1,5}Yuki Kobayashi, ²Bernadeta Dadonaite, ^{2,3}Neeltje van Doremalen,
5 ⁴Yoshiyuki Suzuki, ²Wendy S. Barclay* and ⁵Oliver G. Pybus*
6

7 **Authors' institutional affiliations and addresses:**

8 ¹Nihon University Veterinary Research Center, Fujisawa, Kanagawa 252-0880, Japan

9 ²Section of Virology, Department of Medicine, Imperial College London, London SW7
10 2AZ, United Kingdom

11 ³ Laboratory of Virology, Division of Intramural Research, National Institute of Allergy
12 and Infectious Diseases, National Institutes of Health, Hamilton, MT, 59840, USA

13 ⁴Graduate School of Natural Sciences, Nagoya City University, Nagoya, Aichi
14 467-8501, Japan

15 ⁵Department of Zoology, University of Oxford, Oxford OX1 3PS, United Kingdom
16

17 **Address for correspondence:** Yuki Kobayashi, Nihon University Veterinary Research

18 Center, 1866 Kameino, Fujisawa, Kanagawa 252-0880, Japan

19 TEL: +81-466-84-3382; FAX: +81-466-84-3380

20 E-MAIL: suzuki.yukikb@nihon-u.ac.jp

21 * These authors contributed equally

22 **Running title:** Conserved RNA secondary structures in IAV M segment

- 23 **Key words:** influenza A virus, M segment, RNA secondary structure, packaging signal,
- 24 stem loop

25 **Abstract**

26 As well as encoding viral proteins, genomes of RNA viruses harbor secondary and
27 tertiary RNA structures that have been associated with functions essential for successful
28 replication and propagation. Here, we identified stem-loop structures that are extremely
29 conserved among 1,884 M segment sequences of influenza A virus (IAV) strains from
30 various subtypes and host species using computational and evolutionary methods. These
31 structures were predicted within the 3' and 5' ends of the coding regions of M1 and M2,
32 respectively, where packaging signals have been previously proposed to exist. These
33 signals are thought to be required for the incorporation of a single copy of eight
34 different negative-strand RNA segments (vRNAs) into an IAV particle. To directly test
35 the functionality of conserved stem-loop structures, we undertook reverse genetic
36 experiments to introduce synonymous mutations designed to disrupt secondary
37 structures predicted at three locations and found them to attenuate infectivity of
38 recombinant virus. In one mutant, predicted to disrupt stem loop structure at nucleotide
39 positions 219-240, attenuation was more evident at increased temperature and was
40 accompanied by an increase in the production of defective virus particles. Our results
41 suggest that the conserved secondary structures predicted in the M segment are involved
42 in the production of infectious viral particles during IAV replication.

43

44

45 **Introduction**

46 Influenza A virus (IAV) is a member of the family *Orthomyxoviridae* and possesses a
47 genome comprising eight segmented, negative-strand RNA molecules (vRNAs), all of
48 which are necessary for productive infection. Each vRNA sequence is composed of one
49 or more open reading frames (ORFs) flanked by untranslated regions (UTRs) that are
50 19-58 nt long. The distal 12 and 13 nucleotides of the 3' and 5' UTRs (U12 and U13,
51 respectively) are highly conserved among different IAV strains and different vRNA
52 segments,^{1,2} and contribute to the formation of a panhandle structure via long-range
53 interactions.^{3,4} The proximal parts of the UTRs contain segment-specific conserved
54 nucleotide sequences that extend into the ORF termini.⁵ In complete virions, the vRNA
55 exists in the form of viral ribonucleoprotein (vRNP), which comprises single molecules
56 of the polymerase subunits PB2, PB1, and PA bound to the panhandle structure, whilst
57 polymerized nucleoproteins (NPs) are attached to the phosphate-sugar backbone of
58 vRNA, thereby leaving vRNA bases exposed.^{4,6-10}

59 Fully infectious IAV virions are enveloped and approximately 100 nm in diameter and
60 are thought to contain a single copy of each of the eight different vRNAs (selective
61 packaging).¹¹⁻¹³ Reverse genetic studies of IAV that employ site-directed mutagenesis
62 have shown that mutations in UTR sequences, as well as synonymous mutations in
63 'segment-specific conserved sequences', reduce the efficiency with which vRNAs are
64 incorporated into progeny particles, suggesting that these sequences may contain the
65 signals controlling the selective packaging (packaging signals).^{5,14-20} A recent study has
66 further demonstrated that conserved sequences in the UTRs may be important for

67 segment incorporation whilst those in ORF termini may regulate bundling of the eight
68 different vRNA segments.²¹ However, the mechanism by which these sequences
69 contribute to the selective packaging of IAV segments is not fully understood.
70 Experimental studies based on RNA hybridization have demonstrated RNA/RNA
71 interactions between different vRNA segments, although some of the interacting regions
72 did not overlap with segment-specific conserved sequences described above.²²⁻²⁵
73 RNA-like string structures that interconnect eight vRNPs were observed within the IAV
74 particle by electron tomography.^{12,22} In addition, kissing loop structures have been
75 predicted at the interacting regions between two vRNA segments, although these
76 intersegment interactions may be strain specific.²⁴ These observations suggest that the
77 selective packaging of IAV may occur through RNA/RNA interactions between
78 different vRNA segments mediated by vRNA secondary structures.
79 In this study we investigate vRNA secondary structures in the M segment of IAV. The
80 M segment is the seventh RNA segment and encodes the M1 and M2 proteins, an
81 mRNA3 and a variety of other proteins that are expressed only by certain IAV strains.
82 M1 is translated from unspliced mRNA whilst M2 and mRNA3 RNAs are differentially
83 spliced.²⁶ Incorporation of the M segment vRNA into the IAV virion has been suggested
84 to be critical to the packaging process^{18,27} and *in silico* analyses have predicted
85 secondary RNA structures in the positive and negative sense RNA sequences of the M
86 segment.²⁸⁻³¹ Here, we use evolutionary and bioinformatic methods to identify stem
87 loop structures in the M segment that are highly conserved among 1,884 IAV strains and
88 that are located in regions previously proposed to contain packaging signals.^{18,20} These

89 *in silico* findings were then followed up with reverse genetics experiments in order to
90 investigate the functional significance of these regions. Specifically, we introduced
91 synonymous mutations intended to disrupt the predicted stem loop structures. We
92 observed that the infectivity of a mutant virus bearing mutations in one of the predicted
93 RNA structures was significantly attenuated and this was accompanied by an increase in
94 the production of defective virus particles. Our results suggest that secondary structures
95 predicted within M segment vRNA may be involved in the production of infectious IAV
96 particles.

97

98 **Results**

99 **Identification and prediction of RNA secondary structures.** The M segment
100 nucleotide sequences of 1,884 IAV strains were analyzed to identify conserved
101 secondary RNA structures (Supplementary Tables 1 and 2). These strains belonged to
102 88 subtypes and were isolated from human (n=443), swine (n=172), avian (n=1,205),
103 and other (n=64) host species. We conducted a sliding window search of these
104 sequences for regions that potentially form secondary structures in either positive
105 (+RNA) or negative (-RNA) sense RNA sequences, using the method implemented in
106 SSE with a window size of 150 nucleotides (see Materials & Methods for details).³² The
107 windows with negative z -score values were predicted to form secondary structures.
108 When we focused on windows with z -scores < -1 , secondary structures were predicted
109 to be formed in both strands at nucleotide positions 115-354 and 835-1014 (Fig. 1B and
110 Table 1; position numbers are given for +RNA), which were located in the M1 and M2

