

# Passage Adaptation Correlates With the Reduced Efficacy of the Influenza Vaccine

Hui Chen,<sup>1,2</sup> Jacob Josiah Santiago Alvarez,<sup>1,3</sup> Sock Hoon Ng,<sup>4</sup> Rasmus Nielsen,<sup>5,6</sup> and Weiwei Zhai<sup>1,2,7,8,9</sup>

<sup>1</sup>Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing; <sup>2</sup>Human Genetics, Genome Institute of Singapore, A\*STAR; <sup>3</sup>Department of Biological Science, National University of Singapore, and <sup>4</sup>Defence Medical and Environmental Research Institute, DSO National Labs, Singapore; <sup>5</sup>Department of Integrative Biology, and

<sup>6</sup>Department of Statistics, University of California–Berkeley; <sup>7</sup>School of Biological Sciences, Nanyang Technological University, and <sup>8</sup>National Cancer Center, Singapore; and <sup>9</sup>Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming

**Background.** As a dominant seasonal influenza virus, H3N2 virus rapidly evolves in humans and is a constant threat to public health. Despite sustained research efforts, the efficacy of H3N2 vaccine has decreased rapidly. Even though antigenic drift and passage adaptation (substitutions accumulated during vaccine production in embryonated eggs) have been implicated in reduced vaccine efficacy (VE), their respective contributions to the phenomenon remain controversial.

**Methods.** We utilized mutational mapping, a powerful probabilistic method for studying sequence evolution, to analyze patterns of substitutions in different passage conditions for an unprecedented amount of H3N2 hemagglutinin sequences ( $n = 32\,278$ ).

**Results.** We found that passage adaptation in embryonated eggs is driven by repeated convergent evolution over 12 codons. Based on substitution patterns at these sites, we developed a metric, adaptive distance (AD), to quantify the strength of passage adaptation and subsequently identified a strong negative correlation between AD and VE.

**Conclusions.** The high correlation between AD and VE implies that passage adaptation in embryonated eggs may be a strong contributor to the recent reduction in H3N2 VE. We developed a computational package called MADE (Measuring Adaptive Distance and vaccine Efficacy based on allelic barcodes) to measure the strength of passage adaptation and predict the efficacy of a candidate vaccine strain. Our findings shed light on strategies for reducing Darwinian evolution within the passaging medium in order to potentially restore an effective vaccine program in the future.

**Keywords.** H3N2 influenza virus; vaccine efficacy; passage adaptation; mutational mapping.

Influenza viruses infect 5%–15% of the total population, resulting in an estimated half a million deaths annually [1, 2]. Vaccination against influenza viruses is the primary strategy for prevention and control of influenza outbreaks [3]. Despite decades of effort, vaccine efficacy (VE) against influenza viruses, particularly that of the H3N2 subtype, has rapidly dropped to 10% or lower, calling into question the effectiveness of the current vaccination program [4, 5]. Consequently, understanding and controlling the sharp decline in VE is a pressing public health challenge.

Mismatches between circulating and vaccine strains critically affect VE [6]. Two major sources of this mismatch are antigenic drift [7, 8] and passage adaptation [9]. Antigenic drift refers to the continued evolution of the influenza virus in host humans, and passage adaptation relates to substitutions accumulated while propagating viral strains in the culturing

medium during vaccine production. Even though the importance of antigenic drift and passage adaptation in VE has been proven in independent studies [9, 10], their relative contributions remain unknown, and conclusions drawn between individual studies are discordant [11–13]. The World Health Organization (WHO) enforces a consistent and rigorous vaccine program that matches the vaccine isolate to the circulating strains throughout the years. However, recent declines in VE may not be caused by antigenic drift but might instead be driven by passage adaptation during vaccine production.

Currently, embryonated eggs [14] or mammalian cell lines (eg, Madin Darby canine kidney cell line [MDCK] [15]) are mainly used for propagating influenza viruses. The receptors for viral entry into the host cell are sialic acids found on the surface of cells, which have 2 forms of sialic acid linkages to galactose. The alpha-2,3 glycosidic-bond-linked sialic acids (SA $\alpha$ -2,3Gal) are found mostly in the intestinal tracts of birds, while the alpha-2,6 glycosidic-bond-linked residues (SA $\alpha$ -2,6Gal) predominate in the upper respiratory tract of humans [16]. Growing human influenza viruses that preferentially bind to the SA $\alpha$ -2,6Gal receptor can therefore cause strong adaptive evolution during virus culture in embryonated eggs, which mainly contain the alternative SA $\alpha$ -2,3Gal form [17, 18]. One recent study revealed a steady increase in passage adaptation of the H3N2 virus in

Received 30 April 2018; editorial decision 1 December 2018; accepted 13 December 2018; published online December 18, 2018.

