Molecular Influenza Surveillance with the FluSurver in GISAID

Sebastian Maurer-Stroh*, Catherine Smith*, Alan Hay

*Contact: flusurver@gisaid.org

Bioinformatics Institute (BII), Agency for Science Technology & Research, Singapore
GISAID Database Technical Group (DTC) and Scientific Advisory Council (SAC), Germany

1. Background

Interest in new outbreaks as well as regular surveillance of circulating seasonal strains produce a constant flow of influenza sequences that need to be analyzed and interpreted for epidemiological and phenotypic features. Several steps in typical influenza sequence analysis can be automated and we have been actively developing the free online analysis pipeline FluSurver over the last 4 years to facilitate identification and interpretation of mutations in influenza sequences.

2. Materials and Methods

The FluSurver is accessible as webserver (http://flusurver.bii.a-star.edu.sg/) or directly from GISAID (beta) where users can upload their sequences for fully automated analysis which includes ultra-fast database searches with TACHYON, alignment with MAFFT, structural modeling with MODELLER, structure viewing with JMOL and several PERL scripts to link mutations to in-house derived databases. These include geographic and temporal frequency of occurrence as well as co-occurrence of mutations, curated literature annotations for >250 known mutation effects such as drug resistance, host receptor specificity, virulence, antigenic drift and antibody escape mutants. We also show the position of the mutation(s) in structural models and highlight if mutations are close to common drug, host receptor or antibody binding sites or if a glycosylation motif is lost or created through a mutation.

Notable recent additions/changes are: updated reference strains, ultra-fast database searches using TACHYON, added city-level detail to map view, passage history bias information for selected mutations, integration into GISAID (beta), expansion of help and tutorial sections for guidance of using the results in publications and to avoid over-interpretation.

3. Results

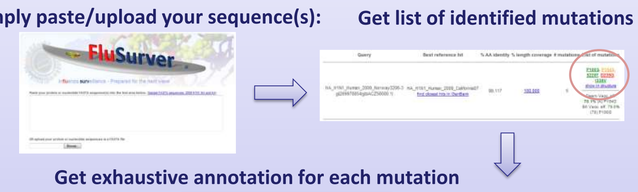
The FluSurver has already been instrumental in the discovery of new influenza strain variants with altered antiviral susceptibility, host specificity, glycosylation and antigenic properties. To showcase the usefulness and speed of analysis made possible with our tool, we use the recent H7N9 outbreak sequences deposited in GISAID as new case study. After submitting the sequences of the first 3 human cases, it only takes a few seconds to identify the critical HA-Q226L and PB2-E627K mutations that are known adaptations facilitating infection of mammalian hosts. Further browsing the alignment results identifies absence of a multi-basic HA cleavage site suggesting low pathogenicity in birds, a small stalk deletion in NA of unclear significance, common resistance to amantadines through M2-S31N, normal neuraminidase drug-sensitivity at NA-H274, one isolated strain with NA-R292K which is associated with drug-treatment induced resistance in closely related avian N9 strains. Another common feature of the early H7N9 outbreak sequences is a truncation of NS1 through an early stop codon resulting in removal of the PDZ-binding motif which is normally used for binding and interfering with host proteins. When including the avian and environmental isolates in the analysis, one can see that while the closest avian precursor strains from 2011-2012 did not have the Q226L host specificity mutation, surprisingly the recent outbreak-linked avian samples already had this mutation which must therefore also be fit to circulate in birds to some extent. At the same time, PB2-E627K was only found in human samples suggesting a stepwise acquisition of host adaptation factors. The FluSurver analysis also highlights unusually diverse strains, e.g. A/Shanghai/1/2013 compared to A/Anhui/1/2013 and A/Shanghai/2/2013, which would indicate multiple introduction events from birds to humans instead of sustained human to human transmission in the beginning of the outbreak.

