Interdiscip Sci Comput Life Sci (2012) 4: 282–290
DOI: 10.1007/s12539-012-0136-7

Genetic Ancestor of External Antigens of Pandemic Influenza A/H1N1 Virus

Ramaiah ARUNACHALAM, Kanniah PAULKUMAR, Gurusamy ANNADURAI*
(Environmental Nanotechnology Division, Sri Paramakalyani Centre for Environmental Sciences, Manonmaniam Sundaranar University, Alwarkurichi 627412, Tamilnadu, India)

Received 9 June 2011 / Revised 5 March 2012 / Accepted 29 August 2012

Abstract: The aim of the present investigation was to discover the genetic relationships of 2009 pandemic novel influenza A/H1N1 virus (NIV) external antigens Hemagglutinin (HA) and Neuraminidase (NA) with other influenza viruses by performing phylogenetic, comparative and statistical analyses. Phylogenetic trees of these two antigens show that the sequences of the NIV viruses are relatively homogeneous and these were derived from several viruses circulating in swine. The phylogenetic tree of HA shows that NIV had the closest relationship with North-American pig lineages whereas NA had with European pig lineages. In both segments, NIVs had the closest genetic relationship with swine influenza virus lineages. It strongly suggests that pigs are the most possible animal reservoir. Comparative analysis shows that among clade A, NIVs had very low genetic divergence as well as high similarity and also suffered strong purifying selection whereas neighbor clade B shows moderate values when compared to those of clades C-F. It indicates that classical swine influenza viruses present in clade B might be an ancestor of NIVs external antigens. The process of re-assortment occurred in classical swine influenza viruses. The mutation sites exclusively fixed in the NIV of swine and human along with vaccine strain provide an important suggestion for disease diagnosis and vaccine research.
Key words: influenza A/H1N1 virus, surface antigens, phylogeny, genetic relationship.

1 Introduction

Influenza A viruses are RNA viruses belonging to the family of *Orthomyxovirus* (Palese and Young, 1982) and they are divided into different subtypes based on two major surface glycoproteins/external antigens like HA and NA (Ma *et al.*, 2010; Webster *et al.*, 2006). There are sixteen known types of HA (H1 to H16) and nine of NA (N1 to N9), all found in waterfowl. Only H1, H2, H3 and N1, N2 are known to have caused epidemic disease in humans (Tamuri *et al.*, 2009; Arunachalam *et al.*, 2012a). However, complete role of the HA in viral infection remains unclear. In general, infection by the influenza virus begins when viral HA binds to the target cell surface receptors containing sialic acid through receptor mediated-endocytosis. On the other hand, NA is responsible for releasing the virus from infected cells (Rogers *et al.*, 1983; Al-Majhdi, 2007). Subtypes H1, H2 and H3 belong to influenza A viruses circulating primarily in human. Influenza viruses are not considered highly pathogenic, however, rapid evolution based on the antigenic changes in surface glycoproteins causes the recurrence of influenza A virus and accounts for major public health concerns (Zambon, 2001; Al-Majhdi, 2007). The nature of the influenza A virus segmented genome allows for antigenic shift or re-assortment that leads to the generation of new viruses. Flu viruses undergo mutations when they spread from place to place, which leads to gradual changes in the HA and/or NA proteins (antigenic drift) (Arunachalam *et al.*, 2012a).

Reports on widespread transmission of 2009 pandemic NIVs among human in Mexico, the United States, Europe and elsewhere highlight this is an ever-present threat to global public health (Babakir-Mina *et al.*, 2009; CDC, 2009a and 2009b). Swine is hypothesized to act as a mixing vessel for the re-assortment of avian, swine and human influenza viruses (Smith *et al.*, 2009; Ding *et al.*, 2009; Shinde *et al.*, 2009) and might play an important role in the emergence of new influenza viruses capable of causing a human pandemic (Babakir-Mina *et al.*, 2009; CDC, 2009a and 2009b; Ding *et al.*, 2009; Van-Reeth and Nicoll, 2009; Scholtissek, 1990; Webby and Webster, 2001). Most of the avian influenza A viruses that can result in sporadic human infection via animals to human transmission, lack the ability of human to human transmission (Ding

*Corresponding author.
E-mail: gannadurai@hotmail.com

et al., 2009; Kuiken *et al.*, 2006; Webby and Webster, 2001). However, interestingly NIVs show strong ability to transmit from human to human (CDC, 2009a). Thus, it is necessary to investigate the genetic ancestor of NIVs to provide an important suggestion for disease diagnosis and vaccine research.

