



Introduction to Mutation Interpretation with FluSurver

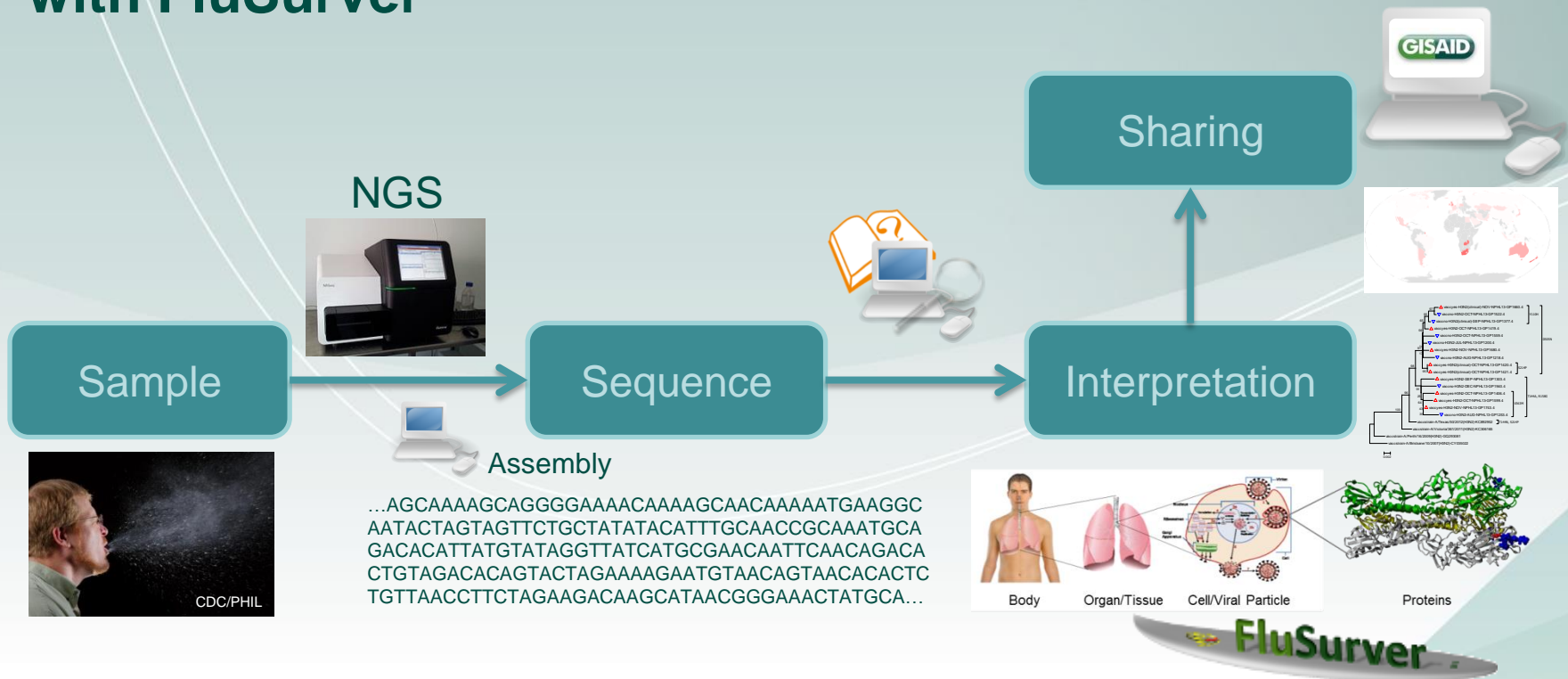
Dr. Sebastian Maurer-Stroh

*Programme Director Human Infectious Diseases,
Bioinformatics Institute, A*STAR, Singapore*

GISAID Database Technical Group



Making full use of influenza sequences with FluSurver



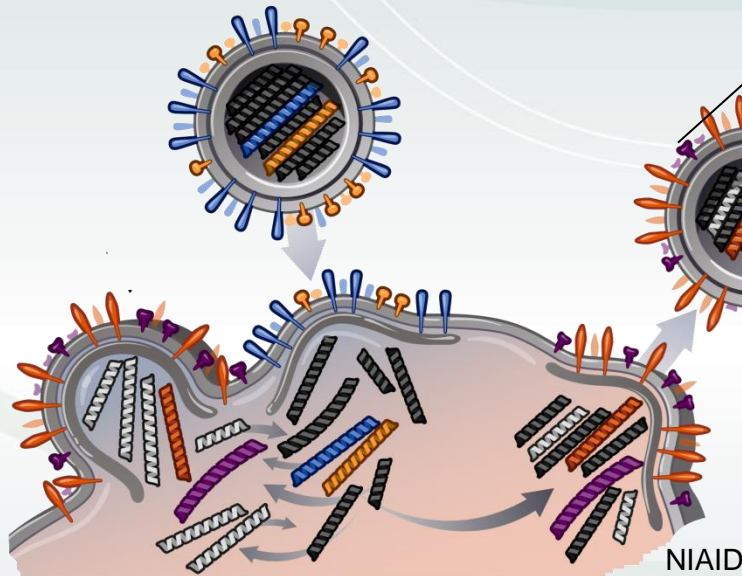
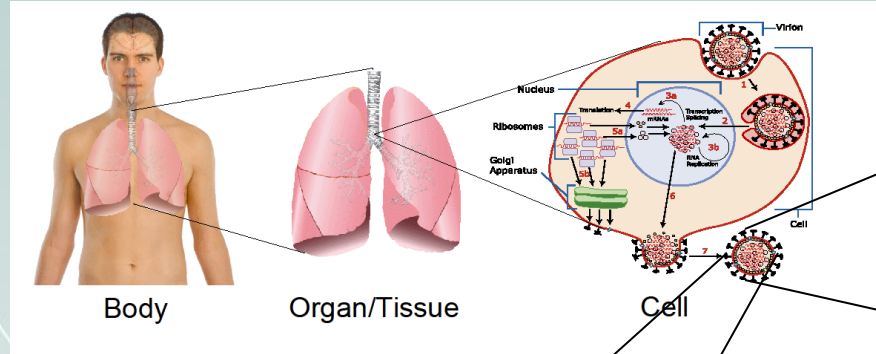
Reduced cost of and easier access to sequencing gives us:

- More sequences (also complete genomes)
- More detail of genetic evolution
- More questions on how to use/interpret/analyze sequences

