finkResearch Article submitted *to Virus Evolution*:

Differential Disease Severity and Whole Genome Sequence Analysis for Human Influenza A/H1N1pdm Virus in 2015-2016 Influenza Season

Hsuan Liu[1,#], Yu-Nong Gong[2,3,#], Kathryn Shaw-Saliba[1,4], Thomas Mehoke[5], Jared Evans[5], Zhen-Ying Liu[6], Mitra Lewis[4], Lauren Sauer[4], Peter Thielen[5], Richard Rothman[4], Kuan-Fu Chen[6,7,8,†], Andrew Pekosz[1, 4, 9,†]

[1]W. Harry Feinstone Department of Molecular Microbiology and Immunology, The Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205 USA
[2]Research Center for Emerging Viral Infections, College of Medicine, Chang Gung University, Taoyuan, Taiwan
[3]Department of Laboratory Medicine, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan
[4]Department of Emergency Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205 USA
[5]Research and Exploratory Development Department, Johns Hopkins Applied Physics Laboratory, Laurel, Maryland 20723 USA
[6]Department of Emergency Medicine, Chang Gung Memorial Hospital, Keelung, Taiwan
[7]Clinical Informatics and Medical Statistics Research Center, Chang Gung University, Taoyuan, Taiwan
[8]Community Medicine Research Center, Chang Gung Memorial Hospital, Keelung, Taiwan
[9]Department of Environmental Health and Engineering, The Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205 USA

[#] These authors contributed equally to the manuscript.

[†] Correspondence should be addressed to Kuan-Fu Chen, Department of Emergency Medicine, Chang Gung Memorial Hospital, Keelung, Taiwan, kfchen@cgmh.org.tw, or Andrew Pekosz, W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, 21205, USA, Phone: +1-410-502-9306, Fax: +1-410-955-0105, email: apekosz1@jhu.edu

Abstract: 232/350

Main Text: 3571/10000

Abstract:

During the 2015-16 winter, the US experienced a relatively mild influenza season compared to Taiwan which had a higher number of total and severe cases. While H1N1pdm viruses dominated global surveillance efforts that season, the global distribution of genetic variants and their contributions to disease severity have not been investigated. Samples collected from influenza A positive patients by the Johns Hopkins Center of Excellence for Influenza Research and Surveillance (JH-CEIRS) active surveillance in the emergency rooms in Baltimore, Maryland, USA and northern Taiwan between November 2015 and April 2016, were processed for influenza A virus whole genome sequencing. In Baltimore, the majority of the viruses were the H1N1pdm clade 6B.1 and no H1N1pdm clade 6B.2 viruses were detected. In northern Taiwan, more than half of the H1N1pdm viruses were clade 6B.1 and 38% were clade 6B.2, consistent with the global observation that most 6B.2 viruses circulated in Asia and not North America. Whole virus genome sequence analysis identified two genetic subgroups present in each of the 6B.1 and 6B.2 clades and one 6B.1 intraclade reassortant virus. Clinical data showed 6B.2 patients had more disease symptoms including higher crude and inverse probability weighted odds of pneumonia than 6B.1 patients, suggesting 6B.2 circulation may contribute to the severe flu season in Taiwan. Local surveillance efforts linking H1N1pdm virus sequences to patient clinical and demographic data improve our understanding of influenza circulation and disease potential.

Key words: influenza, H1N1pdm, whole genome sequencing, disease severity, Taiwan, Baltimore

66 **Introduction:**

67 The 2009 pandemic H1N1 Influenza A virus (H1N1pdm) has supplanted the previous
68 human H1N1 viruses to become the seasonal human H1N1 virus. In the 2015-16
69 influenza season, the northern hemisphere experienced its second global H1N1pdm-
70 dominant year since 2009. The 2015-16 Influenza activity in the US was considered
71 moderate, with lower numbers of total cases and a later peak epidemic when compared
72 with the previous three influenza seasons [1]. However, in Taiwan, the 2015-16 season
73 was much more serious [2], with an earlier start and higher numbers of total as well as
74 severe, influenza cases compared to the prior influenza season 2014-15 [3].
75
76 Influenza A viruses (IAV) are subdivided into subtypes based on the antibody response
77 to the viral glycoproteins, hemagglutinin (HA) and neuraminidase (NA). Viruses are
78 further divided into genetic clades based on HA sequences in order to monitor for
79 mutations that might lead to antigenic drift. The 2015-16 northern hemisphere vaccine
80 strain of H1N1pdm, A/California/7/2009, is defined as H1 clade 1. Prior to the 2015-16
81 season, H1N1pdm was dominated by clade 6B. The global rise of H1 clade 6B.1
82 carrying amino acid changes S84N, S162N and I216T from clade 6B was detected in
83 human surveillance efforts starting in June 2015 and is now the dominant H1N1 clade.
84 The H1 clade 6B.2 viruses (V152T, V173I of HA1 and E164G and D174E of HA2)
85 emerged in July 2015, peaked in January 2016, were found primarily in Asia and
86 disappeared from surveillance efforts at the end of 2016 [4, 5]. The majority of 6B.1 and
87 6B.2 viruses were antigenically similar to the recommended components of the 2015-16
88 Northern Hemisphere H1N1pdm influenza vaccine A/California/7/2009 [1, 3, 6, 7],

89 although antigenic difference between egg-adapted and circulating virus strains may
90 have contributed to increased numbers of cases [8].
91
92 The genome of IAV consist of 8 segments of negative-sense RNA. In addition to HA
93 mutations, mutations in the other 7 viral gene segments also result in genetic variants
94 which can alter virus fitness, subvert pre-existing immunity and/or reduce vaccine
95 effectiveness. Any of these changes could potentially alter disease severity or total
96 number of cases [9]. Next-generation sequencing provides a sensitive and rapid method
97 for sequencing of viruses directly from human specimens without any need for virus
98 isolation or culture – both of which are known to select for viruses bearing additional
99 mutations [10].
100
101 Most current global influenza surveillance systems lack connections between viral
102 sequences and detailed patient demographic and clinical data. The Johns Hopkins
103 Center of Excellence for Influenza Research and Surveillance (JH-CEIRS) has been
104 actively performing human influenza surveillance since 2014. Surveillance efforts
105 involve the connection of viral genome data and isolates with patient demographic and
106 clinical data to gain more insights into how virus genetic variation can affect clinical
107 disease. We hypothesized that sequence differences in influenza viruses could be
108 associated with distinct clinical outcomes. This study involved concurrent, active
109 influenza surveillance during the 2015-16 season in Baltimore, USA and northern
110 Taiwan to compare and contrast differential epidemics in the same year between the

111 two geographic regions and to analyze possible disease severity with influenza viral
112 genotypes.
113
114 **Materials and Methods:**
115 **Active surveillance enrollment and sample collection**
116 Institutional Review Boards at the Johns Hopkins University School of Medicine and
117 Chang Gung Memorial Hospital provided ethical approval of the study (IRB00052743).
118 From November 2015 to April 2016, active surveillance was performed in the adult
119 emergency departments of the Johns Hopkins Medical Institutions (JHMI) at the East
120 Baltimore and Bayview campuses in Baltimore, Maryland, USA and Chang Gung
121 Memorial Hospitals (CGMH) in the northern Taiwan metropolitan area, including
122 hospitals in the Taipei, Linkou, and Keelung Branches. Influenza-like illness was defined
123 as documented or reported ≥38°C fever and any of one of three respiratory symptoms
124 including cough, headache and/or sore throat within the past 7 days. Exclusion criteria
125 included subjects who are unable to speak and understand English (in US) or Mandarin
126 (in Taiwan), unable to provide written informed consent, currently incarcerated, or
127 previously enrolled in the study during the same influenza season. Patients were
128 approached by trained clinical coordinators who obtained written, informed consent
129 before collected specimens and demographic and clinical data using a standard
130 questionnaire. Data was confirmed by examination of the patient's electronic health
131 record. All data was de-identified and stored in a secure REDCap database [11, 12].
132
133 **Diagnosis and subtyping**

134 Nasopharyngeal swabs or nasal washes of patients collected at the time of enrollment
135 were tested for influenza A virus infection using the Cepheid Xpert Flu/RSV test
136 (Cepheid, Sunnyvale, CA). Samples that were influenza A virus positive were further
137 subtyped by qRT-PCR with specific H1 or H3 primers and probes according to USA
138 Centers for Disease Control and Prevention (CDC) protocols. Only H1N1pdm positive
139 samples were moved on to whole genome sequencing (WGS) analysis.
140

141 **Whole genome sequencing**

142 Viral RNA was isolated from the influenza A positive clinical samples using the MagMax
143 Viral RNA isolation reagent (Thermo Fisher) on an Eppendorf epMotion 5073 liquid
144 handling workstation. Samples were processed for WGS using multi-segment PCR [10]
145 and then prepared for deep sequencing using the Nextera XT library preparation
146 reagent (Illumina). Samples were sequenced on the Illumina NextSeq platform. Raw
147 paired-end data were first processed through Trim galore!
148 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with a quality score of
149 30 and an adapter sequence of CTGTCTCTTATACACATCT, only retaining pairs that
150 both passed through this quality control step. The quality-trimmed reads were then
151 aligned to the influenza vaccine strain reference sequence using bowtie2 (version 2.1.0)
152 [13] with the '--very-sensitive-local' option and converted to sorted BAM files using
153 samtools [14]. Only sequences of all 8 gene segments passed quality controls were
154 included in the study. The sequences were submitted to NCBI GenBank database and
155 accession numbers are listed in the Table S1.
156

157 **Phylogenetics and sequence analysis**

158 73 H1N1pdm clinical samples which yielded sufficient WGS coverage of the IAV
159 genome were included in sequence analysis. In addition to these genomes, we
160 retrieved two reference genomes of vaccine strains from the Global Initiative on Sharing
161 All Influenza Data (GISAID, https://www.gisaid.org/) for further analyses. Nucleotide
162 sequences of the longest open reading frames (ORFs) in each gene segment were
163 used to build phylogenetic trees, including PB2, PB1, PA, HA, NP, NA, M1, and NS1, as
164 well as, their alternative splicing (M2 and NS2) and ribosomal frameshift (PA-X)
165 products. Nucleotide sequences of these ORFs (except for PA-X) in each strain were
166 further concatenated to generate one single sequence for constructing a whole-genome
167 phylogenetic tree. Phylogenetic trees for investigating their genetic or genomic
168 relationships were based on maximum likelihood (ML) analysis, generated with RAxML
169 (version 8.2.12) [15] under the GTRGAMMA model with 1000 bootstrap replicates. Their
170 time scales based on sample collection dates were calibrated using the coalescent
171 model in TreeTime [16] after building ML trees. Clade 6B, 6B.1, and 6B.2 were defined
172 by the rules of HA amino acid mutations of global surveillance groups from nextflu [4].
173 Phylogenetic groups in other gene segments were based on their tree branches to
174 separate annotated clades (e.g., 6B.1 and 6B.2). Inconsistent positioning of strains, or
175 clade groups on individual trees was used to identify influenza reassortment. Individual
176 gene segment clades were annotated and visualized using the ggtree R package [17].
177

178 All H1N1pdm genomes (n=2873) isolated from September, 2015 to May, 2016 were
179 downloaded from GISAID as of January 2020, and were further aligned using MAFFT

180 tool [18] under the default conditions. Moreover, we retrieved all NP sequences
181 (n=20,477) of H1N1pdm isolated after the year of 2009 (inclusive) from GISAID as of
182 January 2020 to investigate the population of viruses encoding an upstream translation
183 initiation codon (AUG) in the NP gene segment. 20,470 NP sequences were used for
184 further analyses, after aligning and removing 7 sequences with unexpected gaps.
185

186 **Glycosylation prediction**

187 Potential N-linked glycosylation sites of the 6B.1 and 6B.2 HA and NA proteins were
188 predicted by the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc).
189

190 **Statistical analysis of demographic and clinical data**

191 To identify a contribution of virus clades and genetic subgroups to clinical symptoms,
192 we applied Chi-square or Fisher's exact tests for categorical variables and the rank-sum
193 test for continuous variables when appropriate. The two-tailed statistical significance
194 was set at p-value <0.05. When clinical or demographic category reached statistical
195 significance in a univariate analysis, confounding effects were then analyzed. In addition
196 to adjusting for possible confounding effects between demographics, disease severity
197 and clinical outcomes, a propensity score-based weighting was applied. Briefly, we
198 applied those potential confounders to generate the propensity (or probability) of being
199 infected with different subgroups of H1N1pdm viruses by using machine learning based
200 modeling methods [19]. By using the inverse probability of treatment weighting (IPTW)
201 method, patients with higher propensity will have lower weights in the final logistic
202 regression model to adjust for confounding factors. The weights create a pseudo-