4. Conclusions

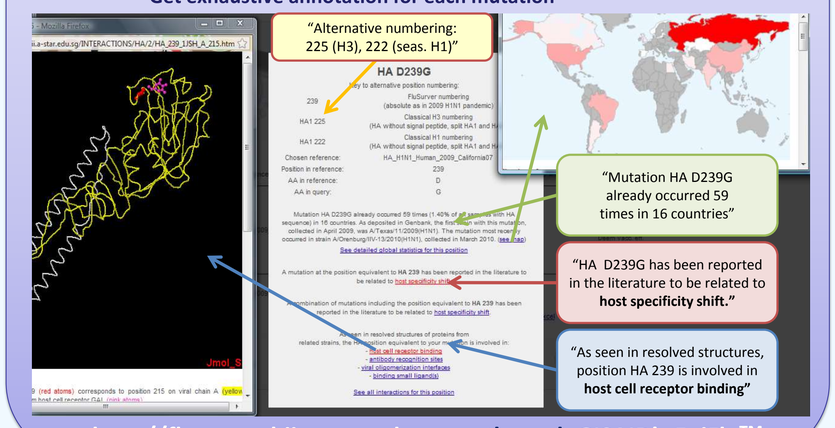
The FluSurver allows researchers, clinician scientists and surveillance labs to rapidly screen their influenza sequences for potentially interesting mutations to identify candidates for phenotypic changes or special epidemiological relevance. In the case of H7N9, it can be used for a solid initial characterization as well as to continue monitoring if additional human host adaptation mutations would occur in the future. The full potential of FluSurver can be realized through the integration with the GISAID database.

FluSurver Basics

Simply paste/upload your sequence(s): → Get list of identified mutations

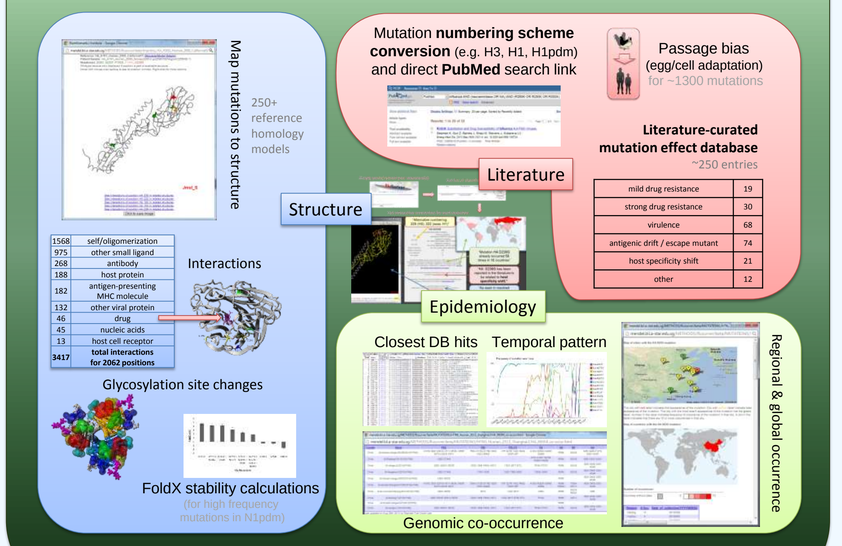


Get exhaustive annotation for each mutation



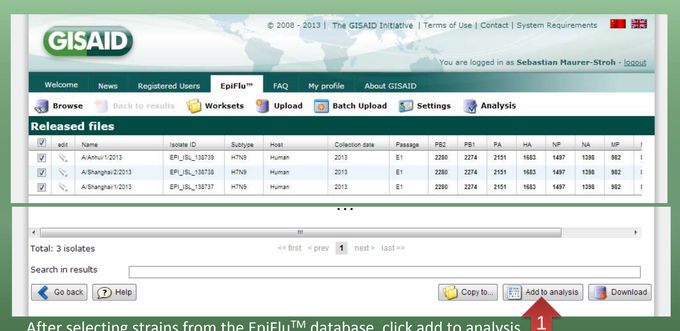
http://flusurver.bii.a-star.edu.sg or through GISAID's EpiFlu™

