VIROLOGICA SINICA 2016, 31 (5): 441–443
DOI: 10.1007/s12250-016-3802-9

LETTER

Large discrepancy between the two-way rNHT distances in hemagglutinin-inhibition assay

Dear Editor,

The influenza viruses cause continual epidemics in human society. As is reported by the World Health Organization (WHO), each year the seasonal influenza viruses, i.e., human influenza A (H1N1), A (H3N2) and B viruses, infected 5%~15% of the world's population, leading to about 3 to 5 million cases of severe illness and about 250000 to 500000 deaths worldwide (WHO, 2014). Vaccination is currently the most effective way to fight against it. Due to the frequent mutations on the HA protein, the virus often changes its antigen, which may lead to the ineffectiveness of the influenza vaccines (Carrat and Flahault, 2007; Taubenberger and Kash, 2010).

Understanding the antigenic changes of the virus is helpful in developing vaccines against the influenza virus. Currently, the hemagglutinin-inhibition (HI) assay is most widely used in determining the antigenic characteristics of influenza virus (Hirst, 1941; Ndifon et al., 2009; Ndifon, 2011). In each HI assay, a HI table is generated, each entry H^{ij} of which represents the HI titer of strain i relative to antisera raised against strain j (Figure 1A). The standard measure of the antigenic difference between virus strains i and j is the reciprocal of the normalized HI titer of i relative to antisera raised against j: $rNHT^{ij} = H^{jj}/H^{ij}$, or vice versa, $rNHT^{ji}$ (Ndifon et al., 2009). Usually, only one-way rNHT distance (either $rNHT^{ij}$ or $rNHT^{ji}$) is used in determining the antigenic variant in influenza surveillance. Few studies have investigated the relationship between the pairwise two-way rNHT distances (Ndifon et al., 2009), i.e., the $rNHT^{ij}$ and $rNHT^{ji}$ distance. Here, by collecting large amounts of HI data of seasonal influenza virus, i.e., influenza A (H1N1), A (H3N2) and B virus (Supplementary Materials), we systematically analyzed the relationship between the pairwise two-way rNHT distances for influenza viruses.

Although there are median correlations between the pairwise two-way rNHT distances for all three (sub) types (with Pearson Correlation Coefficient (PCC) ranging from 0.61 to 0.68, Supplementary Figure S1), surprisingly, large discrepancies were observed between the pairwise two-way rNHT distances, as is shown in Figure 1B. The median differences between the pairwise two-way rNHT distances all equal to 4, while the mean differences range from 11 to 12 for all three (sub)types.

Generally, a strain is said to be antigenically drifted relative to another strain if the rNHT distance is greater than or equal to 4 (WHO, 2002; Ndifon et al., 2009). We found that among a total of 456, 1286 and 753 pairs of two-way rNHT distances for A (H1N1), A (H3N2) and B virus, 72%, 59% and 70% of these pairs agree with each other for the three (sub)types respectively. For the other pairs of two-way rNHT distances, one rNHT distance is greater than or equal to 4, but the other rNHT distance is less than 4, or vice versa. In some cases, the larger rNHT distance is much larger than the smaller one (Figure 1B and Supplementary Figure S1). This suggests that one should be cautious to determine the antigenic variant based on one-way rNHT distance only, which was usually done in routine influenza surveillance.

We next attempted to investigate the factors behind the difference of the pairwise two-way rNHT distances. The fold difference in rate of homogenous titers (the HI titer of a strain relative to the antisera raised against itself) for virus strains i and j was found to correlate most strongly with the differences between the pairwise two-way rNHT distances in all three subtypes (Figure 1C). To some extent, the homogenous titer for a strain reflects its ability of binding to the red blood cells, which was reported to significantly influence the titers of the virus in HI assays. Therefore, the difference of homogenous titers for strain i and j could in part represent the different ability of the viruses binding to the red blood cells, which further influenced the difference between the pairwise two-way rNHT distances. Moreover, this could be validated by the significantly positive correlations between the differences of receptor-binding sites and that of the pairwise two-way rNHT distances (Figure 1C).

