Virus Genes (2011) 43:27–32
DOI 10.1007/s11262-011-0616-6

Amino acid sequence analysis and identification of mutations in the NS gene of 2009 influenza A (H1N1) isolates from Kenya

Gachara George · Symekher Samuel · Mbithi John · Simwa James · Ng'ayo Musa · Magana Japheth · Bulimo Wallace

Received: 23 December 2010 / Accepted: 18 April 2011 / Published online: 3 May 2011

Abstract Although the important role of the nonstructural (NS) gene of influenza A virus in virulence and replication is well-established, the knowledge about the extent of variation in the NS gene of 2009 influenza A (H1N1) viruses in Kenya and Africa is scanty. This study analysed the NS gene of 31 isolates from Kenya in order to obtain a more detailed knowledge about the genetic variation of NS gene of 2009 influenza A (H1N1) isolates from Kenya. A comparison with the vaccine strain and viruses isolated elsewhere in Africa was also made. The amino acid sequences of the non-structural protein, NS1 of the viruses from this study and the vaccine strain revealed 18 differences. Conversely, the nuclear export protein (NEP) of the isolates in this study had 11 differences from the vaccine strain. Analysis of the NS1 protein showed only one fixed amino acid change I123V which is one of the characteristics of clade 7 viruses. In the NEP, the amino acid at position 77 was the most mutable with 9 (39%) of all mutations seen in this protein. A mutation A115T which is a characteristic of clade 5 viruses was noted in the isolates from Lagos, Nigeria. The study shows a substantial number of mutations in the NS gene that has not been reported elsewhere and gives a glimpse of the evolution of this gene in the region.

Keywords Influenza virus · Nonstructural gene · H1N1 · Mutations

G. George (✉) · M. John
Department of Medical Laboratory Sciences, Kenyatta University, P.O. Box 43844, Nairobi 00100, Kenya
e-mail: ggachara@gmail.com

G. George · S. Samuel · S. James · M. Japheth · B. Wallace
National Influenza Centre, CVR, KEMRI, P.O. Box 54840, Nairobi 00200, Kenya

N. Musa
Centre for Microbiology Research, KEMRI, P.O. Box 54840, Nairobi 00200, Kenya

The nucleotide sequence data obtained in this study has been submitted to the GenBank database and are available under accession numbers; HM855241–HM855264 and HQ165789–HQ165795.

The eighth vRNA segment of the influenza A virus directs the synthesis of two mRNAs. The first of these encodes the non-structural (NS) protein, NS1, while the other is derived from splicing of the NS1 mRNA is translated into a protein that localizes to the cell nucleus and which was originally named NS2 [1, 2] but has recently been renamed the nuclear export protein (NEP).

The NS1 protein is translated directly from the mRNA [3] and consists of 124–237 amino acids (aa), depending on the virus strain [4, 5]. In the currently circulating 2009 pandemic H1N1 virus, the protein is only 219 aa long, though viruses with 230 aa-long NS1 have been isolated [6]. The NS1 protein contains two functional domains: the N-terminal RNA-binding domain (residues 1–73) and the C-terminal effector domain (residues 73–237) [7]. Although NS1 is a multifunctional protein, one of its main functions is to suppress type I IFN production by the host [6].

The NS2/NEP protein is translated from the NS2/NEP mRNA into a 121 aa-long protein [3, 8]. The protein may promote the formation of a stable export complex of new viral RNP [9]. In association with the matrix protein 1 (M1), it interacts with cellular export factor (CEF1) and mediates the nuclear export of viral ribonucleoprotein

(vRNP) complexes by connecting the cellular export machinery with vRNPs [9].

Although the important role of the non-structural (NS) gene of influenza A in virulence and replication of the virus is well-established, the knowledge about the extent of variation in the NS gene of 2009 influenza A (H1N1) viruses in Kenya and Africa has not been described. This study analysed the NS gene of 31 isolates from Kenya in order to obtain a more detailed knowledge about its phylogeny and genetic variation. It also compared the mutations found in the NS1 and NEP from strains isolated in Africa with the vaccine strain.