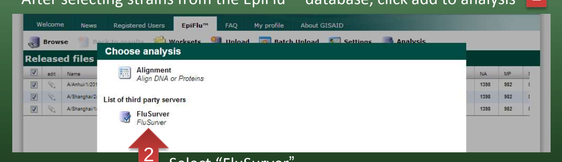
Summary of FluSurver features 2013



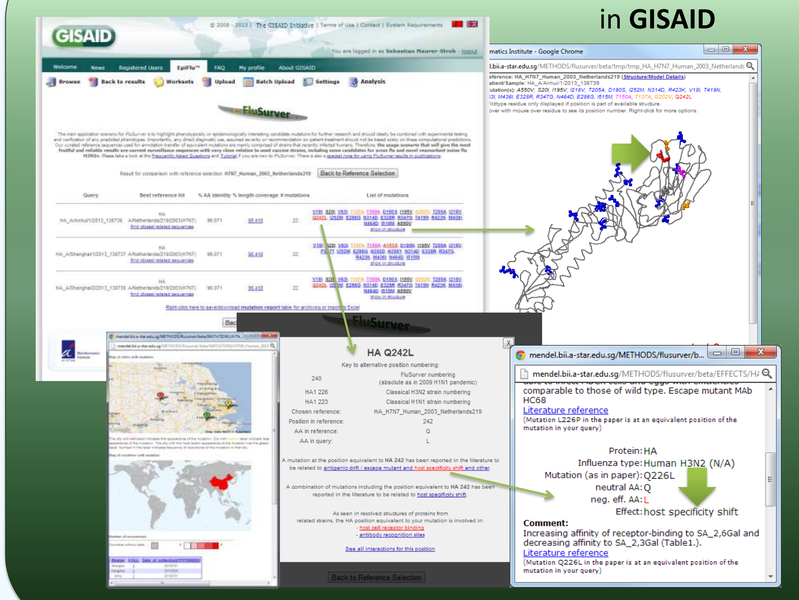
http://flusurver.bii.a-star.edu.sg or through GISAID's EpiFlu™

NEW (beta test): Analyze sequences with FluSurver directly from the **GISAID** platform





H7N9 mutation analysis example with FluSurver in **GISAID**



First 3 H7N9 strains as seen in FluSurver: Summary of major host specificity mutations

Protein-Position(s)	Anhui/1	Shanghai/2	Shanghai/1	Role
HA-226	L	L	Q	Receptor-binding
HA-186	V	V	G	Receptor-binding
HA-138	A	A	S	Receptor-binding
PB2-627	K	K	K	Replication-efficiency

- H7N9 is presumably a **low pathogenic** strain **in birds** as all 3 HAs lack a multibasic cleavage site

```
Query: 301 LPFQNIDSRAVGKCPRYVKQRSLLLATGMKNVPEIPK--GRGLFGAIAGFIENGWEGLID 358
           LPFQNI+SRAVGKCPRYVKQ SLLLATGMKNVPEIPK   RGLFGAIAGFIENGWEGLID
Sbjct: 308 LPFQNIHSRAVGKCPRYVKQESLLLATGMKNVPEIPKRRRRGLFGAIAGFIENGWEGLID 367
```

Low patho (e.g. H7N9) — Single protease cleavage site, localized infection

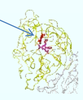
High patho (e.g. H7N7) — Multi protease cleavage site, systemic infection

We acknowledge the authors, originating and submitting laboratories of the sequences from GISAID's EpiFlu™ Database on which this research is based. The list is detailed below.
All submitters of data may be contacted directly via the GISAID website www.gisaid.org

Segment ID	Segment	Country	Collection date	Isolate name	Submitting Lab
EPI439488	PB2	China	2013-Feb-26	A/Shanghai/1/2013	WHO Chinese National Influenza Center
EPI439486	HA	China	2013-Feb-26	A/Shanghai/1/2013	WHO Chinese National Influenza Center
EPI439487	NA	China	2013-Feb-26	A/Shanghai/1/2013	WHO Chinese National Influenza Center
EPI439495	PB2	China	2013-Mar-05	A/Shanghai/2/2013	WHO Chinese National Influenza Center
EPI439500	NA	China	2013-Mar-05	A/Shanghai/2/2013	WHO Chinese National Influenza Center
EPI439502	HA	China	2013-Mar-05	A/Shanghai/2/2013	WHO Chinese National Influenza Center
EPI439504	PB2	China	2013-Mar-20	A/Anhui/1/2013	WHO Chinese National Influenza Center
EPI439507	HA	China	2013-Mar-20	A/Anhui/1/2013	WHO Chinese National Influenza Center
EPI439509	NA	China	2013-Mar-20	A/Anhui/1/2013	WHO Chinese National Influenza Center