111 coding regions, respectively. Further, positions 115-354 contained two subregions with
 112 z -scores < -1 , at around positions 150 and 250 (Fig. 1B). Similar results were obtained
 113 when the sliding-window search was conducted with a window size of 300 nucleotides
 114 (Supplementary Figure 1). Notably, the regions described above were highly conserved
 115 among all 1,884 sequences (regions with $>98\%$ identity are shown in red in Fig. 1C).
 116 Synonymous substitutions also appeared to be more strongly suppressed in these
 117 regions ($d_s < 0.01$) than in the surrounding regions ($d_s > 0.04$; Fig. 1A), suggesting that
 118 functional constraints operate at the nucleotide sequence level. These observations are
 119 consistent with Moss et al.²⁸
 120 Consensus secondary structures that may be formed in these regions were predicted
 121 using RNAalifold.³³ Analyses were undertaken four times separately for human, swine,
 122 and avian IAVs by randomly choosing 10 strains from each host species
 123 (Supplementary Table 3). Within positions 115-354, a stem-loop structure with
 124 multi-branch loops was predicted to be formed at positions 130-217 of +RNA in human,
 125 swine, and avian IAV strains (Supplementary Figure 2). This structure has been reported
 126 to be involved in the splicing of M segment mRNAs.^{28,29,31} In addition, we newly
 127 identified stem-loop structures (named SL3-10(+) in +RNA and SL3-10(−) in −RNA) at
 128 positions 219-240 in M segment sequences from all three host species (Figs. 1, 2A and
 129 3A; Supplementary Figures 2 and 3). When secondary structures at positions 219-240
 130 were predicted for each of the 1,884 M sequences using RNAfold,³⁴ we found an
 131 exceptional level of structural sequence conservation; every set of complementary
 132 base-pairs that comprise the stems of SL3-10(+) and SL3-10(−) was observed in

133 >99.79% of sequences with minimum free energies ranging from -9.2 to -10.6 kcal/mol
 134 and from -12.6 to -15 kcal/mol, respectively (Figs. 2B and 3B). Further, the frequency
 135 at which the entire structures of SL3-10(+) and SL3-10(-) were identified in human,
 136 swine, and avian IAV strains was 96.7 – 99.6% (Table 2).
 137 The above procedure to identify consensus secondary structures was repeated for
 138 positions 835-1014. In this region we newly identified a conserved stem-loop structure,
 139 named SL5-3B(+), at positions 967-994 in the +RNA of human, swine, and avian IAV
 140 strains (Figs 1B and 4A, Supplementary Figure 4). A consensus structure in –RNA in
 141 the same sequence location as SL5-3B(+) was observed only in human and swine IAV
 142 strains, named SL5-3B(-) (Fig. 5A and Supplementary Figure 5). However, when
 143 secondary structures were predicted at positions 967-994 (the location at which
 144 SL5-3B(-) was predicted in human and swine strains) using 1,884 M
 145 sequences SL5-3B(-) was also predicted to be present in avian strains, suggesting that
 146 surrounding sequences could affect the formation of SL5-3B(-). Each complementary
 147 set of base-pairs that comprise the stems of SL5-3B(+) and SL5-3B(-) was observed in
 148 >99.6% of 1,884 sequences (Figs. 4B and 5B). In the prediction of SL5-3B(+)/(-) using
 149 M sequences at nucleotide positions 967-994, the frequency at which the entire
 150 structures of SL5-3B(+) and SL5-3B(-) were observed in human, swine, and avian IAV
 151 strains was 94.9 – 99% (Table 2). In addition, a small stem loop structure named
 152 SL5-2(+) was newly identified at positions 950-964 in the +RNA of IAVs from all three
 153 host species (Fig. 6A and Supplementary Figure 4). Each pair of bases constituting the
 154 stem of SL5-2(+) was observed in >95.54% of 1,884 sequences (Fig. 6B), and the

155 frequency at which the entire SL5-2(+) structure was seen in human, swine, and avian
156 strains was 92-99.6% (Table 2).

157

158 **Effect on virus growth of predicted stem loop disruption.** SL3-10(-) and SL5-3B(-)
159 were identified as highly-conserved secondary RNA structures in the -RNA sequence of
160 the M segment of IAV. Interestingly, SL5-3B(-) and SL3-10(-) were identified within
161 sequence regions where packaging signals have been inferred by reverse genetic studies
162 to be located (yellow boxes in Fig. 1C).^{18,20} In order to examine whether these
163 secondary structures were indeed involved in replication of IAV, three mutant viruses,
164 SL3-10, SL5-3B and SL5-2 were constructed from A/England/195/2009(H1N1)
165 (hereafter termed 195E) by introducing synonymous mutations in the M1 or M2 coding
166 region that disrupt the structures of SL3-10(+)/(-), SL5-3B(+)/(-) and SL5-2(+),
167 respectively (Figs. 2C-6C). It should be noted that SL3-10(+)/(-), SL5-3B(+), and
168 SL5-2(+) were observed in 195E even when nearly full length of M nucleotide
169 sequences (positions 13-1015) were used for the secondary structure prediction
170 (Supplementary Figures 6 and 7), while SL5-3B(-) was observed with the prediction
171 using local nucleotide sequences at positions 967-994 (Fig. 5A).

172 MDCK cells were infected at low multiplicity (MOI=0.001) with SL3-10, SL5-3B,
173 SL5-2, or 195E. The multicycle growth curves of these viruses were compared by
174 titrating the infectious virus in infected cell supernatants collected at 24 h, 48 h, and 74
175 h post infection (pi). We found that titers for SL3-10 and SL5-3B viruses were
176 approximately 5-fold lower than those for 195E after 24 h pi when incubated at 37°C

177 (Fig. 7A), while the titer of SL5-2 was comparable to 195E, indicating that the
178 secondary structures SL3-10 and SL5-3 disrupted in the mutant viruses were important
179 for productive infection of IAV. The differences in titers between SL3-10 and 195E
180 were even greater and statistically significant when the incubation temperature was
181 raised from 37°C to 39°C (Fig. 7B; two-way ANOVA; $p<0.05$). Similarly, we observed
182 significant attenuation of SL3-10 in a single step growth curve obtained at 39°C (Figs.
183 7C and 7D). Since the most significant phenotypic effect was observed in SL3-10, all
184 following analyses were performed using SL3-10 only.

185

186 **Increase in defective particle production of SL3-10 infection.** Bearing in mind the
187 location of the SL3-10 structure, we reasoned that attenuation of the SL3-10 mutant
188 may be due to incomplete packaging of vRNA segments, which in turn would lead to
189 production of semi – infectious (SI) particles, which are non-infectious particles with
190 fewer vRNA segments.³⁵⁻³⁷ When the total number of virus particles released from
191 MDCK cells infected with SL3-10 or 195E at 39°C was quantified using a
192 hemagglutination assay (HA), both viruses produced similar numbers of particles (Fig.
193 8A), but 195E particles exhibited almost 10 times greater infectivity than SL3-10 (Figs.
194 7B and 7D). For the same PFU count, the HA titer of SL3-10 was $\text{Log}2^4$, while 195E
195 did not contain sufficient virus particles to register in the assay (Fig. 8B), indicating that
196 SL3-10 produced defective particles more abundantly than 195E.
197 When the virus particles with the same PFU count were quantified using a fetuin-based
198 capture ELISA using HA antibodies, the OD values in SL3-10 were 3 times higher than

199 those in 195E (Fig. 8C), and a >50 fold dilution of SL3-10 was required to give an
200 absorbance reading equivalent to that of 195E. Taken together, these data suggest that
201 SL3-10 virus preparations produced in MDCK cells at 39°C contained a high number of
202 noninfectious particles, indicating that the conserved secondary structure SL3-10
203 contributes to the production of infectious IAV virions.

204

205 **Discussion**

206 Many RNA viruses with segmented genomes are believed to package a single copy of
207 each genomic segment into the virus particle and the contribution of vRNA secondary
208 structures to the selective packaging of vRNA segments has been suggested for several
209 segmented RNA viruses.³⁸⁻⁴² In IAVs, many studies have shown that selective
210 packaging may occur through direct interactions between different vRNAs.²²⁻²⁵ The
211 proposal that vRNAs form secondary structures in vRNPs was previously controversial
212 because NP can melt secondary structures in vRNAs.⁸ However, several studies have
213 indicated the presence of RNA secondary structures in vRNPs^{7,10,25} and the observation
214 of intermolecular connections among the eight vRNPs within IAV virions suggests that
215 vRNA secondary structures may be involved in RNA/RNA interactions during the
216 selective packaging of IAV segments.²²

217 In our analysis of 1,884 M nucleotide sequences belonging to 88 IAV subtypes, we
218 predicted the existence of highly conserved stem-loop structures (SL3-10(+)/(-) and
219 SL5-3B(+)/(-)), which were important for productive infection of IAV. SL3-10(+)/(-)
220 and SL5-3B(+)/(-) were present in the 5' and 3' ends of the coding regions for M1 and