Comparative genome analysis has become feasible with the availability of a number of completely sequenced genomes. The availability of completely sequenced genomes increases the predictive power in deciphering the hidden information in genome design, function and evolution (Sivashankari and Shanmughavel, 2007). Comparison of complete genomes of organisms by phylogenetic and statistical methods allows for global views on genome evolution. Comparative genomics can not only trace out the evolutionary relationship between organisms, but also indicate differences and similarities within and between species (Arunachalam *et al.*, 2012b). Therefore, the aim of the present investigation was to determine the possible ancestor of NIV external antigens by performing phylogenetic and comparative analyses.

2 Materials and methods

2.1 Phylogenetic reconstruction

The corresponding coding sequences of NIVs HA (totally 89 sequences; 576 bp) and NA (totally 94 sequences; 477 bp) external antigen were retrieved from the NCBI Influenza Virus Resource (http://www.ncbi.nlm.nih.gov/genomes/FLU/SwineFlu.html) on June 10, 2010. Other existing sequences of influenza A viruses sampled from human and swine around the world were also retrieved from GenBank during the period of 1918-2010. The nucleotide sequences were translated into respective peptide sequences by using EMBOSS Transeq (http://www.ebi.ac.uk/Tools/emboss/transeq/index.html). MEGA 4.1 was started with a set of aligned amino acids sequences of HA and NA segment of the NIVs with classical influenza viruses using ClustalW, and the phylogenetic trees were constructed using Neighbor-Joining (NJ) algorithm (Tamura *et al.*, 2007) under *p*-distance model. The sequence alignments were performed under default condition. The gap open and gap extension penalties in the sequence alignments were 15 and 6.66, respectively. The reliability of the trees was evaluated by the bootstrap method with 1000 replications. It should be noted that the *p* distance is known to produce reliable phylogenetic trees, when a large number of closely related sequences is analyzed (Suzuki, 2006). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Phylogenies might have bias based on the statistical model and algorithm used for tree reconstruction. Other two alternative approaches were used to verify the *p*-distance based NJ trees. In the first approach, Jones-Taylor-Thornton (JTT) model (Jones *et al.*, 1992) was used for Maximum Likelihood (ML) tree reconstruction using MEGA5.05 (Tamura *et al.*, 2011) under complete deletion option. In the second approach, simple hill-climbing algorithm was used to adjust tree topology and branch lengths simultaneously. This algorithm starts from an initial tree built by a fast distance-based method and modifies this tree to improve its likelihood in all iteration. Due to this simultaneous adjustment of the topology and branch lengths, only a few iterations are sufficient to reach an optimum (Guindon and Gascuel, 2003).

The genetic diversity/distance (GD) within and between the NIVs and their closest evolutionary relatives were assessed by MEGA4.1 using *p*-distance method (Tamura *et al.*, 2007; Ding *et al.*, 2009).

2.2 Comparative analysis

The GD was calculated between the clades A-F by MEGA4.1 using NJ method under the *p*-distance model (Tamura *et al.*, 2007). Based on the noteworthy results, clades A and B were considered for further comparative and statistical analyses. In pair-wise comparison, bottom sequence HA – A/Ontario/29801/2009(H1N1) and NA – A/California/VRDL12/2009(H1N1) of outbreak clade A was used to calculate the sequence similarity and GD with other sequences present in clades A and B stated in both phylogenetic trees using ClustalW2 (Thompson *et al.*, 1994) and MEGA4.1, respectively. The experimental variables such as similarity and GD were used for perfect negative correlation test to confirm the correlation between the two variables. In perfect negative correlation, the values of the two variables move in opposite direction (Gurumani, 2005). XLSTAT2010 (http://www.xlstat.com/) was used for linear regression analysis to obtain theoretical values using experimental values to verify the agreement. Analysis of variance (ANOVA) was performed at 95% confidence interval.

It is essential to find out the variations/mutations between the NIV of swine and human with vaccine strains viruses for designing the effective vaccines and drugs. The representative strains HA – A/Ontario/29801/2009(H1N1), A/Swine/IL/3910/2010(H1N1) from clade A and A/Swine/NorthCarolina/7386 from clade B, and strains NA - A/California/VRDL12/2009, A/Swine/Alberta/OTH-33-24/2009 from clade A and A/Swine/England/WVL7/1992(H1N1) from clade B with vaccine strain A/California/07/2009(H1N1) were used for phylogenetic tree reconstruction using NJ algorithm (Tamura *et al.*, 2007) under *p*-distance model.