203 population where the weighted virus subgroups have similar distribution across these
204 confounders. In one simulation model, the propensity score-based weighting methods
205 could adjust for most of the confounding effects [20]. Odds ratios of unadjusted and
206 IPTW were calculated using logistic regression analysis.
207
208 **Results:**
209 **H1N1pdm clade 6B.1 predominantly circulated in Baltimore and clade 6B.1 and**
210 **6B.2 co-circulated in northern Taiwan**
211 From November 2015 to April 2016, 261 and 256 symptomatic patients were enrolled in
212 the study from emergency departments in JHMI (Baltimore) and CGMH (northern
213 Taiwan). IAV was found in 32.6% (85/261) of symptomatic patients from Baltimore and
214 57.8% (148/256) from northern Taiwan. About half of the IAV-positive samples were
215 further subtyped, and the demographics and comorbidities were found to be similar. In
216 97.5% (39/40) of samples from Baltimore and 90.1% (64/71) of samples from northern
217 Taiwan H1N1 was the influenza A virus subtype. The epidemic curves of the two
218 surveillance sites were different, with influenza A virus activity starting in January and
219 peaking in mid-March in Baltimore while IAV activity in northern Taiwan started in
220 November and peaked in late January/early February, with peak activity higher than in
221 Baltimore (Fink et al., MedRxiv).
222
223 Seventy-three of the 103 H1N1pdm clinical samples yielded sufficient sequence
224 coverage of the IAV genome. Virus samples were initially analyzed based on HA
225 segment sequences (Figure 1). In Baltimore, 91.7% (22/24) of viruses were clade 6B.1,

226 two were clade 6B, and no 6B.2 viruses were identified. The northern Taiwan samples
227 consisted of 57.1% (28/49) clade 6B.1, 38.8% (19/49) clade 6B.2 and 4.1% (2/49) clade
228 6B viruses. The difference in clade circulation is consistent with global influenza
229 surveillance that H1N1pdm 6B.1 and 6B.2 were dominant in the 2015-16 season and
230 most 6B.2 viruses were found in Asia [5, 21].
231
232 **Identification of four distinct genetic subgroups of H1N1pdm**
233 To better understand the genetic diversity in the identified HA clades, we performed a
234 phylogenetic analysis using concatenated gene segments (Figure 2) and individual ORF
235 trees (Figure S1 and Figure 2). Branch tips of the 10 ORFs from the other seven gene
236 segments were colored by the HA clade of each sample and defining amino acid
237 mutations were noted (Figure S1). For the most part, the sequences of other ORFs
238 clustered together consistent with their HA clade (Figure 2). Amino acid differences
239 between clades 6B.1 and 6B.2 are listed for HA (Table S2), NA (Table S3) and the
240 other ORFs (Table S4).
241
242 Additional amino acid mutations were identified to differentiate viruses into four distinct
243 virus genetic subgroups (Figure 2). The 6B.1 viruses were separated into two groups, a
244 6B.1 parental and a 6B.1 variant defined by unique amino acid changes in NA and NS2
245 (Table S3 and S4). All Baltimore 6B.1 viruses were 6B.1 parental. In northern Taiwan,
246 only 4 were 6B.1 parental, with the remaining clade 6B.1 viruses (23/27) being 6B.1
247 variant. The 6B.2 viruses were also separated into two genetic subgroups - variant-1

248 and variant-2 – that are differentiated by mutations in HA, NA, M1 and M2 (Table S2, S3
249 and S4).
250
251 **Dynamic virus evolution of gene segments and possible reassortment**
252 All segments in two of the four 6B viruses, A/EastBaltimore/0103/2016 and
253 A/Baltimore/0088/2016, were clade 6B. One 6B virus A/Taipei/0045/2016 contained
254 mutations in PB2, PA, NA and M segments consistent with those found in the clade
255 6B.1 and the remaining segments were similar but distinct from clade 6B.1. The other
256 6B virus A/Linkou/0030/2015 contained sequences in the PB2, M and NS segments
257 consistent with the clade 6B.2 viruses and the remaining segments were related to but
258 distinct from clade 6B.2. (Figure 2 and S1). These suggested that all segments of 6B
259 viruses were evolving toward either the clade 6B.1 or clade 6B.2.
260
261 There was also evidence of intraclade reassortment in one northern Taiwan 6B.1
262 sample A/Keelung/0046/2016, as three gene segments of 6B.1 (segments HA, PB2 and
263 PA) and five 6B.2 (segments PB1, NP, NA, M and NS) were detected (Figure 2). While
264 we cannot rule out co-infection with two H1N1 clade viruses, given the number of reads
265 present for each segment, it seems more likely this individual was infected with an
266 intraclade reassortant virus.
267
268 **HA and NA N-linked glycosylation**
269 Eight N-linked glycosylation sites on the HA of A/California/7/2009 were previously
270 determined by sequence analysis and verified by mass spectrometry [22]. The 6B.1 HA

271 sequence predicts an additional N-linked glycosylation site at 162-164 (N-Q-S) on the
272 Sa antigenic site of the head domain due to a S to N mutation at HA-162 (Table S2)
273 [23]. The NA from the A/California/7/2009 vaccine strain has 8 potential N-linked
274 glycosylation sites and 6 sites were experimentally confirmed [22]. Both 6B.1 and 6B.2
275 viruses gain one additional predicted N-linked glycosylation site at 42-44 (N-Q-S) of the
276 NA stalk and lose one at 386-388 (N-F-S) which lies on the lateral surface of the NA
277 head. In addition, 68.4% (13/19) of 6B.2 viruses with NA-S52R mutation lose one
278 glycosylation site at 50-52 (N-Q-S) on the NA stalk (Table S3).
279

280 **Antiviral resistance mutations**

281 All sequences contained the amantadine resistance mutation M2-S31N which is found
282 in the A/Cal/07/2009 virus. One 6B.2 variant-2 sample from northern Taiwan,
283 A/Linkou/0029/2015, contained the NA-H275Y, the most common mutation associated
284 with oseltamivir resistance in N1 NA.
285

286 **An upstream AUG codon lengthens the nucleoprotein (NP) of H1N1pdm**

287 The NP gene segments of all samples in the study have a new, in frame AUG start
288 codon 18 nucleotides upstream of the usual start codon of NP proteins of IAV including
289 human seasonal H1N1 and H3N2 [24, 25], adding 6 amino acids to the amino terminus
290 of the protein with clade 6B.1 (M-S-D-I-E-A) and 6B.2 viruses (M-S-D-I-E-V) differing at
291 amino acid six. This start codon was shown to be utilized and the resulting NP protein
292 capable of supporting viral polymerase activity [26, 27]. The presence of an upstream
293 NP start codon and the amino acid difference at position 6 between 6B.1 and 6B.2 were

294 also present in approximately 70% of GISAID sequences (n=20,470) of H1N1pdm NP
295 since 2009 with an Alanine (A) dominating at position 6.
296
297 **Patients infected with clade 6B.2 had increased symptoms of severe influenza**
298 To determine whether viral genotype contributes to symptoms and disease severity, the
299 demographic and clinical data from patients infected with clade 6B.1 or 6B.2 viruses
300 were first compared by univariate analyses (Table 1). Patients infected with 6B.2 were
301 more likely to be older, female, Asian, and unvaccinated. These patients had a higher
302 prevalence of runny nose and diarrhea and had increased symptoms of severe
303 influenza, including increased likelihood of oxygen supplementation, longer in-hospital
304 stay, and pneumonia diagnosis (by radiological findings). Interestingly, patients infected
305 with clade 6B.1 viruses had a lower number of co-morbidities associated with severe
306 influenza, suggesting a healthier population was being infected with those viruses.
307
308 We further utilized a logistic regression analysis to estimate the propensity of being
309 infected by 6B.2, using potential confounders such as age, gender, comorbidities,
310 obesity, vaccine status, human exposure and travel history (area under the ROC curve:
311 0.798). After IPTW to adjust for the possible confounding effects, infection with clade
312 6B.2 viruses was found to be significantly associated with pneumonia (Odds ratio:
313 3.261, 95% CI: 1.696-4.826, p=0.008) (Table 2).
314
315 The distribution of cases across the four genetic subgroups was then determined
316 (Tables 3 and 4). Patients infected with the 6B.1 parental virus had increased reporting

317 of wheezing, sinus pain, and nausea/vomiting, along with the expected difference in
318 race and ethnicity stemming from the identification of the 6B.1 variant only in northern
319 Taiwan (Table 3). When patients infected with the 6B.2 variant-1 and variant-2 viruses
320 were compared there were no obvious difference in symptoms (Table 4), however, the
321 small sample sizes in the 6B.2 subgroups precluded a definitive characterization of
322 demographic and clinical data between these genotypes. Overall, our data suggested
323 clade 6B.2 circulation in northern Taiwan likely played a significant role in differentiating
324 the influenza epidemics between Baltimore and northern Taiwan in the season.
325

326 **Discussion:**

327 Four distinct genetic subgroups of H1N1pdm viruses were circulating in Baltimore and
328 northern Taiwan during the 2015-16 flu season. While the 6B.1 clade was circulating at
329 both sites, some 6B.1 viruses in northern Taiwan had additional mutations which
330 defined a genetic subgroup we called 6B.1 variant. The two 6B.1 subgroups of viruses
331 did not differ drastically with respect to disease severity. The 6B.2 clade viruses were
332 detected in northern Taiwan but not in Baltimore and consisted of two distinct genetic
333 subgroups. Patients infected with 6B.2 virus showed more and stronger disease
334 symptoms by both uni- and multivariate analysis. This difference in H1N1pdm clade
335 circulation may have contributed to the more severe influenza season experienced in
336 northern Taiwan compared to Baltimore in 2015-16. After adjusting and weighting for
337 age, gender, comorbidities, obesity, vaccine status, human exposure and travel history,
338 people infected with 6B.2 still had a higher odds ratio of pneumonia (Table 2). The data
339 indicate that infection with clade 6B.2 viruses is associated with a higher disease

340 severity at our surveillance sites independent of patient demographics and
341 comorbidities. The presence of the 6B.2 clade in addition to low preexisting immunity to
342 circulating H1N1 in northern Taiwan (reference Fink, 2020 medrxiv) may be the most
343 likely reasons explaining the different numbers and severity of influenza cases between
344 Baltimore and northern Taiwan. The only report for symptoms of 6B.1 compared to 6B.2
345 infection was seen in Israel, where infection with 6B.2 virus was associated with more
346 vomiting and nausea compared to 6B.1-infected patients [28]. In our surveillance, we
347 did see a higher percentage of 6B.2 than 6B.1 patients with vomiting and nausea, but
348 this did not reach statistical significance (Table 1).
349
350 The H1N1pdm 6B.1 and 6B.2 viruses dominated the 2015-16 season in the Northern
351 Hemisphere. However, numbers of cases and disease severity differed between
352 geographic regions. The 2015-16 H1N1pdm season in the US and Canada were
353 moderate compared to the prior season [1, 6]. On the other hand, influenza surveillance
354 in Russia, UK, eastern Europe, the Middle East and Asia (including Taiwan) had
355 increasing numbers of severe cases [3, 7, 29-33]. Reports also showed H1N1pdm
356 impacted the southern hemisphere in the 2016 season, in particular, Brazil and Reunion
357 Island [34-36].
358
359 Serological surveillance indicated that the circulating H1N1pdm strains in 2015-16 were
360 antigenically similar to the vaccine A/California/7/2009 in the US, UK and many other
361 countries [1, 6, 7, 32, 37, 38]. This suggests that the HA amino acid differences
362 between the circulating 6B.1 or 6B.2 viruses and the vaccine strain were likely not able

363 to explain differential epidemics in different regions of the year even though vaccination
364 coverage might differ. Whether mutations in the other 7 viral genes contributed to the
365 variable disease severity reported in 2015-16 has not been studied carefully. Some
366 mutations in NA and M1 occurred at sites previously associated with altered virus
367 transmission [39, 40], but our initial experiments on human nasal epithelial cell cultures
368 did not detect any replication differences between clade 6B.1 and 6B.2 viruses
369 (reference Fink, 2020 medrxiv).
370
371 We identified one 6B.2 patient in northern Taiwan had NA-H275Y mutation, which is
372 associated with NA drug resistance. In another study of 2015-16 H1N1pdm in Taiwan,
373 one clinical isolate with NA-H275Y was resistant to oseltamivir [41]. The data suggest
374 sporadic identification of oseltamivir resistant H1N1pdm viruses in 2015-16 in Taiwan.
375
376 The clade 6B.1 parental genetic subgroup continued to circulate after 2015-16, and
377 evolved into several 6B.1A subclades through August 2020. It is not clear what drove
378 the extinction of 6B.2 but it is possible that clade 6B.1 viruses were better adapted to
379 infect and spread in humans. It has been suggested that glycosylation is important for
380 influenza A virus adaptation in humans [42] and gaining a potential glycosylation site on
381 the HA head at residue 162 may have given clade 6B.1 viruses an evolutionary
382 advantage. Clade 6B.1 viruses also encode an NS1-E125D mutation which increases
383 NS1 interactions with cellular cleavage and polyadenylation factor 30 (CPSF30) [43].
384 The NS1-E125D mutation of 6B.1 viruses has been suggested to be an important
385 marker for virus adaption to humans [7].

