dS/dN Ratios Do Not Provide A More Consistent Means of Comparison Between HIV-1 Progressor Groups

Zach Goldstein and Shivum Desai
October 11, 2016

Department of Biology
Loyola Marymount University

Outline

1. **Analysis of Markham article and figures indicated there may be a better way to classify HIV-1 progressor groups.**
2. Data was reorganized and reanalyzed using s values, theta values, and the formation of comparative trees.
3. Statistical analysis of data showed no significant difference between subjects.
4. Questioning utility of dS/dN ratio modeling may explain insignificant results.
5. Further research involves reanalysis of Markham data using slope of nucleotide difference and slope of divergence.

Patterns of HIV-1 evolution in individuals with differing rates of CD4 T cell decline.

Markham, R. B., Wang, W., Weisstein, A. E., Wang, Z., Munoz, A., Templeton, A., . . . Yu, X. (1998). Patterns of HIV-1 evolution in individuals with differing rates of CD4 T cell decline. *Proceedings of the National Academy of Sciences, 95*(21), 12568-12573. doi:10.1073/pnas.95.21.12568

- Progression of the disease HIV was analyzed and correlated with CD4 (T-cell) counts.

- dS/dN values were computed for each subject's consensus strain.

- Subjects were categorized into three groups: Rapid progressor, moderate progressor, and nonprogressor. *(Markham et. al, 1998)*

What is dS/dN value?

- Non synonymous mutations (dN) are nucleotide changes that lead to the coding of a different amino acid.
- Synonymous mutations (dS) are nucleotide changes that do not lead to the coding of a different amino acid.
- These ratios were formed for each consensus strain for all subjects.
    - A single averaged ratio was taken from all consensus strains.
    - Each strain was compared to a subsequent strain to observe differences.
- The smaller the ratio, the larger amount of non synonymous mutations.

Outline

1. Analysis of Markham article and figures indicated there may be a better way to classify HIV-1 progressor groups.
2. **Data was reorganized and reanalyzed using sequence analysis and the formation of comparative trees.**
3. Statistical analysis of data showed no significant difference between subjects.
4. Questioning utility of dS/dN ratio modeling may explain insignificant results.
5. Further research involves reanalysis of Markham data using slope of nucleotide difference and slope of divergence.

Markham's data was used to create a purpose for this experiment.

- A study was designed to analyze the dS/dN ratios of each subject.
  - The study involved the usage of Markham's dS/dN values, subjects, and classifications (Rapid progressor and nonprogressor).
- The purpose of this study was to analyze the dS/dN ratios of each subject and use this to show significant difference between the rapid progressors and nonprogressors.

Preparing to reanalyze Markham's data cont.

- Hypothesis #1: It was hypothesized that smaller dS/dN ratios would correlate with the newly classified rapid progressor subjects.
    - Null Hypothesis #1: It was hypothesized that larger dS/dN ratios would correlate with rapid progressor subjects.
- Hypothesis #2: It was hypothesized that newly classified rapid progressor subjects would have higher theta values due to a higher level of genetic diversity.
    - Null Hypothesis #2: It was hypothesized that the rapid progressors would have lower theta values due to their lower level of genetic values.

Reorganizing subjects and dS/dN values.

- dS/dN ratios were ranked in increasing order.
- Two groups of four subjects were selected from the two ends of the spectrum: Nonprogressor and rapid progressor.
- Non-progressor group:
  - Subjects: 4, 9 , 11, 14
  - The nonprogressor group was selected so that all subjects had a dS/dN value of 0.
- Rapid Progressor group:
  - Subjects: 7, 5, 2, 13
  - The rapid progressor subjects had a dS/dN value ranging from 1.3-3.5.

Markham's subjects were reorganized within existing classifications.

Rapid Progressor (Subject #)	Moderate Progressor (Subject #)	Non-progressor (Subject #)
4	**7**	**2**
10	8	12
11	**14**	**13**
15	**5**	
3	**9**	
1	6	

Table 1: Markham's classification of subjects based on CD4 cell counts.

Rapid Progressor (Subject #)	Nonprogressor (Subject #)
7	4
5	9
2	11
13	14

Table 2: Reclassification of subjects based on dS/dN value.

(Markham et. al, 1998)

Theta values and unrooted/rooted trees were calculated to compare results.