3 Results and discussion

3.1 Phylogenetic analysis

Based on the phylogenetic analysis, we have estimated genetic relationships of NIV external antigens by using MEGA4.1 (Tamura *et al.*, 2007). It should be noted that the modified analysis was carried out in coding region of entire NIV genomic segments for discovering their evolution (Arunachalam *et al.*, 2012a). Phylogenetic trees show that the HA segment is derived from North-American swine lineages whereas NA is derived from Eurasian swine lineages (Figs. 1(a) and 1(b)). It is noted that these two segments do not cluster with both human and avian influenza viruses. These two segments are comprised by themselves and play a major role in the formation of re-assortant NIVs. It implies that HA and NA segments have the potential to evolve from influenza viruses present in clade B and provide the ability to transmit from human to human after animal-to-human cross species transmission. Phylogenetic trees (Figs. 1(a) and 1(b)) show that outbreak NIV clade A had well-supported branch with clade B containing swine-origin influenza virus lineages. It is supported by the bootstrap value of 93. In the HA tree, clade A contains few sequences of pandemic swine influenza virus which is very close to American swine influenza viruses present in the neighboring clade B. It could indicate the gradual evolution of pandemic swine influenza virus HA segment from North-American swine influenza viruses and recently well adapted in humans for pandemic outbreak. As mentioned above, NA tree also indicate that NIV may come from Eurasian swine lineages, not from American pig lineages, which is supported by the bootstrap value of 89. These results indicate that NIV had a short evolutionary history among human and it had a long evolutionary history before introduction into human. It clearly suggests that the movement of domestic pigs between Eurasia and North-America appears to have mixing of diverse swine influenza viruses, leading to the multiple re-assortment events associated with the genesis of the NIVs. The NIV not only accomplishes the cross-species transmission from pig to human but also gains the ability to spread efficiently among humans (Smith *et al.*, 2009; Ding *et al.*, 2009). It suggests that Eurasian and American swine lineages might be an ancestor of NIVs NA and HA gene, respectively. This result is well consistent with recent observations of Smith *et al.* (2009), Ding *et al.* (2009) and Dawood *et al.* (2009). These findings suggest that 2009 pandemic NIV genesis happened in pigs due to the reassortment of genes HA and NA from classical swine influenza viruses. Our analyses show that each external protein of the NIV is nested within well-established swine influenza lineages. The most parsimonious interpretation of these results is that the progenitor of the NIV pandemic originates from pigs (Smith *et al.*, 2009).

The p-distance based NJ tree of these two segments is evaluated by JTT model based ML trees that shows no significant variations in NA segment (figure not shown). In the NJ tree of HA genes, the strain A/swine/1931(H1N1)/USA is positioned in clade C (Fig. 1(a)) whereas in ML tree the strain does not fall into clade C, and alternatively it is positioned between the clade B and clade D (figure not shown). It is noted that in the NJ tree, the strain A/swine/1931(H1N1)/USA is placed in close to the seasonal human influenza strains isolated from USA, indicating that the A/swine/1931(H1N1)/USA strain is an ancestor of these human influenza strains. In the ML tree, the strain A/swine/1931(H1N1)/USA is placed between the swine and human influenza strains isolated from USA. It clearly indicates that the genetic materials of the A/swine/1931(H1N1)/USA strain are still continuously circulating in swine and also due to the process of cross-species transmission they are transferred from swine to human and cause outbreak. Moreover, maximum likelihood based simple hill-climbing algorithm HA and NA trees show no significant variations when compared with NJ trees (data not shown). It suggests that NJ and simple hill-climbing algorithm execute more similar topologies than those of ML trees.

3.2 Comparative and statistical analyses

Phylogenetic trees of the two external antigens show that the clade A has very short branch length with the nearest clade B when compared to the neighboring clades C-F (Figs. 1(a) and 1(b)). It suggests that HA and NA segments are derived from influenza viruses present in clade B due to the gradual evolution followed by re-assortment process which occurred very recently. Furthermore, independent genetic distance is calculated for clades A-F, showing that the mean of genetic distance (0.005-0.006) of the outbreak NIV in clade A is noticeably smaller than those in the other clades B-F (0.039-0.060; 0.025-0.073; 0.048-0.062; 0.064-0.067; 0.097-0.102) (Fig. 2(a)). Very low genetic distances are found within an outbreak clade A, indicating that all pandemic outbreak NIVs appear homogeneous and clustered together (bootstrap value 100), although it is distinct from classical human (H1N1) influenza viruses. The comparison of mean of genetic distances of outbreak clade A with other clades B-F show (Fig. 2(b)) that in the HA and NA trees, the distance between clades A and B ranges from 0.069-0.073 whereas the distance between clade A and other clades C-F is 0.103-0.186, 0.166-0.197, 0.174-0.359 and 0.541-0.557, respectively. Thus, the net distance (Fig. 2(c)) shows that the distance between the outbreak clade A and clade B ranges from 0.040-0.046 whereas that between clade A and other clades C-F is 0.087-0.147, 0.139-0.163, 0.139-0.323 and 0.487-0.506, respectively. These net distance is significant for the mean of genetic distance analysis (Fig. 2(b)). Calculating GD between clades is based on

Fig. 1 Phylogenetic trees of external antigen HA and NA segments of NIV. These trees were constructed by MEGA4.1 using the NJ method under *p*-distance model. The reliability of the trees was evaluated by the bootstrap method with 1000 replications. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). All sequences were divided into six groups, A-F. (a) HA tree; (b) NA tree.