386

387 Our study is limited in several ways. The different race distributions between our
388 surveillance sites make it difficult to adjust for in this and any study conducted across
389 different geographic regions. The identification of 4 genetic subgroups reduced the
390 power of our study to detect demographic and clinical differences in our populations.
391 Increasing surveillance and whole genome sequencing efforts could dramatically
392 improve the ability to identify novel virus variants that are causing altered disease in
393 humans. Given these limitations, it is still clear that there is significant H1N1 genetic
394 diversity across the two surveillance sites and this genetic diversity contributes to the
395 differing number of influenza cases and disease severity. Influenza is a global disease,
396 but surveillance at the local level is needed to fully understand virus circulation and
397 disease potential.

398

399 **Conflicts of interests**

400 The authors declare no conflict of interest.

401

402 **Author contributions**

403 H.L., R.R., K.F.C., and A.P. conceived of the experimental questions; Y.N.G. performed
404 phylogenetic tree analyses; K.S.S., and K.F.C. ran demographic and clinical analyses;
405 T.M., J.E., and P.T. sequenced samples and cleaned the raw data; Z.Y.L., M.L., and
406 L.S. oversaw enrollment and collection of samples from patients; H.L., Y.N.G., K.S.S.,
407 and K.F.C. analyzed data; H.L., and A.P. wrote the manuscript; all authors edited and
408 reviewed the manuscript prior to submission.

Funding

This work was supported by the NIH/NIAID Center of Excellence in Influenza Research and Surveillance contract HHS N2772201400007C (R.R., K.F.C., and A.P.), the Richard Eliasberg Family Foundation and the Research Center for Emerging Viral Infections from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan and the Ministry of Science and Technology (MOST), Taiwan MOST 109-2634-F-182-001 (Y.N.G.).

Acknowledgements

The authors thank the patients who enrolled and participated in the JH-CEIRS surveillance study. We are grateful for the efforts of the clinical coordination team at JHMI who collected samples. We thank members of the Pekosz lab for feedback on this work.

References:

1. Davlin, S.L., et al., *Influenza Activity - United States, 2015-16 Season and Composition of the 2016-17 Influenza Vaccine.* MMWR Morb Mortal Wkly Rep, 2016. **65**(22): p. 567-75.
2. Gong, Y.N., et al., *Centennial review of influenza in Taiwan.* Biomed J, 2018. **41**(4): p. 234-241.
3. Taiwan_CDC. *Taiwan Influenza Express 2015-16 Flu Season 2016 Week 20*. 2016; Available from: https://www.cdc.gov.tw/En/File/Get/VCNkIvtTOiuP1vo1gJG89A.
4. Neher, R.A. and T. Bedford, *nextflu: real-time tracking of seasonal influenza virus evolution in humans.* Bioinformatics, 2015. **31**(21): p. 3546-8.
5. Bedford, T. and R. Neher. *Seasonal influenza circulation patterns and projections for 2017-2018*. 2017; Available from: https://nextflu.org/reports/feb-2017/.
6. Chambers, C., et al., *Interim estimates of 2015/16 vaccine effectiveness against influenza A(H1N1)pdm09, Canada, February 2016.* Euro Surveill, 2016. **21**(11): p. 30168.

437 7. Komissarov, A., et al., *Rapid spread of influenza A(H1N1)pdm09 viruses with a new set of specific*
438 *mutations in the internal genes in the beginning of 2015/2016 epidemic season in Moscow and*
439 *Saint Petersburg (Russian Federation).* Influenza Other Respir Viruses, 2016. **10**(4): p. 247-53.
440 8. Garretson, T.A., et al., *Identification of human vaccinees that possess antibodies targeting the*
441 *egg-adapted hemagglutinin receptor binding site of an H1N1 influenza vaccine strain.* Vaccine,
442 2018. **36**(28): p. 4095-4101.
443 9. Bouvier, N.M. and P. Palese, *The biology of influenza viruses.* Vaccine, 2008. **26 Suppl 4**: p. D49-
444 53.
445 10. Zhou, B., et al., *Single-reaction genomic amplification accelerates sequencing and vaccine*
446 *production for classical and Swine origin human influenza a viruses.* J Virol, 2009. **83**(19): p.
447 10309-13.
448 11. Harris, P.A., et al., *Research electronic data capture (REDCap)--a metadata-driven methodology*
449 *and workflow process for providing translational research informatics support.* J Biomed Inform,
450 2009. **42**(2): p. 377-81.
451 12. Harris, P.A., et al., *The REDCap consortium: Building an international community of software*
452 *platform partners.* J Biomed Inform, 2019. **95**: p. 103208.
453 13. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012.
454 **9**(4): p. 357-9.
455 14. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16):
456 p. 2078-9.
457 15. Stamatakis, A., *RAxML version 8: a tool for phylogenetic analysis and post-analysis of large*
458 *phylogenies.* Bioinformatics, 2014. **30**(9): p. 1312-3.
459 16. Sagulenko, P., V. Puller, and R.A. Neher, *TreeTime: Maximum-likelihood phylodynamic analysis.*
460 Virus Evol, 2018. **4**(1): p. vex042.
461 17. Yu, G., et al., *ggtree: an r package for visualization and annotation of phylogenetic trees with*
462 *their covariates and other associated data.* Methods Ecol Evol, 2017. **8**(1): p. 28-36.
463 18. Katoh, K. and D.M. Standley, *MAFFT multiple sequence alignment software version 7:*
464 *improvements in performance and usability.* Mol Biol Evol, 2013. **30**(4): p. 772-80.
465 19. Lee, B.K., J. Lessler, and E.A. Stuart, *Improving propensity score weighting using machine*
466 *learning.* Stat Med, 2010. **29**(3): p. 337-46.
467 20. Stuart, E.A., *Matching methods for causal inference: A review and a look forward.* Stat Sci, 2010.
468 **25**(1): p. 1-21.
469 21. World_Health_Organization, *Recommended composition of influenza virus vaccines for use in*
470 *the 2016-2017 northern hemisphere influenza season.* Wkly Epidemiol Rec, 2016. **91**(10): p. 121-
471 32.
472 22. She, Y.M., et al., *Topological N-glycosylation and site-specific N-glycan sulfation of influenza*
473 *proteins in the highly expressed H1N1 candidate vaccines.* Sci Rep, 2017. **7**(1): p. 10232.
474 23. Altman, M.O., et al., *Human Influenza A Virus Hemagglutinin Glycan Evolution Follows a*
475 *Temporal Pattern to a Glycan Limit.* mBio, 2019. **10**(2).
476 24. Reid, A.H., et al., *Novel origin of the 1918 pandemic influenza virus nucleoprotein gene.* J Virol,
477 2004. **78**(22): p. 12462-70.
478 25. Gorman, O.T., et al., *Evolution of influenza A virus nucleoprotein genes: implications for the*
479 *origins of H1N1 human and classical swine viruses.* J Virol, 1991. **65**(7): p. 3704-14.
480 26. Wise, H.M., et al., *An alternative AUG codon in segment 5 of the 2009 pandemic influenza A*
481 *virus is a swine-derived virulence motif.* bioRxiv, 2019: p. 738427.
482 27. Wanitchang, A., et al., *Atypical characteristics of nucleoprotein of pandemic influenza virus H1N1*
483 *and their roles in reassortment restriction.* Arch Virol, 2011. **156**(6): p. 1031-40.

484 28. Friedman, N., et al., *A(H1N1)pdm09 influenza infection: vaccine inefficiency.* Oncotarget, 2017.
485 **8**(20): p. 32856-32863.
486 29. Tjon-Kon-Fat, R., et al., *The potential risks and impact of the start of the 2015-2016 influenza*
487 *season in the WHO European Region: a rapid risk assessment.* Influenza Other Respir Viruses,
488 2016. **10**(4): p. 236-246.
489 30. Newitt, S., et al., *Rapid risk assessment during the early weeks of the 2015-2016 influenza*
490 *season in Ukraine.* Influenza Other Respir Viruses, 2018. **12**(2): p. 241-249.
491 31. Public_Health_England_(PHE). *Summary of UK surveillance of influenza and other seasonal*
492 *respiratory illnesses 26 May 2016 – Week 21 report (up to week 20 data).* PHE weekly national
493 influenza report. 2016; Available from:
494 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_dat
495 a/file/681012/Weekly_national_influenza_report_week_21_2016.pdf.
496 32. Pebody, R., et al., *Effectiveness of seasonal influenza vaccine in preventing laboratory-confirmed*
497 *influenza in primary care in the United Kingdom: 2015/16 mid-season results.* Euro Surveill,
498 2016. **21**(13).
499 33. National_Institute_of_Infectious_Diseases_Japan, *Influenza 2015/16 season, Japan.* Infectious
500 Agents Surveillance Report, 2016. **37**(11): p. 211-213.
501 34. Santos, K.C., et al., *Molecular epidemiology of influenza A(H1N1)PDM09 hemagglutinin gene*
502 *circulating in Sao Paulo State , Brazil: 2016 anticipated influenza season.* Rev Inst Med Trop Sao
503 Paulo, 2017. **59**: p. e9.
504 35. Cardoso, A.M., et al., *Investigation of an outbreak of acute respiratory disease in an indigenous*
505 *village in Brazil: Contribution of Influenza A(H1N1)pdm09 and human respiratory syncytial*
506 *viruses.* PLoS One, 2019. **14**(7): p. e0218925.
507 36. Filleul, L., et al., *A major impact of the influenza seasonal epidemic on intensive care units,*
508 *Reunion, April to August 2016.* Euro Surveill, 2016. **21**(47).
509 37. Broberg, E., et al., *Predominance of influenza A(H1N1)pdm09 virus genetic subclade 6B.1 and*
510 *influenza B/Victoria lineage viruses at the start of the 2015/16 influenza season in Europe.* Euro
511 Surveill, 2016. **21**(13).
512 38. Korsun, N., et al., *Antigenic and genetic characterization of influenza viruses circulating in*
513 *Bulgaria during the 2015/2016 season.* Infect Genet Evol, 2017. **49**: p. 241-250.
514 39. Lv, J., et al., *Amino acid substitutions in the neuraminidase protein of an H9N2 avian influenza*
515 *virus affect its airborne transmission in chickens.* Vet Res, 2015. **46**: p. 44.
516 40. Bialas, K.M., E.A. Desmet, and T. Takimoto, *Specific residues in the 2009 H1N1 swine-origin*
517 *influenza matrix protein influence virion morphology and efficiency of viral spread in vitro.* PLoS
518 One, 2012. **7**(11): p. e50595.
519 41. Gong, Y.N., et al., *Population dynamics at neuraminidase position 151 of influenza A*
520 *(H1N1)pdm09 virus in clinical specimens.* J Gen Virol, 2019. **100**(5): p. 752-759.
521 42. Tate, M.D., et al., *Playing hide and seek: how glycosylation of the influenza virus hemagglutinin*
522 *can modulate the immune response to infection.* Viruses, 2014. **6**(3): p. 1294-316.
523 43. Hale, B.G., et al., *Inefficient control of host gene expression by the 2009 pandemic H1N1*
524 *influenza A virus NS1 protein.* J Virol, 2010. **84**(14): p. 6909-22.
525

526

527 **Figure Legends.**

528 **Figure 1. H1N1pdm clade 6B.1 and clade 6B.2 circulated in Baltimore and**
529 **northern Taiwan in the 2015-16 season.** Maximum likelihood (ML) tree of HA gene
530 coding sequences from 73 H1N1pdm viruses of the surveillance was generated using
531 RAxML under GTRGAMMA model with 1000 bootstrap replicates. Their time scales
532 based on sample collection dates, noted in each sample, were calibrated using
533 coalescent model in TreeTime after building ML tree. Reference sequences are
534 H1N1pdm vaccine strains, A/California/07/2009 and A/Michigan/45/2015. Tips were
535 colored by location of sample collection. Samples from Baltimore were colored in brown;
536 samples from northern Taiwan were colored in green. Most samples were grouped into
537 two clusters, 6B.1 (in red) and 6B.2 (in blue). Specific amino acid mutations defining
538 major branches of 6B.1 and 6B.2 were indicated on the side of branch.