Correspondence: W. Zhai, 1 Beichen West Road, Institute of Zoology, Chinese Academy of Sciences, Chaoyang District, Beijing 100101, P.R. China (weiweizhai@ioz.ac.cn or zhaiww1@gis.a-star.edu.sg).

Clinical Infectious Diseases® 2019;69(7):1198–204

© The Author(s) 2018. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com. DOI: 10.1093/cid/ciy1065

embryonated eggs [18] that seemed to be correlated with the reduction in VE, further supporting the importance of passage adaptation in the recent reduction in VE [9, 19].

With rapid progress in viral sequencing, quantifying adaptive evolution in large datasets ( $>10\,000$  sequences) is becoming a challenging task for researchers due to the increased computational burden. To overcome this, a probabilistic approach known as “stochastic mutational mapping” [20] has emerged as a powerful and efficient way to sample possible mutational histories of observed data [20, 21]. Using theories derived from the continuous time Markov chain, mutational mapping first infers possible histories of sequence evolution for the observed data. Subsequently, many statistical tests can be constructed to test different hypotheses such as the existence of positive evolution. Simulation studies show that mutational mapping provides a very efficient and accurate method for sequence evolution [18, 20, 21]. Here, we used this approach to explore the correlation between passage adaptation and VE and to elucidate the genetic determinants of recent vaccine competence.

## METHODS

### Sequence Curation and Passage Histories

We analyzed 38 360 influenza H3N2 HA1 sequences from the Global Initiative on Sharing All Influenza Data (GISAID). Sequence data were summarized from the GISAID database with collection dates from 1 January 1968 to 8 July 2016 [22]. After multiple quality-control steps, 32 278 sequences were used for subsequent analysis (Supplementary Material, Supplementary Table 1, and Supplementary Figure 1 and 2). Passage information of these sequences was extracted from the Centers for Disease Control and Prevention/National Center for Biotechnology Information [17] and GISAID databases [22]. We categorized sequences according to their passage histories [18]. A total of 661 egg-passaged sequences and 4544 MDCK-passaged sequences were identified (Figure 1A). Other passage information included CELL (3286), SIAT (3712), Clinical (4046), Monkey (2019), Mix (4088), others (1086), and unknown (8836).

### Phylogenetic Reconstruction and Mutational Mapping

Sequence alignment for 32 278 sequences was performed using multiple sequence comparison by log-expectation [23]. After sequence alignment, the phylogenetic relationships among these sequences were reconstructed using the Randomized Axelerated Maximum Likelihood package under a general time reversible (GTR) model (Figure 1A) [24]. The probabilistic mutational mapping algorithm has been shown to be an efficient approach for sampling evolutionary histories according to their posterior probabilities [20]. Using the inferred maximum likelihood parameters and associated

phylogenetic tree, mutational mapping was carried out as in previous studies [18, 20, 21] (Supplementary Material). The mutational history inference was conducted at the base level and provided an efficient and accurate inference of the mutational events. After mutational mapping, summary statistics such as the number of substitutions that occurred along a given set of branches were extracted from the mutational mapping output.