Nasopharyngeal swabs in 2SP (sucrose–phosphate) virus transport medium were collected from patients who met the case definition criteria for pandemic flu developed by the Ministry of public health. The detection of pandemic influenza A(H1N1) 2009 viruses was done using real-time one step RT-PCR conducted using protocols designed and distributed by WHO/USCDC for detection of influenza A and swine influenza A viruses[10]. The positive samples were then cultured in MDCK cells for virus isolation. 31 samples representing the country's provinces were selected and their RNA extracted using a commercial kit QIAamp® RNA extraction kit (Qiagen, Germany) following the manufacturers protocol. The full length NS gene was amplified using the one-step RT-PCR, using Superscript III One-step RT-PCR kit (Invitrogen, UK) with M13 primers. Amplicons were treated with shrimp alkaline phosphatase–exonuclease I (ExoSapI) (U. S Biological, Swampscott, MA, USA) and sequenced directly.

Contigs assembly and determination of open reading frames was done using BIOEDIT program [11]. Nucleotide sequences, prediction of amino acid sequences, alignments, pairwise, number of amino acid differences and phylogenetic tree construction were completed using MEGA® software version 4.1[12]. The sequences were compared with the vaccine strain and those from other African countries already deposited in Genbank.

We analysed the NS gene sequences of 31 pandemic H1N1 influenza A virus strains isolated in this study together with the vaccine strain and the available sequences from Africa in Genbank. The Kenyan isolates represent samples from the first detection in July 2009 to December 2009 and were derived from all provinces in the country. Only four African countries had sequences in Genbank namely Nigeria ($n = 3$), Senegal ($n = 4$), Ethiopia ($n = 2$), and Mali ($n = 1$). All viruses isolated in this study had 890 nucleotides and there were no deletions or insertions. The NS1 protein had 219 amino acids while the NEP had 121 amino acids. Phylogenetic analysis showed that there are distinct clusters of the virus suggesting co-circulation of multiple sub-lineages. The viruses, however, clustered together irrespective of their geographical origin and isolation date suggesting separate introductions of the virus in the country. Viruses from the other African countries formed distinct clusters (Fig. 1). This is possibly due to the small number of sequences from these countries.

In regard to genetic variation between the isolates from this study and the vaccine strain, there were substitutions at 18 and 11 amino acid positions in the NS1 and NS2/NEP proteins, respectively (Table 1). Analysis of sequence identity and variable sites of the NS coding sequence, NS1 gene, and NS2/NEP gene indicate the genes are highly conserved with an average identity of more than 98% (Data not shown). The NS1 proteins were found to be more conserved than the NS2/NEP proteins. The only fixed amino acid substitution occurred at position 123 in the NS1. This occurred as a result of a non-synonymous substitution at the 367–369 codon from ATC to GTC causing

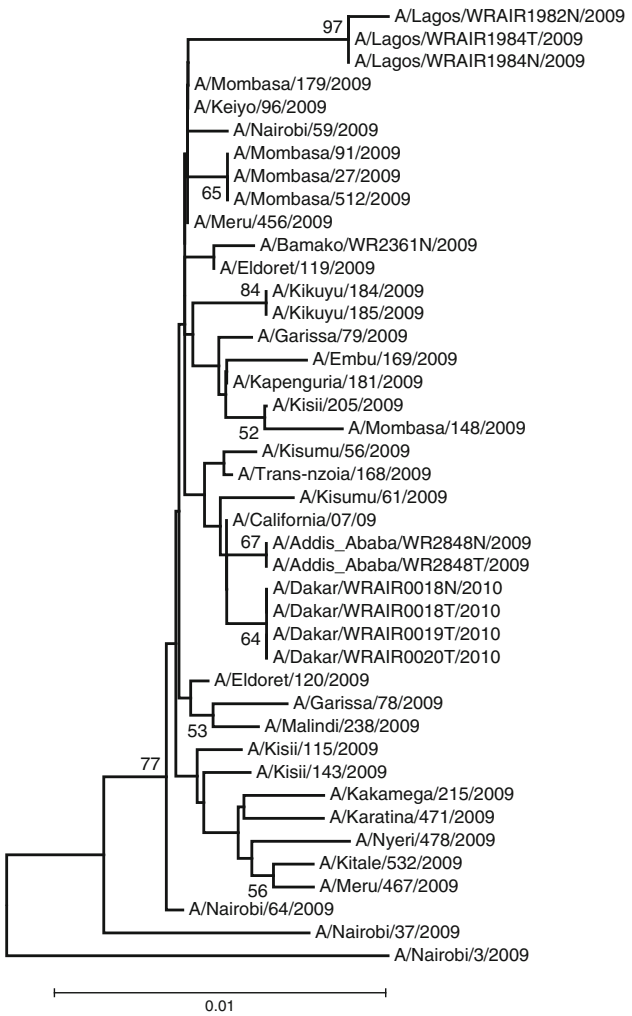

Fig. 1 Phylogenetic relationship of the NS nucleotide coding region of influenza viruses used in this study. The tree was generated by neighbor-joining analysis with Tamura-Nei model, using MEGA 4.0. *Numbers* below key nodes indicate the percentage of bootstrap values of 2,000 replicates