539

540 **Figure 2. Four distinct genetic subgroups of H1N1pdm were identified in**
541 **Baltimore and northern Taiwan in the 2015-16 season.** Maximum likelihood (ML)
542 trees of concatenated gene segments of 73 H1N1pdm viruses of the surveillance were
543 generated using RAxML under GTRGAMMA model with 1000 bootstrap replicates.
544 Their time scales based on sample collection dates, noted in each sample, were
545 calibrated using coalescent model in TreeTime after building ML trees. Reference
546 sequences are H1N1pdm vaccine strains, A/California/07/2009 and
547 A/Michigan/45/2015. Tips were colored by location of sample collection. Samples from
548 Baltimore were colored in brown; samples from northern Taiwan were colored in green.
549 A schematic representation of clade/mutation of each gene segment based on their

550 individual gene tree branches to separate annotated clades (1, 6B, 6B.1, or 6B.2) and
551 amino acid mutations (see Figure S1) was annotated and visualized using the ggtree R
552 package. Mutations (Mut) in NA, M, NS, and HA segments were identified to be able to
553 divide majority of viruses into four distinct genetic subgroups, 6B.1 parental, 6B.1
554 variant, 6B.2 variant-1 and 6B.2 variant-2.
555
556 **Table 1. Univariate analysis of patients infected with H1N1pdm 6B.1 and 6B.2**
557 **viruses of Baltimore and northern Taiwan in the 2015-16 season**
558
559 **Table 2. Logistic regression analysis of pneumonia in patients infected with**
560 **H1N1pdm 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16**
561 **season**
562
563 **Table 3. Univariate analysis of patients infected with H1N1pdm 6B.1 parental and**
564 **6B.1 variant viruses of Baltimore and northern Taiwan in the 2015-16 season**
565
566 **Table 4. Univariate analysis of patients infected with H1N1pdm 6B.2 variant-1 and**
567 **6B.2 variant-2 viruses of Baltimore and northern Taiwan in the 2015-16 season**
568
569 **Supplemental Figure 1. All other segments than HA of H1N1pdm viruses in the**
570 **2015-16 season were differentially evolving toward either the clade 6B.1 or 6B.2.**
571 Phylogenetic trees of the PB2, PB1, PA, PA-X, NP, NA, M1, M2, NS1 and NS2 coding
572 sequences (subfigures from A to J, respectively) of 73 H1N1pdm viruses of the

surveillance were generated using RAxML under GTRGAMMA model with 1000 bootstrap replicates and visualized using the ggtree R package. Branch tips were colored by HA clades, and amino acid mutations on branch were shown. Reference sequences of H1N1pdm vaccine strains, A/California/07/2009 (clade 1), and A/Michigan/45/2015 (clade 6B.1) were included. NP coding sequence numbering begins from upstream start codon.

Supplemental Table 1. GenBank accession numbers of each segment of 73 H1N1pdm strains in this study.

Supplemental Table 2. Amino acid differences (H1 numbering) of HA gene in 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16 season

Supplemental Table 3. Amino acid differences of NA gene in 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16 season

Supplemental Table 4. Amino acid differences of internal genes in 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16 season

Figure 1

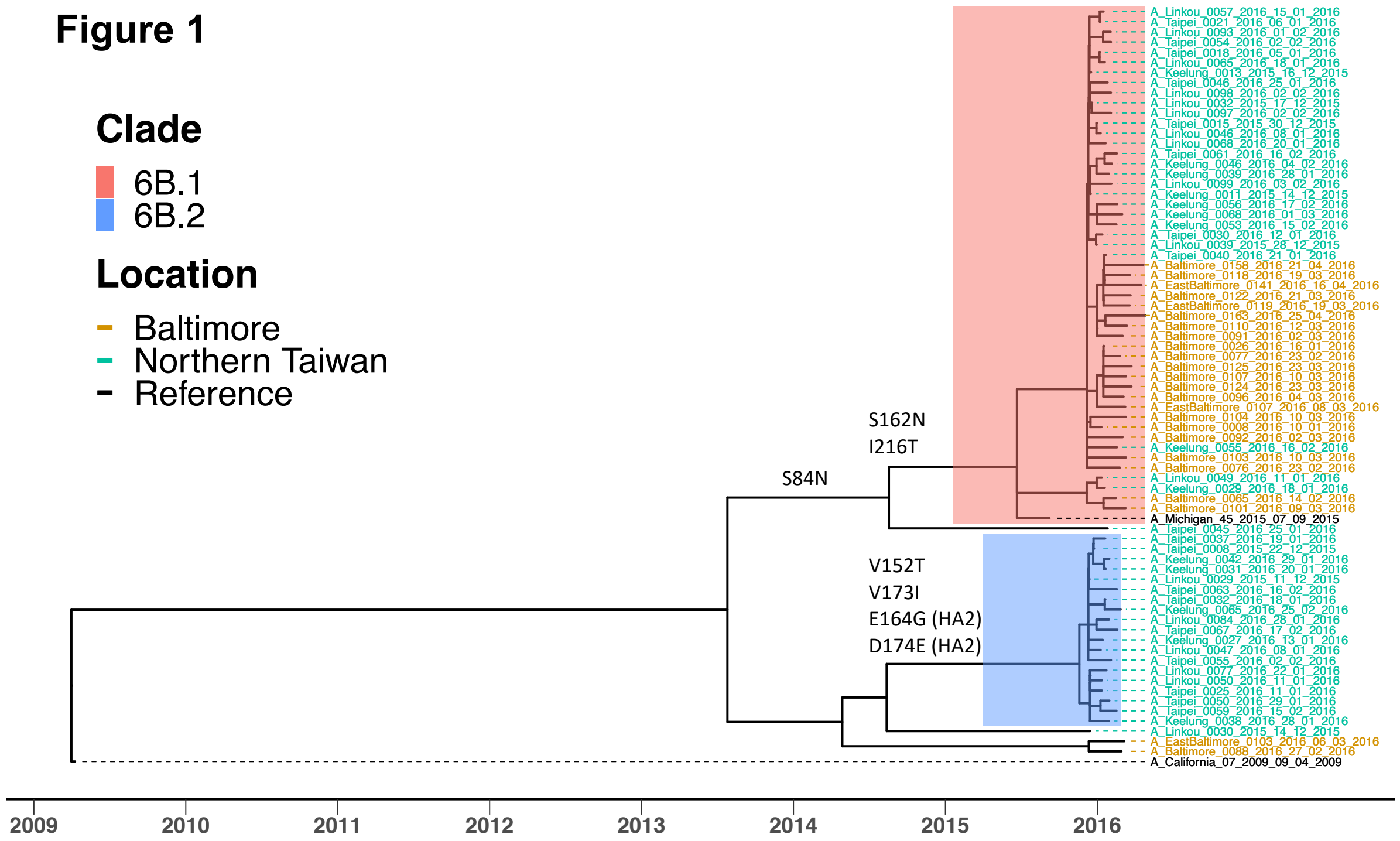

Clade
6B.1
6B.2
Location
Baltimore
Northern Taiwan
Reference
S162N
I216T
S84N
V152T
V173I
E164G (HA2)
D174E (HA2)
A_Linkou_0057_2016_15_01_2016
A_Taipei_0021_2016_06_01_2016
A_Linkou_0093_2016_01_02_2016
A_Taipei_0054_2016_02_02_2016
A_Taipei_0018_2016_05_01_2016
A_Linkou_0065_2016_18_01_2016
A_Keelung_0013_2015_16_12_2015
A_Taipei_0046_2016_25_01_2016
A_Linkou_0098_2016_02_02_2016
A_Linkou_0032_2015_17_12_2015
A_Linkou_0097_2016_02_02_2016
A_Taipei_0015_2015_30_12_2015
A_Linkou_0046_2016_08_01_2016
A_Linkou_0068_2016_20_01_2016
A_Taipei_0061_2016_16_02_2016
A_Keelung_0046_2016_04_02_2016
A_Keelung_0039_2016_28_01_2016
A_Linkou_0099_2016_03_02_2016
A_Keelung_0011_2015_14_12_2015
A_Keelung_0056_2016_17_02_2016
A_Keelung_0068_2016_01_03_2016
A_Keelung_0053_2016_15_02_2016
A_Taipei_0030_2016_12_01_2016
A_Linkou_0039_2015_28_12_2015
A_Taipei_0040_2016_21_01_2016
A_Baltimore_0158_2016_21_04_2016
A_Baltimore_0118_2016_19_03_2016
A_EastBaltimore_0141_2016_16_04_2016
A_Baltimore_0122_2016_21_03_2016
A_EastBaltimore_0119_2016_19_03_2016
A_Baltimore_0163_2016_25_04_2016
A_Baltimore_0110_2016_12_03_2016
A_Baltimore_0091_2016_02_03_2016
A_Baltimore_0026_2016_16_01_2016
A_Baltimore_0077_2016_23_02_2016
A_Baltimore_0125_2016_23_03_2016
A_Baltimore_0107_2016_10_03_2016
A_Baltimore_0124_2016_23_03_2016
A_Baltimore_0096_2016_04_03_2016
A_EastBaltimore_0107_2016_08_03_2016
A_Baltimore_0104_2016_10_03_2016
A_Baltimore_0008_2016_10_01_2016
A_Baltimore_0092_2016_02_03_2016
A_Keelung_0055_2016_16_02_2016
A_Baltimore_0103_2016_10_03_2016
A_Baltimore_0076_2016_23_02_2016
A_Linkou_0049_2016_11_01_2016
A_Keelung_0029_2016_18_01_2016
A_Baltimore_0065_2016_14_02_2016
A_Baltimore_0101_2016_09_03_2016
A_Michigan_45_2015_07_09_2015
A_Taipei_0045_2016_25_01_2016
A_Taipei_0037_2016_19_01_2016
A_Taipei_0008_2015_22_12_2015
A_Keelung_0042_2016_29_01_2016
A_Keelung_0031_2016_20_01_2016
A_Linkou_0029_2015_11_12_2015
A_Taipei_0063_2016_16_02_2016
A_Taipei_0032_2016_18_01_2016
A_Keelung_0065_2016_25_02_2016
A_Linkou_0084_2016_28_01_2016
A_Taipei_0067_2016_17_02_2016
A_Keelung_0027_2016_13_01_2016
A_Linkou_0047_2016_08_01_2016
A_Taipei_0055_2016_02_02_2016
A_Linkou_0077_2016_22_01_2016
A_Linkou_0050_2016_11_01_2016
A_Taipei_0025_2016_11_01_2016
A_Taipei_0050_2016_29_01_2016
A_Taipei_0059_2016_15_02_2016
A_Keelung_0038_2016_28_01_2016
A_Linkou_0030_2015_14_12_2015
A_EastBaltimore_0103_2016_06_03_2016
A_Baltimore_0088_2016_27_02_2016
A_California_07_2009_09_04_2009
2009
2010
2011
2012
2013
2014
2015
2016