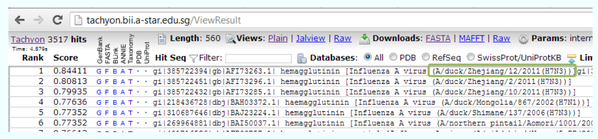
First 3 H7N9 strains as seen in FluSurver: Summary of critical drug sensitivity positions

Protein-Position(s)	Anhui/1	Shanghai/2	Shanghai/1	Role
NA-274	H	H	H	Drug sensitivity (Tamiflu)
NA-292	R	R	K	Drug sensitivity (Tamiflu and Relenza)
NA-58-60	del58-60	del58-60	del58-60	Stalk deletion
M2-31	N	N	N	Drug sensitivity (Amantadine)



Geographic distribution of strains with identified mutations:

Find closest database hits to trace back reassortment events

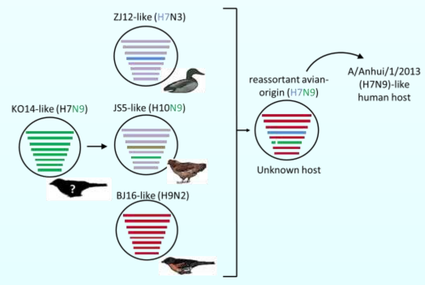


Tachyon is an ultra-fast protein search engine against large databases and can be accessed from inside FluSurver by clicking on "find closest related sequences".

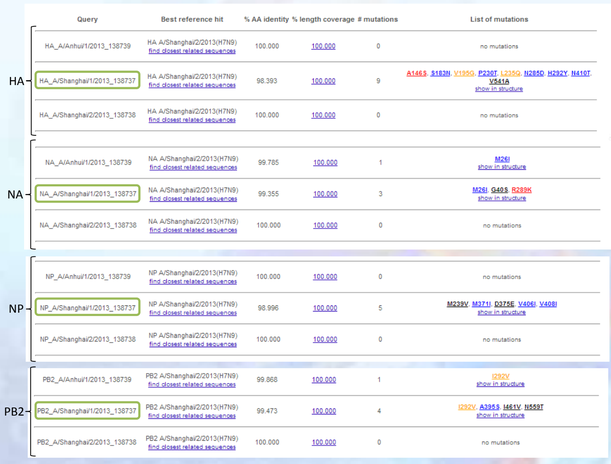
The consensus of best database hits shows the likely origin of the individual segments:
HA = H7N3
NA = N9…H11N9/H7N9/**H10N9**/H9N9?
All other = H9N2

BII's Tachyon takes a few seconds while Bayesian phylogenetic trees take hours to days…

Host and lineage origins of H7N9



Monitor any new mutations or deviating strains in the outbreak

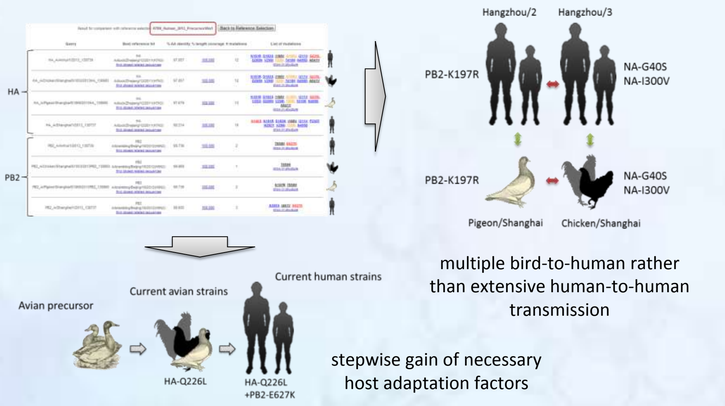


Shanghai/1 is very different from the other two strains in multiple segments!

An independent bird-to-human transmission?

Are there multiple H7N9 reassortant variants?