221 M2, respectively, where packaging signals have been inferred to be located.^{18,20} Since
222 the body temperatures at which IAVs replicate vary between ~33°C and ~40°C among
223 host species⁴³, the stability of a particular vRNA structure may vary in different
224 hosts.^{30,45,46} However, stem base-pairings in SL3-10(+)/(-) and SL5-3(+)/(-) are retained
225 in >99% of 1,884 sequences, and *in silico* analysis predicts these stem-loop structures
226 are stable under different temperature settings (37°C and 40°C) (data not shown),
227 suggesting that they are common and conserved structures within the M segment among
228 IAVs, and may play a role in productive IAV infection even in avian hosts whose body
229 temperature is higher than in mammals. Indeed, our experiments demonstrate that
230 synonymous mutations that disrupt SL3-10(+)/(-) impair the replication of IAV isolate
231 195E. The attenuation of SL3-10 was accompanied by an increase in the production of
232 defective virus particles, probably SI particles. Similar phenomena were observed when
233 synonymous mutations were introduced into other nucleotide sites in the packaging
234 signals, including those of M segment.^{18,21,25}

235 In addition, some nucleotide positions in SL5-3B(-) have been shown to be important
236 for the packaging and accumulation of vRNA segments in A/PR/8/34.¹⁸ Specifically,
237 synonymous mutations in conserved codons encoding H90-V92 in M2 appear to reduce
238 virus replication and vRNA packaging, and these are located within the stem and loop
239 of SL5-3B(-) (bases with green-filled circles in Fig. 5A). We introduced into 195E
240 synonymous mutations at different nucleotide positions to those reported previously and
241 which disrupt SL5-3B(+)/(-). This resulted in a reduction of infectivity in the mutant
242 virus. Our results thus support the idea that formation of secondary structures in vRNA

may be involved in the genome packaging of IAV. Defects in the M segment might impair the incorporation of M or other specific vRNA segments into virus particles.^{18,27} Some previous studies report that the M segment interacts directly with NA and HA segments, although the interacting regions between segments may vary depending on IAV strain,^{22,24,46} and neither SL3-10(-) nor SL5-3B(-) overlap with previously described interacting regions. Further, it is possible that SL3-10(+) affects the expression level of M1 and M2 proteins, which may have an effect on assembly and budding of virus particles.⁴⁷ It would be interesting to analyze the relative copy numbers of all eight vRNA segments within virus particles to further elucidate the functions of the highly-conserved SL5-3B(-) and SL3-10(-) structures.

253

254 **Materials and Methods**

255 *Influenza virus sequences*

A total of 12,984 complete nucleotide sequences for M segment RNA (~1002 nt long) were retrieved from the Influenza Virus Resource on 1st March 2012.⁴⁸ After eliminating multiple sequences isolated from the same host individual and sequences containing ambiguous bases and minor gaps, 1,884 sequences were retained for analysis (Supplementary Tables 1 and 2). Most sequences were isolated from human (n=443), avian (n=1,205) or swine (n=172) hosts and the data set included a total of 88 different IAV subtypes (Supplementary Table 1). Multiple alignment of the data set was undertaken using MAFFT.⁴⁹ Nucleotide positions corresponding to U12 and U13 were removed because they are known to form a panhandle structure.

265

266 ***Searching for nucleotide sequence conservation***

267 Sequence identity at each nucleotide position was calculated from the base frequencies
268 (A, T, C, or G) calculated from the abovementioned multiple alignment. For the
269 non-overlapping coding regions of M1 and M2, synonymous site diversity (d_s) within a
270 sliding window was calculated using the program ADAPTSITE⁵⁰; the window was 5
271 codons wide and was moved in steps of 1 codon. The d_s value for each window was
272 calculated by dividing the total number of synonymous differences by the number of
273 synonymous sites between each pair of sequences, and the average of these pairwise
274 values was shown in Fig. 1A.

275

276 ***Identification and prediction of RNA secondary structures***

277 Nucleotide positions in the M segment RNAs that possibly contain secondary structural
278 elements were predicted by calculating folding energies along the sequence. For both
279 the negative and positive sense sequences of all M segments, we calculated the
280 minimum free energy (MFE) within a sliding window using the UNAFOLD program
281 implemented in the SSE package.^{32,51} The window sizes for this analysis were 150 or
282 300 nucleotides, and the step size was 30 nucleotides. To obtain z -scores of MFE for
283 each window, we randomized the nucleotide sequence within each window.
284 Randomization was replicated 100 times using the dinucleotide model.
285 In order to identify putative secondary structures that were conserved among different
286 IAV strains, we randomly chose 10 sequences each from human, swine, and avian hosts.

287 This subset of sequences was then used to predict consensus secondary structures with
288 RNAalifold.³³ Further, the secondary structures of each of the 1,884 M segments were
289 predicted using RNAfold.³⁴

290

291 ***Tissue culture***

292 Madin Darby canine kidney (MDCK) cells were cultured in Dulbecco's Modified
293 Eagle's Medium (DMEM), supplemented with 10% fetal calf serum (FCS), 1%
294 penicillin-streptomycin (P/S) mix, and 1% of nonessential amino acids (NEAA)
295 (Gibco-Life technologies) and incubated at either 37°C or 39°C in a 5% CO₂
296 atmosphere.

297

298 ***Construction of reverse genetics viruses***

299 The reverse genetics viruses were generated using a 12 plasmid reverse genetics
300 system synthesized (GeneArt) directly from the A/England/195/2009 whole-genome
301 sequence as previously described.⁵² The mutants were generated by site-directed
302 mutagenesis (Stratagene Lightning mutagenesis kit) of A/England/195/2009
303 segment 7 plasmid sequence. The plasmids were sequenced to confirm the presence
304 of the desired mutations. The 12 plasmids were transfected into 293T cells which
305 were subsequently co-cultured with MDCK cells. Viruses generated were passaged
306 one time in MDCK cells at 37°C in DMEM in the presence of TPCK treated trypsin
307 (Worthington), 1% NEAA and 1% P/S mix and titrated in triplicate by plaque assay on
308 MDCK cells.

309

310 ***Cell infection***

311 Confluent MDCK cells were washed with PBS, then inoculated with 100 µl of virus
312 at the appropriate MOI (0.001 or 1) and incubated at 37°C for 1 h. After incubation
313 the inoculum was removed and cells were washed with PBS (for single-step growth
314 curves an additional acid wash was included with PBS at pH=4.8). The cells were
315 incubated with DMEM supplemented with 1% NEAA and 1% P/S at appropriate
316 temperature. 300 µl of cell supernatant was taken from each well at required time
317 point and replaced by 300 µl of DMEM. The viral yield was assessed by plaque assay
318 on MDCK cells.

319

320 ***Plaque assay***

321 Confluent MDCK cells were washed with PBS then inoculated with 100 µl of
322 tenfold serial dilutions in DMEM of the supernatant from infected cells and
323 incubated at 37°C for 1 h. Subsequently, the inoculum was removed and agar
324 overlay was applied to each well as previously described.⁵³ The plaques formed after
325 three days were then counted.

326

327 ***Hemagglutination assay***

328 Virus stocks were serially 2 fold diluted in 50 µl of PBS in 96 well v bottomed plate
329 starting at 1:2 (2¹). Chicken red blood cells (Harlan laboratories UK Ltd) were
330 diluted to 0.5% in PBS and 100 µl were applied to each well. Plates were incubated

331 for 1 h before pelleted and suspension cells were counted.

332

333 ***Fetuin ELISA***

334 This method was modified from Gambaryan and Matrosovich⁵⁴. 25 ng/ml fetuin
335 solution (Sigma-Aldrich) was prepared in 0.1 M carbonate/bicarbonate buffer
336 (Sigma-Aldrich) and 96-well Greiner medium binding plates were coated with the
337 solution. 195E and SL3-10 viruses normalized to be of equal PFU were diluted
338 2-, 25-, 50-, and 100-fold with DMEM along the plate and incubated for 1 h at
339 4°C. Wells were washed 3 times with PBS/Tween (0.01%) and incubated with mouse
340 anti 195E HA monoclonal antibody (a gift from PHE, Colindale, UK) (1:250) at
341 4°C for 1 h. Washing and incubation steps were repeated with secondary anti
342 mouse-goat horseradish peroxidase antibody (AbD Serotec™) (1:1000). 50 µl of TMB
343 liquid ELISA substrate (Thermo Fisher Scientific) were added to each well, incubated
344 for 30 min, and plates were read at 405 nm wave length on a FluoStar Omega (BMG
345 Labtech).

346

347 **Acknowledgements**

348 This work was supported by a Japan Society for the Promotion of Science grant to YK
349 (24-448). OGP was supported by the European Research Council under the European
350 Commission Seventh Framework Programme (FP7/2007-2013)/European Research
351 Council grant agreement 614725- PATHPHYLODYN.