Figure 2

Clade/Mutation

- 1
- 6B
- 6B.1
- 6B.2
- NA−F74L
- NS2−G22V
- NA−S52R,G96S
- M1−S30N,M2−Y52C
- HA−A261S

Location

- Baltimore
- Northern Taiwan
- Reference

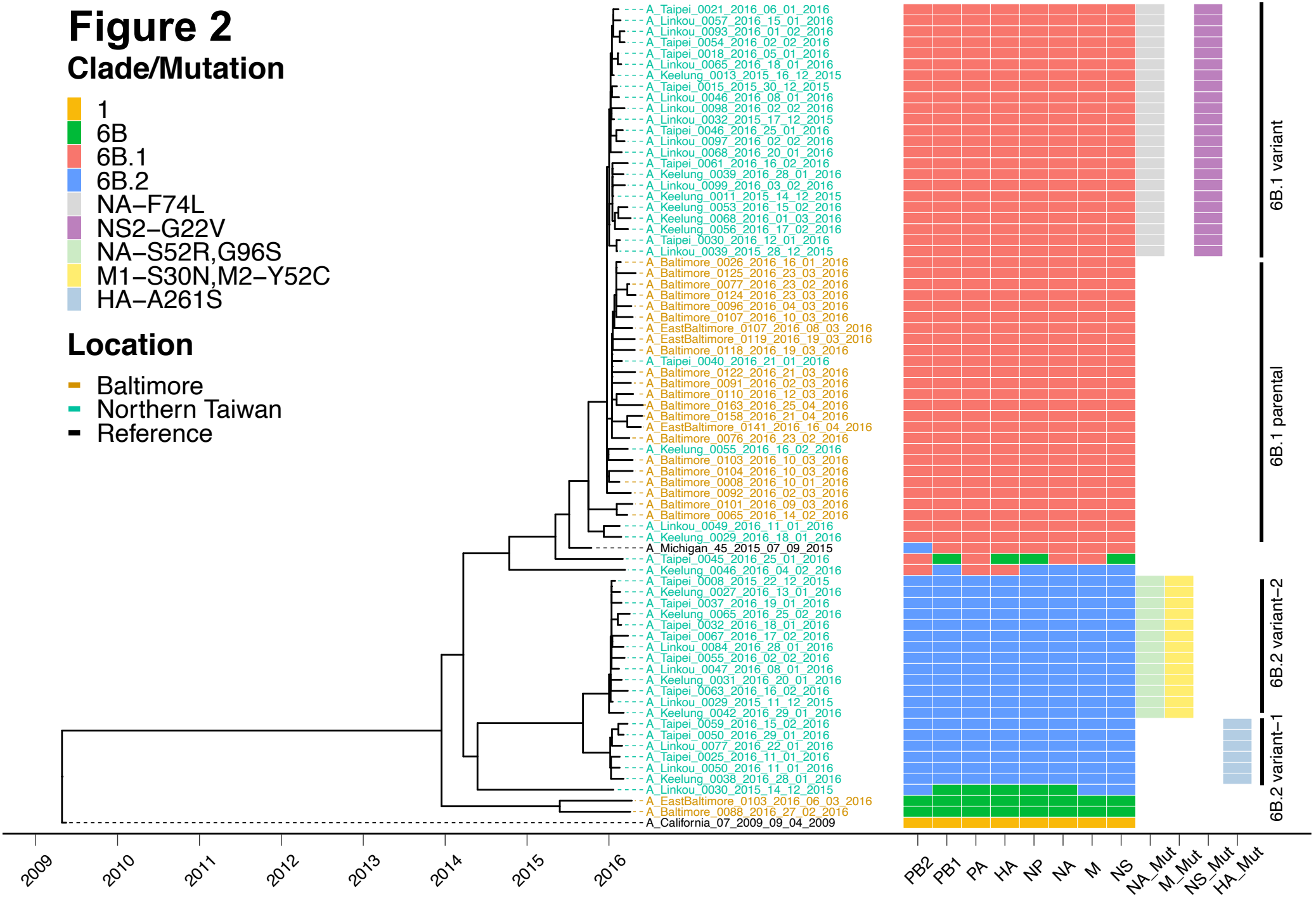
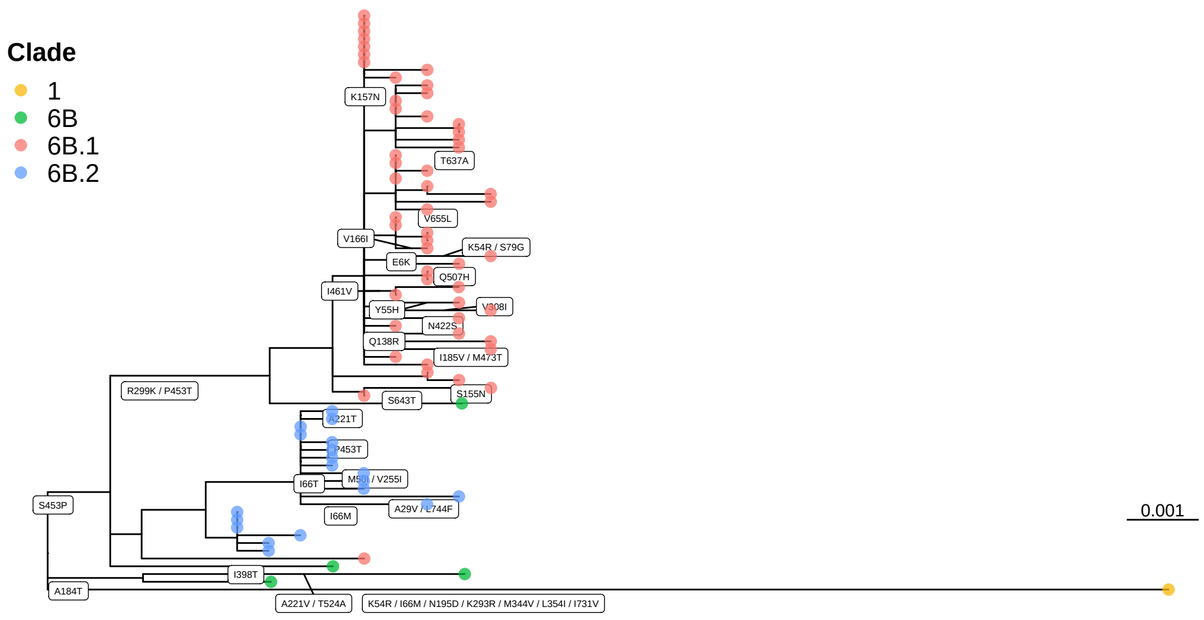
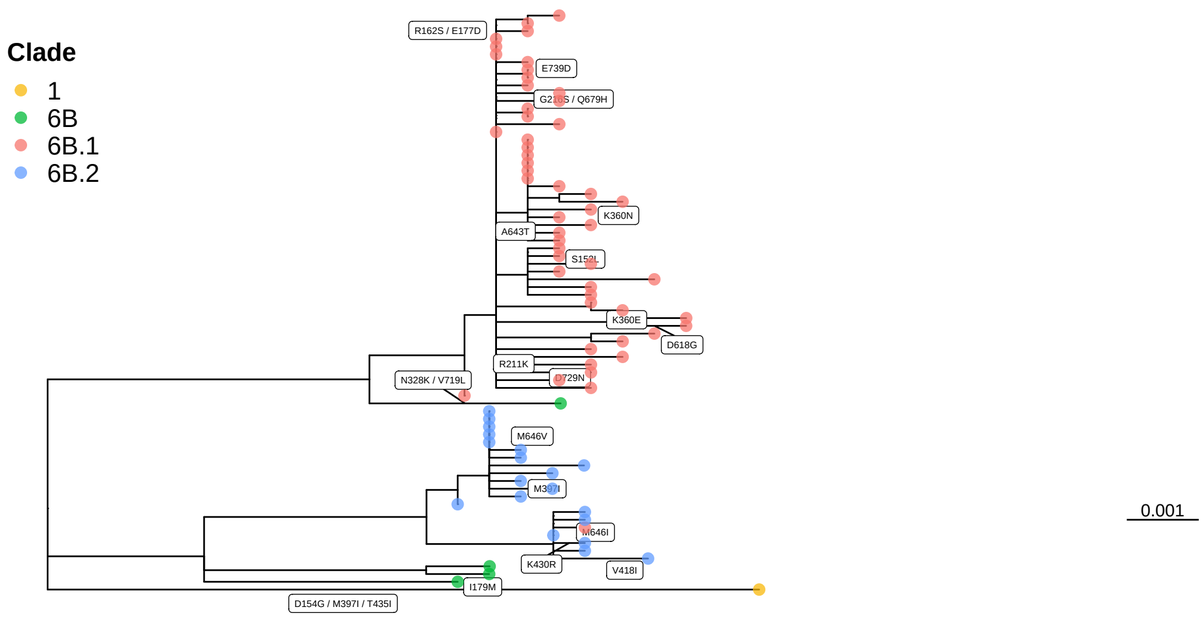