Table 1 Amino acid variation of NS1 and NEP proteins of pandemic H1N1 influenza A virus strains from Kenya compared with the vaccine strain and other African strains

Virus isolate	Amino acid at indicated position in NS1 protein																						
	3	5	21	25	27	29	44	46	48	49	53	55	64	74	91	93	123	125	157	162	180	185	205
A/California/07/09	S	T	R	N	L	D	K	R	N	T	D	E	I	S	S	M	I	E	V	P	V	L	N
A/Kisumu/56/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Kisumu/61/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Garissa/78/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Nairobi/59/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Eldoret/119/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Eldoret/120/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Embu/169/2009	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	L	.	.	.
A/Garissa/79/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	L	.	.	.
A/Kakamega/215/2009	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Kapenguria/181/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	L	.	.	.
A/Karatina/471/2009	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	I	V	.	.	.	.	.	.
A/Keiyo/96/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Kikuyu/184/2009	.	.	.	.	.	.	.	.	.	.	.	K	.	.	.	.	V	.	.	.	.	.	.
A/Kikuyu/185/2009	.	.	.	.	.	.	.	.	.	.	.	K	.	.	.	.	V	.	.	.	.	.	.
A/Kisii/115/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Kisii/143/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Kisii/205/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	L	.	F	.
A/Kitale/532/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Malindi/238/2009	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Meru/456/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Meru/467/2009	.	I	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Mombasa/148/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	L	.	F	.
A/Mombasa/179/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Mombasa/27/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	D
A/Mombasa/512/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	D
A/Mombasa/91/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	D
A/Nairobi/3/2009	.	.	.	.	F	V	R	G	P	P	.	.	.	.	.	.	V	.	I	.	.	.	.
A/Nairobi/37/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	I	.	.	.	.
A/Nairobi/64/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Nyeri/478/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Trans-nzoia/168/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.
A/Addis Ababa/WR2848N/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Addis Ababa/WR2848T/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/Bamako/WR2361N/2009	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	G	.	.	.	.	.
A/Lagos/WRAIR1982N/2009	.	.	.	.	.	.	.	.	.	.	N	.	.	N	F	.	V	.	.	.	.	.	.
A/Lagos/WRAIR1984N/2009	.	.	.	.	.	.	.	.	.	.	N	.	.	N	.	.	V	.	.	.	.	.	.
A/Lagos/WRAIR1984T/2009	.	.	.	.	.	.	.	.	.	.	N	.	.	N	.	.	V	.	.	.	.	.	.
A/Dakar/WRAIR0018N/2010	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.
A/Dakar/WRAIR0018T/2010	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.
A/Dakar/WRAIR0019T/2010	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.
A/Dakar/WRAIR0020T/2010	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.

