Human Influenza A Virus H1N1pdm Full-length Genome Sequencing in USA and Taiwan in 2015-16 Flu Season and Disease Correlation

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[1,9,†]

[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

33

34

Abstract:

During the 2015-16 winter, the US experienced a relatively mild influenza season compared to Taiwan which had a large number of total and severe cases. While H1N1pdm viruses dominated global surveillance efforts in this season, the global distribution of genetic variants and the contribution of genetic to disease severity has not been investigated. Nasal 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 of major hospitals in Baltimore, Maryland, USA (Baltimore sites) and Taipei, Taiwan (Taipei sites) between November 2015 and April 2016, were processed for multi-segment PCR amplification and deep sequencing of the of influenza A virus gene segments. In the Baltimore sites, the majority of the viruses were the H1N1pdm clade 6B.1 and no H1N1pdm clade 6B.2 viruses were detected. In the Taipei sites, more than half of the H1N1pdm viruses were clade 6B.1 and 38% were clade 6B.2. These were consistent with the global observations that most 6B.2 viruses circulated in Asia and not North America. Of the 103 H1N1pdm nasal samples, 73 yielded sufficient coverage to enable the assembly of sequences for all 8 influenza A virus genome segments. Whole genome analysis identified two genetic subgroups present in each of the 6B.1 and 6B.2 clades. One 6B.1 intraclade reassortant virus containing five 6B.2 segments from the Taipei sites was observed. Clinical data showed 6B.2 patients had higher odds of pneumonia than 6B.1 patients. These data on H1N1pdm viral genetic variants which are linked to patient clinical and demographic data, improve our understanding of the impact of viral mutations to influenza disease severity and reaffirm the importance of active influenza surveillance across different geographic areas in the same influenza season.

Key words: influenza, H1N1pdm, whole genome sequencing, disease, clinical outcome

Introduction:

The 2009 pandemic H1N1 Influenza A virus (H1N1pdm) has supplanted the previous human H1N1 viruses to become the seasonal human H1N1 virus. In the 2015-16 influenza season, the northern hemisphere experienced its second global H1N1pdm-dominant year since 2009. The 2015-16 Influenza activity in the US was considered moderate, with lower numbers of total cases and a later peak epidemic when compared with the previous three influenza seasons [1]. In Taiwan, the 2015-16 season was much more serious [2], with an earlier start and higher numbers of severe, complicated influenza compared to the prior influenza season 2014-15 [3].

Influenza A viruses (IAV) are subdivided into serotypes based on the antibody response to the viral glycoproteins, hemagglutinin (HA) and neuraminidase (NA). Viruses are further divided into genetic clades based on HA sequences in order to monitor for mutations that might lead to antigenic drift. The vaccine strain of H1N1pdm, A/California/7/2009, is defined as clade 1. The global rise of H1 clade 6B.1 carrying amino acid changes S84N, S162N and I216T was detected in human surveillance efforts starting in June 2015. All regions of the world are now dominated by 6B.1. The H1 clade 6B.2 viruses (V152T, V173I of HA1 and E164G and D174E of HA2) emerged in July 2015 and have been found primarily in Asia. 6B.2 viruses identification peaked in January 2016 and were isolated less frequently after that, with a complete disappearance from surveillance efforts at the end of 2016 (nextflu.org) [4]. The majority of 6B.1 and 6B.2 viruses were antigenically similar to the reference viruses representing the recommended components of the 2015-16 Northern Hemisphere H1N1pdm influenza vaccine A/California/7/2009 [1, 3, 5, 6], though antigenic difference between egg-adapted and circulating virus strains may have contributed to increased numbers of cases [7].

The genome of influenza A viruses consist of 8 segments of negative-sense RNA. In addition to HA mutations, mutations in the other 7 viral gene segments also result in genetic variants which can alter virus fitness, subvert pre-existing immunity and/or reduce vaccine effectiveness. Any of these changes could potentially increase disease severity or total number of cases, thus leading to more severe epidemics [8]. Next-generation sequencing provides a sensitive and rapid method for sequencing of viruses directly from human specimens without any need for virus isolation or culture – both of which are known to select for viruses bearing additional mutations [9].

The Global Influenza Surveillance and Response System (GISRS) is a systematic worldwide influenza surveillance coordinated by World Health Organization (WHO). Circulating virus strains are selected for serology tests to determine whether a vaccine strain change is needed for the seasonal influenza vaccines. However, the current global influenza surveillance system lacks connections between viral sequences and detailed patient demographic and clinical data. The Johns Hopkins Center of Excellence for Influenza Research and Surveillance (JH-CEIRS) has been actively performing human influenza surveillance since 2014. Surveillance efforts involve the connection of viral genome data with patient demographic and clinical data to be able to gain more insights into how virus genetic variation can affect clinical disease. We hypothesized that differential genomes of influenza viruses could be associated with distinct clinical outcomes. This study involved concurrent, active influenza surveillance during the 2015-16 season in Baltimore, USA and Taipei, Taiwan to compare and contrast differential epidemics in the same year between the two geographic regions and to analyze possible disease severity with influenza viral genotypes.

Materials and Methods:

Active surveillance enrollment and sample collection

Institutional Review Boards at the Johns Hopkins University School of Medicine and Chang Gung Memorial Hospital provided ethical approval of the study (IRB00052743). From November 2015 to April 2016, active surveillance was performed in the adult emergency departments of the Johns Hopkins hospitals (JHH) at the East Baltimore and Bayview campuses in Baltimore, Maryland, USA and Chang Gung Memorial Hospitals (CGMH) in the Taipei, Taiwan metropolitan area, including hospitals in the Taipei, Linkou, and Keelung Branches. Influenza-like illness was defined as documented or reported ≥38◦C fever and any of one of three respiratory symptoms including cough, headache and/or sore throat within the past 7 days. Exclusion criteria included subjects who are unable to speak and understand English (in US) or Mandarin (in Taiwan), unable to provide written informed consent, currently incarcerated, or previously enrolled in the study during the same influenza season. Patients were approached by trained clinical coordinators who obtained written, informed consent before collected specimens and demographic and clinical data using a standard questionnaire. Data was confirmed by examination of the patient's electronic health record. All data was de-identified and stored in a secure REDCap database (https://projectredcap.org/resources/citations/).

Diagnosis and subtyping

Nasopharyngeal swabs or nasal washes of patients collected at the time of enrollment were tested for influenza A virus infection using the Cepheid Xpert Flu/RSV test (Cepheid, Sunnyvale, CA). Samples that were influenza A virus positive were further subtyped by qRT-PCR with specific H1 or H3 primers and probes according to USA Centers for Disease Control and Prevention (CDC) protocols. Only H1N1 positive samples were moved on to whole genome sequencing analysis.

Whole genome sequencing (WGS)

Viral RNA was isolated from the influenza A positive clinical samples using the MagMax Viral RNA isolation reagent (Thermo Fisher) on an Eppendorf epMotion 5073 liquid handling workstation. Samples were processed for whole genome sequencing using multi-segment PCR [9] and then prepared for deep sequencing using the Nextera XT library preparation reagent (Illumina). Samples were sequenced on the Illumina NextSeq platform. Raw paired-end data were first processed through Trim galore! with a quality score of 30 and an adapter sequence of CTGTCTCTTATACACATCT, only retaining pairs that both passed through this quality control step. The quality-trimmed reads were then aligned to the influenza vaccine strain reference sequence using bowtie2 (version 2.1.0) with the '--very-sensitive-local' option and converted to sorted BAM files using samtools. Only sequences of all 8 gene segments passed quality controls were included in the study. The sequences were submitted to NCBI GenBank database. The accession numbers are listed in the Table S1.

