Bioinformation **Volume 5**

open access

www.bioinformation.net Issue 5 **Hypothesis**

Supplementary material:

Table 1: The percentage of expected (**X**) and observed potential N-glycosylation sequons (**NXS** and **NXT**) in viral, archaeal and eukaryotic proteins.

	Virus			*Archaea*			*Eukarya*			*Eukarya*[1]	
	X	**NXS**	**NXT**	**X**	**NXS**	**NXT**	**X**	**NXS**	**NXT**	**NXS**	**NXT**
A	6.26	5.62*	5.98	7.40	7.15	6.68*	6.55	5.57*	5.98*	**9.37***	**7.75***
C	2.03	**2.48***	2.06	0.93	1.05	1.14	1.85	**1.99***	1.85	**4.69***	**3.05***
D	5.46	4.85*	4.80*	6.01	3.63*	4.71*	5.21	4.74*	4.45*	4.85	4.60
E	5.58	4.69*	4.89*	8.44	5.26*	6.61*	6.71	4.83*	5.66*	4.37*	5.86
F	4.22	**4.74***	3.78*	3.60	**4.91***	**4.42***	4.01	**4.18***	**4.13***	4.77	**5.28***
G	5.88	5.96	**6.48***	7.36	6.51*	7.22	6.05	**6.81***	**7.14***	**8.10***	**9.06***
H	2.21	1.65*	1.95	1.75	1.50	1.48	2.40	2.31*	2.08*	2.46	1.84
I	6.37	**8.27***	**7.93***	7.95	**10.97***	**10.25***	5.31	**6.14***	**6.29***	**8.50***	**6.97***
K	6.09	5.84	5.71	7.17	4.98*	7.43	6.13	5.04*	5.73*	1.83*	3.73*
L	9.08	8.54*	8.08*	8.86	**11.56***	9.48	9.62	9.13*	9.21*	6.99*	10.85
M	2.35	2.05*	2.09	2.40	2.11	2.11	2.20	1.66*	1.95*	1.91	1.69
N	5.48	**6.91***	**6.96***	4.16	4.57	4.37	4.76	**7.78***	**6.05***	4.29	2.57*
P	4.80	4.95	4.40*	4.09	4.34	3.87	5.35	5.00*	4.29*	0.48*	0.19*
Q	3.67	3.35	3.45	2.13	1.51*	1.77	4.43	3.59*	4.06*	3.02*	3.73
R	5.04	3.82*	3.89*	5.34	3.66*	4.98	5.16	3.78*	3.98*	2.46*	3.49*
S	7.35	**7.97***	**8.26***	5.52	6.03	4.89*	8.58	**12.16***	**9.89***	**12.07***	7.12*
T	6.36	6.11	**7.14***	4.86	4.86	**5.81***	5.54	**5.78***	**6.72***	6.35	**6.83***
V	6.53	6.80	**7.09***	7.63	**10.22***	8.28	6.01	5.66*	**6.53***	**9.29***	**9.20***
W	1.33	1.24	1.18	0.83	0.79	0.88	1.16	0.95*	1.13	0.71	1.50
Y	3.92	4.15	3.87	3.57	**4.40***	3.64	2.95	2.89	2.88	3.49	**4.70***

[1]Experimentally confirmed N-glycosylated sequons.

*Significantly different ($p < 0.05$, Z test for two proportions) from expected proportion of sequons.
Sequon proportions significantly higher than expected are shown in bold.