### The Convergent and Enrichment Test

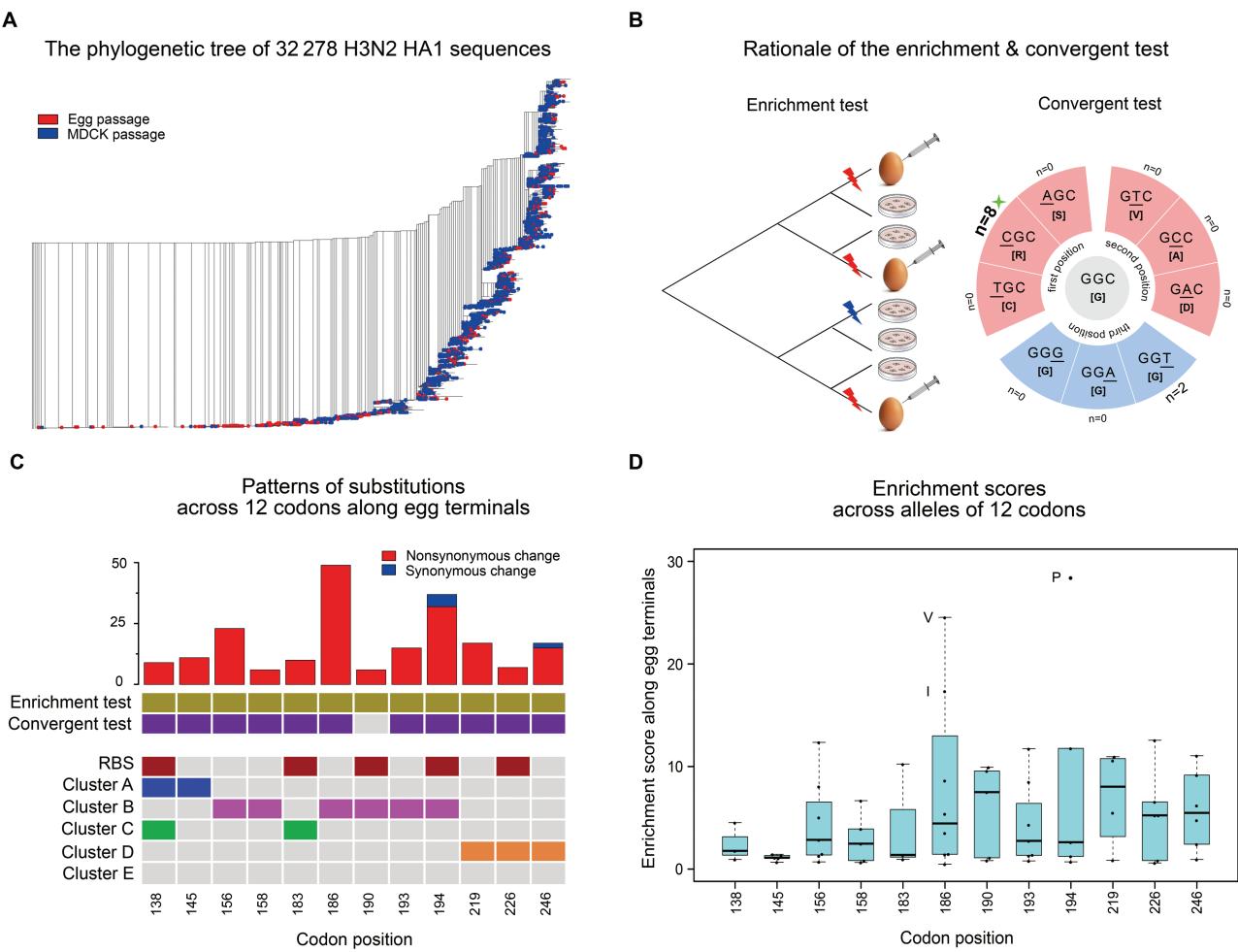
Changes at each of the 3 codon positions were conjugated together to examine the functional consequence of these substitutions. The convergent and enrichment tests survey 2 aspects of the adaptive evolution. The enrichment test examines substitutions in terminal branches and investigates whether non-synonymous substitutions at a given codon tend to happen specifically over terminal branches associated with certain passage histories (Figure 1B, Supplementary Material). The convergent test examines whether a given non-synonymous transition between 2 codon states occurred more frequently than by chance. Using synonymous substitution rates as the baseline, we determined whether a given non-synonymous substitution takes place more often than by chance (Figure 1B, Supplementary Material). After calculating the *P* value from both the enrichment and convergent tests, the Fisher method was used to combine these 2 *P* values, and multiple test correction was performed in R using the *p.adjust()* function. By taking the average *q*-value across 100 replicates from the mutational histories and using a false discovery rate cutoff of 0.01, we identified 12 codon positions that drive egg-passage adaptation.

### Enrichment Scores and Principal Component Analysis

In this study, we define an allele as a specific amino acid at a given codon position. For each allele observed at the 12 codons, we first calculated the proportion of strains that carry the specific allele in the focal set (eg, egg-passaged sequences,  $p_{egg}$ ) and the background set (all sequences in this dataset,  $p_{all}$ ). The enrichment scores (ESs) were then calculated using  $p_{egg}/p_{all}$  for all the alleles across the 12 codons. For each strain, the 12 alleles observed at the 12 positions were combined to define a 12-dimensional vector, indicating enrichment levels at these codons (ie, similar to allelic barcodes). Each of these 32 278 sequences is represented by a data point in the 12-dimensional space. Principal component analysis (PCA) was performed in R using the *princomp()* function to project the 1-dimensional space into the first 2 PCs.

