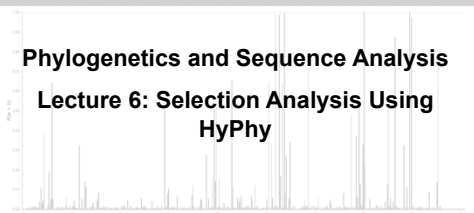


National Institute of Allergy and Infectious Diseases

## Phylogenetics and Sequence Analysis

### Lecture 6: Selection Analysis Using HyPhy



**Kurt Wollenberg, PhD**  
Phylogenetics Specialist  
Bioinformatics and Computational Biology Branch (BCBB)  
Fall 2015

NIH National Institute of Allergy and Infectious Diseases

NIH

NIAD

---

---

---

---

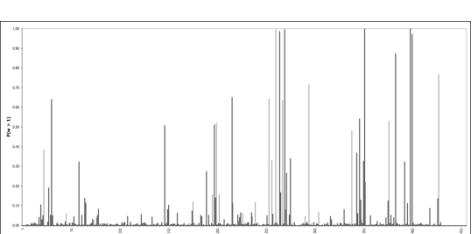
---

---

---

---

$0 \leftarrow 1 \rightarrow \infty$



NIH National Institute of Allergy and Infectious Diseases

NIH

NIAD

---

---

---

---

---

---

---

---

## Course Organization

- Building a clean sequence
- Collecting homologs
- Aligning your sequences
- Building trees
- **Further Analysis**

NIH National Institute of Allergy and Infectious Diseases

NIH

NIAD

---

---

---

---

---

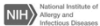
---

---

---

## Lecture Organization

- What is selection?
- What does selection look like?
- HyPhy: using maximum likelihood to quantify selection
- HyPhy input
- Interpreting your output



NIAD

---

---

---

---

---


---

---

---

## What is selection?

- Mutation – the source of all variation
- Substitution – changing one nucleotide to another
  - Synonymous substitution
  - Non-synonymous substitution
    - Conservative substitution



NIAD

---

---

---

---

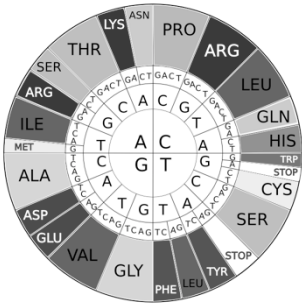
---

---


---

---

## Synonymous/non-synonymous Conservative/non-conservative



Adapted from Miller and Levine



NIAD

---

---

---

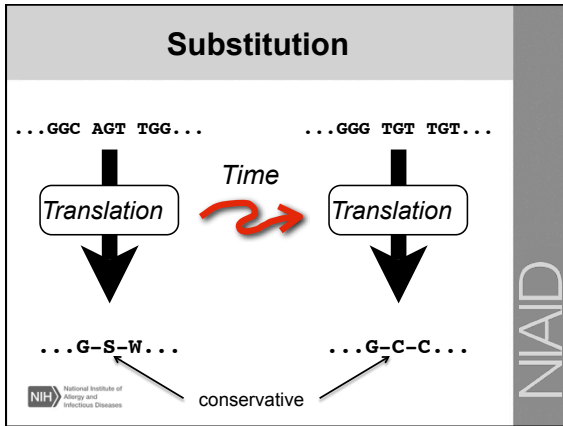
---

---

---

---

---



---

---

---

---

---

---

---

---

### What is selection?

- Differential survival of genotypes
- Positive selection - diversifying
  - Maintain variation
  - Variation can also be maintained by genetic drift
- Negative selection – purifying
  - Remove variation

NIH National Institute of Allergy and Infectious Diseases

NIAID

---

---

---

---

---

---

---

---

### Why is selection important?

- Pathogens
  - Surviving the immune system
- Patients
  - Surviving infection
- Identification of the molecular basis for survival

NIH National Institute of Allergy and Infectious Diseases

NIAID

---

---

---

---

---


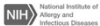
---

---

---

### What is selection analysis?

- Determining the dN/dS ratio
  - Inference of ancestral states
  - Determination of the effect of inferred nucleotide changes on amino acids
- Is dN/dS significantly different from 1?