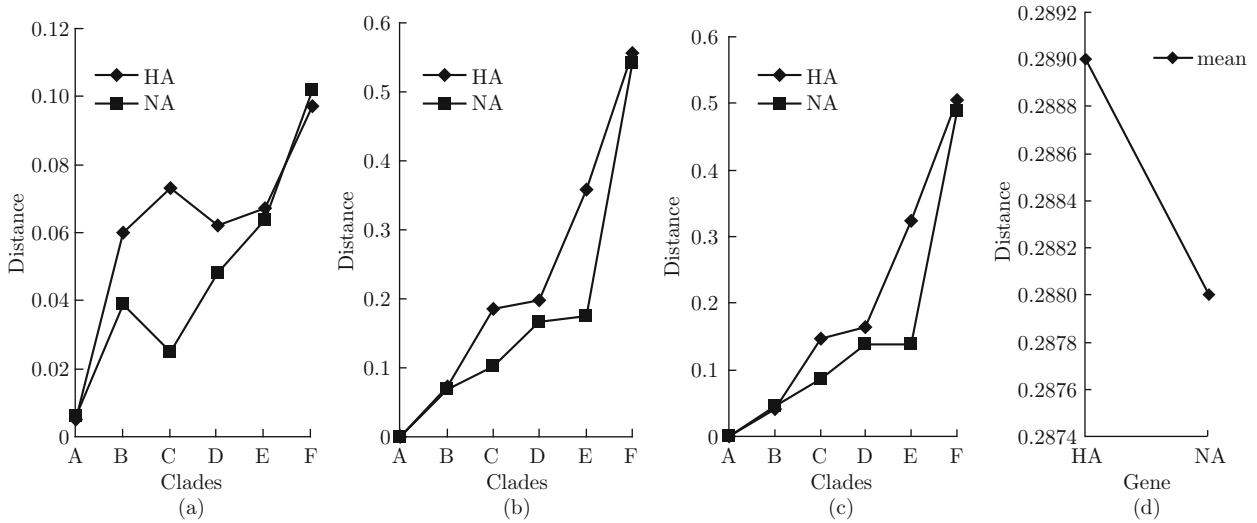

Fig. 2 The clades A-F of HA and NA were used for calculating GD by MEGA4.1 using NJ method under the *p*-distance model. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). (a) GD was calculated within each clade, which indicates difference of base number per site from averaging over all sequence pairs within each group; (b) GD was calculated between the clades, indicating difference of base number per site from averaging over all sequence pairs between groups; (c) Net distance was calculated between the clades, indicating difference of base number per site from estimation of net average between sequence groups; (d) Overall mean was calculated for each tree.

difference of base number per site from averaging overall sequence pairs between groups whereas net GD calculation is carried out based on difference of base number per site from estimation of net average between groups of sequences. Figs. 2(b) and 2(c) show that clade A has very shorter genetic distance with clade B than with other clades C-F. These results indicate that the NIV has a short evolutionary history among humans and before introduction into human it has a long evolutionary history, implying that this virus might have been circulating among animal reservoirs somewhere in the world for a relatively long period of time. NIVs have the closest genetic relationship with swine influenza virus lineages (Figs. 1(a) and 1(b)), suggesting that pigs are the most possible animal reservoir, and as a consequence, the present 2009 outbreak occurred due to re-assortment of swine-origin lineages followed by the cross-species transmission of NIV from pig to human (Dawood *et al.*, 2009). Swine was demonstrated to play a role as a mixing vessel in the co-infection and the re-assortment of various influenza viruses (Ding *et al.*, 2009; Van-Reeth and Nicoll, 2009). It should be noted that there are no significant variation found in overall mean of these two proteins (0.288-0.289) (Fig. 2(d)) in the trees which are typically based on the number of sequences used for analysis.