Phylogenetics and sequence analysis

73 H1N1pdm clinical samples which yielded sufficient WGS coverage of the IAV genome were included in sequence analysis. In addition to these genomes, we retrieved two reference genomes of vaccine strains from the Global Initiative on Sharing All Influenza Data (GISAID, https://www.gisaid.org/) for further analyses. Nucleotide sequences of the longest open reading frames (ORFs) in each gene segment were used to build phylogenetic trees, including PB2, PB1, PA, HA, NP, NA, M1, and NS1, as well as, their alternative splicing (M2 and NS2) and ribosomal frameshift (PA-X) produces. Nucleotide sequences of these ORFs (except for PA-X) in each strain were further concatenated to generate one single sequence for constructing a whole-genome phylogenetic tree. Phylogenetic trees for investigating their genetic or genomic relationships were based on maximum likelihood (ML) analysis, generated by using RAxML (version 8.2.12) [10] under GTRGAMMA model with a bootstrap value of 1000. Their time scales based on sample collection dates were calibrated using TreeTime [11] after building ML trees. 6B clade, and

6B.1, and 6B.2 subclades were defined by the rules of HA amino acid substitutions of global surveillance groups from nextflu (https://github.com/blab/nextflu). Phylogenetic groups in other gene segments were based on their tree branches to separate annotated clades (e.g., 6B.1 and 6B.2). Inconsistent positioning of strains, or clade groups on individual trees was used to identify influenza reassortment. Eventually, individual gene segment clades were annotated and visualized using the ggtree R package [12].

To investigate the genetic groups in public database, all H1N1pdm genomes (n=2873) isolated from September, 2015 to May, 2016 were downloaded from GISAID as of January 2020, and were further aligned using MAFFT tool [13] under the default conditions. Moreover, we retrieved all NP sequences (n=20,477) of H1N1pdm isolated after the year of 2009 (inclusive) from GISAID as of January 2020 to investigate the population of early start codon of NP gene. 20,470 NP sequences were used for further analyses, after aligning and removing 77 sequences with unexpected insertions.

Glycosylation prediction

Potential N-linked glycosylation sites of the 6B.1 and 6B.2 HA and NA proteins were predicted by the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc) to show any changes compared to the vaccine A/California/7/2009.

Statistical analysis of demographic and clinical data

To identify a contribution of virus clades and genetic subgroups to clinical symptoms, we applied Chi-square or Fisher's exact tests for categorical variables and the rank-sum test for continuous variables when appropriate. The two-tailed statistical significance was set at p-value <0.05. When clinical or demographic category reached statistical significance in a univariate analysis, confounding effects were then analyzed. In addition to adjusting for possible confounding effects between demographics, disease severity and clinical outcomes, a propensity score-based weighting was applied. Briefly, we applied those potential confounders to generate the propensity (or probability) of being infected with different subgroups of H1N1pdm viruses by using machine learning based modeling methods [14]. By using the inverse probability of treatment weighting (IPTW) method, patients with higher propensity will have lower weights in the final logistic regression model to adjust for confounding factors. The weights create a pseudo-population where the weighted virus subgroups have similar distribution across these confounders. In one simulation model, the propensity score-based weighting methods could adjust for

most of the confounding effects [15]. Odds ratios of unadjusted and IPTW were calculated using logistic regression analysis.

Results:

Sample collection

From November 2015 to April 2016, 261 and 256 symptomatic patients were enrolled in the study from emergency departments in JHH (Baltimore, US) and CGMH (Taipei, Taiwan). Demographic and clinical data were recorded at the time of enrollment and throughout the patient's hospital stay. Nasopharyngeal swabs or nasal washes were collected at the time of presentation and enrollment. 32.6% (85/261) of symptomatic patients from Baltimore and 57.8% (148/256) from Taipei were IAV-positive. About half of IAV-positive samples were further subtyped. 97.5% (39/40) of samples from Baltimore and 90.1% (64/71) of samples from Taipei were H1N1pdm positive. The epidemic curves of the two surveillance sites were different, with influenza A virus activity starting in January and peaking in mid-March in Baltimore while IAV activity in Taipei started in November and peaked in late January/early February, with peak activity higher than in Baltimore (manuscript in preparation).

HA clades

To understand the differential epidemics of Baltimore and Taipei, H1N1pdm genomes were sequenced directly from clinical samples. Seventy-three of the 103 H1N1pdm clinical samples yielded sufficient coverage of the IAV whole genome. Virus samples were initially analyzed based on HA segment sequences. All samples belonged to clade 6 subgroup B, as defined by amino acid substitutions K163Q (H1 numbering), A256T, K283E in HA1 and E172K in HA2 [16]. Clade 6B.1 (defined by the presence of HA1 mutations S84N, S162N, and I216T) and 6B.2 viruses (defined by the presence of HA1 mutations V152T, V173I and HA2 mutations E164G, D174E) were detected (Figure 1). In Baltimore, 91.7% (22/24) of viruses were clade 6B.1, two were clade 6B and no 6B.2 viruses were identified. The Taipei samples consisted of 57.1% (28/49) clade 6B.1, 38.8% (19/49) clade 6B.2 and two clade 6B viruses. The difference in clade circulation is consistent with global influenza surveillance that H1N1pdm 6B.1 and 6B.2 were dominant in the 2015-16 year and most 6B.2 viruses were found in Asia [17].

Full-length genome analysis and additional genetic differences between clades 6B.1 and 6B.2.

Phylogenetic analysis of 10 open reading frames (ORFs) from the other seven gene segments were also analyzed (Figure S1). Branch tips were colored by the HA clade of each sample and defining amino acid substitutions were noted. For the most part, the sequences of other ORFs clustered together consistent with their HA clade. Fifteen amino acid differences were identified on the other 8 viral ORFs between 6B.1 and 6B.2 viruses (Table 1). Two amino acids K299R (K in 6B.1 and R in 6B.2) and T453P on PB2 were different (Figure S1A). One amino acid K361R on PA was different (Figure S1C). One amino acid A212V on PA-X was different (Figure S1D). One amino acid A6V on NP was different (Figure S1F). Five amino acid I13V, V34I, V67I, M314I and T381I on NA were different (Figure S1G). One amino acid K208Q on M1 was different (Figure S1H). Two amino acids E2D and D125E on NS1 were different (Figure S1J). Two amino acids E2D and I83M on NS2 were different (Figure S1K).

As for the 6B viruses, all segments of two of the four 6B viruses, A/EastBaltimore/0103/2016 and A/Baltimore/0088/2016, were 6B lineage. One 6B virus A/Taipei/0045/2016 contained PB2, PA, NA and M segments from 6B.1 lineage and the rest segments were 6B but more similar to 6B.1 than 6B.2. The other 6B virus A/Linkou/0030/2015 contained PB2 and NS segments of 6B.2 lineage and the rest segments were 6B but more similar to 6B.2 than 6B.1 (Figure 2 and S1). These suggested that every segment of 6B virus which once evolved was either toward to 6B.1 lineage or 6B.2 lineage.

Early start codon of nucleoprotein (NP)

The NP gene segments of all samples have an early start codon 18 nt upstream of the start codon of human seasonal H1N1 and H3N2. This start codon were utilized and obtained polymerase activity [18, 19]. All the NP proteins could have an additional 6 amino acids at their amino terminus, with clade 6B.1 (MSDIEA) and 6B.2 viruses (MSDIEV) differing at amino acid six. The presence of an upstream NP start codon and the difference between 6B.1 and 6B.2 were also present in virus sequences available in GISAID. Approximately 70% of publish sequences (n=20,470) of H1N1pdm NP since 2009 have this early start codon and the amino acid six were mostly Alanine, and 24% (n=4924) without genetic data including A/California/7/2009 in this region may be due to manually trimmed for sequence submission. Among 1932 sequences with entire upstream 18 nucleotides in the year of 2009, 90% contained the early start codon. It suggested most H1N1pdm viruses have this early start codon from the outbreak year of 2009 to now.

Dynamic evolution and possible reassortant/co-infection

One Taipei 6B.1 sample A/Keelung/0046/2016 showed evidence of intraclade reassortment, as three individual gene segments of 6B.1 (segments HA, PB2 and PA) and five 6B.2 (segments PB1, NP, NA, M and NS) were detected in the clinical sample (Figure 2). While we cannot rule out co-infection with two H1N1 clade viruses, it seems more likely this individual was infected with an intraclade reassortant virus.