### Vaccine Isolates and Vaccine Efficacy Data

Given that seed viruses must be generated in eggs in the current WHO-recommended vaccine production protocol [25], we collected the egg-passaged strains in the public database



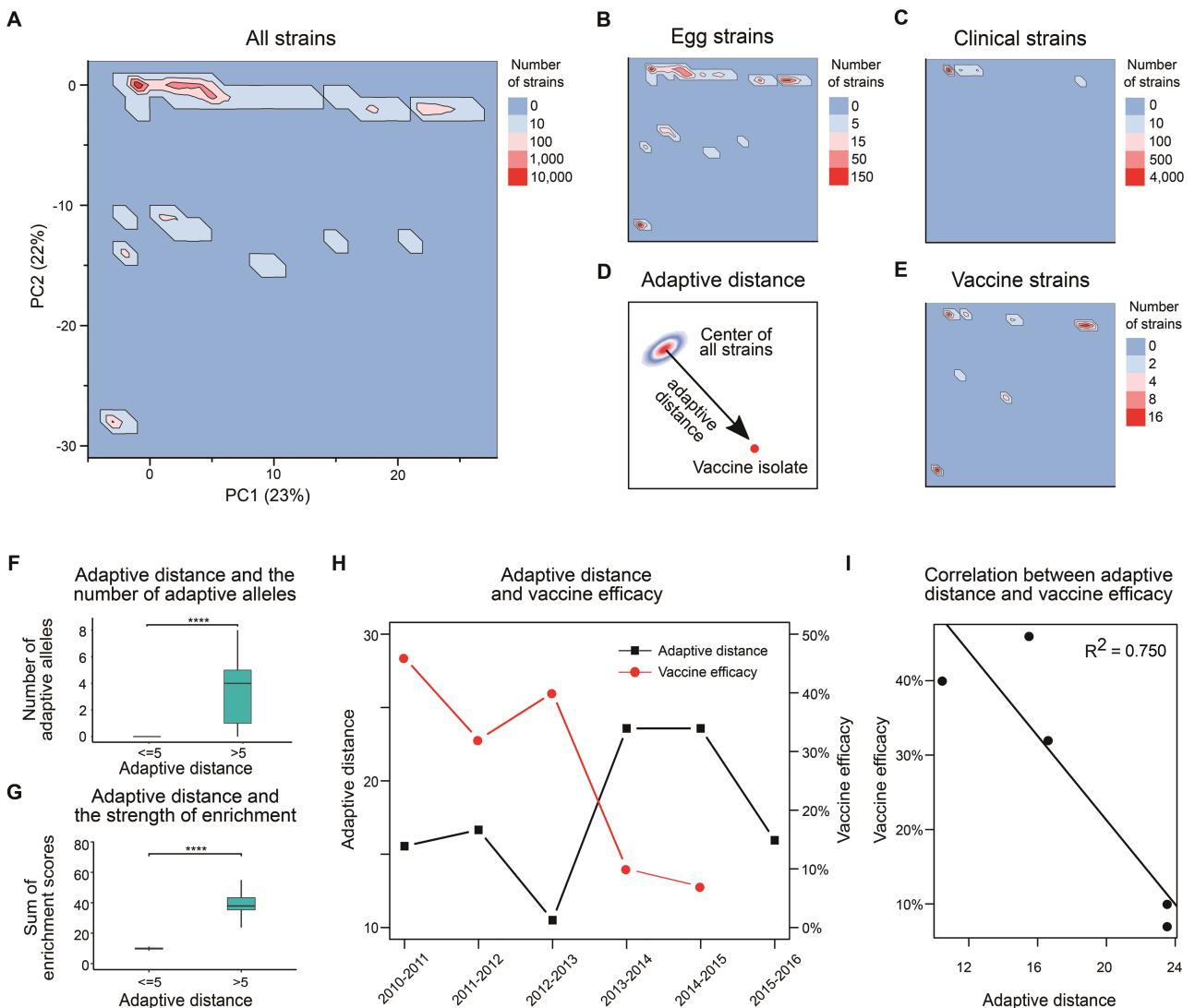
**Figure 1.** Mutational mapping and dynamic convergent evolution. *A*, The phylogenetic tree of the 32 278 HA1 sequences retrieved from the Global Initiative on Sharing All Influenza Data database [22]. Egg and Madin Darby canine kidney (MDCK) cell line passage isolates are labeled on the phylogenetic tree. *B*, Principles of the convergent and enrichment test (see Methods section). The enrichment test determines whether substitutions at a given codon tend to be enriched along terminal branches for the egg-passaged strains ( $n = 661$ ) relative to other terminal branches ( $n = 31 617$ ), and the convergent test determines whether certain nonsynonymous substitutions are observed more frequently than expected by chance along egg terminal branches. *C*, Numbers of nonsynonymous and synonymous changes at the 12 positively selected codons. Significance of the 2 tests ( $P$  value  $\leq .01$ ) are labeled for each codon, and the functional category of the codons are indicated as a heat map beneath the bar plot. *D*, The enrichment scores across 12 codons. Enrichment scores higher than 15 are labeled.

with the same strain name as WHO-recommended vaccine strains. The list of vaccine strains recommended by WHO were curated from WHO and the Influenza Research Database website [26]. When multiple egg-passaged sequences were available for a given strain, we included all of them to represent possible sequence variations during passage in embryonated eggs. The full list of vaccine strains is defined as the vaccine set, and their GISAID isolate names are listed in Supplementary Table 3. In total, 61 sequences from 13 vaccine strains were extracted. In Figure 2H, the VE data were curated from Belongia et al [5], where vaccine efficacies between 2010 and 2015 were calculated by pooling multiple test-negative studies using a random-effects model [5]. The adaptive distance (AD) for a given year was calculated as the mean of the ADs for all the strains of that year in the vaccine set.