352

353 **References**

- 354 1. Robertson JS. 5' and 3' terminal nucleotide sequences of the RNA genome
355 segments of influenza virus. *Nucleic Acids Res.* 1979; 6: 3745-57.
- 356 2. Desselberger U, Racaniello VR, Zazra JJ, Palese P. The 3' and 5'-terminal
357 sequences of influenza A, B and C virus RNA segments are highly conserved and
358 show partial inverted complementarity. *Gene.* 1980; 8: 315-28.
- 359 3. Hsu MT, Parvin JD, Gupta S, Krystal M, Palese P. Genomic RNAs of influenza
360 viruses are held in a circular conformation in virions and in infected cells by a
361 terminal panhandle. *Proc Natl Acad Sci U S A.* 1987; 84: 8140-44.
- 362 4. Arranz R, Coloma R, Chichón FJ, Conesa JJ, Carrascosa JL, Valpuesta JM, et al.
363 The structure of native influenza virion ribonucleoproteins. *Science.* 2012; 338:
364 1634-37.
- 365 5. Gog JR, Afonso Edos S, Dalton RM, Leclercq I, Tiley L, Elton D, et al. Codon
366 conservation in the influenza A virus genome defines RNA packaging signals.
367 *Nucleic Acids Res.* 2007; 35: 1897-907.
- 368 6. Murti KG1, Webster RG, Jones IM. Localization of RNA polymerases on
369 influenza viral ribonucleoproteins by immunogold labeling. *Virology.* 1988; 164:
370 562-6.
- 371 7. Yamanaka K, Ishihama A, Nagata K. Reconstitution of influenza virus
372 RNA-nucleoprotein complexes structurally resembling native viral
373 ribonucleoprotein cores. *J Biol Chem.* 1990; 265: 11151-5.

- 374 8. Klumpp K, Ruigrok RW, Baudin F. Roles of the influenza virus polymerase and
375 nucleoprotein in forming a functional RNP structure. *EMBO J.* 1997; 16: 1248-57.
- 376 9. Elton D, Medcalf L, Bishop K, Harrison D, Digard P. Identification of amino acid
377 residues of influenza virus nucleoprotein essential for RNA binding. *J Virol.* 1999;
378 73: 7357-67.
- 379 10. Zheng W, Olson J, Vakharia V, Tao YJ. The crystal structure and RNA-binding of
380 an orthomyxovirus nucleoprotein. *PLoS Pathog.* 2013; 9: e1003624.
- 381 11. Noda T, Sagara H, Yen A, Takada A, Kida H, Cheng RH, et al. Architecture of
382 ribonucleoprotein complexes in influenza A virus particles. *Nature.* 2006; 439:
383 490-2.
- 384 12. Noda T, Sugita Y, Aoyama K, Hirase A, Kawakami E, Miyazawa A, et al.
385 Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus.
386 *Nat Commun.* 2012; 3: 639.
- 387 13. Chou YY, Vafabakhsh R, Doğanay S, Gao Q, Ha T, Palese P. One influenza virus
388 particle packages eight unique viral RNAs as shown by FISH analysis. *Proc Natl*
389 *Acad Sci U S A.* 2012; 109: 9101-6.
- 390 14. Fujii Y, Goto H, Watanabe T, Yoshida T, Kawaoka Y. Selective incorporation of
391 influenza virus RNA segments into virions. *Proc Natl Acad Sci U S A.* 2003; 100:
392 2002-7.
- 393 15. Fujii K, Fujii Y, Noda T, Muramoto Y, Watanabe T, Takada A, et al. Importance
394 of both the coding and the segment-specific noncoding regions of the influenza A

395 virus NS segment for its efficient incorporation into virions. J Virol. 2005; 79:
 396 3766-74.

397 16. Liang Y, Hong Y, Parslow TG. cis-Acting packaging signals in the influenza virus
 398 PB1, PB2, and PA genomic RNA segments. J Virol. 2005; 79: 10348-10355.

399 17. Liang Y, Huang T, Ly H, Parslow TG, Liang Y. Mutational analyses of packaging
 400 signals in influenza virus PA, PB1, and PB2 genomic RNA segments. J Virol.
 401 2008; 82: 229-36.

402 18. Hutchinson EC, Curran MD, Read EK, Gog JR, Digard P. Mutational analysis of
 403 cis-acting RNA signals in segment 7 of influenza A virus. J Virol. 2008; 82:
 404 11869-79.

405 19. Marsh GA, Rabadán R, Levine AJ, Palese P. Highly conserved regions of
 406 influenza a virus polymerase gene segments are critical for efficient viral RNA
 407 packaging. J Virol. 2008; 82: 2295-304.

408 20. Ozawa M, Maeda J, Iwatsuki-Horimoto K, Watanabe S, Goto H, Horimoto T, et al.
 409 Nucleotide sequence requirements at the 5' end of the influenza A virus M RNA
 410 segment for efficient virus replication. J Virol. 2009; 83: 3384-8.

411 21. Goto H, Muramoto Y, Noda T, Kawaoka Y. The genome-packaging signal of the
 412 influenza A virus genome comprises a genome incorporation signal and a
 413 genome-bundling signal. J Virol. 2013; 87: 11316-22.

- 414 22. Fournier E, Moules V, Essere B, Paillart JC, Sirbat JD, Isel C, et al. A
415 supramolecular assembly formed by influenza A virus genomic RNA segments.
416 Nucleic Acids Res. 2012; 40: 2197-209.
- 417 23. Fournier E, Moules V, Essere B, Paillart JC, Sirbat JD, Cavalier A, et al.
418 Interaction network linking the human H3N2 influenza A virus genomic RNA
419 segments. Vaccine. 2012; 30: 7359-67.
- 420 24. Gavazzi C, Yver M, Isel C, Smyth RP, Rosa-Calatrava M, Lina B, et al. A
421 functional sequence-specific interaction between influenza A virus genomic RNA
422 segments. Proc Natl Acad Sci U S A. 2013; 110: 16604-9.
- 423 25. Gavazzi C, Isel C, Fournier E, Moules V, Cavalier A, Thomas D, et al. An in vitro
424 network of intermolecular interactions between viral RNA segments of an avian
425 H5N2 influenza A virus: comparison with a human H3N2 virus. Nucleic Acids
426 Res. 2013; 41: 1241-54.
- 427 26. Jackson D, Lamb RA. The influenza A virus spliced messenger RNA M mRNA3
428 is not required for viral replication in tissue culture. J Gen Virol. 2008; 89:
429 3097-101.
- 430 27. Gao Q, Chou YY, Doğanay S, Vafabakhsh R, Ha T, Palese P. The influenza A
431 virus PB2, PA, NP, and M segments play a pivotal role during genome packaging.
432 J Virol. 2012; 86: 7043-51.

- 433 28. Moss WN, Priore SF, Turner DH. Identification of potential conserved RNA
434 secondary structure throughout influenza A coding regions. *RNA*. 2011; 17:
435 991-1011.
- 436 29. Moss WN, Dela-Moss LI, Kierzek E, Kierzek R, Priore SF, Turner DH. The 3'
437 splice site of influenza A segment 7 mRNA can exist in two conformations: a
438 pseudoknot and a hairpin. *PLoS One*. 2012; 7: e38323.
- 439 30. Priore SF, Moss WN, Turner DH. Influenza A virus coding regions exhibit
440 host-specific global ordered RNA structure. *PLoS One*. 2012; 7: e35989.
- 441 31. Jiang T, Kennedy SD, Moss WN, Kierzek E, Turner DH. Secondary structure of a
442 conserved domain in an intron of influenza A M1 mRNA. *Biochemistry*. 2014; 53:
443 5236-48.
- 444 32. Simmonds P. SSE: a nucleotide and amino acid sequence analysis platform. *BMC*
445 *Res Notes*. 2012; 5: 50.
- 446 33. Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF. RNAalifold: improved
447 consensus structure prediction for RNA alignments. *BMC bioinformatics*. 2008;
448 9: 474.
- 449 34. Zuker M. Mfold web sever for nucleic acid folding and hybridization
450 prediction. *Nucleic Acids Res*. 2003; 31: 3406-3415.
- 451 35. Brooke CB, Ince WL, Wrammert J, Ahmed R, Wilson PC, Bennink JR, et al. Most
452 influenza a virions fail to express at least one essential viral protein. *J Virol*. 2013;
453 87: 3155-62.