---

---

---

---

---



---

---

---

### What is selection analysis?

- Positive (Diversifying) Selection  
dN/dS > 1
- Negative (Purifying) Selection  
dN/dS < 1



---

---

---

---

---



---

---

---

### Two types of selection analysis

- Selection at sites
  - Are there significant levels of non-synonymous substitution at specific codons?
  - Averaged across lineages.
- Selection along lineages
  - Are there significant levels of non-synonymous substitution in specific groups?
  - Averaged across sequences.



---

---

---

---

---

---

---

---

## Two types of selection analysis

Selection along lineages

- MCI\_01B4fs
- MCI\_01A10
- MCI\_01C1
- MCI\_01A20
- MCI\_01TA1

MCI\_01B4fs

- MCI\_01A10
- MCI\_01C1
- MCI\_01A20
- MCI\_01TA1

Selection at sites (codons)

```

5 100 1
TGCACATA ATCTGATT-----AAT ATCACAGAGA ATACTAATA
TGCAC---A ATCTGACAAA GGCCATTAAG ACCAATGGGA ATGCTAATA
TGCACATA ATCTGATT-----AAT ATCACAGAGA ATACTAATA
TGCACATA ATCTGACAAA GGCCATTAAG ACCAATGGGA ATGCTAATA
TGCACATA ATCTGATT-----AAT ATCACAGAGA ATACTAATA
TGCACATA ATCTGATT-----AAT ATCACAGAGA ATACTAATA

TACCATTACT AAT-----ACCACAT TAAATGCGGG GGGAATAA
TACCATTACT AATGCCACTG AGAGTACCAT TACTAATACC ACTGAATAA
TACCATTACT AAT-----ACCACAT TAAATGCGGG GGGAATAA
TACCATTACT AAT-----ACCACAT TAAATGCGGG GGGAATAA
TACCATTACT AAT-----ACCACAT TAAATGCGGG GGGAATAA
TACCATTACT AAT-----ACCACAT TAAATGCGGG GGGAATAA
    
```

NIAID

---

---

---

---

---

---

---

---

---

---

## Calculating dN/dS

- Estimate transition/transversion rate ratio
  - Transition: A ↔ G C ↔ T
  - Transversion: purine ↔ pyrimidine
  - Based on four-fold degenerate sites at third codon positions and nondegenerate sites.
- Count synonymous and nonsynonymous substitutions for each codon site or along each lineage.
  - Use ML estimates of ancestral codon states.
- Correct counts for multiple hits.
- Fitting the data to a distribution of dN/dS ratios.
  - Calculate the probability of dN/dS falling into a particular class.

NIAID

---

---

---

---

---

---

---

---

---

---

## HyPhy: the software

<http://datamonkey.org>

- Input data format
  - Sequences must be aligned.
  - Fasta, PHYLIP, or Nexus format.
  - Include the tree after the alignment.
- Program options
  - Analysis scripts.

NIAID

---

---

---

---

---

---

---

---

---

---

### HyPhy: the software

#### Standard Analyses

Select a standard analysis to run


- Basic Analyses
- Color Selection Analysis
- Compartmentalization
- Data File Tools
- Miscellaneous
- Model Comparison
- Kernel Analysis Tools
- Molecular Clock
- Phylogeny Reconstruction
- Positive Selection
- Recombination
- Selection/Recombination
- Relative Rate
- Relative Ratio
- Substitution Rates

**Item Description**

Perform codon-based positive selection tests.