## RESULTS

### Data Curation and Mutational Mapping

A total of 32 278 HA1 sequences of the H3N2 influenza virus were downloaded from the GISAID database along with their passage histories. Of these sequences, 661 were from isolates cultured in embryonated eggs and more than 70% had passage annotations (Supplementary Figure 2A). After sequence alignment and the phylogenetic reconstruction (Figure 1A), maximum likelihood estimates of the mutational process (eg, the GTR model) as well as the phylogenetic relationship were constructed. The probabilistic mutational mapping was subsequently employed to sample the nucleotide sequence of the internal nodes recursively from the root of the tree down to the tips, followed by simulation of the mutational history along each branch according to the Markov chain specified by the



**Figure 2.** Adaptive distance and vaccine efficacy. *A*, Principal component plot of the 12-dimensional enrichment scores (ESs) across all isolates ( $n = 32\,278$ ) (principal component analysis [PCA] map). Each strain is a dot in the PCA plot. The clusters of the strains are plotted as the heat map in the PCA plot. *B*, PCA plot of the egg-passaged strains ( $n = 661$ ). *C*, PCA plot of the clinical strains ( $n = 4046$ ). *D*, Adaptive distance (AD) measures the distance between the center of the major cluster and the strain of interest. *E*, PCA plot of the vaccine strains between 2010 and 2016 ( $n = 61$ ). *F*, The number of adaptive alleles (defined as ESs  $\geq 2$ ) for strains with low and high ADs. *G*, The sum of ESs across 12 codon positions for strains with low and high ADs. *H*, ADs for the vaccine strains between 2010 and 2016. Vaccine efficacy (VE) data is plotted using the y-axis on the right-hand side. *I*, Linear regression between VE and AD.  $R^2$  is measured to be 0.750 ( $P$  value = .037).

GTR model (see the Methods section). Since passage adaptation adds extra changes to the original sequences and further extends the terminal branches [27], we merely focused on substitutions that occurred on terminal branches.

#### Positively Selected Codons Under Egg Passage Adaptation and Their Functional Roles

Similar to our previous work [18], in order to identify codon positions that drive passage adaptation, we coupled the inferred substitutions at the 3 codon positions and used 2 statistical tests to target different aspects of the mutational pattern across the phylogenetic tree. The first test (the enrichment test) investigates whether substitutions at a given codon tend to be enriched

along terminal branches for the egg-passaged strains ( $n = 661$ ) relative to other terminal branches ( $n = 31\,617$ ; Figure 1B and the Methods section). The second test (the convergent test) examines whether certain nonsynonymous substitutions are observed more frequently than expected by chance along egg terminal branches ( $n = 661$ ; Figure 1B). We note that the tests for such patterns, using probabilistic mutational mapping, are specifically constructed to take uncertainty in mutational inference into account by incorporating mutation biases, multiple hits, and branch lengths of the phylogeny.

After adjusting for multiple testing and taking into consideration uncertainty across multiple replicates (see the Methods section), we identified 12 codon positions driven by positive

selection in embryonated eggs (Figure 1C). Interestingly, most of these codons are found to be significant by both tests ( $P$  value  $\leq .01$ ). This suggests that passage adaptation tends to be both highly convergent and enriched in embryonated eggs. Across these 12 codons, the total ratio of nonsynonymous to synonymous changes along all the egg terminal branches was 200 to 7. The high nonsynonymous to synonymous ratio indicates exceptionally strong positive selection compared to many other instances of previously reported adaptive evolution detected in natural systems [28]. Of all substitutions across the 12 codons, amino acid positions 156, 186, and 194 contributed most of the nonsynonymous changes (Figure 1C). By correlating these 12 identified codon positions with functional domains (eg, antigenic cluster A-E or receptor binding sites) on the hemagglutinin gene, we observed a strong overlap between antigenic sites (especially antigenic cluster B [9]) and the 12 codons. This suggests that receptor binding during viral infection or host cell immune response may drive passage adaptation in embryonated eggs.