Glycosylation of HA and NA

Glycosylation on HA proteins has been proposed to prevent antibody recognition and viruses often accumulate secondary mutations at antigenic or receptor binding sites to adapt to the presence of new glycans [20]. Glycosylation on the NA stalk has been shown to be important for NA activity and virulence [21] and glycosylation near the NA enzyme sites (NA-245) alters NA enzymatic activity while also preventing the binding of antibodies (Powell et al., BioRxiv). There are 8 potential N-linked glycosylation sites on the HA of A/California/7/2009. 6B.1 HA gains an additional glycosylation site at 162-164 (N-Q-S) on the Sa antigenic site of head domain due to S to N mutation at HA-162 (Table 1) (manuscript in preparation). The NA from the A/California/7/2009 vaccine strain has 8 potential N-linked glycosylation sites. Both 6B.1 and 6B.2 viruses gain one additional site at 42-44 (N-Q-S) of the NA stalk and lose one at 386-388 (N-F-S) which lies on the lateral surface of NA head. 68.4% (13/19) of 6B.2 viruses with NA-S52R mutation lose one glycosylation site at 50-52 (N-Q-S) on the NA stalk (Table 1).

Antiviral resistance mutations

Mutations that are known to decrease antiviral susceptibility were also analyzed. The M2 S31N mutation conferring resistance to adamantane antivirals was detected in almost all current isolates of circulating human H3N2 [22]. H1N1pdm acquired the M segment from a Eurasian adamantane-resistant swine influenza virus which also has the M2 S31N mutation [23]. All of the 6B.1 and 6B.2 in our surveillance have M2 S31N indicating the viruses were no longer susceptible to adamantane. As for mutations associated with resistant to anti-NA drugs, one 6B.2 sample from Taipei, A/Linkou/0029/2015, contained the NA-H275Y mutation associated with oseltamivir resistance. Antiviral resistance to NA inhibitors was not tested in our study. In another study of 2015-16 H1N1pdm in Taiwan, one clinical isolate with NA-H275Y was resistant to oseltamivir [24]. The data suggest sporadic identification of oseltamivir resistant H1N1pdm viruses in 2015-16 in Taiwan.

Four distinct genetic subgroups

To better understand subgroup genetic diversity in the identified HA clades, we performed a phylogenetic analysis using concatenated gene segments (Figure 2). Each gene from the 6B.1 or 6B.2 lineages, which may or may not be the same as its HA clade, was determined by individual ORF tree (Figure S1) and listed on the right of the concatenated gene segment phylogenetic tree (Figure 2). Four distinct virus genetic subgroups were present in this dataset (Figure 2 and Table 1). 6B.1 viruses were separated into two groups; one defined by two substitutions (NA-F74L and NS2-G22V) which we refer to as 6B.1 variant virus compared to viruses without these substitutions, which we refer to as 6B.1 parental virus. All Baltimore 6B.1 viruses were 6B.1 parental. In Taipei, only 4 were 6B.1 parental, with the remaining clade 6B.1 viruses (23/27) being 6B.1 variant.

The 6B.2 viruses were also separated into two genetic subgroups with 6 viruses (identified with the HA-A261S mutation) named 6B.2 variant-1 and 13 viruses (defined by four substitutions: NA-S52R and -G96S, M1-S30N and M2-Y52C) named 6B.2 variant-2 virus.

Demographic and clinical analysis based on virus genetics

The demographic and clinical data from patients infected with clade 6B.1 or 6B.2 viruses were first compared by univariate analyses (Table 2). There were no 6B.2 viruses at the Baltimore sites. Patients infected with 6B.2 were more likely to be older, female, Asian, and unvaccinated. These patients had a higher prevalence of runny nose and diarrhea and had increased symptoms of severe influenza, including increased likelihood of oxygen supplementation, longer in-hospital stay, and pneumonia diagnosis (by radiological findings) (p=0.05). Interestingly, patients infected with clade 6B.1 viruses had a lower number of co-morbidities associated with severe influenza, suggesting a healthier population was being infected with those viruses.

We further utilized a logistic regression analysis to estimate the propensity of being infected by 6B.2, using potential confounders such as age, gender, comorbidities, obesity, vaccine status, human exposure and travel history (area under the ROC curve: 0.798). After inverse probability of treatment weighting (IPTW) to adjust for the possible confounding effects, infection with clade 6B.2 viruses was found to be significantly associated with pneumonia (Odds ratio: 3.261, 95% CI: 1.696-4.826, p=0.008) (Table 3).

315 The distribution of cases across the four H1N1 genetic subgroups was then determined (Tables 4 and 5).
316 When patients infected with the 6B.1 parental or variant viruses were compared (Table 4), more
317 patients infected with 6B.1 parental were from Baltimore and all the 6B.1 variant-infected patients were
318 from Taipei. Patients infected with the 6B.1 parental viruses reported having increased wheezing, sinus
319 pain, and nausea/vomiting. There were no differences between the two 6B.1 subgroups with respect to
320 severe symptoms. When patients infected with the 6B.2 variant-1 and variant-2 viruses were compared
321 there were no obvious difference in symptoms (Table 5). About 30% of patients in 6B.2 genetic
322 subgroups had pneumonia. No patients infected with 6B.2 variant-1 viruses required oxygen-
323 supplement or admission while 23% (3/13) of 6B.2 variant-2 patients required oxygen supplementation
324 and 15% (2/15) were hospitalized but these differences did not reach statistical significance. However,
325 the small sample sizes in the 6B.2 subgroups precluded a definitive characterization of demographic and
326 clinical data.
327
328
329 **Discussion:**
330 Four distinct genetic subgroups of H1N1pdm viruses were circulating in Baltimore and Taipei during the
331 2015-16 flu season. While the 6B.1 clade was circulating at both sites, some 6B.1 viruses in Taipei had
332 additional mutations which defined a genetic subgroup we called 6B.1 variant. The two 6B.1 groups of
333 viruses did not differ drastically with respect to disease severity. The 6B.2 clade viruses were only
334 detected in Taipei but two distinct genetic subgroups were detected. When compared to 6B.1 viruses,
335 6B.2 virus-infected patients showed more and stronger disease symptoms by both uni- and multivariate
336 analysis. This difference in H1N1pdm clade circulation may have contributed to the more severe
337 influenza season experienced in Taipei compared to Baltimore in 2015-16.
338
339 **The global circulation of 2015-16 H1N1pdm clades**
340 The H1N1pdm 6B.1 and 6B.2 viruses dominated the 2015-16 season in the Northern Hemisphere.
341 However, numbers of cases and disease severity differed between geographic regions. The 2015-16
342 H1N1pdm season in the US and Canada were moderate compared to the prior season [1, 5]. On the
343 other hand, surveillance in Russia and some eastern European countries had increasing numbers of
344 severe cases with evidence of 6B.1 circulation in the 2015-16 season [6, 25, 26]. National UK viral
345 surveillance showed excess all cause mortality in adults and higher influenza admission rate in the 2015-
346 16 season with predominant H1N1pdm 6B.1 and small numbers of 6B.2 [27, 28]. High numbers of

H1N1pdm cases with 6B.1 and 6B.2 co-circulation occurred in the middle east counties, Iran and Israel, and in Asia, Taiwan and Japan [3, 29-31]. Reports also showed H1N1pdm impacted the southern hemisphere in the 2016 season. In Sao Paulo State and a village in Brazil, an earlier influenza season with confirmed H1N1pdm 6B.1 as the dominant strain were reported [34, 35]. Reunion Island suffered the highest number of ICU hospitalization since 2009 and most were infected by H1N1pdm during April to August 2016 [36]. Unlike elsewhere with dominant H1N1pdm, a late-peak influenza season with predominant H3N2 and Flu B and only a small number of H1N1pdm cases were in northern China [32, 33].