OK Cancel

NIAID




---

---

---

---

---

---

---

---

### HyPhy: the software

#### Positive Selection Analyses

Select a standard analysis to run


- Phylogeny Reconstruction
- Positive Selection
  - BranchSiteFEL M
  - BranchSiteFEL2 M
  - BranchSiteFEL3 M
  - BranchSiteFEL4 M
  - BranchSiteFEL5 M
  - BranchSiteFEL6 M
  - BranchSiteFEL7 M
  - BranchSiteFEL8 M
  - BranchSiteFEL9 M
  - BranchSiteFEL10 M
  - BranchSiteFEL11 M
  - BranchSiteFEL12 M
  - BranchSiteFEL13 M
  - BranchSiteFEL14 M
  - BranchSiteFEL15 M
  - BranchSiteFEL16 M
  - BranchSiteFEL17 M
  - BranchSiteFEL18 M
  - BranchSiteFEL19 M
  - BranchSiteFEL20 M
  - BranchSiteFEL21 M
  - BranchSiteFEL22 M
  - BranchSiteFEL23 M
  - BranchSiteFEL24 M
  - BranchSiteFEL25 M
  - BranchSiteFEL26 M
  - BranchSiteFEL27 M
  - BranchSiteFEL28 M
  - BranchSiteFEL29 M
  - BranchSiteFEL30 M
  - BranchSiteFEL31 M
  - BranchSiteFEL32 M
  - BranchSiteFEL33 M
  - BranchSiteFEL34 M
  - BranchSiteFEL35 M
  - BranchSiteFEL36 M
  - BranchSiteFEL37 M
  - BranchSiteFEL38 M
  - BranchSiteFEL39 M
  - BranchSiteFEL40 M
  - BranchSiteFEL41 M
  - BranchSiteFEL42 M
  - BranchSiteFEL43 M
  - BranchSiteFEL44 M
  - BranchSiteFEL45 M
  - BranchSiteFEL46 M
  - BranchSiteFEL47 M
  - BranchSiteFEL48 M
  - BranchSiteFEL49 M
  - BranchSiteFEL50 M
  - BranchSiteFEL51 M
  - BranchSiteFEL52 M
  - BranchSiteFEL53 M
  - BranchSiteFEL54 M
  - BranchSiteFEL55 M
  - BranchSiteFEL56 M
  - BranchSiteFEL57 M
  - BranchSiteFEL58 M
  - BranchSiteFEL59 M
  - BranchSiteFEL60 M
  - BranchSiteFEL61 M
  - BranchSiteFEL62 M
  - BranchSiteFEL63 M
  - BranchSiteFEL64 M
  - BranchSiteFEL65 M
  - BranchSiteFEL66 M
  - BranchSiteFEL67 M
  - BranchSiteFEL68 M
  - BranchSiteFEL69 M
  - BranchSiteFEL70 M
  - BranchSiteFEL71 M
  - BranchSiteFEL72 M
  - BranchSiteFEL73 M
  - BranchSiteFEL74 M
  - BranchSiteFEL75 M
  - BranchSiteFEL76 M
  - BranchSiteFEL77 M
  - BranchSiteFEL78 M
  - BranchSiteFEL79 M
  - BranchSiteFEL80 M
  - BranchSiteFEL81 M
  - BranchSiteFEL82 M
  - BranchSiteFEL83 M
  - BranchSiteFEL84 M
  - BranchSiteFEL85 M
  - BranchSiteFEL86 M
  - BranchSiteFEL87 M
  - BranchSiteFEL88 M
  - BranchSiteFEL89 M
  - BranchSiteFEL90 M
  - BranchSiteFEL91 M
  - BranchSiteFEL92 M
  - BranchSiteFEL93 M
  - BranchSiteFEL94 M
  - BranchSiteFEL95 M
  - BranchSiteFEL96 M
  - BranchSiteFEL97 M
  - BranchSiteFEL98 M
  - BranchSiteFEL99 M
  - BranchSiteFEL100 M
- Recombination

**Item Description**

Perform codon-based positive selection tests.

OK Cancel

NIAID




---

---

---

---

---

---

---

---

### HyPhy: the software

#### Quick Selection Detection


- Single Ancestor Counting
- Weighted Ancestor Counting
- Sample Ancestral States
- Process Sampled Ancestral States
- One rate FEL
- Two rate FEL
- Rate Distribution
- Full site-by-site LRT
- Multiple FEL
- BSM co-evolution
- MEME

**Item Description**

Fixed effects site-by-site likelihood estimation, dS adjusted across sites.