#### Adaptive Distance and Vaccine Efficacy

As these 12 identified codons harbor mostly nonsynonymous substitutions, the classic  $d_N/d_S$  metric will give infinite estimates, hence making it ineffective in measuring the strength of passage adaptation. For highly adaptive substitutions that are strongly selected for in embryonated eggs, they will tend to be enriched in egg-passaged isolates while uncommon in the non-egg isolates. To measure the extent of adaptive convergent evolution for a specific amino acid at a given codon, we first calculated the proportion of strains that carry the amino acid at the focal codon in egg-passaged isolates (denoted as  $p_{\text{egg}}$  in 661 egg strains) and in all isolates (denoted as  $p_{\text{total}}$  in all 32 278 sequences), respectively. We then defined the ES for that allele (eg, amino acid V at position 186) as the ratio of  $p_{\text{egg}}$  to  $p_{\text{total}}$ . Higher ESs indicate strong adaptive convergent evolution driving the selection of certain alleles (ie, amino acid) in embryonated eggs. By plotting allelic ESs across all 12 identified codons, we found that a few alleles at certain codons are specific to egg-passaged isolates and have very high ESs (Figure 1D).

For each isolate in the database, the allelic states of the 12 codons were used to define a DNA barcode, with each allele having an associated ES that measures the strength of passage adaptation at that position. In order to understand the multidimensional distribution of the adaptive evolution, we projected the 12-dimensional ESs for all existing sequences onto a 2-dimensional space using PCA (Figure 2A). In the PCA map, each isolate is a data point projected from the 12-dimensional space. Interestingly, we observed a major cluster harboring most of the isolates at one corner of the PCA map. Scattered islands of sequences were also found along the PC1 and PC2 axis. Since influenza strains driven by strong adaptive evolution carry alleles with high ESs, they tend to deviate from the major cluster in

the PCA map. By plotting the distribution of the egg-passaged isolates, we clearly showed that most of the distantly located data points in Figure 2A are indeed egg-passaged sequences (Figure 2B). Clinical strains or isolates passaged in cell lines (eg, MDCK; Supplementary Figure 2) reside in the major cluster and bear little evidence of passage adaptation (Figure 2C).

In order to measure the extent of adaptive evolution, we defined the AD of a target isolate as its distance from the major cluster (Figure 2D). Using a well-studied metaanalysis for vaccine efficacies [5] (2010–2015), as well as a list of curated vaccine isolates (Supplementary Table 3), we found that a large proportion of these sequences were located far away from the major cluster and were possibly driven by egg-passage adaptation (Figure 2E). When we compared strains with low and high ADs, we found that sequences with higher ADs tended to have both a higher number of adaptive alleles (Figure 2F) as well as alleles with stronger ESs (Figure 2G). Thus, this shows that there is a very strong correlation between AD and the level of egg-passage adaptation. Interestingly, when we plotted the AD for each vaccine isolate as a function of time, there was a rapid increase in AD in the years 2013–2015 followed by a slight drop in the 2015–2016 season (Figure 2H). The increase in AD suggests that the effect of passage adaptation has manifested strongly in the vaccine isolates.

In order to further explore the connection between AD and VE, we extracted VE data compiled by Belongia et al [5] (see the Methods section) and plotted them together with the AD of the vaccine strains. Interestingly, we observed a strong negative correlation between the AD and VE ( $P$  value = .037). Linear regression between these 2 matrices yielded a significant  $R$ -square statistic of 75.0% (Figure 2I), indicating that passage adaptation at these 12 codon positions strongly correlates to the reduction in VE in the past few years.

#### DISCUSSION

The high correlation between VE and AD is concordant with a recent study that showed that egg-adapted vaccine strains elicit specific antibodies during influenza immunization [29]. However, this also raises an interesting question: why does antigenic drift contribute so little to the reduction in VE? The  $R$ -square value (0.812) in the linear regression suggests that antigenic drift only contributes minimally to the total variability in VE. Several factors might explain the importance of passage adaptation in VE. First, antigenic evolution of circulating strains often follows a punctuated equilibrium model, whereby genetic changes only occasionally cause a disproportionately large shift in the antigenic space (averaging every 3 years) [7]. As most of the antigenic changes between years have little effect, they may not contribute strongly to the recent drop in VE. Second, the constant adaptation of the H3N2 influenza virus to the human population has enabled the H3N2 subtype to become well adapted to the human cellular environment. This makes propagating them

in an avian host difficult and often results in lower yields [30], as well as an increasing strength of passage adaptation in the embryonated eggs [18]. The resulting larger antigenic distances [12] between the egg-adapted strains and original isolates may lead to large antigenic jumps and therefore reduced VE.

Since passage adaptation in embryonated eggs often results in the fixation of alleles highly specific to egg-passaged sequences, the allelic status of the 12 codons that we identified provide an important barcode for predicting the footprints of passage adaptation. Based on the principles illustrated in this work, we developed a computational package MADE, (*Measuring Adaptive Distance and vaccine Efficacy based on allelic barcodes*), that will enable vaccine developers to measure the strength of passage adaptation and hence predict the potential efficacy of a vaccine strain based on its nucleotide sequence [31].