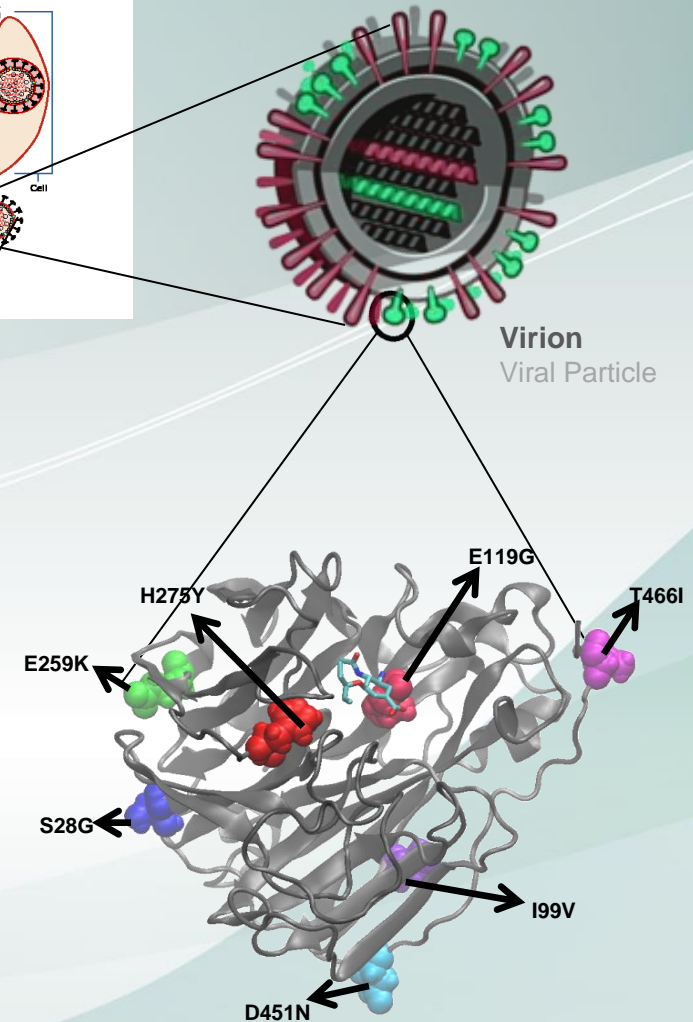
Flu viruses evolve through Reassortment and Mutations

New properties can be:

- Infect new hosts
- More or less severe
- Spread more easily
- Immune escape
- Drug resistance

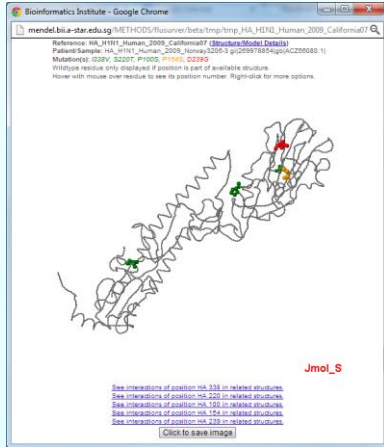


Reassortment = genome segment mixing
e.g. H1N1 to H3N1



Mutations, e.g. NA H275Y

FluSurver for Mutation Interpretation

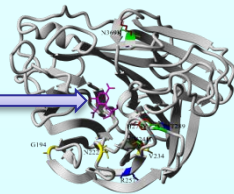


Map mutations to structure

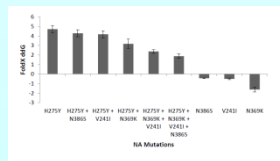
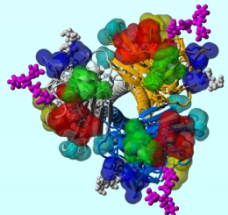
250+
reference
homology
models

1568	self/oligomerization
975	other small ligand
268	antibody
188	host protein
182	antigen-presenting MHC molecule
132	other viral protein
46	drug
45	nucleic acids
13	host cell receptor
3417	total interactions for 2062 positions

Interactions



Glycosylation site changes



FoldX stability calculations
(for high frequency
mutations in N1pdm)

Mutation numbering scheme
conversion (e.g. H3, H1, H1pdm)
and direct **PubMed** search link



Passage bias
(egg/cell adaptation)
for ~1300 mutations

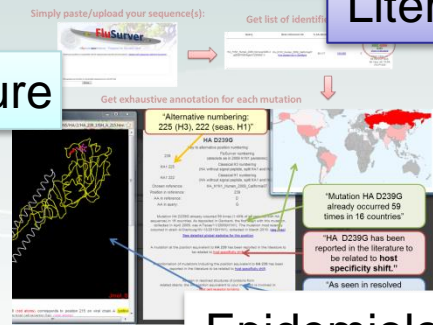
Literature-curated
mutation effect database

~250 entries

mild drug resistance	19
strong drug resistance	30
virulence	68
antigenic drift / escape mutant	74
host specificity shift	21
other	12

Literature

Structure

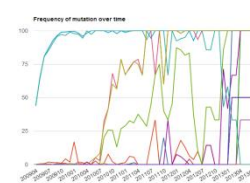


Epidemiology

Closest DB hits

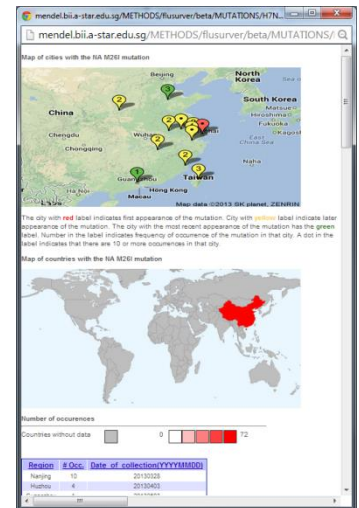
Temporal pattern

Accession	Strain	Year	Country	Region	City	Frequency
A/156/05	A/156/05	2005	China	Beijing	Beijing	100
A/222/05	A/222/05	2005	China	Beijing	Beijing	100
A/222/05	A/222/05	2005	China	Beijing	Beijing	100
A/222/05	A/222/05	2005	China	Beijing	Beijing	100
A/222/05	A/222/05	2005	China	Beijing	Beijing	100



Accession	Strain	Year	Country	Region	City	Frequency
A/156/05	A/156/05	2005	China	Beijing	Beijing	100
A/222/05	A/222/05	2005	China	Beijing	Beijing	100
A/222/05	A/222/05	2005	China	Beijing	Beijing	100
A/222/05	A/222/05	2005	China	Beijing	Beijing	100
A/222/05	A/222/05	2005	China	Beijing	Beijing	100

Genomic co-occurrence



Global occurrence

First steps: find, select and add isolates to analyze from the EpiFlu™ database

The screenshot shows the GISAID EpiFlu™ interface. At the top, there's a navigation bar with links like 'Welcome', 'News', 'Registered Users', 'EpiFlu™', 'FAQ', 'My profile', and 'About GISAID'. Below this is a secondary bar with icons and labels for 'Browse', 'Back to results', 'Worksets', 'Upload', 'Batch Upload', 'Settings', and 'Analysis'. The main section is titled 'Released files' and contains a table of isolates. The table has columns for Name, Isolate ID, Subtype, Host, Collection date, Passage, PB2, PB1, PA, HA, NP, NA, MP, and I. Three isolates are listed, all with subtype H7N9 and host Human. Below the table, there's a pagination bar showing 'Total: 3 isolates' and navigation links like '<< first', '< prev', '1', 'next >', and 'last >>'. There's also a search bar labeled 'Search in results'. At the bottom right, there are three buttons: 'Go back', 'Help', and 'Add to analysis'. A red arrow points to the 'Add to analysis' button.