1 required (1 chosen) OK Cancel

NIAID




---

---

---

---

---

---

---

---

## HyPhy: the software


### Quick Selection Detection

- Single Ancestor Counting
- Weighted Ancestor Counting
- Simple Ancestor States
- Process-Sampled Ancestor States
- One site FEL
- Two site FEL
- Rate Distribution
- Full site-by-site LRT
- Multisite FEL
- BGM co-evolution
- MEME**

**Item Description**

Mixed effects model of evolution to search for evidence of episodic selection at individual sites [EXPERIMENTAL, March 2011]

1 required (1 chosen)



NIAID

---

---

---

---

---

---


---

---

## HyPhy: the software

### The Algorithms

- **REL** – relative effects model – very similar to PAML codeml F61 analysis
- **2pFEL** –2-parameter fixed-effects model
- **MEME** – mixed-effects model



NIAID

---

---

---

---

---

---

---


---

## HyPhy: the software

### The MEME Algorithm

#### Mixed Effects Model of Evolution

- Analyzes levels of positive selection at individual sites while allowing level of positive selection to vary from branch to branch.
- Other models (FEL, REL, etc.) assume level of positive selection to be uniform across all branches.
- Includes site-to-site variation in synonymous substitution rate, just like FEL.



NIAID

---

---

---

---

---

---

---

---

### HyPhy: Input file formats



#### Fasta

No stop codons!

Make sure these match!

```

>MCI_01B4fs
TGCCTAATAATCTGATT-----AATATCACTGAGAATACTAATAATACCATTACT
>MCI_01A10
TGCAC-----AATCTGACAAAGGCTATTAGACCAATGGGAATGCTAATAATACCATTACT
>MCI_01C1
TGCCTAATAATCTGATT-----AATATCACTGAGAATACTAATAATACCATTACT
>MCI_01A20
TGCCTAATAATCTGACAAAGGCTAGTAATGCCACTGAGAAGGCTAATAATACCATTACT
>MCI_01FA1
TGCCTAATAATCTGATT-----AATATCACTGAGAATACTAATAATACCATTACT
((MCI_01B4fs,MCI_01A10),MCI_01C1),(MCI_01A20,MCI_01FA1));
  
```


---

---

---

---

---

---

---

---

---

---

### HyPhy: Input file formats



#### Fasta - interleaved

Make sure these match!

No stop codons!

```

>MCI_01B4fs
>MCI_01A10
>MCI_01C1
>MCI_01A20
>MCI_01FA1
TGCCTAATAATCTGATT-----AATATCACTGAGAATACTAATAATACCATTACT
TGCAC-----AATCTGACAAAGGCTATTAGACCAATGGGAATGCTAATAATACCATTACT
TGCCTAATAATCTGATT-----AATATCACTGAGAATACTAATAATACCATTACT
TGCCTAATAATCTGACAAAGGCTAGTAATGCCACTGAGAAGGCTAATAATACCATTACT
TGCCTAATAATCTGATT-----AATATCACTGAGAATACTAATAATACCATTACT
AAT-----ACCACTATTAATAGCGGGGGAGAAATAA
AATGCCACTGAGAGTACCATTACTACTAATACCCTGAAATAA
AAT-----ACCACTATTAATAGCGGGGGAGAAATAA
AAT-----ACCACTATTAATAGCGGGGGAGAAATAA
AAT-----ACCACTATTAATAGCGGGGGAGAAATAA
((MCI_01B4fs,MCI_01A10),MCI_01C1),(MCI_01A20,MCI_01FA1));
  
```


---

---

---

---

---

---

---

---

---

---

### HyPhy: Input file formats



#### PHLYIP

No stop codons!