- 454 36. Brooke CB. Biological activities of 'noninfectious' influenza A virus particles.
455 Future Virol. 2014; 9: 41-51.
- 456 37. Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW. Influenza A virus
457 nucleoprotein selectively decreases neuraminidase gene-segment packaging while
458 enhancing viral fitness and transmissibility. Proc Natl Acad Sci U S A. 2014;
459 111:16854-49.
- 460 38. Hutchinson EC, von Kirchbach JC, Gog JR, Digard P. Genome packaging in
461 influenza A virus. J Gen Virol; 2010; 91:313-28.
- 462 39. Jouvenet N, Lainé S, Pessel-Vivares L, Mougél M. Cell biology of retroviral RNA
463 packaging. RNA Biol. 2011; 8: 572-80.
- 464 40. Burkhardt C, Sung PY, Celma CC, Roy P. Structural constraints in the packaging
465 of bluetongue virus genomic segments. J Gen Virol. 2014; 95: 2240-50.
- 466 41. Fajardo T Jr, Sung PY, Roy P. Disruption of Specific RNA-RNA Interactions in a
467 Double-Stranded RNA Virus Inhibits Genome Packaging and Virus Infectivity.
468 PLoS Pathog. 2015; 11: e1005321.
- 469 42. Boyce M, McCrae MA, Boyce P, Kim JT. Inter-segment complementarity in
470 orbiviruses: A driver for co-ordinated genome packaging in the Reoviridae? J Gen
471 Virol. 2016; 97: 1145-57.
- 472 43. Massin P, Kuntz-Simon G, Barbezange C, Deblanc C, Oger A, Marquet-Blouin E,
473 et al. Temperature sensitivity on growth and/or replication of H1N1, H1N2 and

474 H3N2 influenza A viruses isolated from pigs and birds in mammalian cells. *Vet*
475 *Microbiol.* 2010; 142: 232-41.

476 44. Brower-Sinning R, Carter DM, Crevar CJ, Ghedin E, Ross TM, Benos PV. The
477 role of RNA folding free energy in the evolution of the polymerase genes of the
478 influenza A virus. *Genome Biol.* 2009; 10: R18.

479 45. Chursov A, Kopetzky SJ, Leshchiner I, Kondofersky I, Theis FJ, Frishman D, et al.
480 Specific temperature-induced perturbations of secondary mRNA structures are
481 associated with the cold-adapted temperature-sensitive phenotype of influenza A
482 virus. *RNA Biol.* 2012; 9: 1266-74.

483 46. Essere B, Yver M, Gavazzi C, Terrier O, Isel C, Fournier E, et al. et al. Critical
484 role of segment-specific packaging signals in genetic reassortment of influenza A
485 viruses. *Proc Natl Acad Sci U S A.* 2013; 110: E3840-8.

486 47. Chen BJ, Leser GP, Jackson D, Lamb RA. The influenza virus M2 protein
487 cytoplasmic tail interacts with the M1 protein and influences virus assembly at the
488 site of virus budding. *J Virol.* 2008; 82: 10059-70.

489 48. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, et al. The
490 influenza virus resource at the National Center for Biotechnology Information. *J*
491 *Virol.* 2008; 82:596-601.

492 49. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid
493 multiple sequence alignment based on fast Fourier transform. *Nucleic acids*
494 *research.* 2002; 30: 3059-66.

- 495 50. Suzuki Y, Gojobori T, Nei M. ADAPTSITE: detecting natural selection at single
496 amino acid sites. *Bioinformatics*. 2001; 17: 660-1.
- 497 51. Markham NR, Zuker M. UNAFold: software for nucleic acid folding and
498 hybridization. *Methods Mol Biol*. 2008; 453: 3-31.
- 499 52. Shelton H, Smith M, Hartgroves L, Stilwell P, Roberts K, Johnson B, et al. An
500 influenza reassortant with polymerase of pH1N1 and NS gene of H3N2 influenza
501 A virus is attenuated in vivo. *J Gen Virol*. 2012; 93:998-1006.
- 502 53. Elleman CJ, Barclay WS. The M1 matrix protein controls the filamentous
503 phenotype of influenza A virus. *Virology*. 2004; 321: 144-53.
- 504 54. Gambaryan AS, Matrosovich MN. A solid-phase enzyme-linked assay for
505 influenza virus receptor-binding activity. *J Virol Methods*. 1992; 39:111-23.
- 506
- 507

508 **Figure legends**

509 **Figure 1. Conserved secondary structures among 1,884 M segment vRNA**

510 **sequences.** (A) Average d_s values in the non-overlapping coding regions of M1 and
511 M2, calculated using ADAPTSITE.⁵⁰ (B) Average z -scores of the minimum free energy
512 for the secondary structures potentially formed by the nucleotide sequences in the
513 windows, calculated using SSE.³² Solid and dashed lines represent the z -scores and the
514 standard deviations for -RNA and +RNA, respectively. Windows with negative z -scores
515 contain sequences that potentially form stable RNA secondary structures. (C)
516 Nucleotide sites with high sequence identity. Nucleotide sites are colored red where >6

contiguous sites with >98% identity are present. Orange, blue, and green boxes show the locations of the SL3-10(+)/(-), SL5-2(+), and SL5-3B(+)/(-) structures, respectively. Yellow boxes represent regions of the M segment where packaging signals have previously been inferred to be located.^{18,20} Grey boxes represent coding regions for M1, and M2. A scale bar for nucleotide positions in M vRNA is shown below. Nucleotide positions are given relative to the +RNA sequence of the M segment A/hvPR8/34(H1N1) (accession number EF190977).

Figure 2. Stem loop structure of SL3-10(+) predicted at nucleotide positions 219-240 of M segment +RNA. (A) Secondary structures were predicted using the M nucleotide sequence of England/195/2009(H1N1) (accession number GQ166660) at temperature setting of 37°C. The background red color of each base represents nucleotide identities of > 98% among 1,884 sequences. Third codon positions are circled in blue. The red arrows and nucleotide sequences indicate the position and nature of synonymous mutations that were introduced to produce SL3-10 mutants. (B) The number (percentage) of sequences that were predicted to form base-pairing (BP) at the stem of SL3-10(+) among 1,884 sequences. (C) Secondary structure predicted at positions 219-240 of SL3-10 M sequence. Bases with synonymous mutations are circled in red.

Figure 3. Stem loop structure of SL3-10(-) predicted at nucleotide positions 219-240 of M segment -RNA. (A) Secondary structures were predicted using the M

nucleotide sequence of England/195/2009(H1N1) (accession number GQ166660) at temperature setting of 37°C. The background red color of each base represents nucleotide identities of > 98% among 1,884 sequences. Third codon positions are circled in blue. The red arrows and nucleotide sequences indicate the position and synonymous mutations, which were introduced to produce SL3-10 mutant. (B) The number (percentage) of sequences that were predicted to form base-pairing (BP) at the stem positions of SL3-10(-) among 1,884 sequences. (C) Secondary structure predicted at positions 219-240 of SL3-10 M sequence. Bases with synonymous mutations are circled in red.

548

Figure 4. Stem loop structure of SL5-3B(+) predicted at nucleotide positions

967-994 of M segment +RNA. (A) Secondary structures were predicted using the M nucleotide sequence of England/195/2009(H1N1) (accession number GQ166660) at temperature setting of 37°C. The background red, pink and white colors of each base represents nucleotide identities of > 98%, 90-98%, and < 90% among 1,884 sequences, respectively. Third codon positions are circled in blue. The red arrows and nucleotide sequences indicate the position and nature of synonymous mutations that were introduced to produce SL5-3B mutants. (B) The number (percentage) of sequences that were predicted to form base-pairing (BP) at the stem positions of SL5-3B(+) among 1,884 sequences. (C) Secondary structure at positions 967-994 of SL5-3B M sequence. Bases with synonymous mutations are circled in red.

560

Figure 5. Stem loop structure of SL5-3B(-) predicted at nucleotide positions 967-994 of M segment -RNA. (A) Secondary structures were predicted using the M nucleotide sequence of England/195/2009(H1N1) (accession number GQ166660) at temperature setting of 37°C. The background red, pink and white colors of each base represents nucleotide identities of > 98%, 90-98%, and < 90% among 1,884 sequences, respectively. Third codon positions are circled in blue. The red arrows and nucleotide sequences indicate the position and nature of synonymous mutations that were introduced to produce SL5-3B mutants. Green-filled circles flanked by nucleotide bases denote the synonymous sites at which mutations affect M segment packaging in A/PR/8/34.¹⁸ (B) The number (percentage) of sequences that were predicted to form base-pairing (BP) at the stem positions of SL5-3B(-) among 1,884 sequences. (C) Secondary structure at positions 967-994 of SL5-3B M sequence. Bases with synonymous mutations are circled in red.