<input checked="" type="checkbox"/>	edit	Name	Isolate ID	Subtype	Host	Collection date	Passage	PB2	PB1	PA	HA	NP	NA	MP	I
<input checked="" type="checkbox"/>		A/Anhui/1/2013	EPI_ISL_138739	H7N9	Human	2013	E1	2280	2274	2151	1683	1497	1398	982	1
<input checked="" type="checkbox"/>		A/Shanghai/2/2013	EPI_ISL_138738	H7N9	Human	2013	E1	2280	2274	2151	1683	1497	1398	982	1
<input checked="" type="checkbox"/>		A/Shanghai/1/2013	EPI_ISL_138737	H7N9	Human	2013	E1	2280	2274	2151	1683	1497	1398	982	1

After selecting strains on the left, click add to analysis

The screenshot shows the 'Choose analysis' dialog box. It has a title bar 'Choose analysis' and a list of options. The first option is 'Alignment' with the subtext 'Align DNA or Proteins'. Below this is a section titled 'List of third party servers' which contains the 'FluSurfer' option. A red arrow points to the 'FluSurfer' option. The background shows the same 'Released files' table as the previous screenshot.

Choose analysis

- Alignment**
Align DNA or Proteins
- List of third party servers**
 - FluSurfer**
FluSurfer

Select "FluSurfer"

GISAID

EpiFlu™ 2.0 – Analysis Tools



GISAID published: 34.593 viruses with 89.498 Sequences

Total count: 113.361 viruses with 393.536 Sequences

Welcome John Doe

[My Settings](#) [Logout](#)

- [Browse](#)
- [Upload](#)
- [Workset Management](#)
- [Administration](#)
- [Registered Users](#)

Personal Worksheet



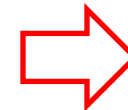
Virus Name



- [Delete Entry](#)
- [Clear List](#)
- [Select All](#)
- [Deselect All](#)

-	Name	Segment	Segment accession #	Length
<input type="checkbox"/>	A/chicken/77/Jiangxi/2014	NS	EPI1880	890
<input type="checkbox"/>	A/chicken/77/Jiangxi/2014	PB1	EPI1874	2274
<input type="checkbox"/>	A/chicken/77/Jiangxi/2014	HA	EPI1876	1686
<input type="checkbox"/>	A/chicken/77/Jiangxi/2014	NP	EPI1877	1501
<input type="checkbox"/>	A/chicken/77/Jiangxi/2014	NA	EPI1878	1413
<input type="checkbox"/>	A/chicken/77/Jiangxi/2014	PB2	EPI1873	2280
<input type="checkbox"/>	A/duck/Jiangxi/95/2014	M	EPI1887	982
<input type="checkbox"/>	A/duck/Jiangxi/95/2014	PA	EPI1883	2151
<input type="checkbox"/>	A/duck/Jiangxi/95/2014	NS	EPI1888	823
<input type="checkbox"/>	A/duck/Jiangxi/95/2014	HA	EPI1884	1704
<input type="checkbox"/>	A/duck/Jiangxi/95/2014	NP	EPI1885	1497
<input type="checkbox"/>	A/duck/Jiangxi/95/2014	PB1	EPI1882	2274
<input type="checkbox"/>	A/duck/Jiangxi/95/2014	PB2	EPI1881	2280
<input type="checkbox"/>	A/duck/Jiangxi/95/2014	NA	EPI1886	1380
<input type="checkbox"/>	A/Galicia/1786/2014	HA	EPI1907	1040
<input type="checkbox"/>	A/Hong Kong/308/2014	PB2	EPI498034	2280
<input type="checkbox"/>	A/Hong Kong/308/2014	NA	EPI498036	1401
<input type="checkbox"/>	A/Hong Kong/308/2014	PB1	EPI498035	2274
<input type="checkbox"/>	A/Hong Kong/308/2014	PA	EPI498033	2151
<input type="checkbox"/>	A/Hong Kong/308/2014	M	EPI498032	982

[Back](#) [1](#) [2](#) [3](#) ... [5](#) [Forward](#)



[Export selected](#)

[Blast Nucleotide](#)

[Blast Protein](#)

[Analyze with FluSurver](#)

[Align Sequences](#)

[View Tree](#)





The main application scenario for FluSurver is to highlight phenotypically or epidemiologically interesting candidate mutations for further research and should ideally be combined with experimental testing and verification of any predicted phenotypes. Importantly, any direct diagnostic use, assumed severity or recommendation on patient treatment should not be based solely on these computational predictions. Our curated reference sequences used for annotation transfer of equivalent mutations are mainly comprised of strains that recently infected humans. Therefore, **the usage scenario that will give the most fruitful and reliable results are current surveillance sequences with very close relation to used vaccine strains, including some candidates for avian flu and novel reassortant swine flu H3N2v.** Please take a look at the [Frequently Asked Questions](#) and [Tutorial](#) if you are new to FluSurver. There is also a [special note for using FluSurver results in publications](#).