Whole genome sequencing and disease outcome comparison may explain differential epidemics

Most influenza surveillance focuses on HA segment sequencing in order to monitor for the emergence of variants that differ from the sequence of the vaccine strain. Serology can then be utilized to identify whether the mutations have led to antigenic drift of the H1 HA protein. Serological surveillance reported that the circulating H1N1pdm strain in 2015-16 were antigenically similar to the vaccine A/California/7/2009 in the US, UK and many other countries [1, 5, 6, 27, 37, 38] (manuscript in preparation). This suggests that the HA amino acid differences between the circulating 6B.1 or 6B.2 clade viruses and the vaccine strain were likely not able to explain differential epidemics in different regions of the year even though vaccination coverage might differ in different regions. Whether mutations in the other 7 viral genes contributed to the variable disease severity reported in 2015-16 has not been studied carefully.

Only a few studies reported full genome sequences of the 2015-16 circulating H1N1pdm strains. Fifty-five isolates and 22 clinical samples in Russia, 5 virus isolates in eastern India, and 2 viruses in Bulgaria were whole-genome sequenced and reported [6, 38, 39], all of which were clade 6B.1 viruses. In the GISAID database, 6B.2 viruses circulated mainly in Asia, primarily in Japan, Singapore, Indonesia, Taiwan, and Shanghai (China). However, clade 6B.2 viruses were also identified in the USA, England, Scotland, France, Denmark, Norway, Ukraine, Australia and Brazil, albeit in very low numbers. In the 2015-16 season, only 126 whole genome sequences of clade 6B.2 viruses were present according to 2873 H1N1pdm whole genomes retrieved from GISAID as of January, 2020. Despite the genome sequence differences between the circulating H1N1pdm 6B.1 and 6B.2 strains and the vaccine strain, the relationship of the viral whole genome sequences with disease outcome has not yet been investigated.

Our surveillance in Baltimore, USA and Taipei, Taiwan led to the identification and sequencing of both 6B.1 and 6B.2 viruses in the 2015-16 year. 6B.1 and 6B.2 have 7 amino acid differences on HA and 15 amino acid differences in the other 8 viral proteins, along with a larger number of synonymous mutations. The data revealed that not only did the HA gene evolve significantly during the 2015-16 year but the other 7 genes also accumulated mutations to form distinct genetic subgroups (Figure 2 and S1). We hypothesized the two viral strains may relate to differential disease outcome and infection with 6B.2 was associated with increased disease symptoms. The only report for symptoms of 6B.1 and 6B.2 infection was seen in Israel, where infection with 6B.2 virus was associated with more vomiting and nausea compared to 6B.1-infected patients [30]. In our surveillance, we did see higher percentage of 6B.2 than 6B.1 patients with vomiting and nausea, but this did not reach statistical significance (Table 2). There was increased reports of runny nose and diarrhea in patients infected with 6B.2 compared to 6B.1 in our study. Reporting symptoms may be subjective. Hence, we considered disease severity using more objective criteria, such as oxygen supplementation, intubation, pneumonia, inpatient admission or death. A higher percentage of 6B.2-infected patients required oxygen, had a longer hospital stay and had pneumonia. After adjusting and weighting for age, gender, comorbidities, obesity, vaccine status, human exposure and travel history, people infected with 6B.2 still had a higher odds ratio of pneumonia (Table 3). The data indicate that infection with clade 6B.2 viruses is associated with a higher disease severity at our surveillance sites independent of patient demographics and comorbidities. The lack of clade 6B.2 infections in Baltimore coupled with the large number present in Taipei is likely one of the reasons that the influenza season is moderate in the US but severe in Taiwan.

Viral gene mutations known to be associated with transmission

Fifteen amino acid differences in 8 other virus open reading frames were identified between 6B.1 and 6B.2 viruses but none of these has been associated with increased virus pathogenicity. However, some within previously known mutations affecting transmission. There are five amino acid differences on the NA protein with 314 and 381 located on the enzymatic head domain. These mutations are located close to two of four NA mutations identified in avian H9N2 influenza viruses which were shown to affect air-born transmission in chicken [40]. The M1-208 amino acid located between amino acids 207 and 209, which affect H1N1pdm virus morphology and the efficiency of virus spread in vitro [41]. Effects of the amino acid differences between 6B.1 and 6B.2 on pathogenicity will require to further experimental confirmation but our initial experiments on human nasal epithelial cell cultures did not detect any replication differences between clade 6B.1 and 6B.2 viruses (manuscript in preparation).

411

412 The 6B.1 viruses were separated into two genotypes, 6B.1 variant with 2 amino acid substitutions, NA-
413 F74L and NS2-G22V, defining the subgroup. Patients infected with 6B.1 parental or variant had no
414 obvious different in disease outcome (Table 4). However, it cannot rule out that 6B.1 in other regions,
415 for example in Russia, might have substitutions in the genome which could lead to more severe
416 outcome than others. It requires further comparison of viral whole genome from the other regions. The
417 clade 6B.2 viruses also were separated into two genetic subgroups but no differences in patient
418 demographics or symptoms were identified between these subgroups.

419

420 **H1N1pdm 6B.1 and 6B.2 virus evolution**

421 Severe and fatal outbreaks occurred in early 2015 in Nepal and India [42]. A study in Taiwan showed
422 patients with laboratory-confirmed H1N1pdm 6B, 6B.1 and 6B.2 in 2013-14 and 2015-16 had higher risk
423 for influenza-related complications compared to patients from seasons where 6B, 6B.1 and 6B.2 viruses
424 did not dominate [43]. Both suggested 6B viruses, ancestors of 6B.1 and 6B.2, may have caused severe
425 outcomes. The emergence and maintenance of clade 6B.1 and 6B.2 viruses in different parts of the
426 world may help explain the differential impact of H1N1pdm epidemics in the 2015-16 season.

427

428 Only 6B.1 is still circulating in human populations. It is not clear what drove the extinction of 6B.2 but it
429 is possible that clade 6B.1 viruses were better adapted to infect and spread in humans. It has been
430 suggested that glycosylation is important for influenza A virus adaptation in humans [20] and gaining a
431 potential glycosylation site on the HA head at residue 162 may have given clade 6B.1 viruses an
432 evolutionary advantage. Clade 6B.1 viruses also encode an NS1-E125D mutation which increases NS1
433 interactions with cellular cleavage and polyadenylation factor 30 (CPSF30) [44]. The NS1-E125D
434 mutation of 6B.1 viruses has been suggested to be an important marker for virus adaption to humans
435 [6].

436

437 **Limitation of the study**

438 Influenza vaccination, overall higher in the US than in Taiwan, could also affect baseline neutralization
439 antibody titers, thus leading to different magnitudes of epidemics (manuscript in preparation).
440 Furthermore, the different race distributions between our surveillance sites are difficult to adjust for in
441 this and any study conducted across different geographic regions. Finally, the identification of 4 genetic
442 subgroups reduced the power of our study to detect demographic and clinical differences in our

443 populations. Increasing surveillance and whole genome sequencing efforts could dramatically improve
444 the ability to identify novel virus variants that are causing altered disease in humans.
445
446
447 **Conflicts of interests**
448 The authors declare no conflict of interest.
449
450 **Author contributions**
451 H.L., R.R., K.F.C., and A.P. conceived of the experimental questions; Y.N.G. performed phylogenetic
452 analyses; K.S.S., and K.F.C. ran demographic and clinical analyses; T.M., J.E., and P.T. sequenced samples
453 and cleaned the raw data; Z.Y.L., M.L., and L.S. oversaw enrollment and collection of samples from
454 patients; H.L., Y.N.G., K.S.S., and K.F.C. analyzed data; H.L., and A.P. wrote the manuscript; all authors
455 edited and reviewed the manuscript prior to submission.
456
457 **Funding**
458 This work was supported by the NIH/NIAID Center of Excellence in Influenza Research and Surveillance
459 contract HHS N2772201400007C (R.R., K.F.C., and A.P.), the Richard Eliasberg Family Foundation and the
460 Research Center for Emerging Viral Infections from The Featured Areas Research Center Program within
461 the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan and
462 the Ministry of Science and Technology (MOST), Taiwan MOST 108-3017-F-182-001 (Y.N.G.).
463
464 **Acknowledgements**
465 The authors thank the patients who enrolled and participated in the JHCEIRS surveillance study. We are
466 grateful for the efforts of the clinical coordination team at JHMI who collected samples. We thank
467 members of the Pekosz lab for feedback on this work.
468
469
470 **References:**
471
472 1. Davlin, S.L., et al., *Influenza Activity - United States, 2015-16 Season and Composition of the*
473 *2016-17 Influenza Vaccine.* MMWR Morb Mortal Wkly Rep, 2016. **65**(22): p. 567-75.
474 2. Gong, Y.N., et al., *Centennial review of influenza in Taiwan.* Biomed J, 2018. **41**(4): p. 234-241.