It is of note that there are potential limitations in our study. First, the mutational mapping and the 2 statistical tests for enrichment and convergence implemented in this study are approximations of a full codon-based likelihood model. Nonetheless, it is worth pointing out that previous simulation studies have shown that statistical inference is quite accurate in cases of short-term evolution such as in influenza viruses (see [Supplementary Material](#) for a discussion). Second, in addition to passage adaptation, we were not able to consider some factors that may affect the measurement of VE, such as similarities between vaccine strains and circulating strains [32], the time between vaccination and infection, the age of the infected individuals, the immunogenicity variation across individuals [33], the dynamics of immune memory [34], as well as the study design for measuring VE. Moreover, although test-negative design (TND) [35, 36] is a very popular approach for measuring VE, it may be affected by selection bias such as in seeking medical treatment and exposure misclassification [37]. Additionally, TND also relies on several important assumptions such as independence between vaccination and exposure or susceptibility to infection, as well as dichotomous protection (ie, all-or-nothing) from vaccination. Last, the strong correlation between AD and VE is based on observations from a short time period. In future studies with more extensive VE data, researchers will be able to further examine the relationship between passage adaptation and VE ([Supplementary Figure 3](#)).

The observations from our study suggest that strong passage adaptation in embryonated eggs will continue to impede the progress of any current H3N2 vaccine program. One solution may be to switch vaccine production to a cell-based system. A natural question is: will passage adaptation also occur in cell lines? Interestingly, when an analogous analysis on the MDCK medium was conducted, we observed that the allelic ESs for MDCK-passaged isolates were much lower than those of egg-passaged isolates and that the ADs are negligible in most cases and are not correlated with VE. A cell-based vaccine production program and possibly other effective strategies are urgently needed to help increase the global protection against influenza [38].

## Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

## Notes

**Author contributions.** W. Z. conceived and supervised the work. H. C. performed the analysis. J. J. S. A., S. H. N., and R. N. participated in the discussions and data interpretation. H. C., J. J. S. A., S. H. N., R. N., and W. Z. wrote the paper.

**Acknowledgments.** The authors thank Sebastian Maurer-Stroh, Jianjun Liu, Chung-I Wu, Paola Florez de Sessions, and Swaine Chen for helpful discussions. They also thank Cheryl Zi Jin Phua and Chu Quan Tan for their help with the manuscript.

**Financial support.** W. Z. is supported in part by National Key R&D program of China (grant 2018YFC1406902 and 2018YFC0910400).

**Potential conflicts of interest.** All authors: no reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

## References

- Nelson MI, Holmes EC. The evolution of epidemic influenza. *Nat Rev Genet* 2007; 8:196–205.
- Stöhr K. Influenza-WHO cares. *Lancet Infect Dis* 2002; 2:517.
- Centers for Disease Control and Prevention. Prevention and control of influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2010. *MMWR* 2010; 59(No. RR-2).
- Osterholm MT, Kelley NS, Sommar A, Belongia EA. Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. *Lancet Infect Dis* 2012; 12:36–44.
- Belongia EA, Simpson MD, King JP, et al. Variable influenza vaccine effectiveness by subtype: a systematic review and meta-analysis of test-negative design studies. *Lancet Infect Dis* 2016; 16:942–51.
- Belongia EA, Kieke BA, Donahue JG, et al. Effectiveness of inactivated influenza vaccines varied substantially with antigenic match from the 2004–2005 season to the 2006–2007 season. *J Infect Dis* 2009; 199:159–67.
- Smith DJ, Lapedes AS, de Jong JC, et al. Mapping the antigenic and genetic evolution of influenza virus. *Science* 2004; 305:371–6.
- Chambers BS, Parkhouse K, Ross TM, Alby K, Hensley SE. Identification of hemagglutinin residues responsible for H3N2 antigenic drift during the 2014–2015 influenza season. *Cell Rep* 2015; 12:1–6.
- Skowronski DM, Janjua NZ, De Serres G, et al. Low 2012–13 influenza vaccine effectiveness associated with mutation in the egg-adapted H3N2 vaccine strain not antigenic drift in circulating viruses. *PLoS One* 2014; 9:e92153.
- Boni MF. Vaccination and antigenic drift in influenza. *Vaccine* 2008; 26 Suppl 3:C8–14.
- Carat F, Flahault A. Influenza vaccine: the challenge of antigenic drift. *Vaccine* 2007; 25:6852–62.
- Robertson J. Clinical influenza virus and the embryonated hen's egg. *Rev Med Virol* 1993; 3:97–106.
- Skowronski DM, Sabaiduc S, Chambers C, et al. Mutations acquired during cell culture isolation may affect antigenic characterisation of influenza A(H3N2) clade 3C.2a viruses. *Euro Surveill* 2016; 21:30112.
- Burnet F. Growth of influenza virus in the allantoic cavity of the chick embryo. *Aust J Exp Biol Med Sci* 1941; 19: 291–5.
- Madin SH, Darby NB Jr. Established kidney cell lines of normal adult bovine and ovine origin. *Proc Soc Exp Biol Med* 1958; 98:574–6.
- Rogers GN, Pritchett TJ, Lane JL, Paulson JC. Differential sensitivity of human, avian, and equine influenza A viruses to a glycoprotein inhibitor of infection: selection of receptor specific variants. *Virology* 1983; 131:394–408.
- Bush RM, Smith CB, Cox NJ, Fitch WM. Effects of passage history and sampling bias on phylogenetic reconstruction of human influenza A evolution. *Proc Natl Acad Sci USA* 2000; 97:6974–80.
- Chen H, Deng Q, Ng SH, Lee RT, Maurer-Stroh S, Zhai W. Dynamic convergent evolution drives the passage adaptation across 48 years' history of H3N2 influenza evolution. *Mol Biol Evol* 2016; 33:3133–43.