Result for comparison with reference selection: H7N7_Human_2003_Netherlands219

[Back to Reference Selection](#)

Query	Best reference hit	% AA identity	% length coverage	# mutations	List of mutations
HA_A/Anhui/1/2013_138739	HA A/Netherlands/219/2003(H7N7) find closest related sequences	96.071	98.418	22	V18I , S20I , V63I , T137A , T150A , D190S , I195V , G202V , T205A , I218V , Q242L , I252M , E286G , N314D , E328R , R347G , T419N , R423K , M436I , N464D , I515M , A550V show in structure
HA_A/Shanghai/1/2013_138737	HA A/Netherlands/219/2003(H7N7) find closest related sequences	96.071	98.418	22	V18I , S20I , V63I , T137A , T150A , A153S , D190N , I195V , T205A , I218V , P237T , I252M , E286G , N292D , H299Y , N314D , E328R , R347G , R423K , M436I , N464D , I515M show in structure
HA_A/Shanghai/2/2013_138738	HA A/Netherlands/219/2003(H7N7) find closest related sequences	96.071	98.418	22	V18I , S20I , V63I , T137A , T150A , D190S , I195V , G202V , T205A , I218V , Q242L , I252M , E286G , N314D , E328R , R347G , T419N , R423K , M436I , N464D , I515M , A550V show in structure

[Right-click here to save/download mutation report table for archiving or import to Excel](#)

[Back to Reference Selection](#)

For each of the query sequences, users may proceed to look at the alignment to the reference strain, get more information on each mutation, generate a structural view of all the mutations ("show in structure")...

Analysis – FluSurfer for Mutation Interpretation

GISAID
GISAID published: 26,732 viruses with 40,071 Sequences
Total count: 35,932 viruses with 70,581 Sequences

Welcome John Doe
My Settings Logout

Browse Upload Workset Management Administration Registered Users

Personal Worksheet

FluSurfer

The main application scenario for FluSurfer is to highlight phenotypically or epidemiologically interesting candidate mutations for further research and should ideally be combined with experimental testing and verification of any predicted phenotypes. Importantly, any direct diagnostic use, assumed severity or recommendation on patient treatment should not be based solely on these computational predictions. Our curated reference sequences used for annotation transfer of equivalent mutations are mainly comprised of strains that recently infected humans. Therefore, the usage scenario that will give the most fruitful and reliable results are current surveillance sequences with very close relation to used vaccine strains, including some candidates for avian flu and novel reassortant swine flu H3N2v. Please take a look at the [Frequently Asked Questions](#) and [Tutorial](#) if you are new to FluSurfer. There is also a [special note for using FluSurfer results in publications](#).

Result for comparison with reference selection: H7N7_Human_2003_Netherlands219 [Back to Reference Selection](#)

Query	Best reference hit	% AA identity	% length coverage	# mutations	List of mutations
HA_A/Anhui/1/2013_138739	HA A/Netherlands/219/2003(H7N7) find closest related sequences	96.071	96.418	22	V18I S20I V63I T137A T150A D190S I195V Q202V T205A I218V Q242L I252M E286G N314D E328R R347G T419I R423K M436I M464D I515M A550V show in structure
HA_A/Shanghai/1/2013_138737	HA A/Netherlands/219/2003(H7N7) find closest related sequences	96.071	96.418	22	V18I S20I V63I T137A T150A A153S D190N I195V T205A I218V P237T I252M E286G N314D H292D H299Y I314D E328R R347G R423K M436I M464D I515M show in structure
HA_A/Shanghai/2/2013_138738	HA A/Netherlands/219/2003(H7N7) find closest related sequences	96.071	96.418	22	V18I S20I V63I T137A T150A D190S I195V Q202V T205A I218V Q242L I252M E286G N314D E328R R347G T419I R423K M436I M464D I515M A550V show in structure

[Right-click here to save/download mutation report table for archiving or import to Excel](#)

matics institute - Google Chrome

l.bii.a-star.edu.sg/METHODS/flusurver/beta/tmp/tmp_HA_H7N7_Human_2003_Netherlands219

reference: HA_H7N7_Human_2003_Netherlands219 ([Structure/Model Details](#))
atient/Sample: HA_A/Anhui/1/2013_138739
utation(s): A550V, S20I, I195V, I218V, T205A, D190S, I252M, N314D, R423K, V18I, T419N, I3I, M436I, E328R, R347G, N464D, E286G, I515M, T150A, T137A, Q202V, Q242L
*ldtype residue only displayed if position is part of available structure.
over with mouse over residue to see its position number. Right-click for more options.

Map of cities with the HA Q235L mutation

The city with **red** label indicates first appearance of the mutation. City with **yellow** label indicate latest appearance of the mutation. The city with the most recent appearance of the mutation has the **green** label. Number in the label indicates frequency of occurrence of the mutation in that city. A dot in the label indicates that there are 10 or more occurrences in that city.

Map of countries with the HA Q235L mutation

FluSurfer

HA Q242L

Key to alternative position numbering:

	FluSurfer numbering (absolute as in 2009 H1N1 pandemic)	Classical H3N2 strain numbering	Classical H1N1 strain numbering
240			
HA1 226			
HA1 223			
Chosen reference:	HA_H7N7_Human_2003_Netherlands219		
Position in reference:	242		
AA in reference:	Q		
AA in query:	L		

A mutation at the position equivalent to HA 242 has been reported in the literature to be related to [antigenic drift / escape mutant and host specificity shift and other](#).

A combination of mutations including the position equivalent to HA 242 has been reported in the literature to be related to [host specificity shift](#).

As seen in resolved structures of proteins from related strains, the HA position equivalent to your mutation is involved in:

- [host cell receptor binding](#)
- [antibody recognition sites](#)

[See all interactions for this position](#)

[Back to Reference Selection](#)

mendel.bii.a-star.edu.sg/METHODS/flusurver/b...

mendel.bii.a-star.edu.sg/METHODS/flusurver/beta/EFFECTS/H...

able to infect both cells and eggs with efficiencies comparable to those of wild type. Escape mutant MAb HC68
[Literature reference](#)
(Mutation L226P in the paper is at an equivalent position of the mutation in your query)

Protein: HA
Influenza type: Human H3N2 (N/A)
Mutation (as in paper): Q226L
neutral AA: Q
neg. eff. AA: L
Effect: host specificity shift

Comment:
Increasing affinity of receptor-binding to SA₂,6Gal and decreasing affinity to SA₂,3Gal (Table1).
[Literature reference](#)
(Mutation Q226L in the paper is at an equivalent position of the mutation in your query)

Analysis – FluSurver for Mutation Interpretation



The screenshot displays the FluSurver web application interface. The main panel shows a table of mutations with columns for Query, Best reference ID, % AA identity, % length coverage, and # mutations. A 3D protein structure is shown on the right, with a green arrow pointing to a specific mutation site. Below the table, there are maps of the world and a detailed view of the HA Q242L mutation, including its key to alternative position numbering, chosen reference, and a comment about its effect on receptor-binding.