- Theta values were calculated from the number of position where there was at least one nucleotide difference across all that subject's clones (S value).
  - This value was calculated to show a significant difference between rapid progressors and nonprogressors.
  - Rapid progressors presumed to have higher theta values (more genetic diversity) due to the increased presence of nonsynonymous mutations.
- Unrooted and rooted trees were generated for the purpose of presenting visual evidence of the significant difference between rapid progressors and nonprogressors in terms of genetic similarity.

Outline

1. Analysis of Markham article and figures indicated there may be a better way to classify HIV-1 progressor groups.
2. Data was reorganized and reanalyzed using sequence analysis and the formation of comparative trees.
3. **Statistical analysis of data showed no significant difference between subjects.**
4. Questioning utility of dS/dN ratio modeling may explain insignificant results.
5. Further research involves reanalysis of Markham data using slope of nucleotide difference and slope of divergence.

Statistical results showed no correlation between dS/dN values and subject groups.

- The theta values calculated for each subject were not the expected numbers for those subjects in the nonprogressor and rapid progressor groups.
  - Ex. Subject 4 (NP-group): 15.8 Subject 5 (RP-group): 13.6
- The theta values calculated for each subject were averaged for both the nonprogressor and rapid progressor groups.
  - The resulting two theta values were measured for statistical significance using a t-test and no statistically significant difference was found.
  - P-value (significance= <0.05) of 0.146

There is no significant difference between Θ values of reclassified groups.

Subject #	2	5	7	13
"S" Value	36	58	55	25
"Θ" Value	9.6	13.6	12.7	6.6

Table 3: Calculated s-values and theta values of reclassified non progressor subjects.

Average "Θ" Value: 10.63

Subject #	4	9	11	14
"S" Value	70	69	41	77
"Θ" Value	15.8	14.6	10.2	15.7

Table 4: Calculated s-values and theta values of reclassified rapid progressor subjects.