Compare avian and human samples



multiple bird-to-human rather than extensive human-to-human transmission

stepwise gain of necessary host adaptation factors

We acknowledge the authors, originating and submitting laboratories of the sequences from GISAID's EpiFlu™ Database on which this research is based. The list is detailed below.
All submitters of data may be contacted directly via the GISAID website www.gisaid.org

Segment ID	Segment	Country	Collection date	Isolate name	Originating Lab	Submitting Lab	Authors
EPI440682	PB2	China	2013-Apr-03	A/Chicken/Shanghai/S1053/2013	Harbin Veterinary Research Institute	Harbin Veterinary Research Institute	
EPI440685	HA	China	2013-Apr-03	A/Chicken/Shanghai/S1053/2013	Harbin Veterinary Research Institute	Harbin Veterinary Research Institute	
EPI440698	PB2	China	2013-Apr-02	A/Pigeon/Shanghai/S1069/2013	Harbin Veterinary Research Institute	Harbin Veterinary Research Institute	
EPI440701	HA	China	2013-Apr-02	A/Pigeon/Shanghai/S1069/2013	Harbin Veterinary Research Institute	Harbin Veterinary Research Institute	
EPI442710	HA	China	2013-Mar-25	A/Hangzhou/2/2013	Hangzhou Center for Disease Control and Prevention	Hangzhou Center for Disease Control and Prevention	Jing-Cao, Pan
EPI442709	PB2	China	2013-Mar-25	A/Hangzhou/2/2013	Hangzhou Center for Disease Control and Prevention	Hangzhou Center for Disease Control and Prevention	Jing-Cao, Pan
EPI442713	HA	China	2013-Apr-02	A/Hangzhou/3/2013	Hangzhou Center for Disease Control and Prevention	Hangzhou Center for Disease Control and Prevention	Jing-Cao, Pan
EPI442712	PB2	China	2013-Apr-02	A/Hangzhou/3/2013	Hangzhou Center for Disease Control and Prevention	Hangzhou Center for Disease Control and Prevention	Jing-Cao, Pan

Collaborative influenza research where our FluSurver has already been useful:

Guarnaccia T, Carolan LA, Maurer-Stroh S, Lee RT, Job E, Reading PC, Petrie S, McCaw JM, McVernon J, Hurt AC, Kelso A, Mosse J, Barr IG, Laurie KL. *Antigenic Drift of the Pandemic 2009 A(H1N1) Influenza Virus in a Ferret Model.* PLoS Pathog. 2013 May;9(5):e1003354. doi: 10.1371/journal.ppat.1003354. Epub 2013 May 9. PubMed PMID: 23671418.

Maurer-Stroh S, Lee RT, Gunalan V, Eisenhaber F. *The highly pathogenic H7N3 avian influenza strain from July 2012 in Mexico acquired an extended cleavage site through recombination with host 28S rRNA.* Virol J. 2013 May 1;10:139. doi: 10.1186/1743-422X-10-139. PubMed PMID: 23635025;

Job ER, Deng YM, Barfod KK, Tate MD, Caldwell N, Reddiex S, Maurer-Stroh S, Brooks AG, Reading PC. *Addition of Glycosylation to Influenza A Virus Hemagglutinin Modulates Antibody-Mediated Recognition of H1N1 2009 Pandemic Viruses.* J Immunol. 2013 Mar 1;190(5):2169-77. doi: 10.4049/jimmunol.1202433. Epub 2013 Jan 30. PubMed PMID: 23365085.

Simões EA, Patel C, Sung WK, Lee CW, Loh KH, Lucero M, Nohynek H, Nai G, Thien PL, Koh CW, Chan YS, Ma J, Maurer-Stroh S, Carosone-Link P, Hibberd ML, Wong CW; IVAC Consortium. *Pathogen chip for respiratory tract infections.* J Clin Microbiol. 2013 Mar;51(3):945-53. doi: 10.1128/JCM.02317-12. Epub 2013 Jan 9. PubMed PMID: 23303493.

Kuznetsov V, Lee HK, Maurer-Stroh S, Molnar MJ, Pongor S, Eisenhaber B and Eisenhaber F. *How bioinformatics influences health informatics: Usage of biomolecular sequences, expression profiles and automated microscopic image analyses for clinical needs and public health.* Health Information Science and Systems, 2013 1:2

Hurt AC, Hardie K, Wilson NJ, Deng YM, Osbourn M, Leang SK, Lee RT, Iannello P, Gehrig N, Shaw R, Wark P, Caldwell N, Givney RC, Xue L, Maurer-Stroh S, Dwyer DE, Wang B, Smith DW, Levy A, Booy R, Dixit R, Merritt T, Kelso A, Dalton C, Durrheim D, Barr IG. *Characteristics of a Widespread Community Cluster of H275Y Oseltamivir-Resistant A(H1N1)pdm09 Influenza in Australia.* J Infect Dis. 2012 Jul;206(2):148-57. Epub 2012 May 4. PubMed PMID: 22561367.