Figure 6. Stem loop structure of SL5-2(+) predicted at nucleotide positions 950-964 of M segment +RNA. (A) Secondary structures were predicted using the M nucleotide sequence of England/195/2009(H1N1) (accession number GQ166660) at temperature setting of 37°C. The background red, pink and white colors of each base represents nucleotide identities of > 98%, 90-98%, and < 90% among 1,884 sequences, respectively. Third codon positions are circled in blue. The red arrows and nucleotide sequences indicate the position and nature of synonymous mutations that were introduced to produce SL5-2 mutant. (B) The number (percentage) of sequences that

583 were predicted to form base-pairing (BP) at the stem positions of SL5-2(+) among 1,884
584 sequences. (C) Secondary structure at positions 950-964 of SL5-2 M sequence. Bases
585 with synonymous mutations are circled in red.

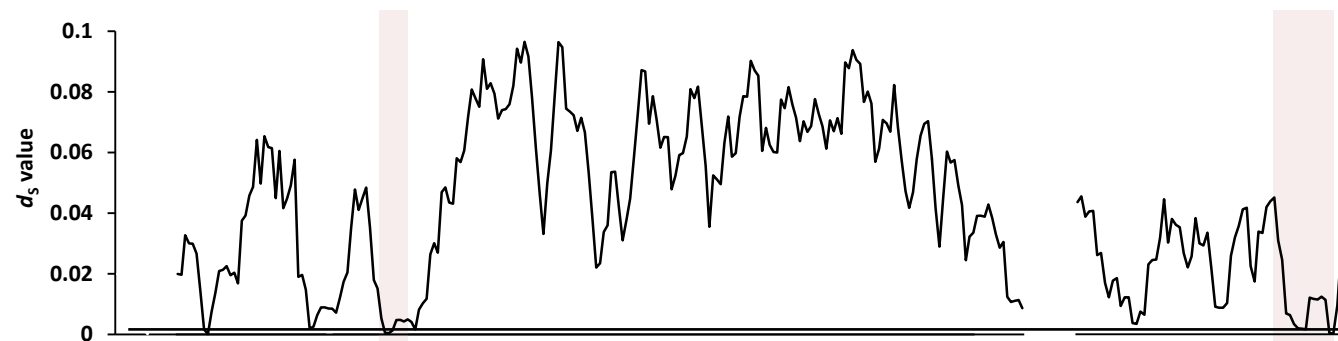
586

587 **Figure 7. Growth rates of SL3-10, SL5-3B, SL5-2 and 195E in MDCK cells at**
588 **different temperatures.** PFU counts in MDCK infections (MOI = 0.001) with
589 SL3-10, SL5-3B, SL5-2 and 195E at (A) 37°C and (B) 39°C. PFU counts in single-step
590 MDCK infections (MOI = 1) with SL3-10 and 195E at (C) 37°C and (D) 39°C.

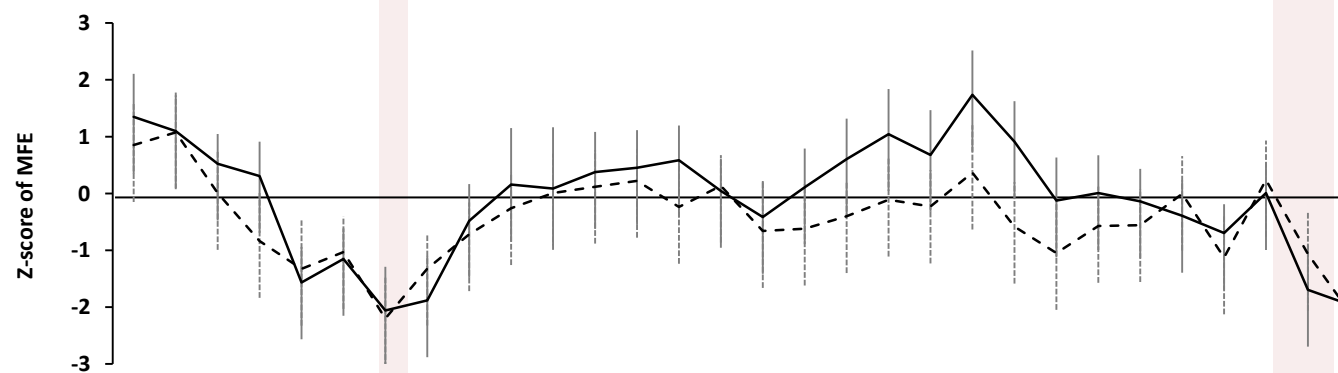
591

592 **Figure 8. Quantification of virus particles produced by infection with isolates**
593 **SL3-10 and 195E.** Viral particle counts assessed by hemagglutination of chicken red
594 blood cells in supernatant harvested 9 hours post infection following MOI=1 infection
595 of MDCK cells incubated at 39°C. (A) HA titer of unadjusted supernatants. HA titration
596 performed in triplicate showed no variation between triplicates, data representative of 2
597 independent experiments. (B) HA titer normalized to the PFU of the harvested samples
598 of SL3-10 and 195E. HA titration performed in triplicate showed no variation between
599 triplicates, data representative of 2 independent experiments. (C) Fetuin ELISA to detect
600 virus particles using normalized PFU samples of SL3-10 and 195E. y-axis indicates the
601 OD values following capture of virus particles onto the ELISA plate coated with fetuin,
602 and detection using HA antibody, and thus indicates the quantity of virus particles in
603 each viral sample. x-axis indicates the dilution of SL3-10 and 195E viruses with
604 DMEM.

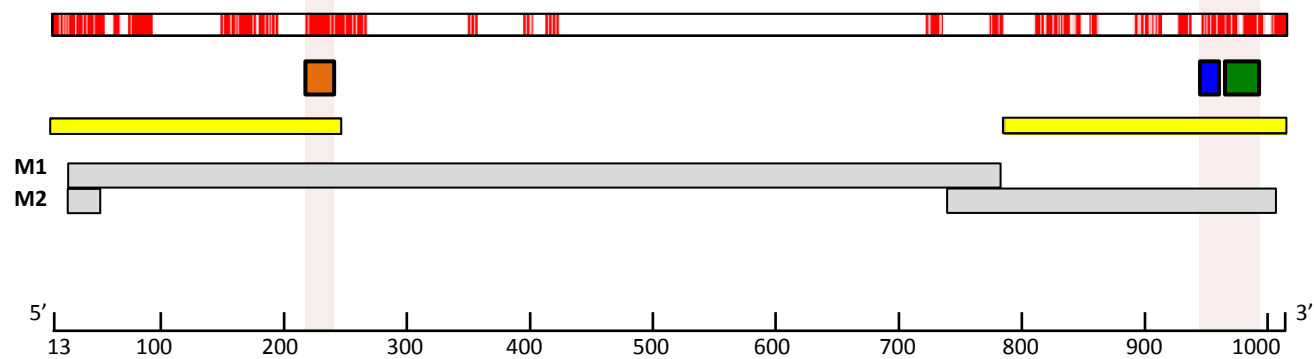
A



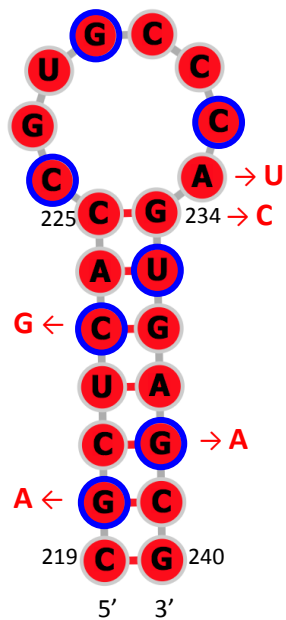
B



C



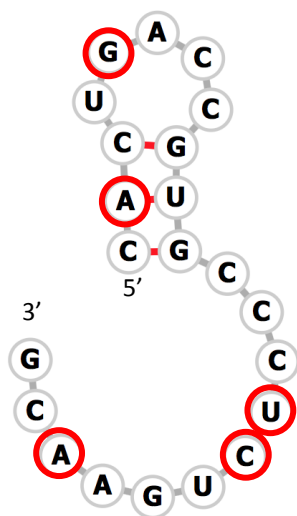
A



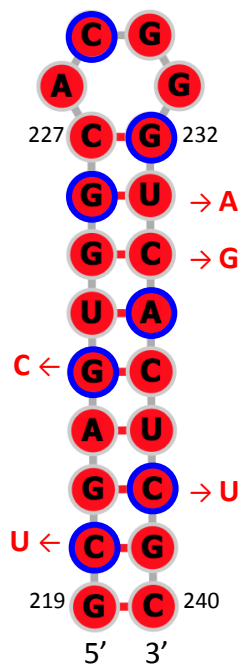
B

BP positions	BP(%)						Non-BP (%)
	GC	CG	AU	UA	GU	UG	
225-234	0	1881 (99.84)	0	0	0	0	3 (0.16)
224-235	0	0	1877 (99.63)	0	1 (0.05)	0	6 (0.32)
223-236	0	1884 (100)	0	0	0	0	0
222-237	0	0	0	1883 (99.95)	0	1 (0.05)	0
221-238	0	1862 (98.83)	0	0	0	0	22 (1.17)
220-239	1873 (99.42)	0	0	0	0	0	11 (0.58)
219-240	0	1880 (99.79)	0	0	0	0	4 (0.21)