Table 1 continued

Virus isolate	Amino acid at indicated position in NEP protein											
	3	5	77	78	87	98	103	108	114	115	118	120
A/California/07/09	S	T	R	W	L	T	L	E	R	A	F	L
A/Kisumu/56/2009	.	.	.	.	.	.	.	.	.	.	.	R
A/Kisumu/61/2009	.	.	K	.	.	.	.	.	.	.	.	.
A/Garissa/78/2009	.	.	K	.	S	.	.	.	G	.	.	.
A/Nairobi/59/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Eldoret/119/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Eldoret/120/2009	.	.	K	.	.	.	.	.	.	.	.	.
A/Embu/169/2009	T	.	.	.	.	.	.	.	.	.	.	.
A/Garissa/79/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Kakamega/215/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Kapenguria/181/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Karatina/471/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Keiyo/96/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Kikuyu/184/2009	.	.	.	.	.	.	.	D	.	.	.	.
A/Kikuyu/185/2009	.	.	.	.	.	.	.	D	.	.	.	.
A/Kisii/115/2009	.	.	M	.	.	.	.	.	.	.	.	.
A/Kisii/143/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Kisii/205/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Kitale/532/2009	.	.	M	.	.	.	.	.	.	.	.	.
A/Malindi/238/2009	.	.	.	.	.	.	.	.	G	.	.	.
A/Meru/456/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Meru/467/2009	.	I	M	.	.	.	.	.	.	.	.	.
A/Mombasa/148/2009	.	.	.	.	.	.	.	D	.	.	.	.
A/Mombasa/179/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/27/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/512/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Mombasa/91/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Nairobi/3/2009	.	.	M	.	.	.	.	.	.	.	.	.
A/Nairobi/37/2009	.	.	.	G	.	K	V	.	.	.	Y	.
A/Nairobi/64/2009	.	.	M	.	.	.	.	.	.	.	.	.
A/Nyeri/478/2009	.	.	K	.	.	.	.	.	.	.	.	R
A/Trans-nzoia/168/2009	.	.	.	.	.	.	.	.	.	.	.	R
A/Addis Ababa/WR2848N/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Addis Ababa/WR2848T/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Bamako/WR2361N/2009	.	.	.	.	.	.	.	.	.	.	.	.
A/Lagos/WRAIR1982N/2009	.	.	.	.	.	.	.	.	.	T	.	.
A/Lagos/WRAIR1984N/2009	.	.	.	.	.	.	.	.	.	T	.	.
A/Lagos/WRAIR1984T/2009	.	.	.	.	.	.	.	.	.	T	.	.
A/Dakar/WRAIR0018N/2010	.	.	.	.	.	.	.	.	.	.	.	.
A/Dakar/WRAIR0018T/2010	.	.	.	.	.	.	.	.	.	.	.	.
A/Dakar/WRAIR0019T/2010	.	.	.	.	.	.	.	.	.	.	.	.
A/Dakar/WRAIR0020T/2010	.	.	.	.	.	.	.	.	.	.	.	.

an isoleucine to valine substitution. The only two isolates lacking this mutation A/Kisumu/56/2009 and A/Kisumu/61/2009 represent the first two introductions of the pandemic in the country. The second introductions of the virus in the country represented by A/Garissa/78/2009 and A/Nairobi/64/2009 had this I123V mutation. The further confirms the separate introductions of the virus in the country.

Analysis of the NEP/NS2 does not reveal any fixed amino acid change. The amino acid at position 77 was the most mutable with 9 (39%) of all mutations seen in this protein occurring at this position resulting in R77M and R77K mutants. This change was due to two non-synonymous substitutions at the 229–231 codon. Regarding variability of the sequences from other African countries, the Nigerian isolates had three unique mutations D53N and S74N in the NS1 and A115T in the NEP while the Senegalese isolates were unique in containing isoleucine at position 123 while possessing an V180I mutation in the NS1 protein.

It is known that the virulence of Influenza A virus is a multigenic trait, and determinants may differ among animal species. Recent findings on the 2009 influenza A H1N1 have indicated that a single amino acid substitution in the NS1 protein do not obviously alter the pathogenicity of the virus, but co-mutation of two amino acid residues increase virus pathogenicity in mice [13]. Molecular analysis is therefore important for the monitoring of modifications in virus genome related to pathogenesis and susceptibility to antiviral drugs.

In the NS1 protein, the mutation D92E known for high virulence in human is absent in all H1N1 pdm viruses, as well as the Kenyan isolates. The PDZ ligand domain at the C-terminus of NS1 implicated in pathogenicity of the 1918 H1N1 virus is also absent since all the Kenyan isolates had a truncated NS1 due to the presence of a stop codon at position 220. Analysis of the NS1 protein shows only one fixed amino acid change I123V which is one of the characteristics of clade 7 viruses that have dominated since July 2009[14]. It has also been noted previously that there were four introductions of the virus in the country [15]. Of these introductions, 2 (50%) of the isolates had this change indicating that it is possible that two clades of the virus were introduced in the country. Other changes noted in more than one isolate were P162L, L185F, and N205D which are all in the effector domain of the virus. Among the strains from other African countries, the lack of the I123V mutation among the Ethiopian and Senegalese strains is unique. This is so considering that these isolates were collected in December 2009, January and February 2010, respectively, in the second wave of the pandemic during which time this mutation was dominant worldwide. Since Senegal reported its first cases around this time, this may be presumably due to founder effect. However, we cannot account for the lack of this mutation in the Ethiopian strains 6 months after her first case.