3. TaiwanCDC, *Taiwan Influenza Express*. 2016, Taiwan CDC: Taipei, Taiwan.
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. 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.
6. Komissarov, A., et al., *Rapid spread of influenza A(H1N1)pdm09 viruses with a new set of specific mutations in the internal genes in the beginning of 2015/2016 epidemic season in Moscow and Saint Petersburg (Russian Federation).* Influenza Other Respir Viruses, 2016. **10**(4): p. 247-53.
7. Garretson, T.A., et al., *Identification of human vaccinees that possess antibodies targeting the egg-adapted hemagglutinin receptor binding site of an H1N1 influenza vaccine strain.* Vaccine, 2018. **36**(28): p. 4095-4101.
8. Bouvier, N.M. and P. Palese, *The biology of influenza viruses.* Vaccine, 2008. **26 Suppl 4**: p. D49-53.
9. Zhou, B., et al., *Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses.* J Virol, 2009. **83**(19): p. 10309-13.
10. Stamatakis, A., *RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.* Bioinformatics, 2014. **30**(9): p. 1312-3.
11. Sagulenko, P., V. Puller, and R.A. Neher, *TreeTime: Maximum-likelihood phylodynamic analysis.* Virus Evol, 2018. **4**(1): p. vex042.
12. Yu, G., et al., *ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data.* Methods Ecol Evol, 2017. **8**(1): p. 28-36.
13. Katoh, K. and D.M. Standley, *MAFFT multiple sequence alignment software version 7: improvements in performance and usability.* Mol Biol Evol, 2013. **30**(4): p. 772-80.
14. Lee, B.K., J. Lessler, and E.A. Stuart, *Improving propensity score weighting using machine learning.* Stat Med, 2010. **29**(3): p. 337-46.
15. Stuart, E.A., *Matching methods for causal inference: A review and a look forward.* Stat Sci, 2010. **25**(1): p. 1-21.
16. European_Centre_for_Disease_Prevention_and_Control, *Influenza virus characterisation summary Europe, May 2016*. 2016, Stockholm: ECDC.

17. World_Health_Organization, *Recommended composition of influenza virus vaccines for use in the 2016-2017 northern hemisphere influenza season.* Wkly Epidemiol Rec, 2016. **91**(10): p. 121-32.
18. Wanitchang, A., et al., *Atypical characteristics of nucleoprotein of pandemic influenza virus H1N1 and their roles in reassortment restriction.* Arch Virol, 2011. **156**(6): p. 1031-40.
19. Wise, H.M., et al., *An alternative AUG codon that produces an N-terminally extended form of the influenza A virus NP is a virulence factor for a swine-derived virus.* bioRxiv, 2019: p. 738427.
20. Tate, M.D., et al., *Playing hide and seek: how glycosylation of the influenza virus hemagglutinin can modulate the immune response to infection.* Viruses, 2014. **6**(3): p. 1294-316.
21. Wu, C.Y., et al., *Influenza A surface glycosylation and vaccine design.* Proc Natl Acad Sci U S A, 2017. **114**(2): p. 280-285.
22. Simonsen, L., et al., *The genesis and spread of reassortment human influenza A/H3N2 viruses conferring adamantane resistance.* Mol Biol Evol, 2007. **24**(8): p. 1811-20.
23. Garten, R.J., et al., *Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans.* Science, 2009. **325**(5937): p. 197-201.
24. Gong, Y.N., et al., *Population dynamics at neuraminidase position 151 of influenza A (H1N1)pdm09 virus in clinical specimens.* J Gen Virol, 2019. **100**(5): p. 752-759.
25. Tjon-Kon-Fat, R., et al., *The potential risks and impact of the start of the 2015-2016 influenza season in the WHO European Region: a rapid risk assessment.* Influenza Other Respir Viruses, 2016. **10**(4): p. 236-246.
26. Newitt, S., et al., *Rapid risk assessment during the early weeks of the 2015-2016 influenza season in Ukraine.* Influenza Other Respir Viruses, 2018. **12**(2): p. 241-249.
27. Pebody, R., et al., *Effectiveness of seasonal influenza vaccine in preventing laboratory-confirmed influenza in primary care in the United Kingdom: 2015/16 mid-season results.* Euro Surveill, 2016. **21**(13).
28. Public_Health_England_(PHE), *Summary of UK surveillance of influenza and other seasonal respiratory illnesses 26 May 2016 – Week 21 report (up to week 20 data).* in *PHE weekly national influenza report.* 2016.
29. Mohebbi, A., et al., *Molecular epidemiology of the hemagglutinin gene of prevalent influenza virus A/H1N1/pdm09 among patient in Iran.* Virus Res, 2019. **259**: p. 38-45.
30. Friedman, N., et al., *A(H1N1)pdm09 influenza infection: vaccine inefficiency.* Oncotarget, 2017. **8**(20): p. 32856-32863.

31. National_Institute_of_Infectious_Diseases_Japan, *Influenza 2015/16 season, Japan.* Infectious Agents Surveillance Report, 2016. **37**(11): p. 211-213.
32. World_Health_Organization, *Review of global influenza activity, October 2015-October 2016.* Wkly Epidemiol Rec, 2016. **91**(51-52): p. 604-22.
33. Zhang, L., et al., *The 2015-2016 influenza epidemic in Beijing, China: Unlike elsewhere, circulation of influenza A(H3N2) with moderate vaccine effectiveness.* Vaccine, 2018. **36**(33): p. 4993-5001.
34. Santos, K.C., et al., *Molecular epidemiology of influenza A(H1N1)PDM09 hemagglutinin gene circulating in Sao Paulo State , Brazil: 2016 anticipated influenza season.* Rev Inst Med Trop Sao Paulo, 2017. **59**: p. e9.
35. Cardoso, A.M., et al., *Investigation of an outbreak of acute respiratory disease in an indigenous village in Brazil: Contribution of Influenza A(H1N1)pdm09 and human respiratory syncytial viruses.* PLoS One, 2019. **14**(7): p. e0218925.
36. Filleul, L., et al., *A major impact of the influenza seasonal epidemic on intensive care units, Reunion, April to August 2016.* Euro Surveill, 2016. **21**(47).
37. Broberg, E., et al., *Predominance of influenza A(H1N1)pdm09 virus genetic subclade 6B.1 and influenza B/Victoria lineage viruses at the start of the 2015/16 influenza season in Europe.* Euro Surveill, 2016. **21**(13).
38. Korsun, N., et al., *Antigenic and genetic characterization of influenza viruses circulating in Bulgaria during the 2015/2016 season.* Infect Genet Evol, 2017. **49**: p. 241-250.
39. Mukherjee, A., et al., *Genetic Characterization of Circulating 2015 A(H1N1)pdm09 Influenza Viruses from Eastern India.* PLoS One, 2016. **11**(12): p. e0168464.
40. Lv, J., et al., *Amino acid substitutions in the neuraminidase protein of an H9N2 avian influenza virus affect its airborne transmission in chickens.* Vet Res, 2015. **46**: p. 44.
41. Bialas, K.M., E.A. Desmet, and T. Takimoto, *Specific residues in the 2009 H1N1 swine-origin influenza matrix protein influence virion morphology and efficiency of viral spread in vitro.* PLoS One, 2012. **7**(11): p. e50595.
42. Nakamura, K., et al., *Characterization of influenza A(H1N1)pdm09 viruses isolated from Nepalese and Indian outbreak patients in early 2015.* Influenza Other Respir Viruses, 2017. **11**(5): p. 399-403.
43. Hsieh, Y.C., et al., *Clinical characteristics of patients with laboratory-confirmed influenza A(H1N1)pdm09 during the 2013/2014 and 2015/2016 clade 6B/6B.1/6B.2-predominant outbreaks.* Sci Rep, 2018. **8**(1): p. 15636.

