

Snake Sisters

How I use genes, models, and morphology to understand
what we know about squamate evolution



Genevieve G. Mount
LSU, Dept. of Biological Sciences &
Museum of Natural Science

LSU Computational Biology Seminar Series for Undergraduates
November 10th 5 PM LSA 101

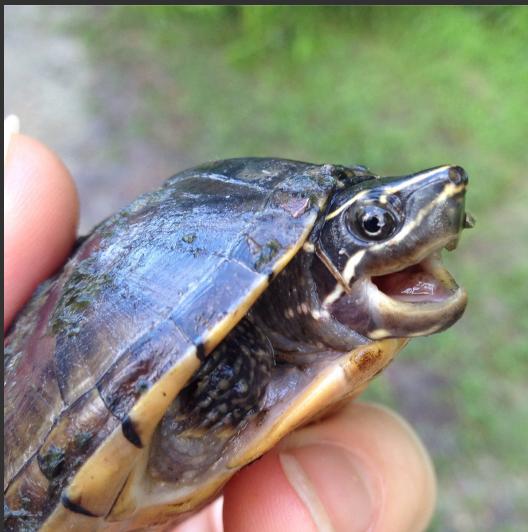
Pop Quiz



Pop Quiz: Who is sister to snakes?



A) Toad



B) Turtle



C) Lizard

Pop Quiz: Who is sister to snakes?



B) Mole Lizards (Amphisbaenians)



©2012 ADAM G. CLAUSE

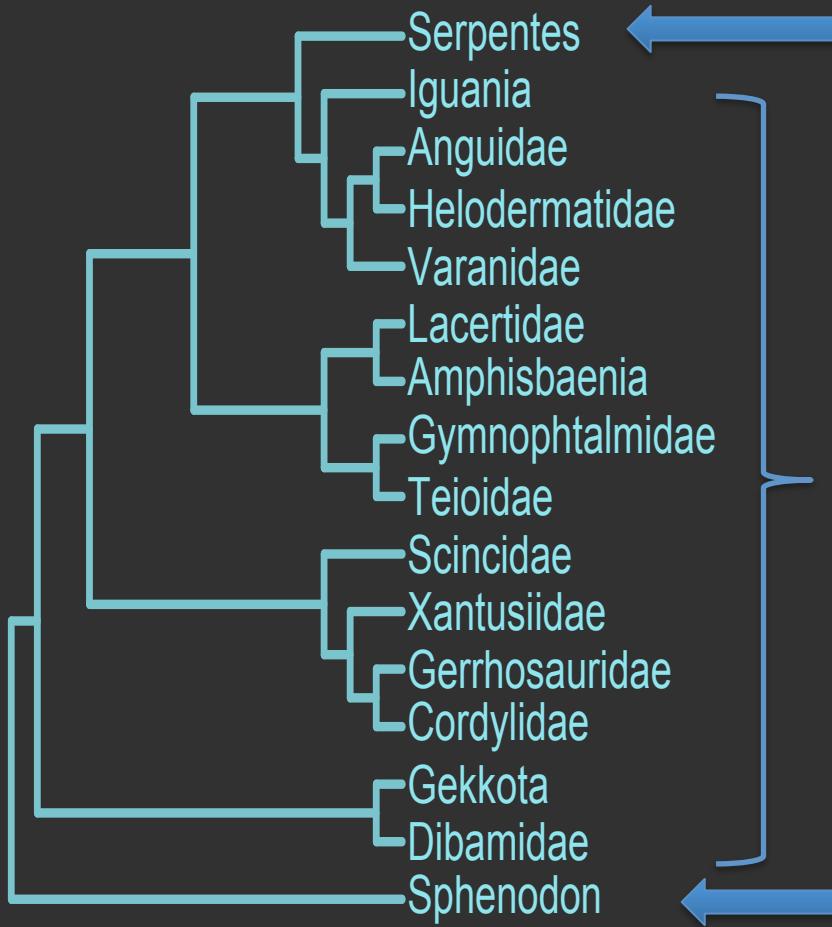
A) Iguanas



C) Skinks



Squamate Tree of Life



Snakes



Lizards

Tuataras (look like lizards but aren't lizards)