In addition, we found the differences between the pairwise two-way rNHT distances for a pair of virus strains i and j correlated positively with the length of the time period between the isolation time of this pair of viruses (Figure 1C). Interestingly, we found that for all three (sub)types of influenza viruses, the rNHT distance with the virus generating homogenous titer isolated in earlier

A

Antisera against	i	j
Strain i	H^{ii}	H^{ij}
Strain j	H^{ji}	H^{jj}

$$rNHT^{ji} = \frac{H^{ii}}{H^{ji}} \qquad rNHT^{ij} = \frac{H^{jj}}{H^{ij}}$$

C

Factors	H1N1	H3N2	B
Homo-titer	0.43	0.44	0.39
Time period	0.21	0.27	0.19
HA1	0.28	0.32	0.26
Epitopes	0.27	0.35	0.28
RBSt	0.34	0.13	0.29
N-gly	0.26	0.15	0.22

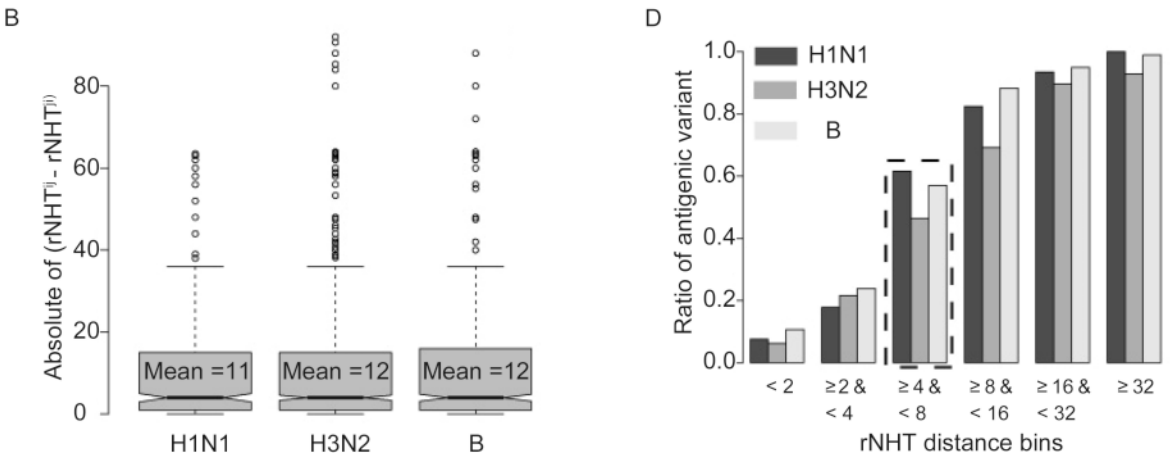

Figure 1. The discrepancies of pairwise two-way rNHT distances for influenza A (H1N1), A (H3N2) and B virus. (A) The HI table and the calculation of pairwise two-way rNHT distances. (B) The distribution of the differences between the pairwise two-way rNHT distances in influenza A (H1N1), A (H3N2) and B viruses. (C) Correlations (Spearman Correlation Coefficient) between the differences of pairwise two-way rNHT distances and fold differences in rate of the homologous titers (Homo-titer), the length of isolation time interval (Time period), the number of amino acid mutations on HA1 protein (HA1), antigenic epitopes, receptor-binding sites (RBS) and N-glycosylation sites (N-gly) between pairs of viruses for influenza A (H1N1), A (H3N2) and B virus. (D) The ratio of antigenic variant in different bins of rNHT distance. The dashed box refers to the bin with the largest uncertainty in determining the antigenic variant for influenza virus.

time is significantly larger than that with the virus generating homogenous titer isolated later (Supplementary Figure S2). This suggests that more antisera raised against the virus isolated in earlier time are needed to neutralize the virus isolated later than the reverse.

We further analyzed whether the differences on HA1 protein between virus strains contributed to the discrepancies of pairwise two-way rNHT distances. The number of amino acid mutations on HA1, antigenic epitopes, receptor-binding sites and N-glycosylation sites were all observed to correlate positively with the differences between the pairwise two-way rNHT distances (Figure 1C). Moreover, the changes of 60, 72 and 50 amino acid positions on HA1 protein for influenza A (H1N1), A (H3N2) and B virus respectively also correlated significantly with the differences between the pairwise two-way rNHT distances (Supplementary Table S2). Most of them were located in antigenic epitopes for influenza A (H1N1) and A (H3N2) viruses. Taken together, the differences on HA1 protein could explain 46%, 45% and 42% of the variance of the differences between the pairwise two-way rNHT distances for influenza A (H1N1), A (H3N2) and B virus, respectively (Supplementary Table S3).