Figure S1

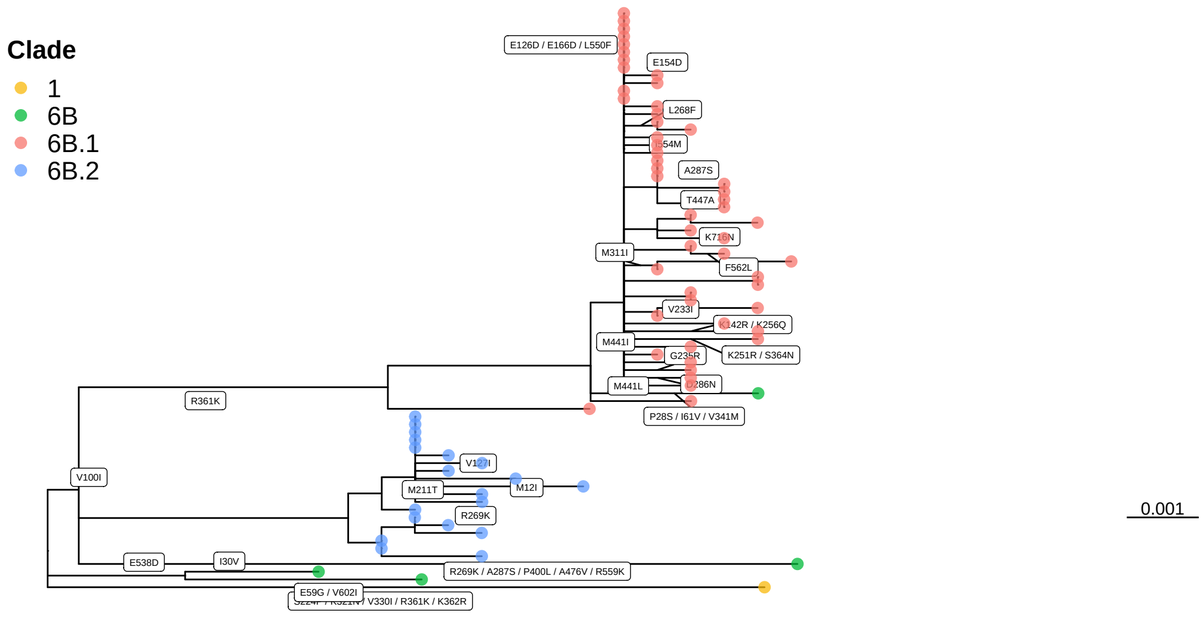
A. PB2



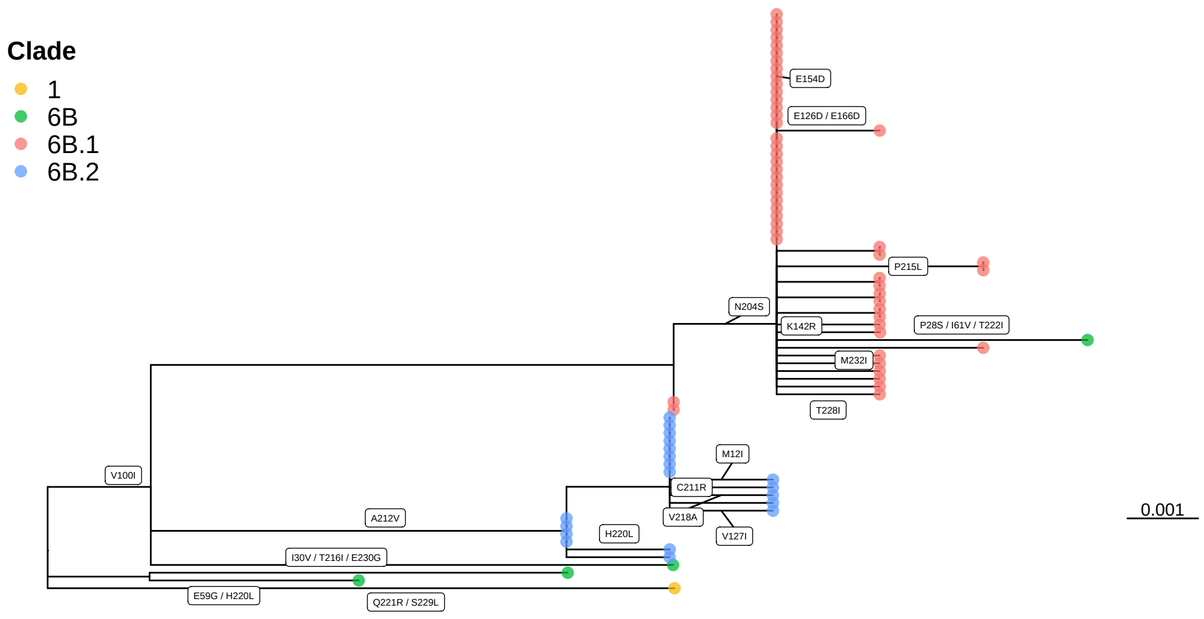
B. PB1



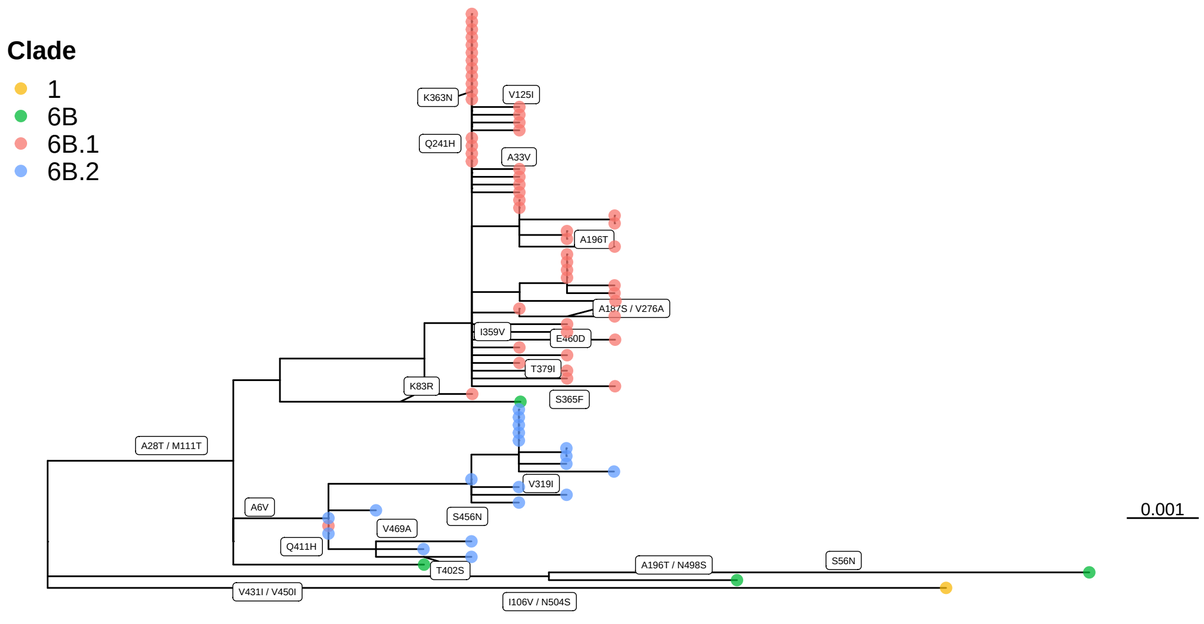
C. PA



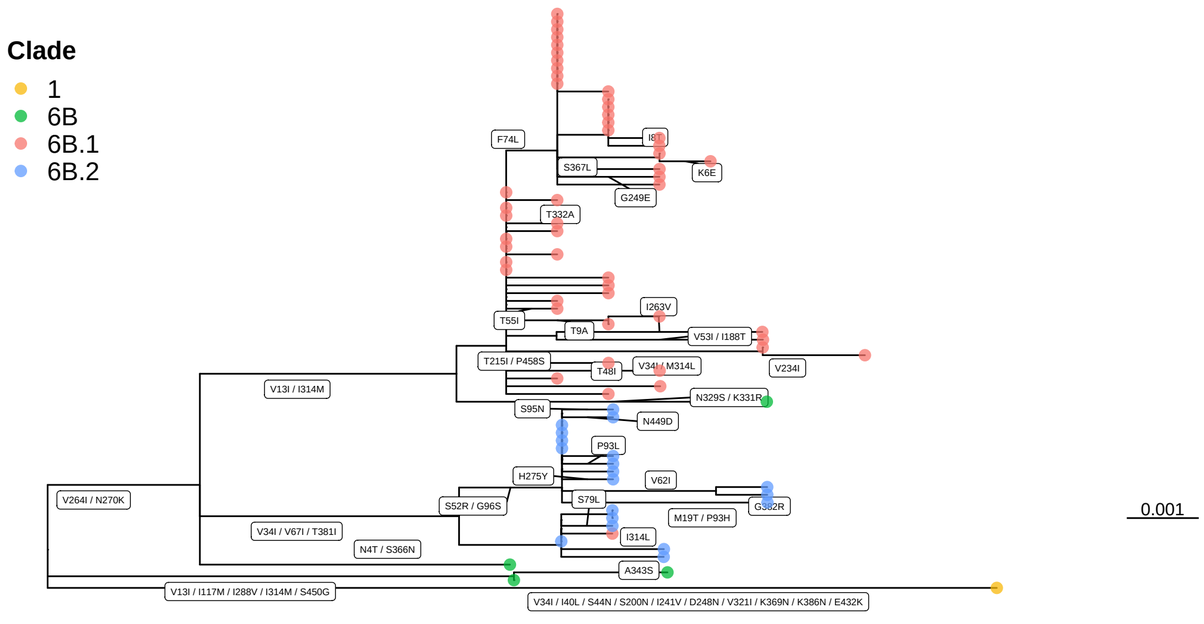
D. PA−X



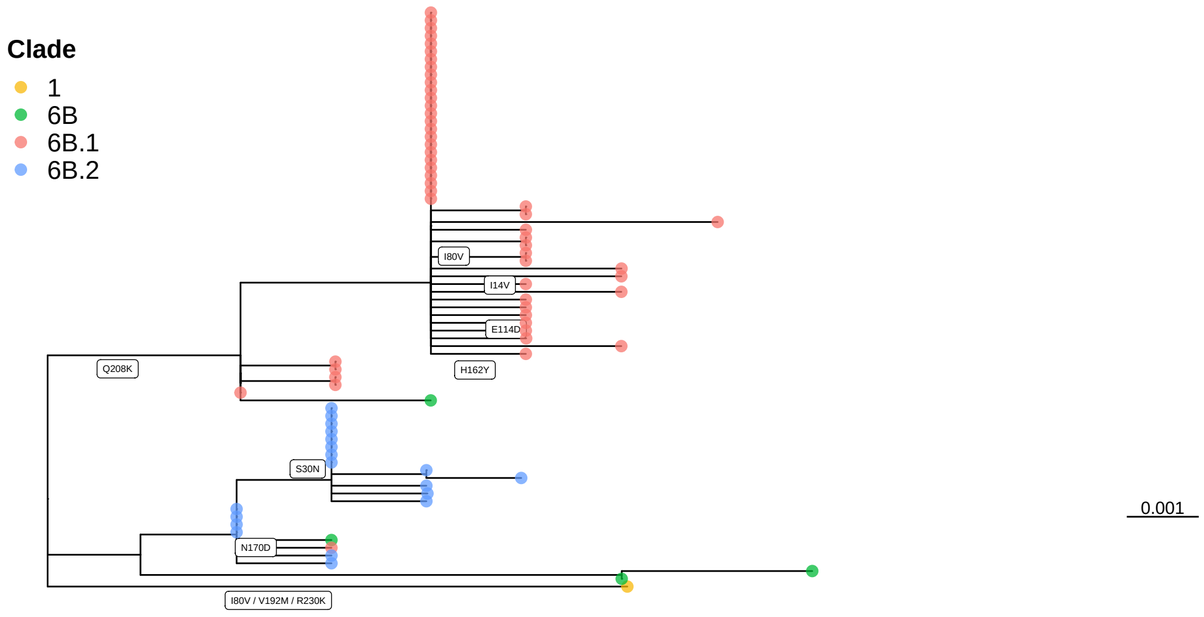
E. NP



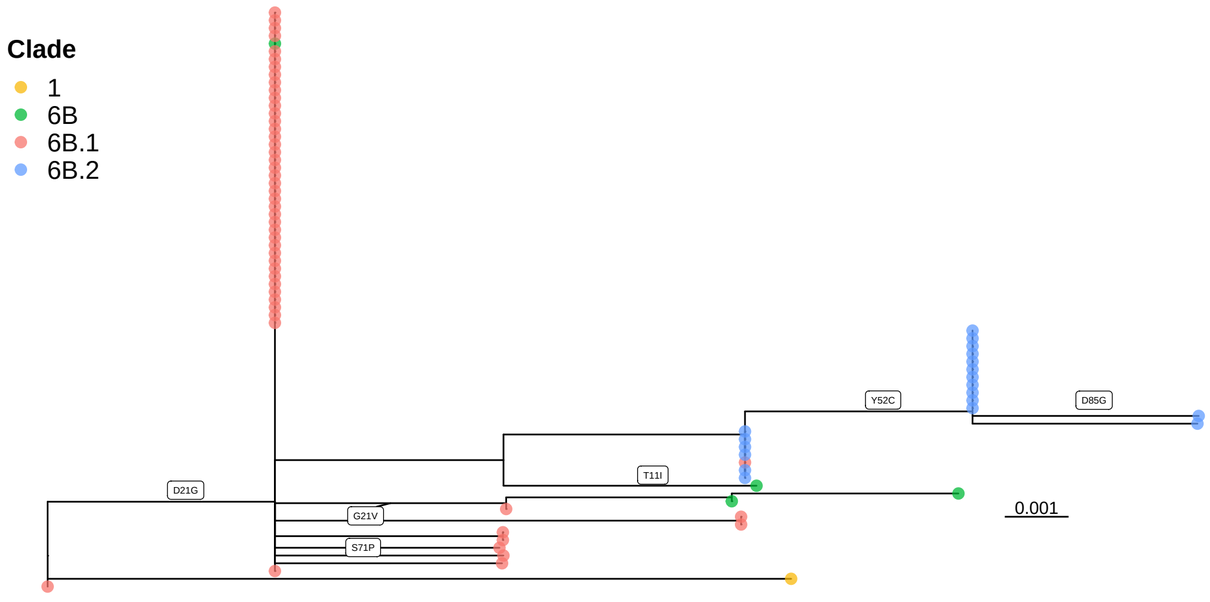
F. NA



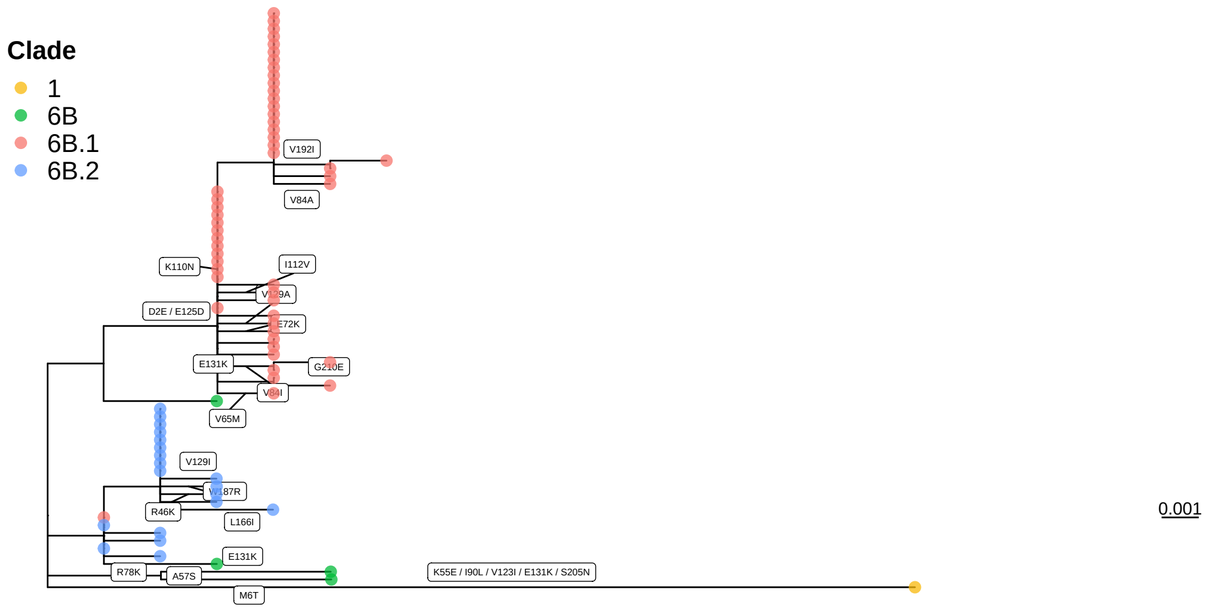
G. M1



H. M2



I. NS1



J. NS2

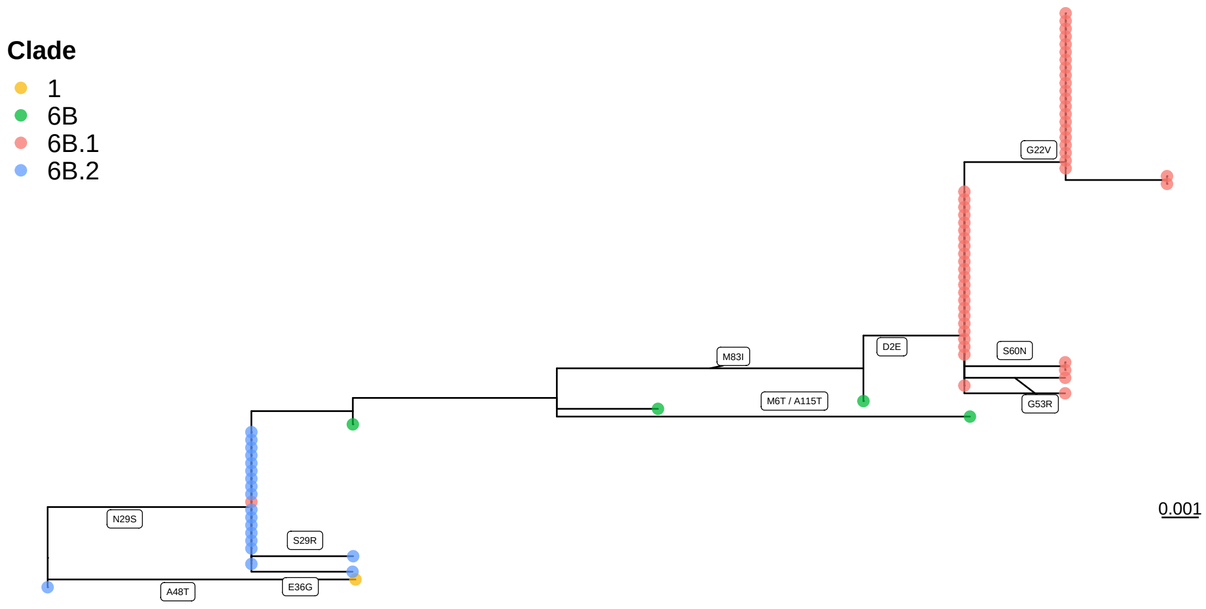


Table 1. Univariate analysis of patients infected with H1N1pdm 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16 season

