

The Markham et al.HIV-1 *env* Sequence Dataset

Richard Markham and his colleagues (1998), published some research on the pattern of HIV evolution and the rate of CD4 T-cell decline in the Proceedings of the National Academy of Sciences. In addition to the journal article they submitted 666 nucleotide sequences to the GenBank database. They studied a 285 base pair region of the *env* gene. The gene product, membrane protein gp120, binds to the CD4 receptor site on T-lymphocytes and is involved with the entry of the virus into those cells. Markham et al. followed the evolution of this viral gene sequence in 15 subjects by collecting blood samples at six month intervals for up to four years. For each visit all the forms of the gene (clones) were sequenced and CD4 T-cell counts were made. This data set provides a rich resource for looking closely at the patterns of change in HIV over time.

Summary of the data set

Subjects: 15

Number of visits: 3-9

Number of clones per visit: 2-18

Total number of sequences available: 666

CD4 cell counts for each visit

Subject	Total Number of Visits	Total Number of Clones	Visit Number	Number of Clones	CD4 Count ¹
1	3	42	1	13	464
			2	16	305
			5	13	15
2	3 ²	24	1	6	715
			3	9	825
			4	9	830
3	5	39	1	4	819
			3	10	375
			4	9	265
			5	10	100
			6	6	45
4	4	47	1	3	1028
			2	13	710
			3	18	470
			4	13	135
5	5	43	1	8	749
			2	12	770
			3	11	650
			4	7	550
			5	5	700

Subject	Total Number of Visits	Total Number of Clones	Visit Number	Number of Clones	CD4 Count¹
6	7	54	1	3	405
			2	3	225
			3	9	350
			4	12	390
			5	9	475
			7	9	400
			9	9	560
7	5	43	1	10	1072
			2	7	735
			3	8	330
			4	9	270
			5	9	310
8	7	49	1	5	538
			2	5	800
			3	7	605
			4	6	510
			5	6	625
			6	10	515
			7	10	250
9	8	64	1	5	489
			2	5	485
			3	8	440
			4	11	370
			5	9	365
			6	8	665
			7	10	555
			8	8	270
10	5	49	1	7	833
			2	6	850
			4	16	420
			5	10	150
			6	10	15
11	4	32	1	7	753
			2	6	600
			3	10	270
			4	9	175
12	6	37	1	4	772
			2	4	780
			3	5	1285
			4	6	1030
			5	10	1395
			8	8	850

Subject	Total Number of Visits	Total Number of Clones	Visit Number	Number of Clones	CD4 Count ¹
13	5	26	1	4	671
			2	2	825
			3	7	835
			4	7	770
			5	6	975
14	9	77	1	6	523
			2	6	580
			3	6	570
			4	10	595
			5	7	460
			6	11	420
			7	10	460
			8	9	450
			9	12	350
15	4	40	1	12	707
			2	9	250
			3	9	75
			4	10	15

Summary table of information available on the subjects studied in Markham et. al (1998).

¹ The CD4 count for time 1 is reported in Table 1. of Markham et. al (1998), the others values are estimated from the Figure 1. of the same publication.

² The paper reports 5 visits for subject 2, only 3 visits were identified in the GenBank records.

NOTE: Some of the visit numbers are not sequential. In all cases visit 1 represents the first time the subject was evaluated. The subsequent time points (visits 2 through 9) represent six month intervals from the initial visit. Thus, if a subject missed their six-month appointment their visits would be numbered 1, 3, 4, etc.

References

Markham RB, Wang WC, Weisstein AE, Wang Z, Munoz A, Templeton A, Margolick J, Vlahov D, Quinn T, Farzadegan H, Yu XF (1998). Patterns of HIV-1 evolution in individuals with differing rates of CD4 T cell decline. *Proc. Natl. Acad. Sci.* 95(21):12568-73.

Pub Med ID: 98445411

<http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=9770526&dopt=Abstract>

PNAS online: Vol. 95, Issue 21, 12568-12573, October 13, 1998

<<http://www.pnas.org/cgi/content/full/95/21/12568>>

GenBank sequences: [AF016760-AF016825](#) and [AF089109-AF089708](#)

<http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Link&db=Nucleotide&dbFrom=PubMed&from_uid=9770526>