569 44. Hale, B.G., et al., *Inefficient control of host gene expression by the 2009 pandemic H1N1*
570 *influenza A virus NS1 protein.* J Virol, 2010. **84**(14): p. 6909-22.
571
572

Figure Legends.

Figure 1. Phylogenetic tree of HA gene sequences from influenza H1N1pdm viruses obtained in the 2015-16 season. 6B.1 and 6B.2 clades with specific substitutions defining major branches indicated. Tips are colored by location of sample collection.

Figure 2. Phylogenetic tree of concatenated H1N1pdm segments with a schematic representation of four genetic subgroups and the mutations that define each.

Table 1. Amino acid differences of H1N1 6B.1 and 6B.2 collected from Baltimore and Taipei during 2015-16 flu season

Table 2. Univariate analysis of patients infected with H1N1 6B.1 and 6B.2 viruses of Baltimore and Taipei in 2015-16 season

Table 3. Logistic regression analysis of pneumonia in patients infected with H1N1 6B.1 and 6B.2 viruses of Baltimore and Taipei in 2015-16 season

Table 4. Univariate analysis of patients infected with H1N1 6B.1 parental and 6B.1 variant viruses of Baltimore and Taipei in 2015-16 season

Table 5. Univariate analysis of patients infected with H1N1 6B.2 variant-1 and 6B.2 variant-2 viruses of Baltimore and Taipei in 2015-16 season

Supplemental Figure 1. Phylogenetic trees of all gene segments from influenza H1N1pdm viruses obtained in Baltimore and Taipei in the 2015-16 season. Phylogenetic trees of the PB2, PB1, PA, PA-X, HA, NP, NA, M1, M2, NS1 and NS2 coding sequences (A to K, respectively) with branch tips colored by HA clades. Amino acid mutations if any of each branch are shown. NP coding sequence numbering begins from upstream start codon.

Supplemental Table 1. Influenza A virus strains and associated GenBank accession numbers of each segment.

Figures

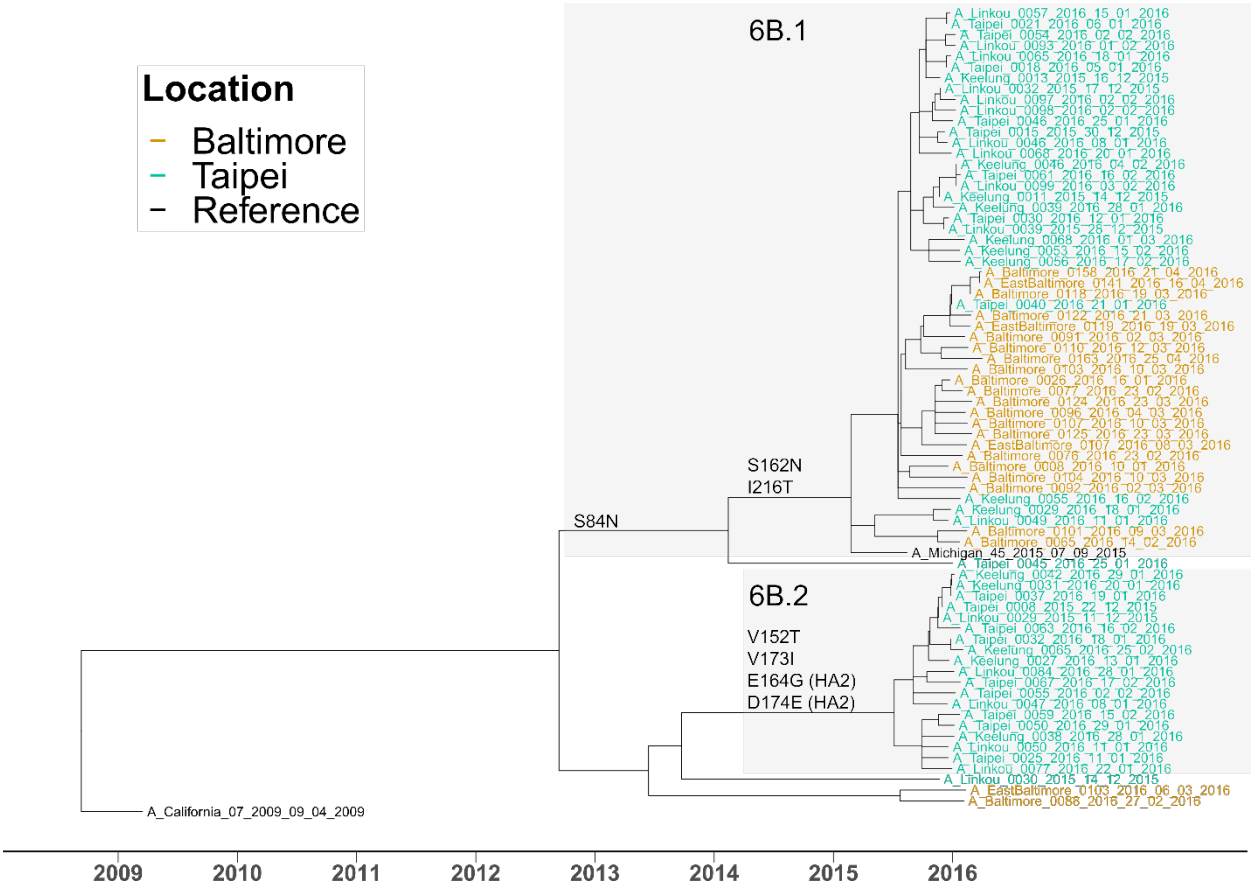


Figure 1. Phylogenetic tree of HA gene sequences from influenza H1N1pdm viruses obtained in the 2015-16 season. 6B.1 and 6B.2 clades with specific substitutions defining major branches indicated. Tips are colored by location of sample collection.