Characteristic	6B.1 (n=49) N (%) / median (IQR)	6B.2 (n=19) N (%) / median (IQR)	p-value
Demographics			
Site			
JHMI	***22 (44.9)***	***0 (0)***	***<0.0001****
CGMH	27 (55.1)	19 (100.0)	
Age			
In years	34 (28-44)	37 (32-52)	0.20[a]
18-49	40 (81.6)	12 (63.2)	0.12
50-64	***6 (12.2)***	***7 (36.8)***	***0.04****
65+	3 (6.1)	0 (0)	0.55
Gender			
Female	31 (63.3)	14 (73.7)	0.57
Male	18 (36.7)	5 (26.3)	
Ethnicity			
Non-Hispanic or Latino	47 (95.9)	19 (100.0)	0.99
Hispanic or Latino	2 (4.1)	0 (0)	
Race			
Black or African American	***16 (32.7)***	***0 (0)***	***<0.01****
White	1 (2.0)	0 (0)	0.99
Asian	***27 (55.1)***	***19 (100.0)***	***<0.01****
Other	5 (10.2)	0 (0)	0.56
Vaccination and exposure			
Human exposure	11 (22.4)	2 (10.5)	0.31
Travel in last 30 days	9 (18.4)	6 (31.6)	0.33
Influenza vaccine	8 (16.3)	0 (0)	0.09
Co-morbidities and high risk			
No co-morbidities	***26 (53.1)***	***7 (36.8)***	***<0.01****
One co-morbidity	15 (30.6)	8 (42.1)	0.4
Multiple co-morbidities	8 (16.3)	4 (21.1)	0.73
Immunosuppressive medication	1 (2.0)	0 (0)	0.99
Pregnant	1 (2.0)	0 (0)	0.99
Disease severity			
NEWS score	3 (0-6)	3 (0-6)	0.81[a]
Oxygen supplementation	2 (4.1)	3 (15.8)	0.13
BiPAP/CPAP	1 (2.0)	0 (0)	0.99
Intubated	0 (0)	0 (0)	0.99
Pneumonia	***7 (14.3)***	***7 (36.8)***	***0.05****
Admitted or Observation	6 (12.2)	2 (10.5)	0.99
LOS	7.0 (1-12)	10.0 (5-15)	0.14[a]

	ICU	0 (0)	0 (0)	0.99
	Death	0 (0)	0 (0)	0.99
	Follow up due to sequala	3 (6.1)	1 (5.3)	0.99
	Viral co-infection	0 (0)	0 (0)	0.99
Symptoms				
	fever	47 (95.9)	19 (100.0)	0.99
	cough	48 (98.0)	19 (100.0)	0.99
	sputum	32 (65.3)	14 (73.7)	0.77
	increased sputum	17 (34.7)	6 (31.6)	0.78
	shortness of breath	35 (71.4)	10 (52.6)	0.16
	wheezing	16 (32.7)	10 (52.6)	0.16
	headache	40 (81.6)	18 (94.7)	0.26
	sore throat	39 (79.6)	15 (78.9)	0.99
	body aches	43 (87.8)	19 (100.0)	0.17
	runny nose	***37 (75.5)***	***19 (100.0)***	***0.02****
	sinus pain	13 (26.5)	1 (5.3)	0.091
	fatigue	43 (87.8)	18 (94.7)	0.66
	chest pain	30 (61.2)	8 (42.1)	0.18
	chills	38 (77.6)	16 (84.2)	0.74
	loss of appetite	36 (73.5)	16 (84.2)	0.53
	nausea/vomiting	28 (57.1)	12 (63.2)	0.79
	diarrhea	***10 (20.4)***	***9 (47.4)***	***0.04****
	conjunctivitis	11 (22.4)	2 (10.5)	0.33

Statistical analysis was done using R software.
Categorical variables were compared using Chi-square or Fisher's exact tests.
[a] Continuous variables were compared with rank-sum test.
* Statistical significance was set at $p<0.05$.

Table 2. Logistic regression analysis of pneumonia in patients infected with H1N1pdm 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16 season

		Unadjusted			IPTW**		
Clade	**Pneumonia**	**Odds Ratio**	**95% CI**	**p-value**	**Odds Ratio**	**95% CI**	**p-value**
6B.1	14.3% (7/49)	1	-	-	1	-	-
6B.2	36.8% (7/19)	3.500	1.629-5.372	0.046*	3.261	1.696-4.826	0.008*

*p<0.05, logistic regression analysis.

** IPTW (inverse probability of treatment weighting) results were done by propensity score weighting with recursive partitioning algorithm (area under the ROC curve: 0.798) for age, gender, comorbidities, obesity, vaccine status, human exposure and travel history, then followed by logistic regression analysis.

Table 3. Univariate analysis of patients infected with H1N1pdm 6B.1 parental and 6B.1 variant viruses of Baltimore and northern Taiwan in the 2015-16 season

Characteristic			6B.1 Parental (n=26) N (%) / median (IQR)	6B.1 Variant (n=23) N (%) / median (IQR)	p-value
Demographics					
	Site				
		JHMI	***22 (84.6)***	***0 (0)***	***<0.0001****
		CGMH	4 (15.4)	23 (100.0)	
	Age				
		In years	33 (29-44)	35 (29-44)	0.75[a]
		18-49	20 (76.9)	20 (87.0)	0.47
		50-64	5 (19.2)	1 (4.3)	0.19
		65+	1 (3.8)	2 (8.7)	0.59
	Gender				
		Female	14 (53.8)	17 (73.9)	0.24
		Male	12 (46.2)	6 (26.1)	
	Ethnicity				
		Non-Hispanic or Latino	24 (92.3)	23 (100.0)	0.49
		Hispanic or Latino	2 (7.7)	0 (0)	
	Race				
		Black or African American	***16 (61.5)***	***0 (0)***	***<0.0001****
		White	1 (3.8)	0 (0)	0.99
		Asian	***4 (15.4)***	***23 (100.0)***	***<0.0001****
		Other	5 (19.3)	0 (0)	0.08
	Vaccination and exposure				
		Human exposure	4 (15.4)	7 (30.4)	0.36
		Travel in last 30 days	3 (11.5)	6 (26.1)	0.35
		Influenza vaccine	6 (23.1)	2 (8.7)	0.33
	Co-morbidities and high risk				
		No co-morbidities	11 (42.3)	15 (65.2)	0.15
		One co-morbidity	8 (30.8)	7 (30.4)	0.99
		Multiple co-morbidities	7 (26.9)	1 (4.3)	0.05
		Immunosuppressive medication	1 (3.8)	0 (0)	0.99
		Pregnant	1 (3.8)	0 (0)	0.99
Disease severity					
	NEWS score		3 (0-6)	3 (0-6)	0.57[a]
	Oxygen supplementation		2 (7.7)	0 (0)	0.49
	BiPAP/CPAP		1 (3.8)	0 (0)	0.49
	Intubated		0 (0)	0 (0)	0.99
	Pneumonia		3 (11.5)	4 (17.4)	0.69
	Admitted or Observation		4 (15.4)	2 (8.7)	0.67
	LOS		6.5 (1-12)	6.0 (3-9)	0.67[a]

	ICU	0 (0)	0 (0)	0.99
	Death	0 (0)	0 (0)	0.99
	Follow up due to sequala	2 (7.7)	1 (4.3)	0.99
	Viral co-infection	0 (0)	0 (0)	0.99
Symptoms				
	fever	24 (92.3)	23 (100.0)	0.49
	cough	25 (96.2)	23 (100.0)	0.99
	sputum	15 (57.7)	17 (73.9)	0.37
	increased sputum	10 (38.5)	7 (30.4)	0.77
	shortness of breath	20 (76.9)	15 (65.2)	0.53
	wheezing	***12 (46.2)***	***4 (17.4)***	***0.04****
	headache	23 (88.5)	17 (73.9)	0.27
	sore throat	19 (73.1)	20 (87.0)	0.3
	body aches	24 (92.3)	19 (82.6)	0.4
	runny nose	17 (65.4)	20 (87.0)	0.1
	sinus pain	***11 (42.3)***	***2 (8.7)***	***0.01****
	fatigue	22 (84.6)	21 (91.3)	0.67
	chest pain	19 (73.1)	11 (47.8)	0.08
	chills	21 (80.8)	17 (73.9)	0.73
	loss of appetite	20 (76.9)	16 (69.6)	0.75
	nausea/vomiting	***19 (73.1)***	***9 (39.1)***	***0.02****
	diarrhea	7 (26.9)	3 (13.0)	0.3
	conjunctivitis	8 (30.8)	3 (13.0)	0.18

Statistical analysis was done using R software.
Categorical variables were compared using Chi-square or Fisher's exact tests.
[a] Continuous variables were compared with rank-sum test.
* Statistical significance was set at $p<0.05$.

Table 4. Univariate analysis of patients infected with H1N1pdm 6B.2 variant-1 and 6B.2 variant-2 viruses of Baltimore and northern Taiwan in the 2015-16 season

Characteristic	6B.2 Variant-1 (n=6) N (%) / median (IQR)	6B.2 Variant-2 (n=13) N (%) / median (IQR)	p-value
Demographics			
Site			
JHMI	0 (0)	0 (0)	0.99
CGMH	6 (100.0)	13 (100.0)	
Age			
In years	34 (29-48)	44 (34-54)	0.20[a]
18-49	4 (66.7)	8 (61.5)	0.99
50-64	2 (33.3)	5 (38.5)	0.99
65+	0 (0)	0 (0)	0.99
Gender			
Female	3 (50.0)	11 (84.6)	0.26
Male	3 (50.0)	2 (15.4)	
Ethnicity			
Non-Hispanic or Latino	6 (100.0)	13 (100.0)	0.99
Hispanic or Latino	0 (0)	0 (0)	
Race			
Black or African American	0 (0)	0 (0)	0.99
White	0 (0)	0 (0)	0.99
Asian	6 (100.0)	13 (100.0)	0.99
Other	0 (0)	0 (0)	0.99
Vaccination and exposure			
Human exposure	0 (0)	2 (15.4)	0.83
Travel in last 30 days	3 (50.0)	3 (23.1)	0.32
Influenza vaccine	0 (0)	0 (0)	0.99
Co-morbidities and high risk			
No co-morbidities	3 (50.0)	4 (30.8)	0.77
One co-morbidity	2 (33.3)	6 (46.2)	0.99
Multiple co-morbidities	1 (16.7)	3 (23.1)	0.99
Immunosuppressive medication	0 (0)	0 (0)	0.99
Pregnant	0 (0)	0 (0)	0.99
Disease severity			
NEWS score	4 (2-6)	3 (0-5)	0.54[a]
Oxygen supplementation	0 (0)	3 (23.1)	0.52
BiPAP/CPAP	0 (0)	0 (0)	0.99
Intubated	0 (0)	0 (0)	0.99
Pneumonia	2 (33.3)	5 (38.5)	0.99
Admitted or Observation	0 (0)	2 (15.4)	0.99
LOS	***0 (0-0)***	***10.0 (5-15)***	***<0.01****

	ICU	0 (0)	0 (0)	0.99
	Death	0 (0)	0 (0)	0.99
	Follow up due to sequala	0 (0)	1 (7.7)	0.99
	Viral co-infection	0 (0)	0 (0)	0.99
Symptoms				
	fever	6 (100.0)	13 (100.0)	0.99
	cough	6 (100.0)	13 (100.0)	0.99
	sputum	5 (83.3)	9 (69.2)	0.99
	increased sputum	3 (50.0)	3 (23.1)	0.32
	shortness of breath	2 (33.3)	8 (61.5)	0.35
	wheezing	2 (33.3)	8 (61.5)	0.35
	headache	6 (100.0)	12 (92.3)	0.99
	sore throat	6 (100.0)	9 (69.2)	0.26
	body aches	6 (100.0)	13 (100.0)	0.99
	runny nose	6 (100.0)	13 (100.0)	0.99
	sinus pain	0 (0)	1 (7.7)	0.99
	fatigue	6 (100.0)	12 (92.3)	0.99
	chest pain	2 (33.3)	6 (46.2)	0.99
	chills	5 (83.3)	11 (84.6)	0.99
	loss of appetite	6 (100.0)	10 (76.9)	0.52
	nausea/vomiting	3 (50.0)	9 (69.2)	0.62
	diarrhea	4 (66.7)	5 (38.5)	0.35
	conjunctivitis	0 (0)	2 (15.4)	0.99

Statistical analysis was done using R software.
Categorical variables were compared using Chi-square or Fisher's exact tests.
[a] Continuous variables were compared with rank-sum test.
* Statistical significance was set at $p<0.05$.

Supplemental Table 1. GenBank accession numbers of each segment of 73 H1N1pdm strains in this study.