Make sure these match!

```

5 100 I
MCI_01B4fs TGCACAAEA ATCTGATT-----AAT ATCACTGAGA ATACTAAEA
MCI_01A10 TGCAC-----A ATCTGACAAA GGCTATTAG ACCAATGGGA ATGCTAAEA
MCI_01C1 TGCACAAEA ATCTGATT-----AAT ATCACTGAGA ATACTAAEA
MCI_01A20 TGCACAAEA ATCTGACAAA GGCTAGTAAT GCCACTGAGA AGCTAAEA
MCI_01FA1 TGCACAAEA ATCTGATT-----AAT ATCACTGAGA ATACTAAEA
TACCATTACT AAT----- --ACCACTAT TAATAGCGGG GGAGAAATAA
TACCATTACT AATGCCACTG AGATTACTCAE TACTAATACC ACTGAAEA
TACCATTACT AAT----- --ACCACTAT TAATAGCGGG GGAGAAATAA
TACCATTACT AAT----- --ACCACTAT TAATAGCGGG GGAGAAATAA
TACCATTACT AAT----- --ACCACTAT TAATAGCGGG GGAGAAATAA
1
((MCI_01B4fs,MCI_01A10),MCI_01C1),(MCI_01A20,MCI_01FA1));
  
```


---

---

---

---

---

---

---

---

---

---



## HyPhy: Input file formats

**Nexus**

```


#NEXUS
Begin data;
  Dimensions ntax=5 nchar=100;
  Format datatype=DNA gap=- interleave;
MATRIX
MCI_01B4fs  TGCACAAATA ATCTGATT- - - - - AAT ATCACTGAGA ATACTAAATA
MCI_01A10  TGCACAAATA ATCTGATT- - - - - AAT ATCACTGAGA ATACTAAATA
MCI_01C1   TGCACAAATA ATCTGATT- - - - - AAT ATCACTGAGA ATACTAAATA
MCI_01A20  TGCACAAATA ATCTGAAA  GGCTATTAAG ACCAATGGGA ATGCTAAATA
MCI_01TAL  TGCACAAATA ATCTGATT- - - - - AAT ATCACTGAGA ATACTAAATA

TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
TACCATTACT AATGCACTG AGATATCACT TACTATATCC ACTGAAATA
TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
;
End;
Begin trees;
  TREE tree=((MCI_01B4fs,MCI_01A10),MCI_01C1),(MCI_01A20,MCI_01TAL));
End;

```

Annotations on the slide:

- Arrows pointing to the sequence labels: "Make sure these match!"
- Arrow pointing to a gap: "No stop codons!"



NIAID

---

---

---

---

---

---

---

---

## HyPhy: Input file formats

**GAPS!**

Sequence data

```

5 100 I
MCI_01B4fs  TGCACAAATA ATCTGATT- - - - - AAT ATCACTGAGA ATACTAAATA
MCI_01A10  TGCACAAATA ATCTGAAA  GGCTATTAAG ACCAATGGGA ATGCTAAATA
MCI_01C1   TGCACAAATA ATCTGATT- - - - - AAT ATCACTGAGA ATACTAAATA
MCI_01A20  TGCACAAATA ATCTGAAA  GGCTATTAAG ACCAATGGGA ATGCTAAATA
MCI_01TAL  TGCACAAATA ATCTGATT- - - - - AAT ATCACTGAGA ATACTAAATA


TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
TACCATTACT AATGCACTG AGATATCACT TACTATATCC ACTGAAATA
TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA
TACCATTACT AAT----- --ACCCTAT TAATAGCGGG GGAGAAATA

```

Annotations on the slide:

- Squares highlighting gaps: "GAPS!"

- Gaps must be in multiples of three and preserve the reading frame!
- ML models cannot analyze gapped codon sites – they will be ignored for the analysis.



NIAID

---

---

---

---


---

---

---


---

## HyPhy: data input



Annotations on the slide:

- Red boxes highlighting alignment options: "Choose a sequence alignment" (Marking/Clustal)
- Red boxes highlighting the "Upload" button.