The clades A and B were considered for further statistical analysis because of short branch length and the presence of low GD between these two clades (Figs. 2(b) and 2(c)). In both trees, the bottom/root sequence of outbreak clade A (HA – A/Ontario/29801/2009(H1N1); NA – A/California/VRDL12/2009(H1N1)) was used as a query to estimate the pairwise GD and similarity with outbreak clade A and neighboring clade B sequences. In these tests, query sequence has very low GD and high similarity with outbreak clade A sequences when compared to clade B viruses (Figs. 3(a) and 3(b)). In sequence similarity analysis, the high similarity was observed among the outbreak clade A viruses whereas the moderate similarity was observed in the neighboring clade B viruses. It indicates that the NIVs have high similarity and very low GD among the outbreak clade A viruses when compared to clade B viruses. It suggests that the clade A NIV isolates are homogeneous and the clade B viruses may be an ancestor of outbreak NIVs. The genetic distance and similarity values of these sequences were used for perfect negative correlation test. In perfect negative correlation, the values of the two variables moved in opposite direction (Gurumani, 2005). In support of this rule, the correlation between GD and sequence similarity with effective relationship is negative one. As the GD increases, the sequence similarity declines, and the GD decreases, the similarity increases (Fig. 3). The results also show that individual (pair-wise) GD is very low among the outbreak clade A viruses whereas that in the neighboring clade B viruses is higher. Values of $R^2 > 0.9$ and < 1 indicate model terms are significant. The R^2 values of the two segments are significant. It indicates a perfect correlation/agreement occurred between the distance and similarity values (Figs. 3(a) and 3(b)). Estimation of the gaps in genetic surveil-

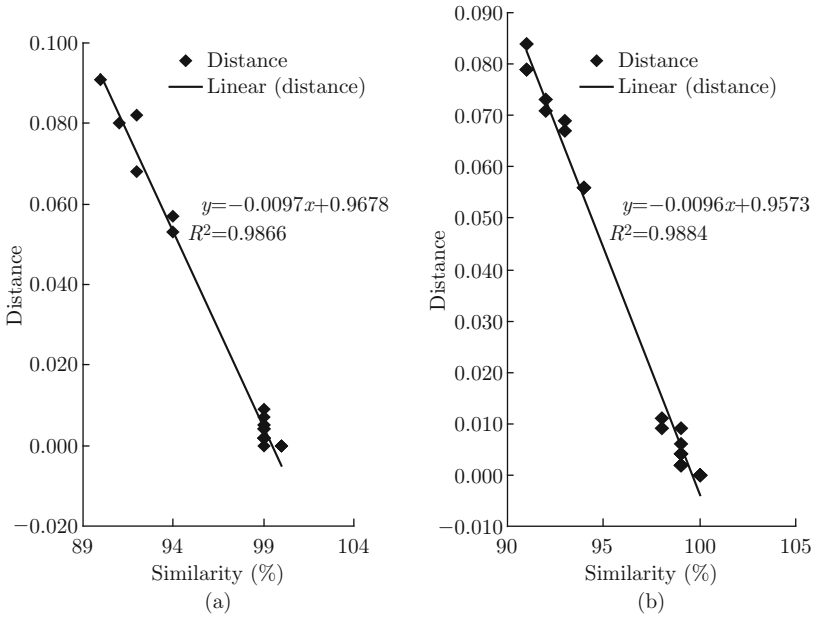

Fig. 3 Similarity and GD were used for perfect negative correlation test. (a) HA; (b) NA.