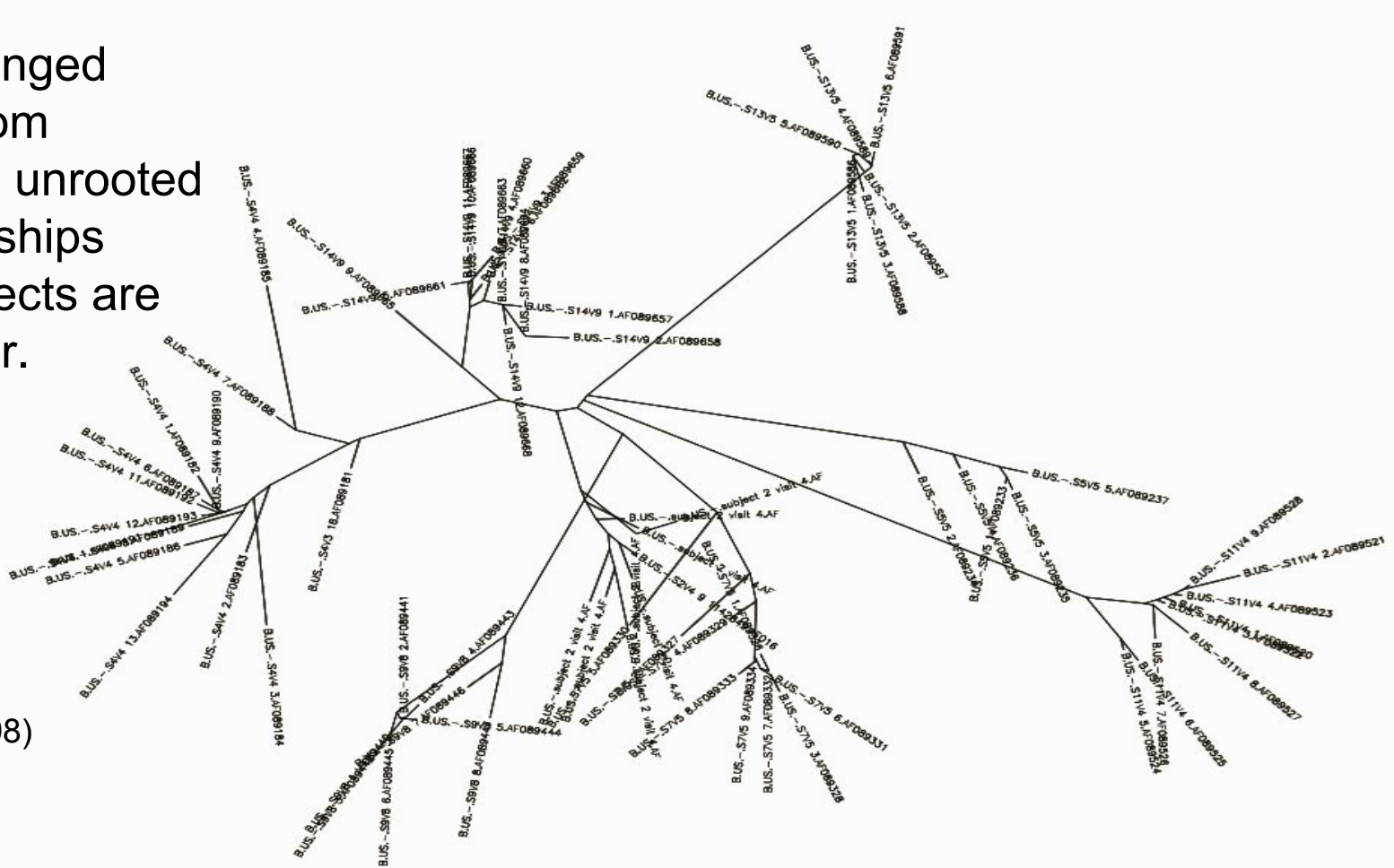
Average "Θ" Value: 14.08

Statistical Significance (<0.05): P-Value: 0.146

*Θ Values calcultaed using mathisfun.com (Subramaniam, 1998)

Interpretation of rooted and unrooted trees.

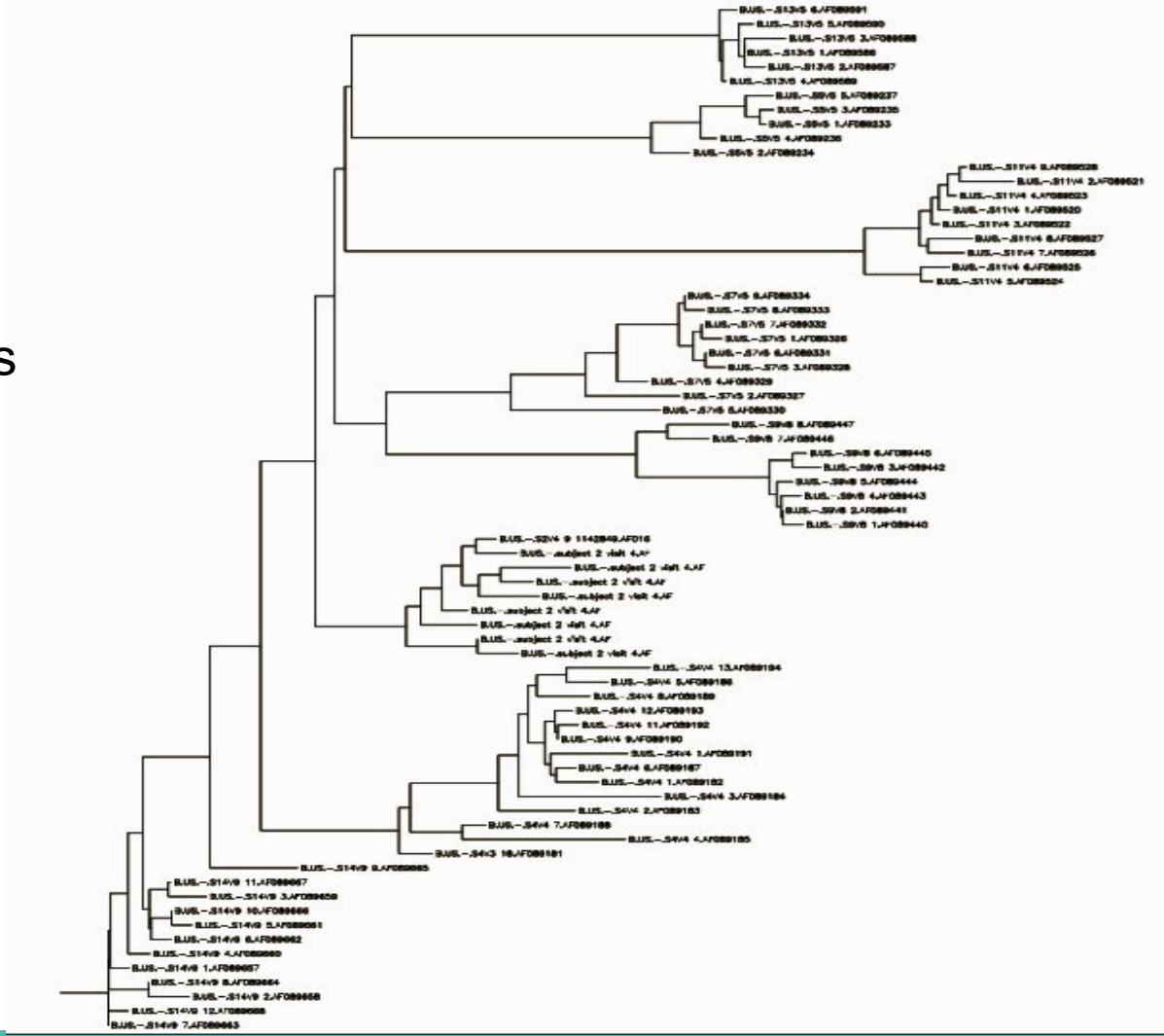
Figure 1: Arranged sequences from subjects in an unrooted tree. Relationships between subjects are difficult to infer.