NIAID

---

---

---

---

---

---

---

---

## Codon frequency models

0: 1/61 – All codons equally probable

1: F1x4 - Codon frequencies based on overall nucleotide frequencies ( $f_A, f_C, f_G, f_T$ )

2: F3x4 - Codon frequencies based on nucleotide frequencies at each codon site ( $f_{A1}, f_{C1}, f_{G1}, f_{T1}, f_{A2}, f_{C2}, f_{G2}, f_{T2}, f_{A3}, f_{C3}, f_{G3}, f_{T3}$ )

3: F61 - Codon frequencies based on frequency of each codon in the data ( $f_{AAA}, f_{AAC}, f_{AAG}, f_{AAT}, f_{ACA}, f_{ACC}, f_{ACG}, f_{ACT}, f_{AGA}, f_{AGC}, f_{AGG}, f_{AGT}, \dots$ )

F3x4 is used in the MEME analysis

NIAID

---

---

---

---

---

---

---

---

---

---

## HyPhy: data output

NIAID

---

---

---

---

---

---

---

---

---

---

## HyPhy: data output

NIAID

---

---

---

---

---

---

---

---

---

---

### HyPhy: data output

Codon	alpha	beta1	p1	beta2	p2	LRT	p-value	q-value	Log(L)
125	0.0000	0.0000	0.7253	13.6010	0.2747	8.0176	0.0081	0.3258	-36.6957
153	0.9139	0.0000	0.8927	49.5335	0.1073	8.7797	0.0055	0.2580	-22.6063
163	0.6646	0.0000	0.8478	20.0270	0.1522	5.4309	0.0304	0.8559	-25.8135
188	0.0000	0.0000	0.9873	1893.3900	0.0127	9.7907	0.0033	0.1853	-12.7060
211	0.7245	0.0000	0.9660	4586.8100	0.0340	11.9000	0.0011	0.1062	-22.5928
218	0.8903	0.0000	0.9662	304.8150	0.0338	13.3192	0.0006	0.1555	-17.2640
234	0.0000	0.0000	0.8160	263.5910	0.1840	10.1891	0.0027	0.1893	-33.7469
237	0.0969	0.0969	0.7191	17.6831	0.2809	12.3769	0.0009	0.1252	-35.2664
239	0.6719	0.0000	0.8254	13.6736	0.1746	5.9541	0.0232	0.7269	-22.8152
264	0.0000	0.0000	0.9370	8.8091	0.0630	6.6746	0.0160	0.5655	-11.4430

**Column Headings**

**Codon** The index of the codon in the alignment  
**n** The maximum likelihood estimate (MLE) of the synonymous substitution rate  $\alpha$   
 **$\beta^1$**  The maximum likelihood estimate (MLE) of the non-synonymous rate for the branch class with  $\beta = \alpha$   
**Pr[ $\beta = \alpha$ ]** The MLE of the proportion of sites evolving at  $\beta^1$   
 **$\beta^2$**  The MLE of the unconstrained  $\beta$  non-synonymous rate  
**Pr[ $\beta = \beta^2$ ]** The MLE of the proportion of sites evolving at  $\beta^2$   
**LRT** Likelihood ratio test statistic for  $\beta^1 = \alpha$  (null) versus  $\beta^1$  unrestricted (alternative)  
**p-value** The p-value for the LRT test, using the mixture distribution :  $0.33 \chi^2_1 + 0.30 \chi^2_2 + 0.37 \chi^2_3$   
**q-value** The q-value for independent tests (upper bound on the false discovery rate), derived from the corresponding p-value using Storey's procedure

NIAD

---

---

---

---

---

---

---

---

---

---

---

---

### HyPhy: further analyses

Partitions	Sequences	Codons	Genetic Code	The SMC Uploaded	Express in
1	17	282	Universal	90101314	Full (282)
Nucleotide selection			D3136	D320	
Amnuaaa selection			D3136	D320	
Neighbor Joining Tree			RM003	D320	
Line Trends			RM003	D320	