C



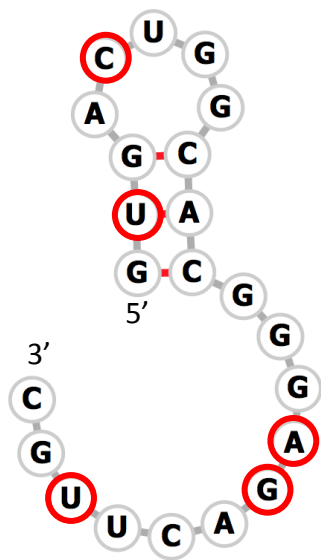
A



B

BP positions	BP(%)						Non-BP (%)
	GC	CG	AU	UA	GU	UG	
227-232	0	1880 (99.79)	0	0	0	0	4 (0.21)
226-233	0	0	0	0	1883 (99.94)	0	1 (0.05)
225-234	1881 (99.84)	0	0	0	3 (0.16)	0	0
224-235	0	0	0	1877 (99.63)	0	6 (0.32)	0
223-236	1884 (100)	0	0	0	0	0	0
222-237	0	0	1883 (99.94)	0	0	0	1 (0.05)
221-238	1862 (98.83)	0	0	0	22 (1.17)	0	0
220-239	0	1873 (99.42)	0	0	0	11 (0.58)	0
219-240	1880 (99.79)	0	0	0	4 (0.21)	0	0

C



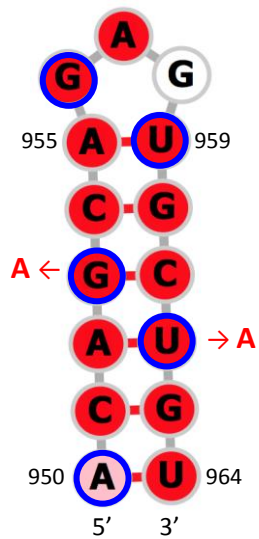
[illegible]

BP	BP(%)						Non-BP
positions	GC	CG	AU	UA	GU	UG	(%)
973-964	0	0	1865 (98.99)	0	15 (0.8)	0	4 (0.21)
974-963	0	0	21 (1.11)	0	1859 (98.67)	0	4 (0.21)
975-962	0	1883 (99.95)	0	0	0	0	1 (0.05)
976-961	0	0	1881 (99.84)	0	1 (0.05)	0	2 (0.11)
977-960	1884 (100)	0	0	0	0	0	0
978-959	0	2 (0.11)	0	1879 (99.73)	0	0	3 (0.16)
979-958	0	362 (19.21)	0	1517 (80.52)	0	0	5 (0.27)
980-957	1880 (99.79)	0	0	0	2 (0.11)	0	2 (0.11)
981-956	0	0	0	1877 (99.63)	0	1 (0.05)	6 (0.32)
982-955	0	0	1876 (99.58)	0	8 (0.42)	0	0

BP	BP(%)						Non-BP
positions	GC	CG	AU	UA	GU	UG	(%)
955-989	1883 (99.95)	0	0	0	0	0	1 (0.05)
957-990	0	0	0	1881 (99.84)	0	2 (0.11)	1 (0.05)
959-991	0	1884 (100)	0	0	0	0	0
961-992	2 (0.11)	0	1879 (99.73)	0	1 (0.05)	0	1 (0.05)
963-993	362 (19.21)	0	1517 (80.52)	0	4 (0.21)	0	1 (0.05)
965-994	0	1880 (99.79)	0	0	0	1 (0.05)	3 (0.16)
967-995	0	0	1877 (99.63)	0	1 (0.05)	0	6 (0.32)
969-996	0	0	0	1876 (99.58)	0	8 (0.42)	0

A diagram of a branched RNA structure. The main strand is a sequence of nucleotides: 5'-U-A-C-G-C-U-A-U-G-U-A-C-A-G-U-U-A-A-3'. A branch is formed by a C-G base pair, where the C is on the main strand and the G is on a side strand. The side strand continues with the sequence: G-U-A-C-A-A. The nucleotides U (at the 5' end of the main strand), G (on the side strand), and C (on the main strand) are highlighted with red circles. The 5' and 3' ends of the main strand are labeled.

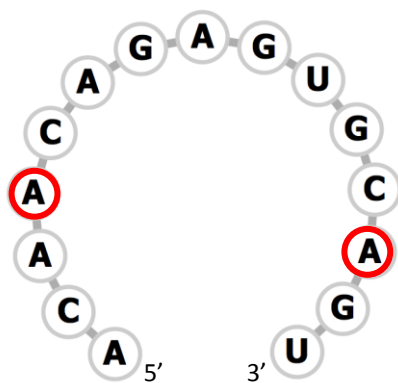
A



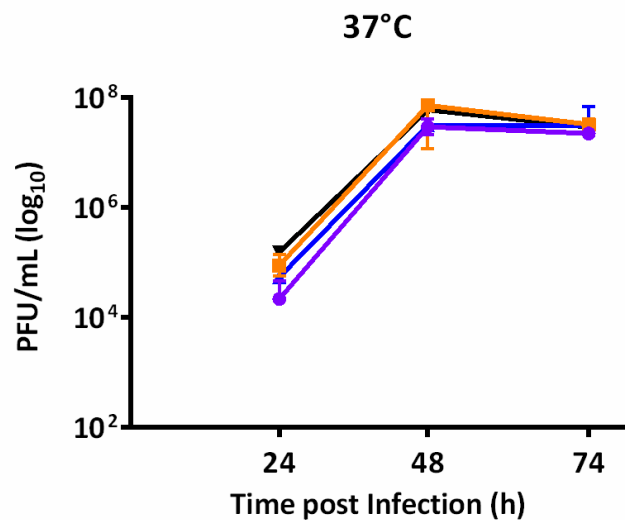
B

BP positions	BP(%)						Non-BP (%)
	GC	CG	AU	UA	GU	UG	
955-959	0	0	1874 (99.47)	0	3 (0.16)	0	7 (0.37)
954-960	0	1881 (99.84)	0	0	0	0	3 (0.16)
953-961	1844 (99.78)	0	0	0	2 (0.11)	0	38 (2.02)
952-962	0	0	1871 (99.31)	0	9	0	4 (0.21)
951-963	0	1876 (99.58)	0	0	0	0	8 (0.42)
950-964	2 (0.11)	0	1796 (99.53)	0	82 (4.35)	0	4 (0.21)