Query	Best reference ID	% AA identity	% length coverage	# mutations
HA_A/Annam/02011_158738	HA_A/Annam/02011_158738	98.071	98.433	22
HA_A/Thailand/02011_158737	HA_A/Thailand/02011_158737	98.071	98.433	22
HA_A/Thailand/02011_158738	HA_A/Thailand/02011_158738	98.071	98.433	22

HA Q242L
Key to alternative position numbering:
240
HA1 228
HA1 223
Chosen reference: HA_H7N7_Human_2003_Netherlands219
Position in reference: 242
AA in reference: Q
AA in query: L
A mutation at the position equivalent to HA 242 has been reported in the literature to be related to [antigenic drift](#), [escape mutants](#) and [drug resistance](#) [prior](#) and [after](#).
A combination of mutations including the position equivalent to HA 242 has been reported in the literature to be related to [host specificity](#) [prior](#) and [after](#).
As seen in sequenced clonates of proteins from related strains, the HA position equivalent to your mutation is involved in: [host cell receptor binding](#) [antigenic drift](#) [antigenic escape](#) [drug resistance](#) [prior](#) and [after](#).
[See all correlations for this position](#)

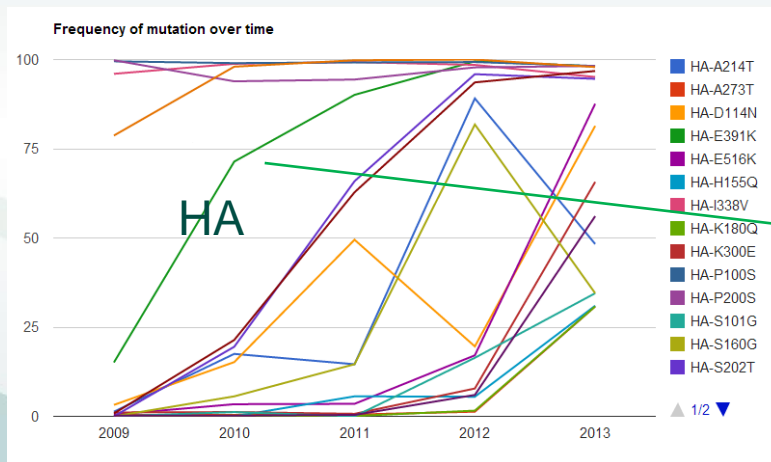
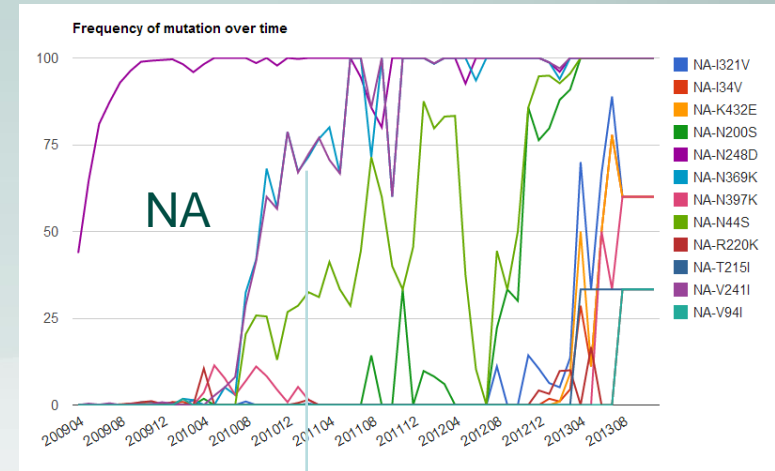
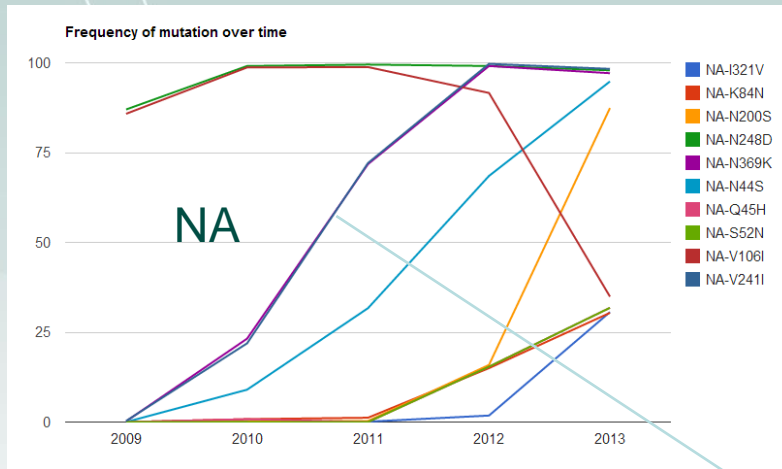
Protein: HA
Influenza type: Human H3N2 (N/A)
Mutation (as in paper): Q226L (N/A)
neutral AA: Q
neg. eff. AA: L
Effect: host specificity shift
Comment: Increasing affinity of receptor-binding to SA₂6Gal and decreasing affinity to SA₂3Gal (Table1.).
[Literature reference](#)
(Mutation Q226L in the paper is at an equivalent position of the mutation in your query)

Important disclaimer:

FluSurver makes it very easy to link mutations with prior literature and potential phenotypic effects.

While we have placed great emphasis on avoiding false positive alerts and provide tutorials, one still needs to read the associated papers and interpret the provided evidence carefully to judge any effect realistically.

Mutation frequency pattern highlights relevant changes



New H275Y permissive mutations

Hurt *et al.* J Infect Dis. 2012 Jul 15;206(2):148-57.

Butler *et al.* PLoS Pathog. 2014 Apr 3;10(4):e1004065.

Change in pH-dependency of fusion

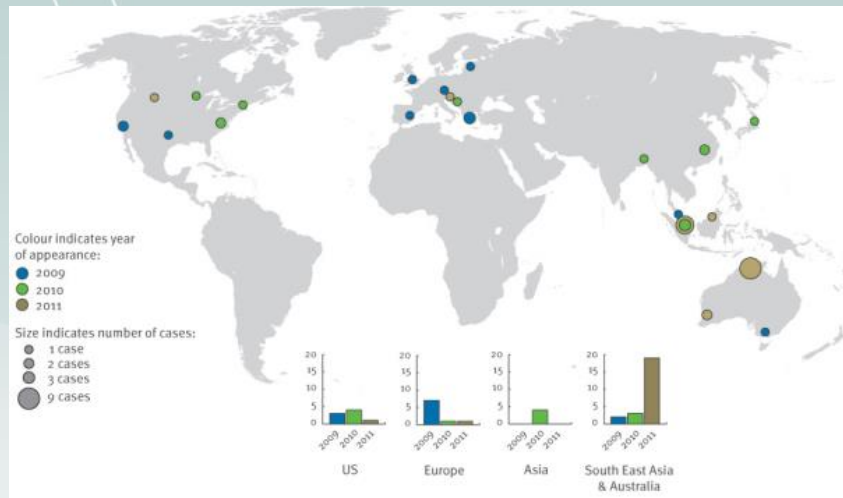
Maurer-Stroh *et al.* PLoS Curr. 2010 Jun 1;2:RRN1162.