<b>Accession</b>	<b>Status</b>
<b>SLAC</b>	No results. Set in the [SLAC] matrix.
<b>Context Model Selection</b>	No results. Set in the [Context] matrix.
<b>ESL</b>	No results. Set in the [ESL] matrix.
<b>FTL</b>	No results. Set in the [FTL] matrix.
<b>DOABLE</b>	No results. Set in the [DOABLE] matrix.
<b>REL</b>	No results. Set in the [REL] matrix.
<b>RELIN</b>	No results. Set in the [RELIN] matrix.
<b>GA-Branch</b>	No results. Set in the [GA-Branch] matrix.
<b>PhyML</b>	No results. PhyML requires the installation of the PH-MLINK, PHMLINK, or PH-MLINK packages.
<b>SpokenanaryGCM</b>	No results. Set in the [SpokenanaryGCM] matrix.
<b>Single Site Reconstruction</b>	No results. Set in the [SSR] matrix.
<b>GLAD</b>	No results. Set in the [GLAD] matrix.
<b>Biological State Reconstruction</b>	No results. Set in the [BSR] matrix.
<b>Evolutionary Programming</b>	No results. Set in the [EP] matrix.
<b>Directional Evolution in Protein</b>	This analysis is not available for codon alignments.
<b>Selection</b>	No results. Set in the [SEL] matrix.
<b>Branch-site REL test for Episodic Diversification</b>	No results. Set in the [BS-REL] matrix.
<b>Model Specific Model of Episodic Diversification</b>	<b>Model Specific Model of Episodic Diversification</b> . 19 sites under episodic diversification (p < 1).
<b>Branch-site REL test for Episodic Diversification</b>	No results. Set in the [BS-REL] matrix.
<b>Fast Unconstrained Bayesian Approximation</b>	No results. Set in the [FUBAP] matrix.
<b>Phylogeny Informed Models of Selection</b>	No results. Set in the [PIM] matrix.
<b>FUBAP Approach to Directional Evolution</b>	This analysis is not available for codon alignments.

NIAD

---

---

---

---

---

---

---

---

---

---

---

---

### Selection Analysis

**Site vs. Lineage Selection**

Generally, if significant positive (or negative) selection occurs at only a few sites, averaging over the length of the sequences will result in no significant positive (or negative) selection being detected.

NIAD

---

---

---

---

---

---

---

---

---

---

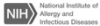

---

---

## Selection Analysis

Significance of Results

Selection analysis only determines if there is a significant **excess** or **lack** of non-synonymous substitution. The relative level of physiochemical similarity among the two amino acids is beyond the capabilities of selection analysis software.


---

---

---

---

---



---

---

---

## References

- ❖ Fundamentals of Molecular Evolution. 2nd edition. Grauer and Li. 2000  
A good introduction to the major concepts in molecular evolution.
- ❖ The Phylogenetics Handbook. P. Lemey, editor. 2005.  
Chapter 14: Theory of and practice of diversifying selection analysis.
- ❖ MEME reference:  
Murrell B, et al. (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, 8(7):e1002764.


---

---

---

---

---

---

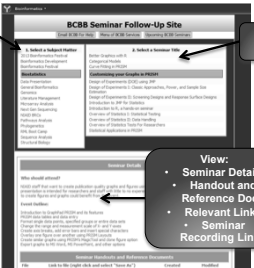
---

---

## Seminar Follow-Up Site

- For access to past recordings, handouts, slides visit this site from the NIH network: <http://collab.niaid.nih.gov/sites/research/SIG/Bioinformatics/>

1. Select a Subject Matter





2. Select a Topic

Recommended Browsers:

- IE for Windows,
- Safari for Mac (Firefox on a Mac is incompatible with NIH Authentication technology)

Login

- If prompted to log in use "NIH" in front of your username


---

---

---

---

---

---

---

---

## Retrieving Slides/Handouts

**1. Select a Subject Matter**

- 2012 Bioinformatics Festival
- Bioinformatics Development
- Bioinformatics Method
- Statistics
- Data Presentation
- General Bioinformatics
- Genomics
- Literature Management
- Microarray Analysis
- Next-Gen Sequencing
- NASC-BISCS
- Pathways Analysis
- Phylogenetics
- RNA-Seq Camp
- Sequence Analysis
- Structural Biology