The R77M and R77K change in the NEP is noted to be too close to Trp 78 which mediates the binding of NEP and M1 to allow nuclear export of viral ribonucleoproteins (vRNPs) replicated in the nucleus to the cytoplasm. However, as shown previously mutations at this position may not affect the ability of NEP to regulate viral RNA levels [16]. Three other amino acid changes in this protein were noted in two or more Kenyan isolates namely E108D, R114G, and L120R. These mutations in the NEP have not been described before elsewhere. A mutation A115T which is a characteristic of clade 5 viruses [17] was noted in the isolates from Lagos, Nigeria.

Sequencing of the HA and NA genes is currently ongoing and has been completed for 15 of the isolates. In the HA gene, the following mutations have become fixed namely; S220T, P100S, and I338V. In the NA gene V106I and N248D have become fixed. All these clade-specific mutations have been described previously [14] and represents the amino acid changes characterizing clade 4, 5, 6, and 7 viruses.

In conclusion, this is the first documented study that shows the evolution of the NS gene of the 2009 influenza A (H1N1) virus in Africa. The study has shown a substantial number of mutations in the NS gene that has not been reported before. These results provide useful molecular data to give a glimpse of the evolution of this gene in the region. A study of the other genes is ongoing to understand better the molecular epidemiology of this virus as influenza virus virulence in polygenic.

Acknowledgments We acknowledge the Director of Kenya Medical Research Institute (KEMRI), Director of Center for Virus Research (CVR), and National Influenza Centre staff for facilitating this study. We wish to also acknowledge Naomi Komadina of the World Health Organization Collaborating Centre for Reference and Research in Influenza, Melbourne, Australia for her guidance while carrying this study.

References

1. S.C. Inglis, T. Barrett, C.M. Brown, J.W. Almond, Proc. Natl. Acad. Sci. USA **76**, 3790 (1979)
2. R.A. Lamb, P.W. Choppin, Proc. Natl. Acad. Sci. USA **76**, 4908 (1979)
3. A.G. Porter, J.C. Smith, J.S. Emtage, Proc. Natl. Acad. Sci. USA **77**, 5074 (1980)
4. D.L. Suarez, M.L. Perdue, Virus Res. **54**, 59 (1998)
5. S. Wang, W.M. Shi, A. Mweene, H.L. Wei, G.R. Bai, J.H. Liu, Virus Genes **31**, 329 (2005)
6. B.G. Hale, R.E. Randall, J. Ortin, D. Jackson, J. Gen. Virol. **89**, 2359 (2008)
7. X.Y. Qian, F. Alonso-Caplen, R.M. Krug, J. Virol. **68**, 2433 (1994)

8. P. Palese, M. Shaw, in *Fields Virology,* 5th edn. ed. by D.M. Knipe, P.M. Howley (Lippincott Williams & Wilkins, Philadelphia, 2001) p. 1647
9. G. Neumann, M.T. Hughes, Y. Kawaoka, EMBO J. **19**, 6751 (2000)
10. CDC protocol of realtime RT-PCR for influenza A (H1N1) (2010), http://www.who.int/csr/resources/publications/swineflu/realtimeptpcr/en/index.html. Accessed 9 Dec 2010
11. T.A. Hall, Nucl. Acids. Symp. Ser **41**, 95 (1999)
12. K. Tamura, J. Dudley, M. Nei, S. Kumar, Mol. Biol. Evol. **24**, 1596 (2007)
13. J. Pu, J. Wang, Y. Zhang, G. Fu, Y. Bi, Y. Sun, J. Liu, Virus Res. **151**(2), 200 (2010)
14. M. Nelson, D. Spiro, D.Wentworth, E. Beck, J. Fan, E. Ghedin, R. Halpin, J. Bera, E. Hine, K. Proudfoot, T. Stockwell, X. Lin, S. Griesemer, S. Kumar, M.Bose, C. Viboud, E. Holmes, K. Henrickson, PLoS Curr. (2009). doi:10.1371/currents.RRN1126
15. C. Tabu, S. Sharif, P. Okoth, J. Kioko, C. Nzioka, P. Muthoka, M. Ope, S. Makama, R. Kalani et al., MMWR Morb Mortal Wkly Rep. **58**(41), 1143 (2009)
16. N.C. Robb, M. Smith, F.T. Vreede, E. Fodor, J. Gen. Virol. **90**, 1398 (2009)
17. V.A. Potdar, M.S. Chadha, S.M. Jadhav, J. Mullick, S.S. Cherian, A.C. Mishra, PLoS ONE. (2010). doi:10.1371/journal.pone.0009693