To facilitate its usage in influenza surveillance, we next determined the confidence level of one-way rNHT distances in determining the antigenic variant of influenza virus. The rAHM distance, which integrated the pairwise rNHT distances (Archetti et al., 1950; WHO, 2002) (see Supplementary Materials), was reported to be more accurate in measuring the antigenic difference. When the rAHM distance between a pair of viruses is greater than or equal to 4, they were regarded as antigenic variants. Using this definition, we determined the ratio of antigenic variants for different bins of rNHT distance (Figure 1D). When the rNHT distance is less than 4, or greater than or equal to 8, the ratio of antigenic variants is mostly larger than 80%. While in the bin of 4~8 ($\geqslant$ 4

& < 8), the rNHT distance has the least confidence in determining the antigenic variants for all three (sub) types. We next attempted to improve the confidence of antigenic variants determination for the rNHT distances in this bin using the HA protein sequences. For influenza A (H3N2) virus, we found the HA protein sequence-based computational method PREDAC-H3 (Du et al., 2012; Peng et al., 2016) achieved a higher accuracy than the rNHT did (0.65 *vs* 0.54) (Supplementary Table S4).

Considering the large discrepancy between the pairwise two-way rNHT distances in the HI assay, determining the antigenic variants using one-way rNHT distance may be misleading. The better way is to determine the pairwise two-way rNHT distances between a pair of viruses. Considering the substantial amount of experimental effort involved in HI assays, an alternative way is to improve the quality of HI data by computational methods. For example, the methods developed by Ndifon could be used to recover unmeasured HI data from the measured data and to minimize noise and non-antigenic variation in the HI data (Ndifon, 2011). Besides, since the influenza virus is reported to evolve at population level, the population-based methods, such as the antigenic map, antigenic cluster or antigenic cartography (Smith et al., 2004; Barnett et al., 2012; Du et al., 2012; Peng et al., 2016), may improve the confidence level of individual one-way rNHT distance. This is particular useful when determining the antigenic relationships for lots of viruses. Last but not the least, the sequence-based computational methods (Liao et al., 2008; Deem and pan, 2009; Du et al., 2012) developed recently, such as the PREDAC-H3 mentioned above (Supplementary Table S4), could also help determine the antigenic relationship between viruses when there is a lack of two-way rNHT distances.

FOOTNOTES

This study was supported by the National Natural Science Foundation (31500126 and 31371338), National Key Plan for Scientific Research and Development of China (2016YFD0500300 and 2016YFC1200200) and the Young Teacher's Development Plan of Hunan University to YS (531107040720). There are no conflicts of interest. This article does not contain any studies with human or animal subjects performed by any of the authors.

Supplementary figures/tables are available on the websites of *Virologica Sinica*: www.virosin.org; link.springer.com/journal/12250.

Yousong Peng[1], Dayan Wang[2], Yuelong Shu[2]✉, Taijiao Jiang[3,4]✉

1. College of Biology, Hunan University, Changsha 410082, China
2. National Institute for Viral Disease Control and Prevention, China CDC, Beijing 102206, China
3. Center of System Medicine, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100005, China
4. Suzhou Institute of Systems Medicine, Suzhou 215123, China

✉Correspondence:
Yuelong Shu, Phone: +86-010-58900851,
Email: yshu@cnic.org.cn
ORCID: 0000-0001-6151-5540
Taijiao Jiang, Phone: +86-010-64888427,
Email: taijiao@ibms.pumc.edu.cn
ORCID: 0000-0002-6071-0122

Published online: 9 August 2016

REFERENCES

Archetti I, Horsfall FL. 1950. J Exp Med, 92: 441–462.
Barnett JL, Yang J, Cai Z, Zhang T, Wan XF. 2012. Bioinformatics, 28: 1292–1293.
Carrat F, Flahault A. 2007. Vaccine, 25: 6852–6862.
Deem MW, Pan KY. 2009. Protein Eng Des Sel, 22: 543–546.
Du X, Dong L, Lan Y, et al. 2012. Nat Commun, 3: 709.
Hirst GK. 1941. Science, 94: 22–23.
Liao YC, Lee MS, Ko CY, et al. 2008. Bioinformatics, 24: 505–512.
Ndifon W. 2011. Influenza Other Respir Viruses, 5: 206–212.
Ndifon W, Dushoff J, Levin SA. 2009. Vaccine, 27: 2447–2452.
Peng Y, Yang L, Li H, et al. 2016. Bioinformatics. DOI: 10.1093/bioinformatics/btw185.
Smith DJ, Lapedes AS, de Jong JC, et al. 2004. Science, 305: 371–376.
Taubenberger JK, Kash JC. 2010. Cell Host Microbe, 7: 440–451.
WHO. 2002. WHO manual on animal influenza diagnosis and surveillance.
WHO. Mar 5th, 2014. http://www.who.int/mediacentre/factsheets/fs211/en/