**2. Select a Seminar Title**

**Building Tree - Phylogenetics I**

- Building Tree - Phylogenetics I
- Building Tree - Phylogenetics II
- Chondrocyte Phenotyping and Sequence Alignment
- Introduction to Phylogenetics and Sequence Assembly
- Using Phylogenetic Quality Phylogenetic Trees
- Making Publication-Quality Tree Figures
- How Tools for Adding Genomic and Functional Context to Your Data
- Phylogenetic Analysis Using BEAST
- Phylogenetic Molecular Evolution I
- Phylogenetic Molecular Evolution II
- Selection Analysis Using SHaPy
- Selection Analysis Using PAML

**Seminar Details**

**Seminar Handouts and Reference Documents**

**Links Relevant to This Seminar**

**Seminar Recording Links**

NIAD  
37

---

---

---

---

---

---

---

---

---

---

---

---

---

---

## Retrieving Slides/Handouts

**1. Select a Subject Matter**

- 2012 Bioinformatics Festival
- Bioinformatics Development
- Bioinformatics Method
- Statistics
- Data Presentation
- General Bioinformatics
- Genomics
- Literature Management
- Microarray Analysis
- Next-Gen Sequencing
- NASC-BISCS
- Pathways Analysis
- Phylogenetics
- RNA-Seq Camp
- Sequence Analysis
- Structural Biology

**2. Select a Seminar Title**

**Building Tree - Phylogenetics I**

- Building Tree - Phylogenetics I
- Building Tree - Phylogenetics II
- Chondrocyte Phenotyping and Sequence Alignment
- Introduction to Phylogenetics and Sequence Assembly
- Using Phylogenetic Quality Phylogenetic Trees
- Making Publication-Quality Tree Figures
- How Tools for Adding Genomic and Functional Context to Your Data
- Phylogenetic Analysis Using BEAST
- Phylogenetic Molecular Evolution I
- Phylogenetic Molecular Evolution II
- Selection Analysis Using SHaPy
- Selection Analysis Using PAML

**Seminar Details**

**Seminar Handouts and Reference Documents**

File	Links to the right click and select "Save As"	Download	Thumbnail
01	10/23/15 11:00 AM	10/23/15 11:00 AM	
02	10/23/15 11:00 AM	10/23/15 11:00 AM	
03	10/23/15 11:00 AM	10/23/15 11:00 AM	
04	10/23/15 11:00 AM	10/23/15 11:00 AM	
05	10/23/15 11:00 AM	10/23/15 11:00 AM	

**Links Relevant to This Seminar**

**Seminar Recording Links**

NIAD  
38

---

---

---

---

---

---

---

---

---

---

---

---

---

---

## Questions?

Consultation & Advice | Software Development | Bioinformatics Resources

ScienceApps@niaid.nih.gov

NIH National Institute of Health and Infectious Diseases

BCBB Bioinformatics and Computational Biosciences Branch  
NIAD's Office of Cyber Infrastructure and Computational Biology

NIAD  
39

---

---

---

---

---

---

---

---

---

---

---

---

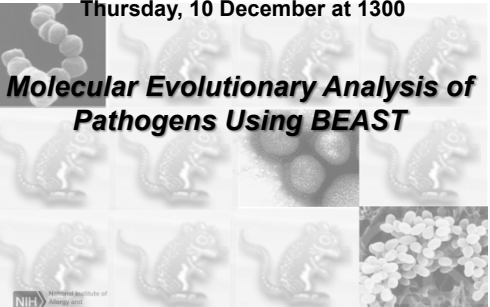
---

---

**Next**

Thursday, 10 December at 1300

***Molecular Evolutionary Analysis of Pathogens Using BEAST***



NIAID  
National Institute of Allergy and Infectious Diseases

NIH

NIAD

40

---

---

---

---

---

---

---

---