Cotter *et al.* PLoS Pathog. 2014 Jan;10(1):e1003831.

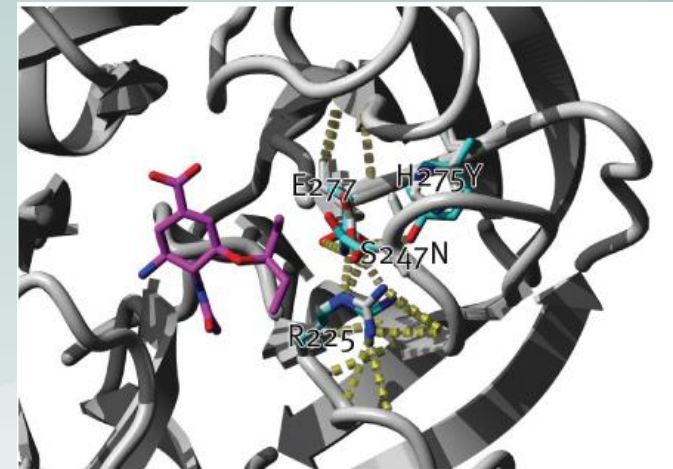
Example H1N1pdm in FluSurver



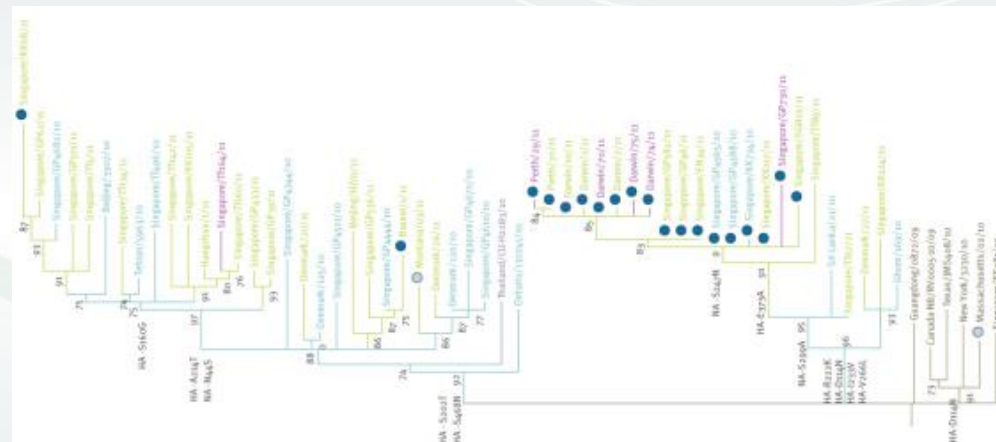
New drug sensitivity altering mutation NA S247N



Global occurrence of new variant



Structural context of mutation

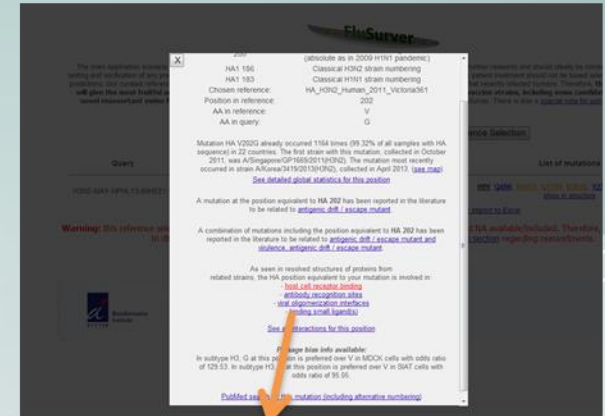
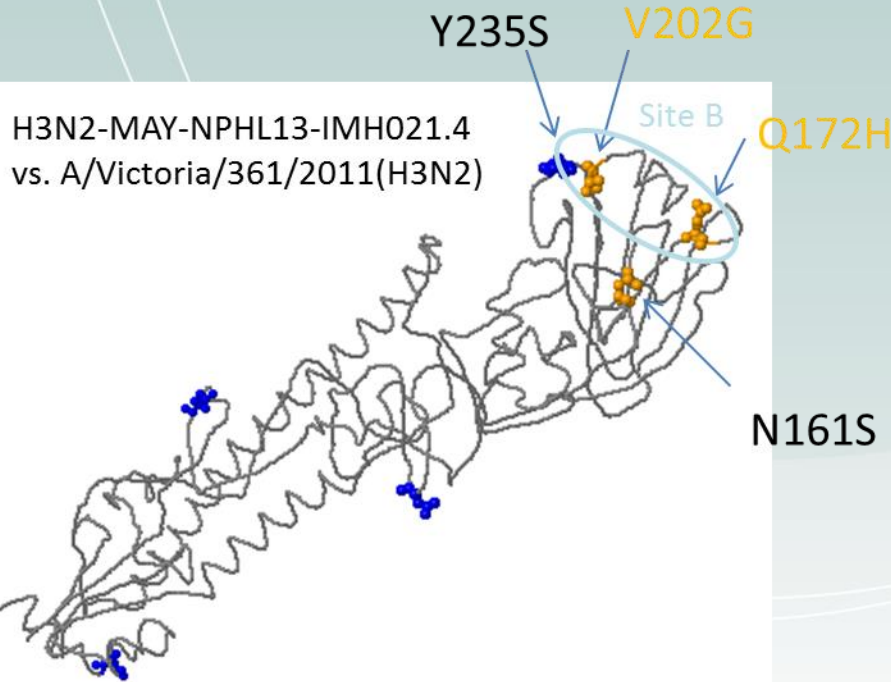


Phylogenetic context of new variant

Found circulating in 10% of samples in Singapore and 30% of samples in Northern Australia in early 2011.

Experimentally measured increase of IC₅₀ for Tamiflu by 6-fold and Relenza by 3-fold but **normally administered dose of drugs still sufficient.**

Current H3N2 strains have HA passage bias mutations in antigenic sites



As seen in resolved structures of proteins from related strains, the HA position equivalent to your mutation is involved in:

- [host cell receptor binding](#)
- [antibody recognition sites](#)
- [viral oligomerization interfaces](#)
- [binding small ligand\(s\)](#)

V202G

[See all interactions for this position](#)

Passage bias info available:

In subtype H3, G at this position is preferred over V in MDCK cells with odds ratio of 129.53. In subtype H3, G at this position is preferred over V in SIAT cells with odds ratio of 95.05.