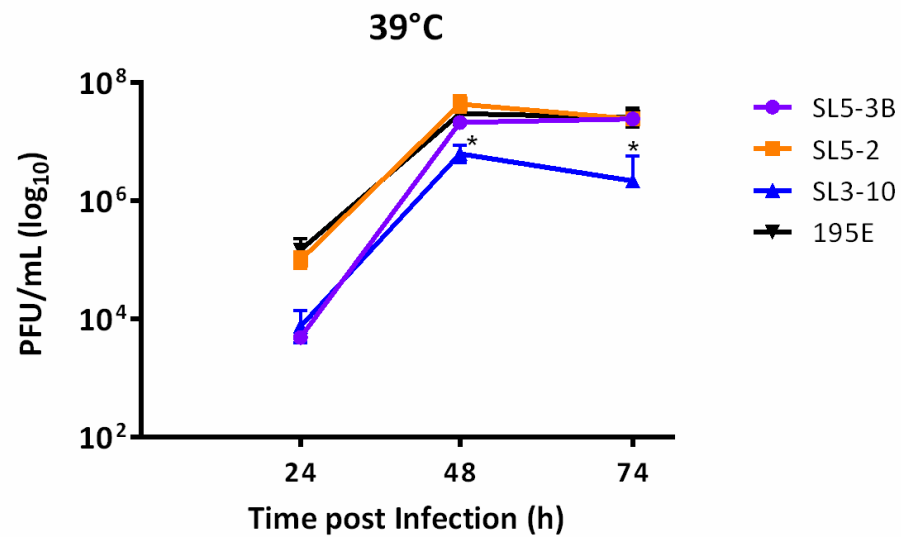
C



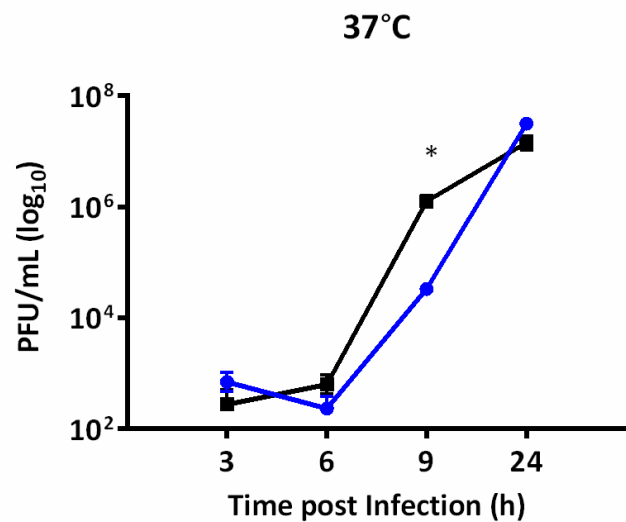
A



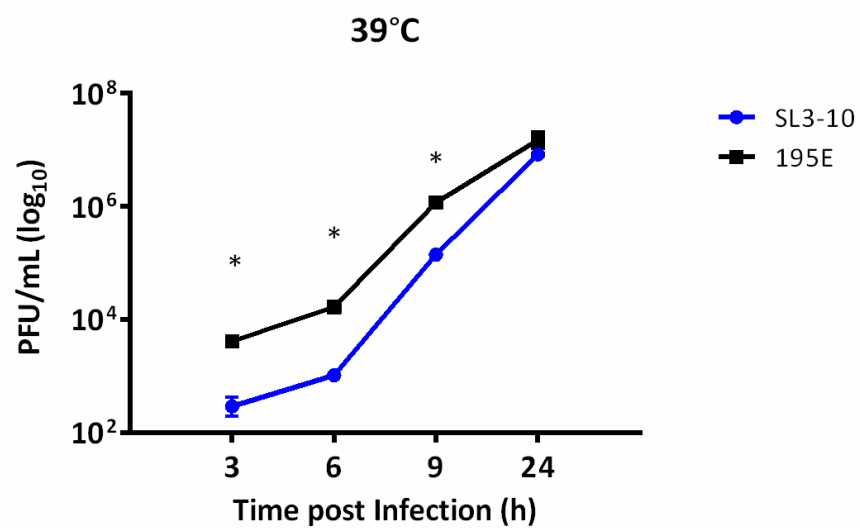
B



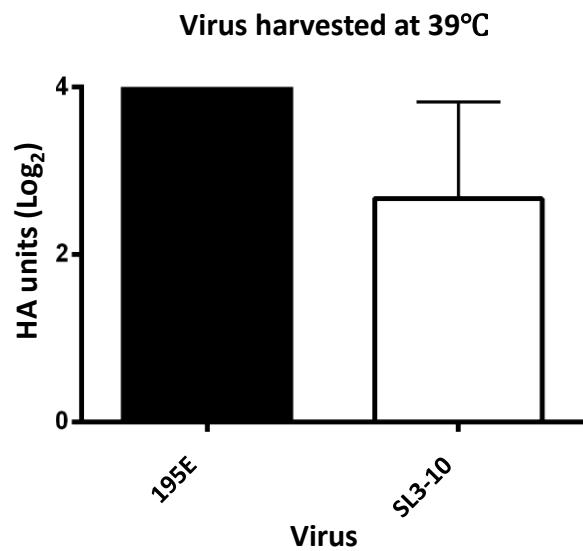
C



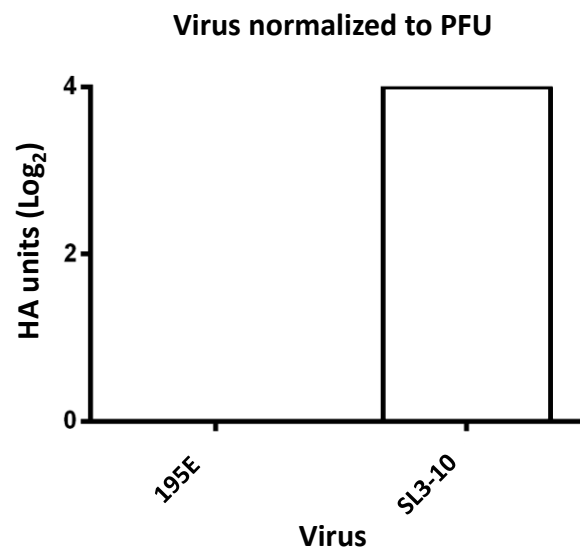
D



A



B



C

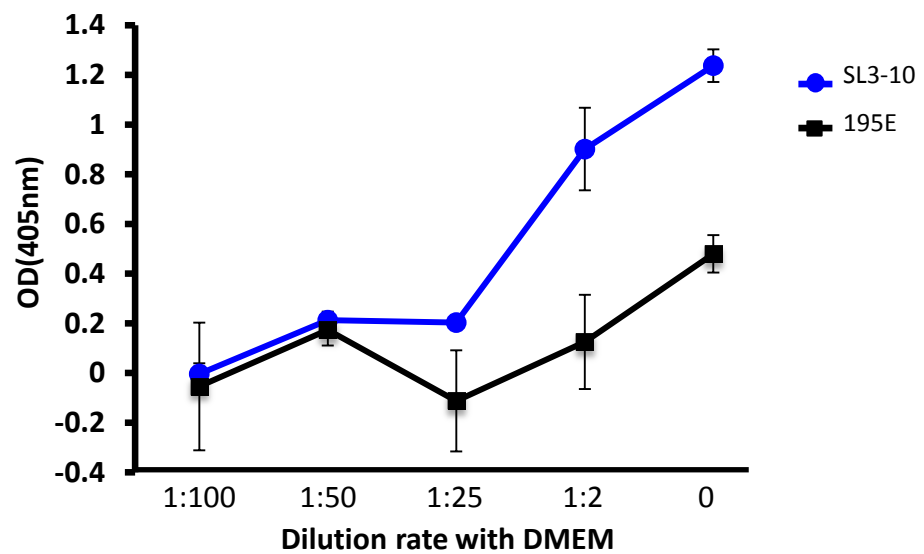


Table 1. M segment positions predicted to form secondary structures.

coding region	nt positions ^a	strand	z-score	
			average	SE
M1	115-264	-	-1.5666	0.6246
		+	-1.3270	0.8544
	145-294	-	-1.1529	0.5390
		+	-1.0313	0.6243
	175-324	-	-2.0586	0.7688
		+	-2.2007	0.7180
	205-354	-	-1.8819	0.7828
		+	-1.3244	0.5862
M2	835-984	-	-1.6945	0.7812
		+	-1.0636	0.7549
	865-1014	-	-1.9571	0.5580
		+	-2.0936	0.5514

^aNucleotide positions of the windows with z-score of <-1.

Table 2. Frequencies of stem-loop structures predicted in M segment RNA.

stem loop	nt positions ^a	strand	frequency(%) ^b		
			human	swine	avian
SL3-10	219-240	–	99.6	98.9	99.5
		+	99.4	98.3	96.7
SL5-3B	967-994	–	97.1	94.9	98.7
		+	97.4	99	98.7
SL5-2	950-964	+	99.6	92	96.3

^aNucleotide positions refer to those of M segment +RNA in A/hvPR8/34(H1N1) (accession number: EF190977).

^bFrequency of entire stem-loop structures predicted in M segment RNAs of human (n=443), swine (n=172), and avian (n=1,250) IAV strains.