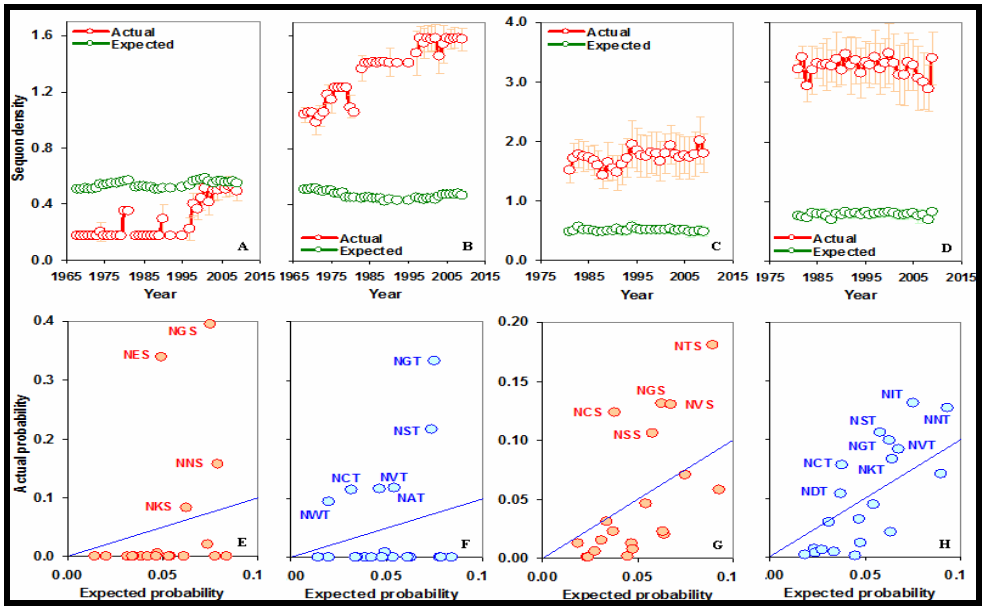

Figure S1 Sequon densities and types in viral envelope glycoproteins. Both NXS (A) and NXT (B) sequon densities have increased considerably in HA of influenza A H3N2 during the period 1968–2009. But NXS(C) and NXT (D) sequon densities have remained nearly same in gp120 of HIV-1 during the period 1981–2009. Actual versus expected probability of sequon types are shown for NXS (E) and NXT (F) in HA of influenza A H3N2 and NXS (G) and NXT (H) in gp120 of HIV-1. Inclined lines indicate the equal probability of actual versus expected sequon types. Only over-represented sequon types are labeled.

Bioinformation Volume 5

open access

www.bioinformation.net Issue 5

Hypothesis

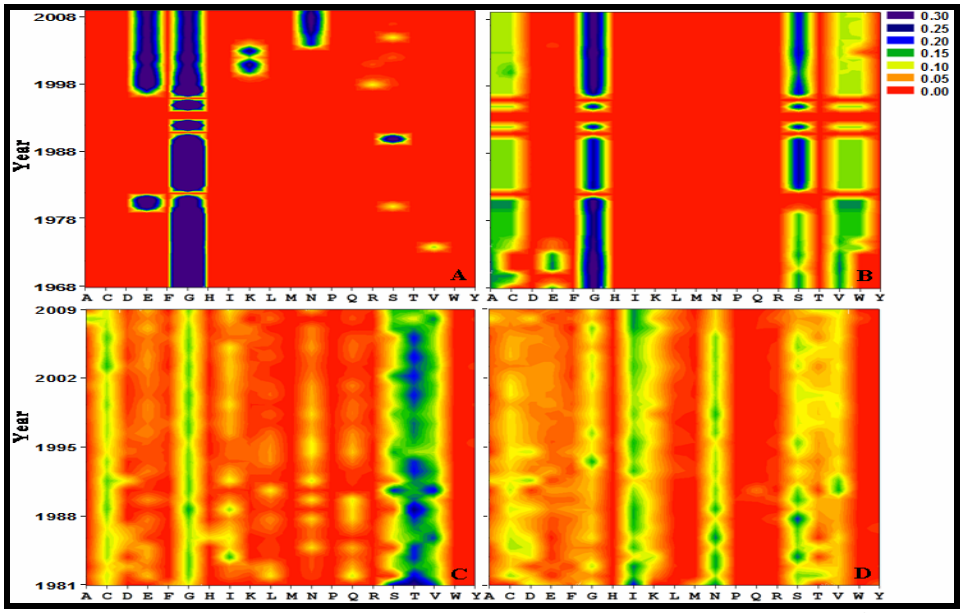

Figure S2 Tracking sequon types in viral envelope glycoproteins. Changes in the sequon types are shown for NXS (A) and NXT (B) in HA of influenza A H3N2 during the period 1968–2009 and for NXS(C) and NXT (D) in gp120 of HIV-1 during the period 1981–2009