Nonsynonymous Amino Acid Mutations Most Affect the Structure of the HIV-1 gp120 protein

Matthew Allegretti and Anindita Varshneya
BIOL 368: Bioinformatics Laboratory
Loyola Marymount University
November 8, 2016

Mutations in Four Subjects' Protein Sequences Were Analyzed to Determine Structural Differences

- Subjects 4, 9, 11, and 14
  - Exist in different progressor groups
- Clones collected from Visit 1 and Visit 4
  - Compare extent of change in structure over time
- Predict that changes between hydrophillic and hydrophibic amino acids will most affect function

Subjects have Identical dS/dN values in different progressor groups

Table 1. Summary data on 15 seroconverters

Subject	No. of observations	Baseline			Annual rate of CD4 T cell decline	Slope of change in intravisit nucleotide differences per clone per year	Slope of divergence (% nucleotides mutated from baseline consensus sequence per year)	Median dS/dN
		CD4	Median intravisit nucleotide differences among clones	Virus copy number ($\times 10^3$)				
Rapid Progressor								
Subject 4	4	1,028	0.90	6.8	−593	4.64	2.09	0.0
Subject 10	5	833	1.71	99.3	−363	3.16	1.00	0.2
Subject 11	4	753	2.27	62.2	−363	1.11	0.32	0.0
Subject 15	4	707	15.16	171.0	−362	−2.94	0.68	0.7
Subject 3	5	819	1.82	302.5	−294	0.53	0.74	1.0
Subject 1	3	464	5.64	307.6	−117	5.10	1.55	0.3
Moderate Progressor								
Subject 7	5	1,072	2.27	317.6	−392	−0.79	1.35	1.3
Subject 8	7	538	1.24	209.0	−92	1.68	1.16	0.5
Subject 14	9	523	1.00	50.9	−51	1.69	0.60	0.0
Subject 5	5	749	2.50	260.6	−41	0.06	0.50	1.4
Subject 9	8	489	9.49	265.0	−11	1.58	1.21	0.0
Subject 6	7	405	2.82	321.4	52	1.92	0.82	0.4
Nonprogressor								
Subject 2	5	715	1.64	21.6	30	1.32	0.49	1.8
Subject 12	6	772	2.80	5.1	44	0.62	0.13	0.9
Subject 13	5	671	0.87	1.7	53	0.53	0.28	3.5

Proportionally More Amino Acids Mutated Over Time Than Nucleic Acids

	% Change in AA		% Change in NA	
	Visit 1	Visit 4	Visit 1	Visit 4
Subject 4	2.13	18.09	1.06	11.35
Subject 9	4.21	10.53	2.11	6.32
Subject 11	6.25	8.33	3.47	3.47
Subject 14	4.17	9.38	2.08	5.56
Average	4.19	11.58	2.18	6.67

Variability Among Amino Acid Substitutions Was Measured According to Differences in Amino Acids per Position

- Single variability: 1 amino acid among all clones, but different from reference sequence
- Low Variability: 2 different amino acids present at each position among clones
- High Variability: More than 2 different amino acids present at each position among clones

Subject 4 Indicated Single and Low Variability

- Single variability at 11 positions
  - Positions 275, 283, 289, 301, 305, 308, 339, 343, 346, 358, 361
- Low variability at 14 positions
  - Positions 290, 299, 300, 302, 311, 313, 314, 321, 322, 324, 333, 335, 337, 345
    - 321 is a technicality as some clone sequences did not have an amino acid for that position
- High variability at 6 positions
  - 304, 306, 320, 327, 334, 353
    - 320 is a technicality as some clone sequences did not have an amino acid for that position

More Analysis Required For Results

- Variability among subjects 9, 11, and 14
- Implications of low vs. high variability on structure/ function of gp120 protein at V3 region