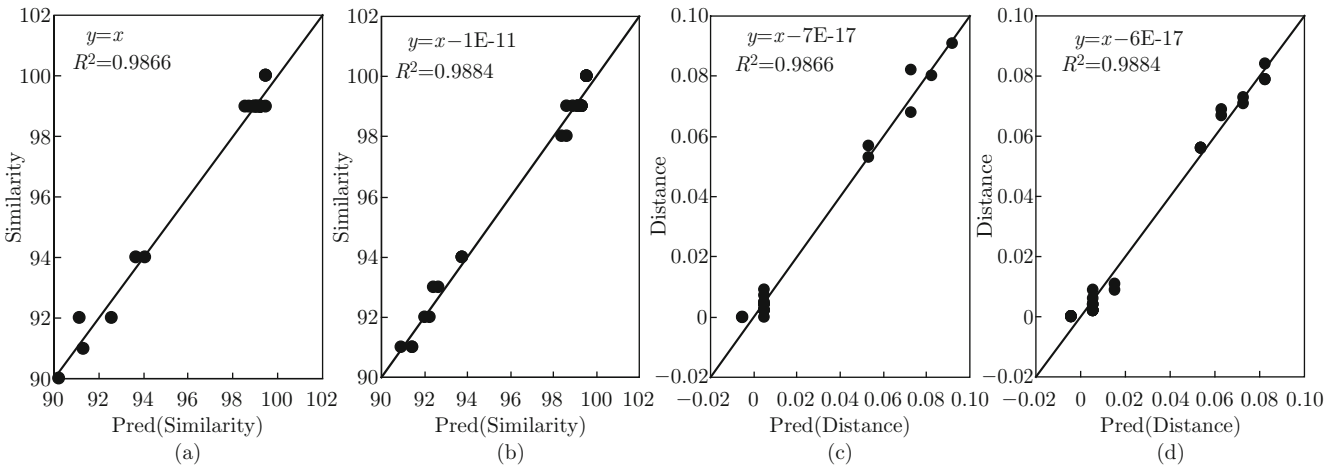

Fig. 4 Linear regression analysis was performed to obtain theoretical values using experimental values of similarity and distance to verify the consistency. (a) Similarity values of HA; (b) Similarity values of NA; (c) Distance values of HA; (d) Distance values of NA.

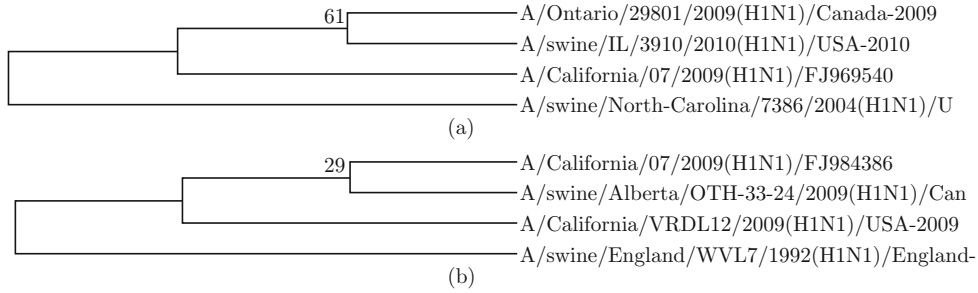

Fig. 5 Phylogenetic trees of external antigen HA and NA segment of the NIVs from swine, human, vaccine strain 2009 and classical swine virus. The trees were constructed by MEGA4.1 using the NJ method under *p*-distance model. The reliability of the trees was evaluated by the bootstrap method with 1000 replications. (a) HA tree; (b) NA tree.

lance/regression line (scatter plot) indicates a long period of unsampled ancestry before the NIV outbreak, suggesting that the re-assortment of swine lineages may have occurred years before emergence in human.

The experimental values of similarity and GD were confirmed by theoretical values which were obtained using linear regression analysis. Experimental and theoretical values were subjected to scatter plot analysis (Figs. 4(a)-4(d)). The consistency of the experimental and theoretical values was determined by calculating R^2 values of each protein segment. The R^2 values of the two segments are significant (Table 1 and Figs. 4(a)-4(d)). The R^2 values of similarity and GD of the two segments obtained from these experiments are shared common values (Table 1). In addition to confirming the result, the variance analysis (ANOVA) was performed at 95% confidence interval. Here the values of '*Prob*>F' was less than 0.0500 indicating model terms are significant. It is shown in Table 1 that the two protein segments show values of < 0.0001, which is statistically significant. It should be noted that the perfect consistency is observed between the experimental and theoretical values. It shows that the experimental values obtained from ClustalW (similarity) and MEGA4.1 (GD) are acceptable.

Table 1 The consistency of theoretical values and experimental values was analyzed using Analysis of Variance (ANOVA) of XLSTAT 2010 at 95% confidence interval

Gene	R^2 Similarity / Distance	Similarity / Distance	
		F	Pr > F
HA	0.9866	1617.618	< 0.0001
NA	0.9884	2300.814	< 0.0001

Values of "Prob > *F*" less than 0.0500 indicate model terms are significant.

3.3 Mutation analysis

The representative strains of HA - A/Ontario/29801/2009(H1N1), A/Swine/IL/3910/2010(H1N1) from clade A and A/Swine/North-Carolina/7386 from clade B, and strains of NA - A/California/VRDL12/2009, A/Swine/Alberta/OTH-33-24/2009 from clade A and A/Swine/England/WVL7/1992 (H1N1) from clade B, and vaccine strain A/California/07/2009(H1N1) were used for phylogenetic tree reconstruction (Figs. 5(a) and 5(b)). It is important to note that in HA segment, NIV of human and swine had an own branch and clustered with A/California/07/2009(H1N1) vaccine strain and North-American classical swine. It suggests that HA segment is derived from cross-species transmission from North-American pig to human. Moreover, the vaccine strain is placed between NIV of swine and classical swine virus. In the NA tree, the vaccine strain had an own branch with NIV of swine. It is essential to confirm the variations/mutations between the entire viruses used in Figs. 5(a) and 5(b). The comparison results of the strains used in the present study show (Figs. 5(a) and 5(b)) that NIVs of human and swine and vaccine strain have 27 widespread mutations in HA antigen at amino acid level (Table 2). NIVs of swine and human share two widespread mutations whereas NIVs of human and vaccine strain have one mutation in common that are highlighted in bold. Single independent mutation in NIVs of swine and vaccine strain was also found (underlined), however, in vaccine strain we found two unknown amino acids. NA segment shows 24 widespread mutations in NIVs of swine and human and vaccine strain, here 3 autonomous mutations in NIVs of swine, 2 mutations in NIVs of human and no mutation was found in vaccine strain (Table 2). The widespread mutations found in NIVs of swine indicate that the gene segments have experienced a long time for genetic evolution among the American and European pigs. The autonomous mutations help the virus to adapt in human system. It provides a potential evidence of selective sweep in early evolution of the NIVs.