19. Zost SJ, Parkhouse K, Gumina ME, et al. Contemporary H3N2 influenza viruses have a glycosylation site that alters binding of antibodies elicited by egg-adapted vaccine strains. *Proc Natl Acad Sci* **2017**; 114: 12578–83.
20. Nielsen R. Mapping mutations on phylogenies. *Syst Biol* **2002**; 51:729–39.
21. Zhai W, Slatkin M, Nielsen R. Exploring variation in the d(N)/d(S) ratio among sites and lineages using mutational mappings: applications to the influenza virus. *J Mol Evol* **2007**; 65:340–8.
22. Bogner P, Capua I, Lipman DJ, Cox NJ. A global initiative on sharing avian flu data. *Nature* **2006**; 442:981.
23. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **2004**; 32:1792–7.
24. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **2006**; 22:2688–90.
25. Beer DG, Kardia SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* **2002**; 8:816–24.
26. Squires RB, Noronha J, Hunt V, et al. Influenza research database: an integrated bioinformatics resource for influenza research and surveillance. *Influenza Other Respir Viruses* **2012**; 6:404–16.
27. McWhite CD, Meyer AG, Wilke CO. Sequence amplification via cell passaging creates spurious signals of positive adaptation in influenza virus H3N2 hemagglutinin. *Virus Evol* **2016**; 2:pii: vew026.
28. Eyre-Walker A. The genomic rate of adaptive evolution. *Trends Ecol Evol* **2006**; 21:569–75.
29. Raymond DD, Stewart SM, Lee J, et al. Influenza immunization elicits antibodies specific for an egg-adapted vaccine strain. *Nat Med* **2016**; 22:1465–9.
30. Lu B, Zhou H, Chan W, Kemble G, Jin H. Single amino acid substitutions in the hemagglutinin of influenza A/Singapore/21/04 (H3N2) increase virus growth in embryonated chicken eggs. *Vaccine* **2006**; 24:6691–3.
31. Chen Hui, Zhai W. MADE: Measuring Adaptive Distance and Vaccine Efficacy based on Allelic Barcodes (Manuscript in preparation).
32. Bonomo ME, Deem MW. Predicting influenza H3N2 vaccine efficacy from evolution of the dominant epitope. *Clin Infect Dis* **2018**; 67:1129–31.
33. Cobey S, Gouma S, Parkhouse K, et al. Poor immunogenicity, not vaccine strain egg adaptation, may explain the low H3N2 influenza vaccine effectiveness in 2012–2013. *Clin Infect Dis* **2018**; 67: 327–33.
34. Lewnard JA, Cobey S. Immune history and influenza vaccine effectiveness. *Vaccines* **2018**; 6. doi: 10.3390/vaccines6020028
35. Skowronski D, Gilbert M, Tweed S, Petric M, Li Y, Mak A. Effectiveness of vaccine against medical consultation due to laboratory-confirmed influenza: results from a sentinel physician pilot project in British Columbia, 2004–2005. *Can Commun Dis Rep* **2005**; 31: 181–91.
36. Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. *Vaccine* **2013**; 31:2165–8.
37. Joseph A, Lewnard CAT, Benjamin JC, Marc L. Quantifying biases in test-negative studies of vaccine effectiveness. *bioRxiv* **2018**; 237503.
38. Hanon E, Van der Most R, Del Giudice G, Rappuoli R. Short-term and mid-term solutions for influenza vaccines. *Lancet Infect Dis* **2018**; 18:832–3.