Q172H

As seen in resolved structures of proteins from related strains, the HA position equivalent to your mutation is involved in:

- [host cell receptor binding](#)
- [antibody recognition sites](#)
- [binding small ligand\(s\)](#)
- is involved in [binding host protein\(s\)](#)
- [viral oligomerization interfaces](#)

[See all interactions for this position](#)

Passage bias info available:

In subtype H3, H at this position is preferred over Q in SIAT cells with odds ratio of 67.59.

Same isolate but different passage
(A/SINGAPORE/22/2012 NPHL: GP1187-2012)

GISAID ID	Submitter	Passage	Mutations relative to Victoria/361
EPI_ISL_128750	WHO CC Melbourne via NPHL	MDCK0, MDCK1	H9Y, Q49R, N161S, Q172H, V202G, Y235S , N294K
EPI_ISL_135838	US CDC via WHO CC Melbourne	E4/E1	H9Y, Q49R, N161S, N294K

FluSurver Acknowledgements

Many current and former colleagues from the A*STAR Bioinformatics Institute (BII) contribute(d) critically to its development and maintenance, including:

Sebastian Maurer-Stroh, Raphael Tze Chuen Lee, Vithiagarun Gunalan, Vachiranee Limviphuvadh, Fernanda L Sirota, Biruhalem Taye, Jianmin Ma, Swe Swe Thet Paing, Narumol Doungpan, Joy Xiang and Frank Eisenhaber.

The FluSurver would be nothing without the valuable feedback and interaction with the influenza research and surveillance community, including especially and in chronological order:

- Genome Institute of Singapore (GIS), Singapore
- INMEGEN Mexico City, Mexico
- Experimental Therapeutics Centre (ETC), Singapore
- Tan Tock Seng Hospital (TTSH), Singapore
- National Public Health Laboratory (NPHL) of the Ministry of Health, Singapore
- IAL Sao Paulo, Brazil
- WHO Collaborating Centre for Reference and Research on Influenza, Australia
- Duke-NUS Emerging Infectious Disease Programme, Singapore
- University of Melbourne, Australia
- Global Initiative for Sharing All Influenza Data (GISAID)
- Federal Office for Agriculture and Food (BLE), Germany
- Health Protection Agency of Canada

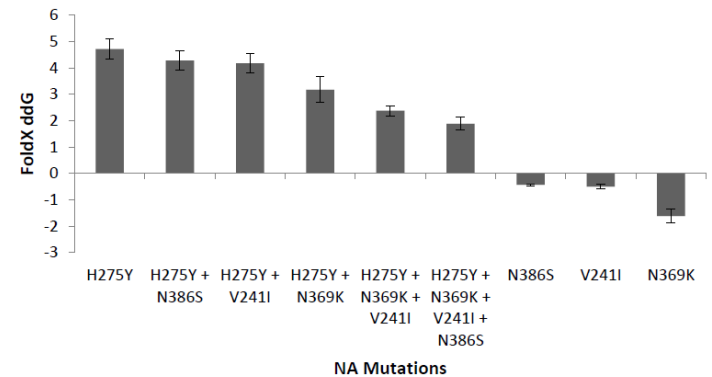
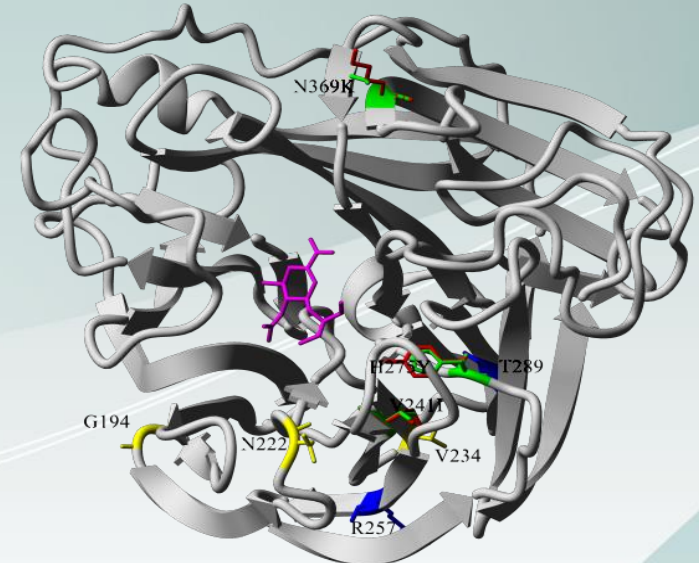
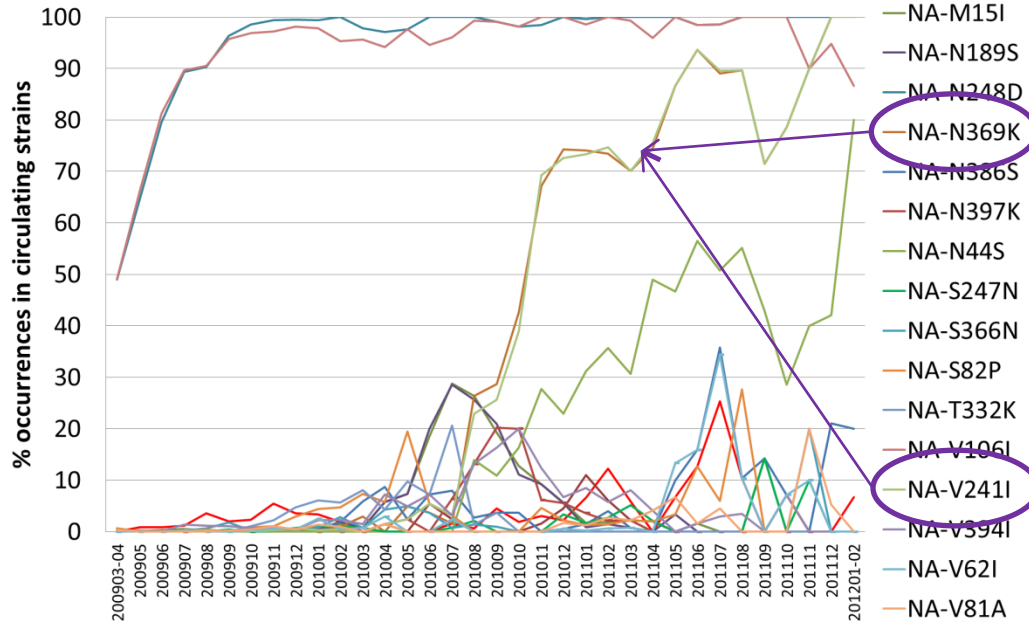
Fishing for Flu Mutations since 2009!