(Subramaniam, 1998)

Interpretation of rooted and unrooted trees.

Figure 2: Arranged sequences from subjects in rooted tree format demonstrating non-distinct relationships between subjects classified in new groups.



(Subramaniam, 1998)

Outline

1. Analysis of Markham article and figures indicated there may be a better way to classify HIV-1 progressor groups.
2. Data was reorganized and reanalyzed using sequence analysis and the formation of comparative trees.
3. Statistical analysis of data showed no significant difference between subjects.
4. **Questioning utility of dS/dN ratio modeling may explain insignificant results.**
5. Further research involves reanalysis of Markham data using slope of nucleotide difference and slope of divergence.

The Relationship between dN/dS and Scaled Selection Coefficients

Spielman, S. J., & Wilke, C. O. (2015). The Relationship between dN/dS and Scaled Selection Coefficients. Molecular Biology and Evolution, 32(4), 1097-1108. doi:10.1093/molbev/msv003

- dS/dN ratio modeling assumes constant selective pressures over time between subjects.
- Confounding variables may cause differences in mutation rates.
- dN/dS ratio threshold =1 is highly sensitive to modeling assumption violations.

OUTLINE

1. Analysis of Markham article and figures indicated there may be a better way to classify HIV-1 progressor groups.
2. Data was reorganized and reanalyzed using sequence analysis and the formation of comparative trees.
3. Statistical analysis of data showed no significant difference between subjects.
4. Questioning utility of dS/dN ratio modeling may explain insignificant results.
5. **Further research involves reanalysis of Markham data using slope of nucleotide difference and slope of divergence.**

Reclassification based on genetic diversity or divergence may support initial hypotheses.

- Re-analyzing subjects based upon slope of change in nucleotide difference or slope of divergence may provide a more accurate method of analysis.
- Groups would still be classified as rapid progressor and nonprogressor, however dS/dN ratios would no longer be important to our investigations.
- T-test analysis of groups based off of diversity/divergence may show significant results.
    - S-values and theta values would be incorporated into this test.

Outline

1. Analysis of Markham article and figures indicated there may be a better way to classify HIV-1 progressor groups.
2. Data was reorganized and reanalyzed using sequence analysis and the formation of comparative trees.
3. Statistical analysis of data showed no significant difference between subjects.
4. Questioning utility of dS/dN ratio modeling may explain insignificant results.
5. Further research involves reanalysis of Markham data using slope of nucleotide difference and slope of divergence.

Acknowledgements

Dr. Dahlquist

Department of Biology at Loyola Marymount University

References

Markham, R. B., Wang, W., Weisstein, A. E., Wang, Z., Munoz, A., Templeton, A., . . . Yu, X. (1998). Patterns of HIV-1 evolution in individuals with differing rates of CD4 T cell decline. *Proceedings of the National Academy of Sciences, 95*(21), 12568-12573. doi:10.1073/pnas.95.21.12568

Spielman, S. J., & Wilke, C. O. (2015). The Relationship between dN/dS and Scaled Selection Coefficients. Molecular Biology and Evolution, 32(4), 1097-1108. doi:10.1093/ molbev/msv003

Subramaniam, S. (1998) The Biology Workbench--a seamless database and analysis environment for the biologist. Nucleic Acid Tools