Table 2 Mutation sites exclusively fixed in the novel influenza A/H1N1 virus from swine and human along with vaccine strain 2009 by using MEGA4.1

Classical virus gene	Swine 2009	Human 2009	Vaccine Strain 2009
HA	T13A	T13A	T13A
	N52Y	N52D	N52D
	L78I	L78I	L78I
	–	–	S100P
	N101S	N101S	N101S
	N114D	N114D	N114D
	T145S	T145S	T145S
	R147K	R147K	R147K
	T158A	T158A	T158A
	N159K	N159K	N159K
	R163K	R163K	R163K
	I178L	I178L	I178L
	N185D	N185D	N185D
	E187G	E187G	E187G
	S220T	**S220T**	–
	R225K	R225K	R225K
	E228K	E228K	E228K
	T233I	T233I	T233I
	–	–	?
	–	–	?
	A241E	A241E	A241E
	L274M	L274M	L274M

Continued

Classical Virus Gene	Swine 2009	Human 2009	Vaccine Strain 2009
	K275E	K275E	K275E
	S278A	S278A	S278A
	S288P	S288P	S288P
	V315I	V315I	V315I
	–	**E319K**	**E319K**
	M331L	M331L	M331L
	I338V	**I338V**	–
	I362V	I362V	I362V
	G391E	G391E	G391E
	S415N	S415N	S415N
	A440S	–	–
	D490N	D490N	D490N
NA	I13V	I13V	I13V
	I19M	I19M	I19M
	S21N	S21N	S21N
	I26V	–	–
	I40L	I40L	I40L
	L46I	L46I	L46I
	K73N	K73N	K73N
	V75A	V75A	V75A
	E77G	E77G	E77G
	T79S	T79S	T79S
	S95G	–	–
	–	V106I	–
	S125P	–	–
	H126P	H126P	H126P
	V163I	V163I	V163I
	S189N	S189N	S189N
	–	N248D	–
	K257R	K257R	K257R
	R260K	R260K	R260K
	V263I	V263I	V263I
	L269M	L269M	L269M
	G286S	G286S	G286S
	R331K	R331K	R331K
	T365I	T365I	T365I
	S369N	S369N	S369N
	V389I	V389I	V389I
	T397N	T397N	T397N
	D398E	D398E	D398E
	M418I	M418I	M418I

'–' represents no mutation; '?' represents unknown mutation.

We concluded that the phylogenetic, comparative and statistical analyses supported that HA and NA of NIVs was re-assorted and evaluated from American and European swine rather than swine from other countries. The closest genetic relationship of the external NIV antigens with swine influenza virus lineages strongly suggest that pigs are the most possible animal reservoir, and as a consequence, the current pandemic outbreak occurred due to the process of re-assortment followed by cross-species transmission of NIV from pig to human. The need of widespread influenza surveillance in humans and systematic swine surveillance allows for the undetected persistence and evolution of the pandemic NIVs. It is important to observe the mutations which have higher impact on functional amino acids of these external antigens and also to study the genetics of the NIVs in the world for selecting the suitable strains along with A/H3N2 and influenza B virus for developing more effective antiviral drugs and vaccines.

Acknowledgements Authors are grateful to Emeritus Prof. K.G. Sivaramakrishnan, Madras Christian College, Chennai, India for his fruitful suggestion.

References

[1] Al-Majhdi, F.N. 2007. Structure of the sialic acid binding site in influenza A virus: Hemagglutinin. J Biol Sci 7, 113–122.

[2] Arunachalam, R., Paulkumar, K., Annadurai, G. 2012a. Phylogenetic analysis of pandemic influenza A/H1N1 virus. Biologia 67, 14–31.

[3] Arunachalam, R., Senthilkumar, B., Senbagam, D., Selvamaleeswaran, P., Rajasekarapandian, M. 2012b. Molecular phylogenetic approach for classification of *Salmonella typhi*. Res J Microbiol 7, 13–22.