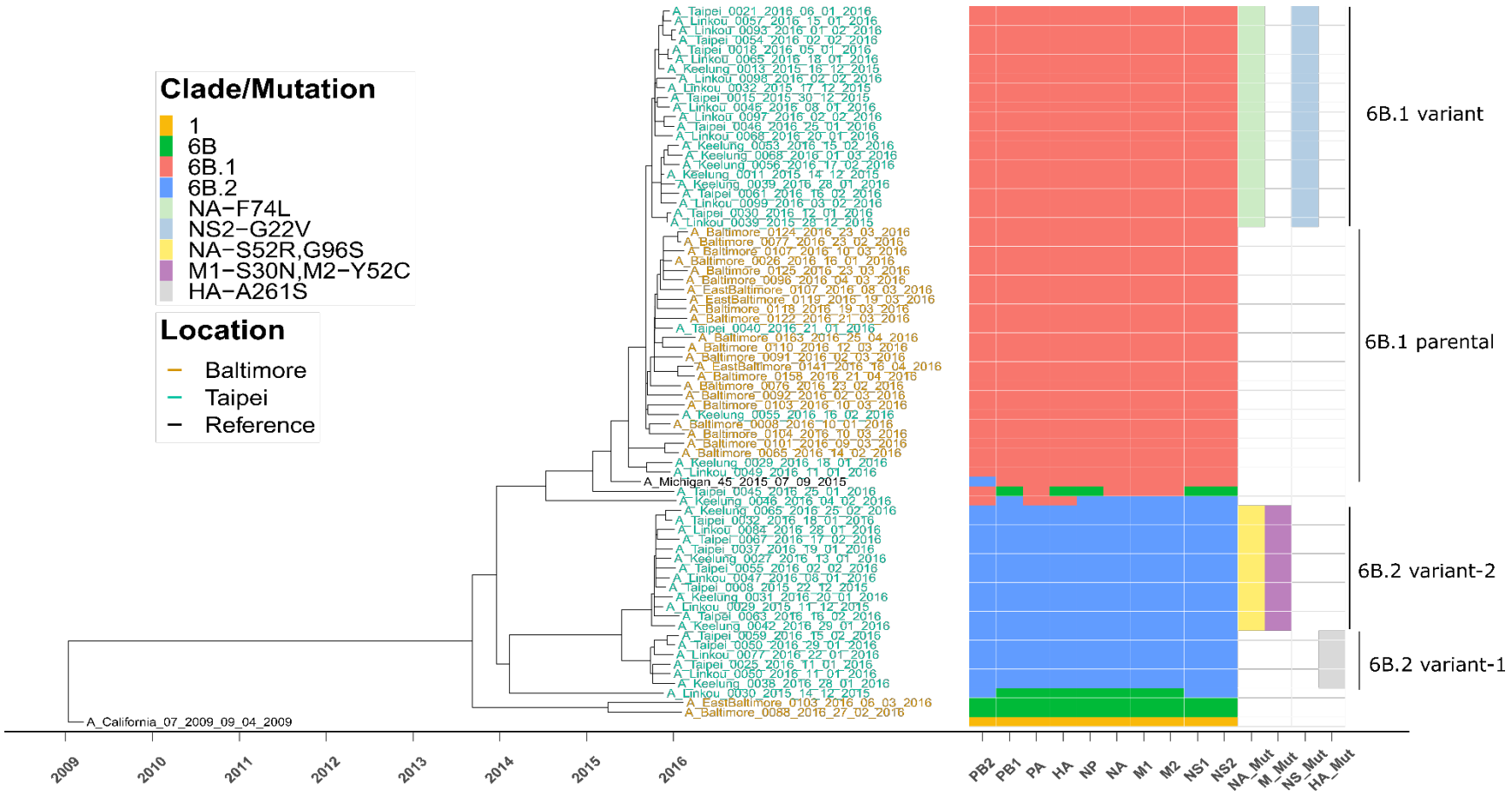
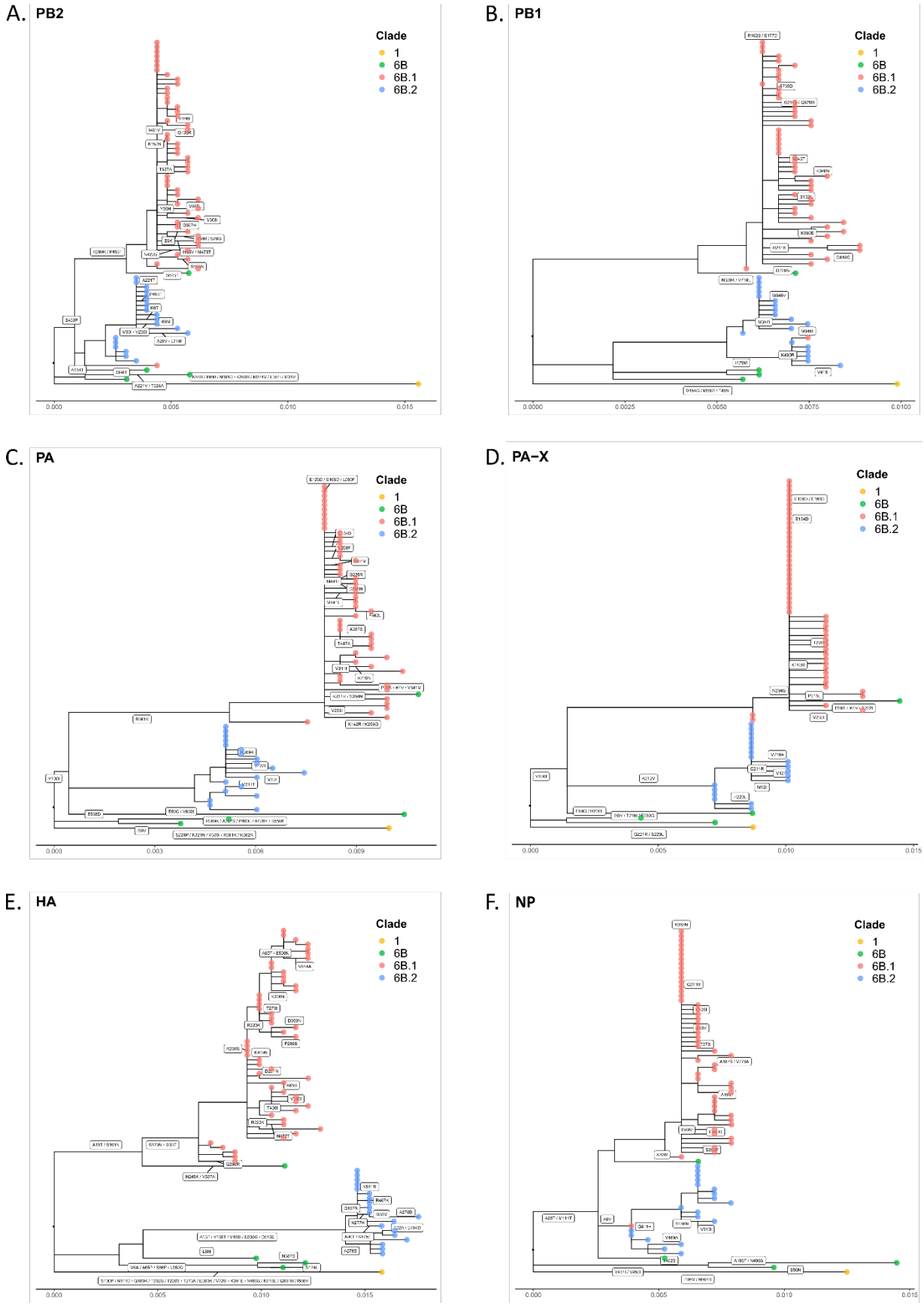
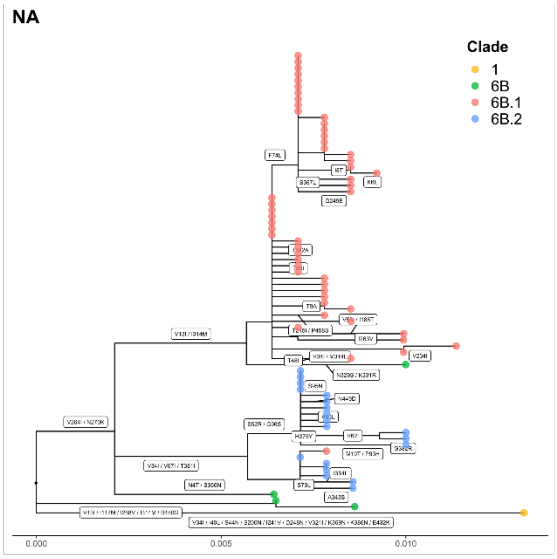


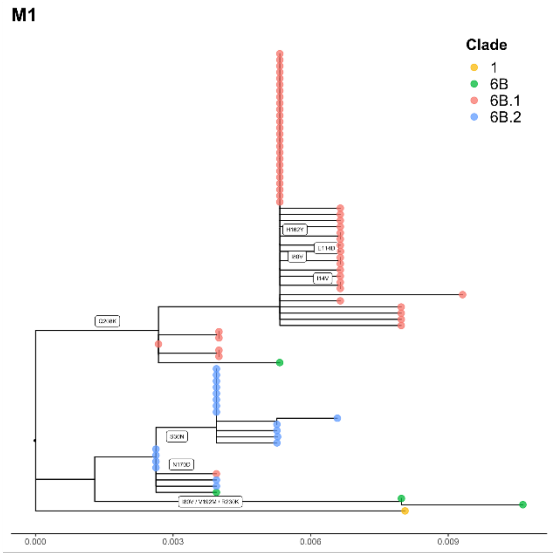
Figure 2. Phylogenetic tree of concatenated H1N1pdm segments with a schematic representation of four genetic subgroups and the mutations that define each.