Strain name	HA clade	PB2	PB1	PA	HA	NP	NA	M	NS
A/Baltimore/0008/2016	6B.1	KY487410	KY487672	KY487343	KY487698	KY487491	KY487604	KY487576	KY487127
A/Baltimore/0026/2016	6B.1	KY949909	KY949764	KY950028	KY950126	KY949939	KY949753	KY950197	KY949800
A/Baltimore/0065/2016	6B.1	KY950125	KY949852	KY950143	KY950052	KY949710	KY949880	KY950094	KY950041
A/Baltimore/0076/2016	6B.1	KY487315	KY487375	KY487419	KY487689	KY487608	KY487361	KY487616	KY487333
A/Baltimore/0077/2016	6B.1	KY949732	KY949997	KY949930	KY950108	KY949912	KY949844	KY950209	KY950085
A/Baltimore/0088/2016	6B	KY487517	KY487300	KY615534	KY487140	KY487313	KY487269	KY487124	KY487551
A/Baltimore/0091/2016	6B.1	KY487652	KY487289	KY487640	KY615553	KY615415	KY487668	KY487464	KY487393
A/Baltimore/0092/2016	6B.1	KY487138	KY487194	KY615404	KY487543	KY487486	KY487463	KY487409	KY487653
A/Baltimore/0096/2016	6B.1	KY487098	KY615465	KY487452	KY615388	KY487164	KY487401	KY487428	KY487191
A/Baltimore/0101/2016	6B.1	KY487518	KY487065	KY487522	KY487503	KY487259	KY615395	KY487263	KY487762
A/Baltimore/0103/2016	6B.1	KY487208	KY487397	KY487290	KY487506	KY487341	KY487256	KY487384	KY487201
A/Baltimore/0104/2016	6B.1	KY487282	KY487502	KY487344	KY487646	KY487471	KY487475	KY487605	KY487212
A/Baltimore/0107/2016	6B.1	KY487556	KY487562	KY487722	KY487472	KY487183	KY487084	KY487451	KY487439
A/Baltimore/0110/2016	6B.1	KY487731	KY615501	KY487168	KY615408	KY487185	KY615483	KY487645	KY487752
A/Baltimore/0118/2016	6B.1	KY615383	KY615428	KY615505	KY615604	KY615376	KY615516	KY487552	KY615550
A/Baltimore/0122/2016	6B.1	KY615526	KY615507	KY487237	KY615559	KY615489	KY615491	KY487461	KY615412
A/Baltimore/0124/2016	6B.1	KY487667	KY487566	KY487101	KY487575	KY487158	KY487434	KY615563	KY487666
A/Baltimore/0125/2016	6B.1	KY487348	KY487415	KY487427	KY487453	KY487458	KY487175	KY487215	KY487497
A/Baltimore/0158/2016	6B.1	KY487156	KY615599	KY487293	KY487430	KY615592	KY487539	KY487525	KY487281
A/Baltimore/0163/2016	6B.1	KY487433	KY487770	KY487188	KY487592	KY487572	KY487523	KY615557	KY487524
A/EastBaltimore/0103/2016	6B	KY487534	KY487611	KY487704	KY487656	KY487615	KY487392	KY487465	KY487599
A/EastBaltimore/0107/2016	6B.1	KY487621	KY487721	KY487292	KY487123	KY487223	KY487136	KY487240	KY487116
A/EastBaltimore/0119/2016	6B.1	KY487252	KY487758	KY487325	KY615447	KY487381	KY487391	KY487596	KY487498
A/EastBaltimore/0141/2016	6B.1	KY615562	KY615541	KY615512	KY487527	KY487701	KY487628	KY487529	KY487696
A/Linkou/0029/2015	6B.2	KY949707	KY950139	KY949697	KY950184	KY950083	KY949706	KY950178	KY950073
A/Linkou/0030/2015	6B	KY949913	KY950117	KY950138	KY949802	KY949767	KY949760	KY949954	KY949636
A/Linkou/0032/2015	6B.1	KY949925	KY949664	KY950068	KY949823	KY950112	KY949903	KY949646	KY950070

A/Linkou/0039/2015	6B.1	KY949951	KY949669	KY949719	KY949680	KY949722	KY949743	KY949988	KY950158
A/Linkou/0046/2016	6B.1	KY949658	KY950088	KY949714	KY949679	KY949884	KY949888	KY949993	KY950131
A/Linkou/0047/2016	6B.2	KY949962	KY950065	KY949782	KY949975	KY950049	KY949999	KY949696	KY949731
A/Linkou/0049/2016	6B.1	KY950066	KY950082	KY950071	KY950223	KY949703	KY949738	KY949901	KY950106
A/Linkou/0050/2016	6B.2	KY949917	KY949704	KY949660	KY949736	KY950061	KY950170	KY949787	KY950194
A/Linkou/0057/2016	6B.1	KY487199	KY615531	KY487496	KY615390	KY615380	KY487310	KY487581	KY487584
A/Linkou/0065/2016	6B.1	KY487225	KY487445	KY487071	KY487702	KY487404	KY487488	KY487321	KY487373
A/Linkou/0068/2016	6B.1	KY487554	KY487719	KY487387	KY615570	KY487505	KY615429	KY487610	KY487555
A/Linkou/0077/2016	6B.2	KY487278	KY615572	KY487598	KY487687	KY487372	KY487383	KY487283	KY487087
A/Linkou/0084/2016	6B.2	KY487285	KY615373	KY487481	KY487767	KY487487	KY487099	KY487376	KY487500
A/Linkou/0093/2016	6B.1	KY487658	KY487319	KY487260	KY487613	KY487717	KY487623	KY487380	KY487067
A/Linkou/0097/2016	6B.1	KY487417	KY615513	KY615499	KY615425	KY615535	KY487639	KY487328	KY487213
A/Linkou/0098/2016	6B.1	KY487303	KY615539	KY487195	KY615441	KY487327	KY487137	KY487294	KY487720
A/Linkou/0099/2016	6B.1	KY487476	KY487764	KY487131	KY487353	KY487597	KY487197	KY487068	KY487072
A/Taipei/0008/2015	6B.2	KY487634	KY487510	KY487631	KY487478	KY487715	KY487143	KY487364	KY487755
A/Taipei/0015/2015	6B.1	KY487198	KY487557	KY487423	KY487273	KY615585	KY487485	KY487507	KY487200
A/Taipei/0018/2016	6B.1	KY487754	KY487607	KY487690	KY487725	KY487167	KY487501	KY487692	KY487378
A/Taipei/0021/2016	6B.1	KY487085	KY487345	KY487106	KY487147	KY615436	KY487459	KY487339	KY487211
A/Taipei/0025/2016	6B.2	KY487148	KY487326	KY487145	KY487209	KY487275	KY487484	KY487119	KY487362
A/Taipei/0030/2016	6B.1	KY615580	KY487726	KY487079	KY487336	KY487277	KY487553	KY487489	KY487386
A/Taipei/0032/2016	6B.2	KY487304	KY487371	KY487083	KY487330	KY487239	KY487718	KY487511	KY487663
A/Taipei/0037/2016	6B.2	KY487264	KY487441	KY487738	KY487649	KY487232	KY615564	KY487243	KY487349
A/Taipei/0040/2016	6B.1	KY487411	KY487538	KY615496	KY487744	KY487241	KY487229	KY487089	KY487257
A/Taipei/0045/2016	6B	KY487296	KY487456	KY487367	KY487602	KY487178	KY487714	KY487426	KY487565
A/Taipei/0046/2016	6B.1	KY487740	KY487531	KY487560	KY487436	KY487504	KY487170	KY487734	KY487635
A/Taipei/0050/2016	6B.2	KY487526	KY487474	KY487110	KY487591	KY487312	KY487102	KY487746	KY487111
A/Taipei/0054/2016	6B.1	KY487736	KY487169	KY487299	KY487403	KY487184	KY487247	KY487449	KY487258
A/Taipei/0055/2016	6B.2	KY487595	KY487249	KY487218	KY487233	KY487660	KY487733	KY487712	KY487528
A/Taipei/0059/2016	6B.2	KY487320	KY487118	KY615435	KY615586	KY615461	KY487324	KY487648	KY487291
A/Taipei/0061/2016	6B.1	KY487379	KY487546	KY487732	KY487248	KY487224	KY487350	KY487636	KY487360

A/Taipei/0063/2016	6B.2	KY487629	KY487090	KY615394	KY487365	KY487571	KY487230	KY487516	KY487174
A/Taipei/0067/2016	6B.2	KY487311	KY615407	KY487470	KY487709	KY487388	KY487675	KY487255	KY487226
A/Keelung/0011/2015	6B.1	KY487600	KY487113	KY487166	KY487251	KY487574	KY487437	KY487699	KY487593
A/Keelung/0013/2015	6B.1	KY950069	KY949665	KY949995	KY949643	KY950031	KY950092	KY949803	KY950115
A/Keelung/0027/2016	6B.2	KY487512	KY487713	KY487086	KY487340	KY487676	KY487231	KY487134	KY487418
A/Keelung/0029/2016	6B.1	KY615478	KY487334	KY615481	KY487499	KY487769	KY615400	KY487400	KY487314
A/Keelung/0031/2016	6B.2	KY487363	KY487573	KY615579	KY615480	KY487589	KY487204	KY487276	KY487444
A/Keelung/0038/2016	6B.2	KY487483	KY487739	KY487614	KY487655	KY487268	KY487466	KY487091	KY487129
A/Keelung/0039/2016	6B.1	KY949952	KY949976	KY950181	KY949887	KY949948	KY949872	KY949898	KY949946
A/Keelung/0042/2016	6B.2	KY487244	KY487186	KY487187	KY487695	KY487274	KY615567	KY487399	KY615457
A/Keelung/0046/2016	6B.1	KY487665	KY487761	KY615468	KY487149	KY487406	KY487271	KY487245	KY487217
A/Keelung/0053/2016	6B.1	KY487594	KY487609	KY487684	KY487727	KY487088	KY487181	KY487177	KY487748
A/Keelung/0055/2016	6B.1	KY949741	KY949863	KY950220	KY950148	KY949733	KY949761	KY949653	KY950042
A/Keelung/0056/2016	6B.1	KY487082	KY487495	KY615527	KY487617	KY487368	KY487416	KY487443	KY487165
A/Keelung/0065/2016	6B.2	KY487674	KY487216	KY615378	KY487219	KY487567	KY615459	KY487075	KY487579
A/Keelung/0068/2016	6B.1	KY487114	KY487742	KY487180	KY615443	KY487374	KY487728	KY487105	KY487228

Supplemental Table 2. Amino acid differences (H1 numbering) of HA gene in 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16 season

Vaccine/ Surveillance strains*	HA1-84	HA1-152	HA1-162	HA1-173	HA1-209	HA1-216	HA1-223	HA1-261	HA2-164	HA2-174
A/California/7/2009	S	V	S	V	T[&]	I	R[&]	A	E	D
A/Michigan/45/2015	N	V	N [Gly +]	V	M[&]	T	R[&]	A	E	D
6B.1 parental (n=26)	N	V	N [Gly +]	V	K	T	Q	A	E	D
6B.1 variant (n=23)	N	V	N [Gly +]	V	K	T	Q	A	E	D
6B.2 variant-1 (n=6)	S	T	S	I	K	I	Q	S	G	E
6B.2 variant-2 (n=13)	S	T	S	I	K	I	Q	A	G	E
6B.1 reassortant (n=1)	N	V	N [Gly +]	V	K	T	Q	A	E	D

[*] Twenty-six strains of 6B.1 parental group include 22 Baltimore and 4 northern Taiwan strains. 6B.1 variant, 6B.2 variant-1 and -2, and 6B.1 reassortant strains are all isolated from northern Taiwan.

[Gly+] Amino acid changes lead to gain of potential N-linked glycosylation site.

[&] Mutations of clade 1 A/California/7/2009 vaccine X-179A or clade 6B.1 A/Michigan/45/2015 vaccine X275. The HA1 sequences of clinical isolates of A/California/7/2009 and A/Michigan/45/2015 were 209K and 223Q.

Supplemental Table 3. Amino acid differences of NA gene in 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16 season

Vaccine/ Surveillance strains*	NA-13	NA-34	NA-52	NA-67	NA-74	NA-96	NA-314	NA-381
A/California/7/2009	V	I	S	V	F	G	I	T
A/Michigan/45/2015	I	V	S	V	F	G	M	T
6B.1 parental (n=26)	I	V	S	V	F	G	M	T
6B.1 variant (n=23)	I	V	S	V	L	G	M	T
6B.2 variant-1 (n=6)	V	I	S	I	F	G	I	I
6B.2 variant-2 (n=13)	V	I	R[Gly-]	I	F	S	I	I
6B.1 reassortant (n=1)	V	I	R[Gly-]	I	F	S	I	I

* Twenty-six strains of 6B.1 parental group include 22 Baltimore and 4 northern Taiwan strains. 6B.1 variant, 6B.2 variant-1 and -2, and 6B.1 reassortant strains are all isolated from northern Taiwan.

[Gly-] Amino acid changes lead to loss of potential N-linked glycosylation site.

Supplemental Table 4. Amino acid differences of internal genes in 6B.1 and 6B.2 viruses of Baltimore and northern Taiwan in the 2015-16 season

Vaccine/ Surveillance strains*	PB2-299	PB2-453	PA-361	PA-X-212	NP-6&	M1-30	M1-208	M2-52	NS1-2	NS1-125	NS2-2	NS2-22	NS2-83
A/California/7/2009	R	S	K	A	A	S	Q	Y	D	E	D	G	M
A/Michigan/45/2015	R	P	K	A	A	S	K	Y	E	D	E	G	I
6B.1 parental (n=26)	K	T	K	A	A	S	K	Y	E	D	E	G	I
6B.1 variant (n=23)	K	T	K	A	A	S	K	Y	E	D	E	V	I
6B.2 variant-1 (n=6)	R	P	R	V	V	S	Q	Y	D	E	D	G	M
6B.2 variant-2 (n=13)	R	P	R	V	V	N	Q	C	D	E	D	G	M
6B.1 reassortant (n=1)	K	T	K	A	V	N	Q	C	D	E	D	G	M

* Twenty-six strains of 6B.1 parental group include 22 Baltimore and 4 northern Taiwan strains. 6B.1 variant, 6B.2 variant-1 and -2, and 6B.1 reassortant strains are all isolated from northern Taiwan.

& NP coding sequence numbering begins from upstream start codon.