[4] Babakir-Mina, M., Dimonte, .S, Perno, C.F., Ciotti, M. 2009. Origin of the 2009 Mexico influenza virus: A comparative phylogentic analysis of the principal external antigens and matrix protein. Arch Virol 154, 1349–1352.

[5] Centers for Disease Control and Prevention (CDC). 2009a. Swine influenza A(H1N1) infection in two children-southern California, March-April 2009. MMWR, Morb Mortal Wkly Rep 58, 400–402.

[6] Centers for Disease Control and Prevention (CDC). 2009b. Update swine influenza A(H1N1) infections - California and Texas, April 2009. MMWR Mrob Mortal Wkly Rep 58, 1–3.

[7] Dawood, F.S., Jain, S., Finelli, L., Shaw, M.W., Lindstrom, S., Garten, R.J., Gubareva, L.V., Xu, X., Bridges, C.B., Uyeki, T.M. 2009. Emergence of a novel swine-origin influenza A(H1N1) virus in humans. N Engl J Med 360, 2605–2615.

[8] Ding, N., Wu, N., Xu, Q., Chen, K., Zhang, C. 2009. Molecular evolution of novel swine-origin A/H1N1 influenza viruses among and before human. Virus Genes 39, 293–300.

[9] Guindon, S., Gascuel, O. 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52, 696–704.

[10] Gurumani, N. 2005. An Introduction to Biostatistics, 2^{nd} Edition. MJP Publishers, Chennai.

[11] Jones, D.T., Taylor, W.R., Thornton, J.M. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8, 275–282.

[12] Kuiken, T., Holme, E.C., McCauley, J., Rimmelzwaan, G.F., Williams, C.S., Grenfell, B.T. 2006. Host species barriers to Influenza virus infections. Science 312, 394–397.

[13] Ma, W., Vincent, A.L., Lager, K.M., Janke, B.H., Henry, S.C., Rowland, R.R.R., Hesse, R.A., Richt, J.A. 2010. Identification and characterization of a highly virulent triple reassortant H1N1 swine influenza virus in the United States. Virus Genes 40, 28–36.

[14] Palese, P., Young, J.F. 1982. Variation of influenza A, B and C viruses. Science 215, 1468–1474.

[15] Rogers, G.N., Paulson, J.C., Daniels, R.S., Skehel, J.J., Wilson, I.A., Wiley, D.C. 1983. Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity. Nature 304, 76–78.

[16] Scholtissek, C. 1990. Pigs as 'mixing vessels' for the creation of new pandemic influenza A viruses. Med Princ Pract 2, 65–71.

[17] Shinde, V., Bridges, C.B., Uyeki, T.M., Shu, B., Balish, A., Xu, X., Lindstrom, S., Gubareva, L.V., Deyde, V., Garten, R.J., Harris, M., Gerber, S., Vagasky, S., Smith, F., Pascoe, N., Martin, K., Dufficy, D., Ritger, K., Conover, C., Quinlisk, P., Klimov, A., Bresee, J.S., Finelli, L. 2009. Triple-reassortant swine influenza A (H1) in humans in the United States, 2005-2009. N Engl J Med 360, 2616–2625.

[18] Sivashankari, S., Shanmughavel, P. 2007. Comparative genomics – A perspective. Bioinformation 1, 376–378.

[19] Smith, G.J., Vijaykrishna, D., Bahl, J., Lycett, S.J., Worobey, M., Pybus, O. 2009. Origin and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459, 1122–1126.

[20] Suzuki, Y. 2006. Natural selection on the Influenza virus genome. Mol Biol Evol 23, 1902–1911.

[21] Tamura, K., Dudley, J., Nei, M., Kumar, S. 2007. MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24, 1596–1599.

[22] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28, 2731–2739.

[23] Tamuri, A.U., Reis, D.M., Hay, A.J., Goldstein, R.A. 2009. Identifying changes in selective constraints: Host shifts in influenza. PLoS Comput Biol 5, e1000564.

[24] Thompson, J.D., Higgins, D.G., Gibson, T.J. 1994. Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22, 4673–4680.

[25] Van-Reeth, K., Nicoll, A. 2009. A human case of swine influenza virus infection in Europe-implication for human health and research. Euro Surveill 14, 19124.

[26] Webby, R.J., Webster, R.G. 2001. Emergence of influenza A viruses. Philos Trans R Soc Lond B Biol Sci 356, 1817–1828.

[27] Webster, R.G., Peiris, M., Chen, H., Guan, Y. 2006. H5N1 outbreaks and enzootic influenza. Emerg Infect Dis 12, 3–8.

[28] Zambon, M.C. 2001. The pathogenesis of influenza in human. Rev Med Virol 11, 227–241.