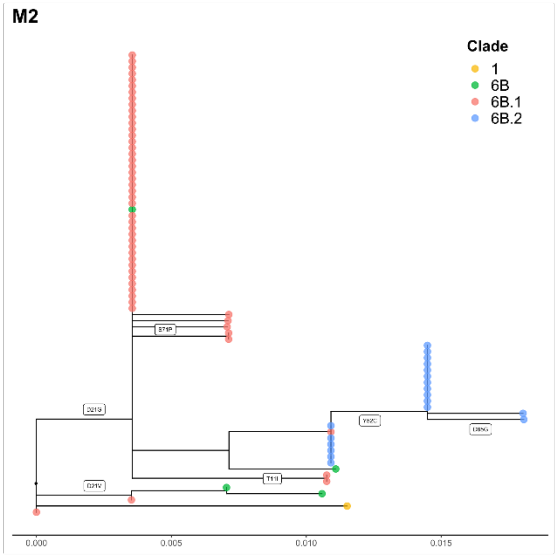
G. NA



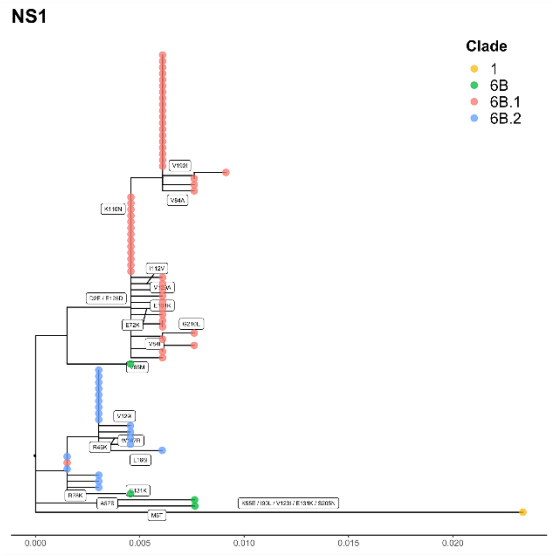
H. M1



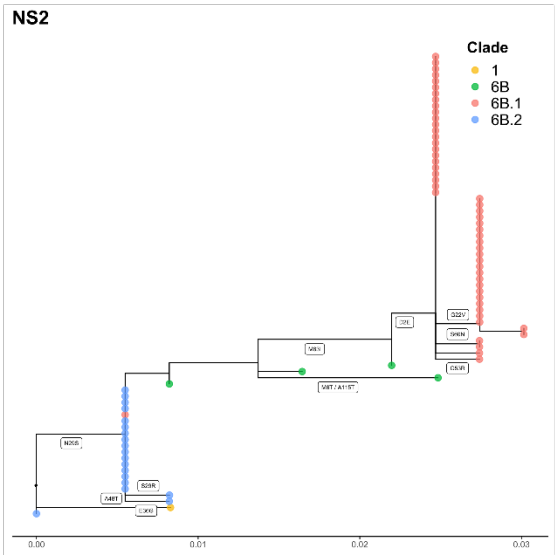
I. M2



J. NS1



K. NS2



Supplemental Figure 1. Phylogenetic trees of all gene segments from influenza H1N1pdm viruses obtained in Baltimore and Taipei in the 2015-16 season. Phylogenetic trees of the PB2, PB1, PA, PA-X, HA, NP, NA, M1, M2, NS1 and NS2 coding sequences (A to K, respectively) with branch tips colored by HA clades. Amino acid mutations if any of each branch are shown. NP coding sequence numbering begins from upstream start codon.

Table 1. Amino acid differences of H1N1 6B.1 and 6B.2 collected from Baltimore and Taipei during 2015-16 flu season

				HA1								HA2	
	Clade	H0 numbering		101	169	179	190	226	233	240	278	508	518
Vaccine		H1 numbering		84	152	162	173	209	216	223	261	164	174
A/California/7/2009	1			S	V	S	V	T[1]	I	R[1]	A	E	D
A/Michigan/45/2015	6B.1			N	V	N[+Gly]	V	M[1]	T	R[1]	A	E	D
Surveillance		Baltimore (n)	Taipei (n)										
6B.1 parental	6B.1	22	4	N	V	N[+Gly]	V	K	T	Q	A	E	D
6B.1 variant		0	23	N	V	N[+Gly]	V	K	T	Q	A	E	D
6B.2 variant-1	6B.2	0	6	S	T	S	I	K	I	Q	S	G	E
6B.2 variant-2		0	13	S	T	S	I	K	I	Q	A	G	E
Reassortment	6B.1	0	1	N	V	N[+Gly]	V	K	T	Q	A	E	D

	NA								NS1		NS2			M1		M2	PB2		PA	PA-X	NP[2]
	13	34	52	67	74	96	314	381	2	125	2	22	83	30	208	52	299	453	361	212	6
A/California/7/2009	V	I	S	V	F	G	I	T	D	E	D	G	M	S	Q	Y	R	S	K	A	A
A/Michigan/45/2015	I	V	S	V	F	G	M	T	E	D	E	G	I	S	K	Y	R	P	K	A	A
6B.1 parental	I	V	S	V	F	G	M	T	E	D	E	G	I	S	K	Y	K	T	K	A	A
6B.1 variant	I	V	S	V	L	G	M	T	E	D	E	V	I	S	K	Y	K	T	K	A	A
6B.2 variant-1	V	I	S	I	F	G	I	I	D	E	D	G	M	S	Q	Y	R	P	R	V	V
6B.2 variant-2	V	I	R[-Gly]	I	F	S	I	I	D	E	D	G	M	N	Q	C	R	P	R	V	V
Reassortment	V	I	R[-Gly]	I	F	S	I	I	D	E	D	G	M	N	Q	C	K	T	K	A	V

Clade B viruses were not included in the table. Differences between clade 6B.1 and 6B.2 identified in the study were highlighted in grey.

Amino acid changes lead to gain or loss of potential N-linked glycosylation site was labeled as +Gly or -Gly.

[1] Mutations of A/California/7/2009 vaccine X-179A or 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.

[2]NP coding sequence numbering begins from upstream start codon.

Table 2. Univariate analysis of patients infected with H1N1 6B.1 and 6B.2 viruses of Baltimore and Taipei in 2015-16 season

		6B.1	6B.2	p-value
		N= 49	N= 19	
Demographics		N (%) / median (IQR)	N (%) / median (IQR)	
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 (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 3. Logistic regression analysis of pneumonia in patients infected with H1N1 6B.1 and 6B.2 viruses of Baltimore and Taipei in 2015-16 season

		Unadjusted				IPTW**		
	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*

IPTW: inverse probability of treatment weighting

*p<0.05, logistic regression analysis.

** IPTW 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 4. Univariate analysis of patients infected with H1N1 6B.1 parental and 6B.1 variant viruses of Baltimore and Taipei in 2015-16 season

		6B.1 Parental	6B.1 Variant	p-value
		N= 26	N= 23	
Demographics		N (%) / median (IQR)	N (%) / median (IQR)	
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 5. Univariate analysis of patients infected with H1N1 6B.2 variant-1 and 6B.2 variant-2 viruses of Baltimore and Taipei in 2015-16 season

		6B.2 Variant-1	6B.2 Variant-2	p-value
		N= 6	N= 13	
Demographics		N (%) / median (IQR)	N (%) / median (IQR)	
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. Influenza A virus strains and associated GenBank accession numbers of each segment.

strain_name	HA clade	segment_1	segment_2	segment_3	segment_4	segment_5	segment_6	segment_7	segment_8
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