Hurt AC, Leang SK, Speers DJ, Barr IG, Maurer-Stroh S. *Mutations I117V and I117M and Oseltamivir Sensitivity of Pandemic (H1N1) 2009 Viruses.* Emerg Infect Dis. 2012 Jan;18(1):109-12. doi: 10.3201/eid1801.111079. PubMed PMID: 22260817.

Hurt AC, Hardie K, Wilson NJ, Deng YM, Osbourn M, Gehrig N, Kelso A. *Community transmission of oseltamivir-resistant A(H1N1)pdm09 influenza.* N Engl J Med. 2011 Dec 29;365(26):2541-2. PubMed PMID: 22204735.

Hurt AC, Lee RT, Leang SK, Cui L, Deng YM, Phuah SP, Caldwell N, Freeman K, Komadina N, Smith D, Speers D, Kelso A, Lin RT, Maurer-Stroh S, Barr IG. *Increased detection in Australia and Singapore of a novel influenza A(H1N1)2009 variant with reduced oseltamivir and zanamivir sensitivity due to a S247N neuraminidase mutation.* Euro Surveill. 2011 Jun 9;16(23). pii: 19884. PubMed PMID: 21679678.

Lee VJ, Yap J, Maurer-Stroh S, Lee RT, Eisenhaber F, Tay JK, Ting PJ, Loh JP, Wong CW, Tan BH, Koay ES, Kelly PM, Hibberd ML. *Investigation of causes of oseltamivir chemoprophylaxis failures during influenza A (H1N1-2009) outbreaks.* J Clin Virol. 2011 Feb;50(2):104-8. Epub 2010 Nov 19. PubMed PMID: 21094080.

Lee RT, Santos CL, de Paiva TM, Cui L, Sirota FL, Eisenhaber F, Maurer-Stroh S. *All that glitters is not gold--founder effects complicate associations of flu mutations to disease severity.* Virol J. 2010 Nov 1;7:297. PubMed PMID: 21040570;

Barr IG, Cui L, Komadina N, Lee RT, Lin RT, Deng Y, Caldwell N, Shaw R, Maurer-Stroh S. *A new pandemic influenza A(H1N1) genetic variant predominated in the winter 2010 influenza season in Australia, New Zealand and Singapore.* Euro Surveill. 2010 Oct 21;15(42). pii: 19692. PubMed PMID: 21034722.

Maurer-Stroh S, Paing SS, Lee RT, Eisenhaber F. *Sporadic human cases of swine-origin influenza before 2009 share the Sa epitope.* Cell Cycle. 2010 Sep 15;9(18):3826-8. Epub 2010 Sep 26. PubMed PMID: 20930526;

Lee VJ, Yap J, Cook AR, Chen MI, Tay JK, Tan BH, Loh JP, Chew SW, Koh WH, Lin R, Cui L, Lee CW, Sung WK, Wong CW, Hibberd ML, Kang WL, Seet B, Tambyah PA. *Oseltamivir ring prophylaxis for containment of 2009 H1N1 influenza outbreaks.* N Engl J Med. 2010 Jun 10;362(23):2166-74. PubMed PMID: 20558367.

Inoue M, Barkham T, Leo YS, Chan KP, Chow A, Wong CW, Tze Chuen Lee R, Maurer-Stroh S, Lin R, Lin C. *Emergence of oseltamivir-resistant pandemic (H1N1) 2009 virus within 48 hours.* Emerg Infect Dis. 2010 Oct;16(10):1633-6. PubMed PMID: 20875299.

Maurer-Stroh S, Lee RT, Eisenhaber F, Cui L, Phuah SP, Lin RT. *A new common mutation in the hemagglutinin of the 2009 (H1N1) influenza A virus.* PLoS Curr. 2010 Jun 1;2:RRN1162. PubMed PMID: 20535229.

Maurer-Stroh S, Ma J, Lee RT, Sirota FL, Eisenhaber F. *Mapping the sequence mutations of the 2009 H1N1 influenza A virus neuraminidase relative to drug and antibody binding sites.* Biol Direct. 2009 May 20;4:18; discussion 18. PubMed PMID: 19457254.