Optional back-up slides for questions...

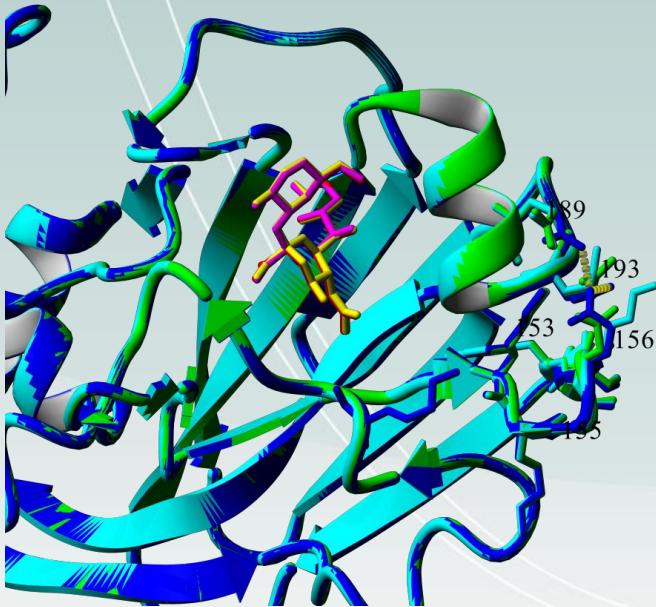
Frequency rise points to role of permissive mutations

Temporal appearance and frequencies of H1N1pdm neuraminidase mutations in Genbank and GISAID



FoldX predicts increase in structural stability for mutations that were increasing in frequency and were fixed in Newcastle strains.

A “stealth” antigenic drift mutation due to passage bias



A new mutation causing vaccine escape in ferret model cannot be found by classical virus culture plus sequencing because it always reverts to wildtype under culture conditions.

HA receptor with bound ligand (pink/yellow) and passage dependent mutations with number labels

Analysis for
Ian Barr
WHO CC

Mutation	No. reported ^a	No. isolates (mutant/wildtype) ^b	Total frequency (%) ^c	Passage history Odds Ratio (mutant vs wildtype)			
				Egg isolate	MDCK cell isolate	MDCK-SIAT1 cell isolate	Original clinical sample
N125D	581	321/7730	3.99	2.05	0.96	0.57	1.04
K153E	19	16/8018	0.20	1.59	6.55	-	-
G155E	133	103/7827	1.30	0.46	4.63	0.68	0.05
N156D	31	22/7980	0.27	-	4.32	0.52	0.22
N156K	22	12/7980	0.15	-	0.43	1.00	3.28
L191I	47	25/8008	0.31	55.31	0.20	-	-
Q223R	133	71/7966	0.88	564.42	0.02	0.00	0.00

The odds ratio, indicating strength of association to passage history, for mutant versus wildtype virus is indicated. Mutations with <10 samples with any passage information were omitted (e.g. K156E). (-) indicates that 10–30 records with passage information were available, and no reports were indicated in this passage history.

^aOccurrence of mutation in all 16740 A(H1N1)pdm09 sequences on GISAID and/or Genbank, regardless of passage history up to December 2012.

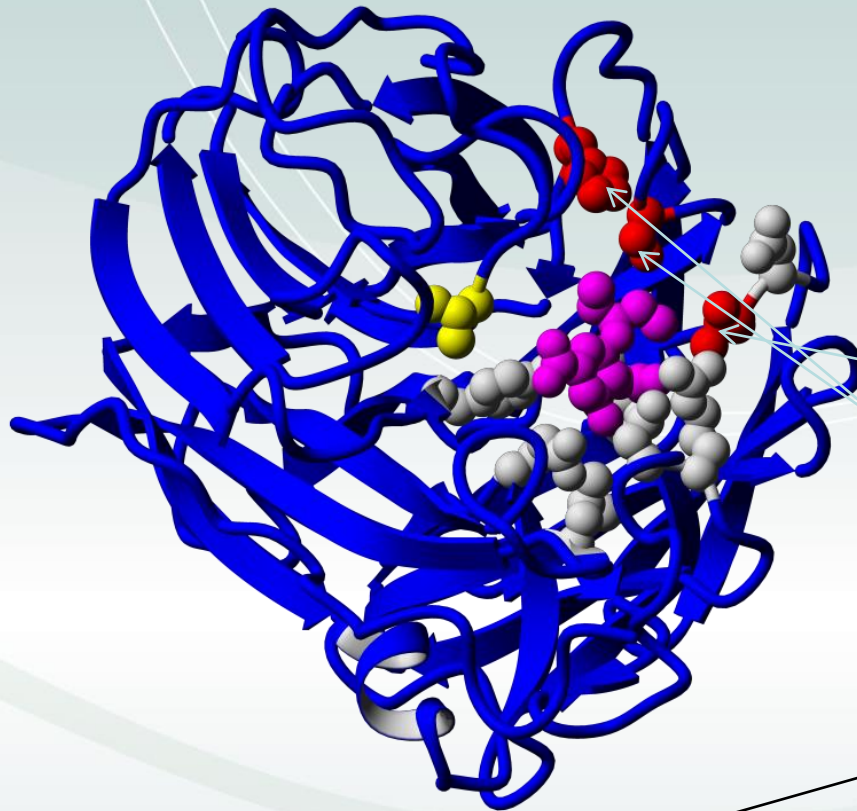
^bOccurrence of mutant or wildtype in all A(H1N1)pdm09 sequences on GISAID with passage history information.

^c% Occurrence of mutant in all A(H1N1)pdm09 sequences on GISAID with passage history information.

doi:10.1371/journal.ppat.1003354.t004

New drug sensitivity altering mutations

Neuraminidase and Tamiflu (pink)



amino acid of NA with distance<5 from G39	position	mutation occurred in H1N1	frequency of the mutation
R	118	-	-
E	119	G	1
		K	1
L	134	S	1
D	151	N	1
R	152	S	1
R	156	-	-
W	179	-	-
S	180	P	1
N	222	K	1
		S	1
		K	2
I	223	V	1
		R	2
L	224	-	-
R	225	-	-
T	226	A	1
E	228	-	-
S	247	N	5
H	275	Y	48
E	277	-	-
E	278	-	-
R	293	-	-
N	295	-	-
G	345	-	-
V	346	I	2
R	368	-	-
Y	402	H	1

Position # occ mut

275 48

247 5

223 5

H5N1 in Laos:

"S246N, identified in two isolates, reduced the sensitivity of isolate A/chicken/Laos/13/08 to oseltamivir by 24-fold as a single mutation" J Gen Virol. 2010 Apr;91(Pt 4):949-59.