Squamates =



Snakes

+

Lizards



What does this have to do with computers?

```
7  for f in *.nex;
8  do
9    base=`basename $f .nex`
10   cd $base
11   rm *garli.conf
12   cp ..garli.conf .
13   nex=$base".nex"
14   model=`tail -n 1 *.jMT.txt | awk '{print $11}'`
15   echo $model $base
16   ratematrix=`echo $model | awk -F'+' '{ print $1}'``
17   echo $ratematrix
18   modelp2=`echo $model | awk -F'+' '{ print $2}'``
19   echo $modelp2
20   modelp3=`echo $model | awk -F'+' '{ print $3}'``
21   echo $modelp3
22   numberModels=`grep '\[model.\]' garli.conf | wc -l`
23   echo "The number of models in your .conf file is $numberModels"
24   if [ "$numberModels" -gt 1 ]
25     then
26       echo "It looks like your starting .conf file has too many models specified."
27       echo "This script does not deal with partitioned data. One model per nexus file, please"
28       echo "Remove additional models from your .conf file and run this script again."
29       exit 1
30   fi
31 cp garli.conf $base".garli.conf"
32 rm garli.conf
33 conf=$base".garli.conf"
34 sed -i.tmp "s/hostname = .*/hostname = $nex/" $conf
35 sed -i.tmp "s/constraintfile = .*/constraintfile = none/" $conf
36 sed -i.tmp "s/ofprefix = .*/ofprefix = $base/" $conf
37 sed -i.tmp "s/outputphyliptree = .*/outputphyliptree = 1/" $conf
38   if [ $ratematrix == "GTR" ]
39     then
40       sed -i.tmp "s/ratematrix = .*/ratematrix = 6rate/" $conf
41       sed -i.tmp "s/statefrequencies = .*/statefrequencies = estimate/" $conf
42     elif [ $ratematrix == "SYM" ]
43     then
44       sed -i.tmp "s/ratematrix = .*/ratematrix = 6rate/" $conf
45       sed -i.tmp "s/statefrequencies = .*/statefrequencies = equal/" $conf
46     elif [ $ratematrix == "JC" ]
47     then
```



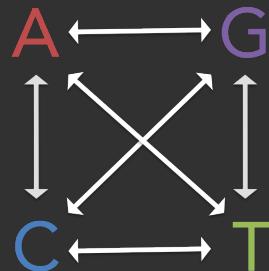
What are phylogenies? How do we build them?

Data



Morphology

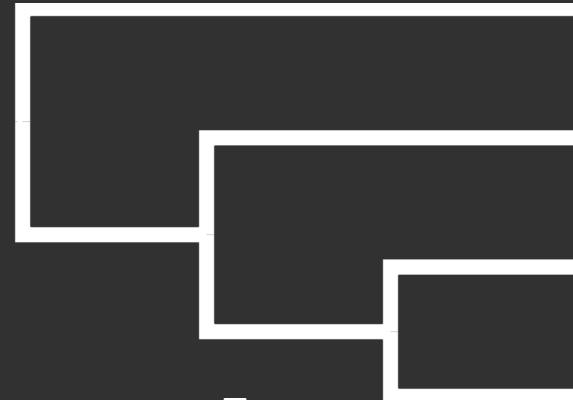
Model



Model of nucleotide evolution

TTCCACACAAACATACGAGCCGGAAGCATAAAAGTGTAAA
CCAACGGCGGGGAGAGGGCGGTTTGCCTATTGGGCCGCT
TGAGTCGTATTAGAGCTTGGCGTAATCATGGTCATAGC
TTCCACACAAACATACGAGCCGGAAGCATAAAAGTGTAAA
TTCCACACAAACATACGAGCCGGAAGCATAAAAGTGTAAA
CCAACGGCGGGGAGAGGGCGGTTTGCCTATTGGGCCGCT
TGAGTCGTATTAGAGCTTGGCGTAATCATGGTCATAGC
TTCCACACAAACATACGAGCCGGAAGCATAAAAGTGTAAA
TTCCACACAAACATACGAGCCGGAAGCATAAAAGTGTAAA
CCAACGGCGGGGAGAGGGCGGTTTGCCTATTGGGCCGCT
TGAGTCGTATTAGAGCTTGGCGTAATCATGGTCATAGC
TTCCACACAAACATACGAGCCGGAAGCATAAAAGTGTAAA

DNA



Tree

Why do we care about phylogenies?

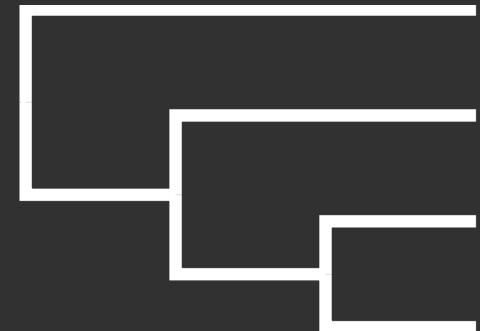
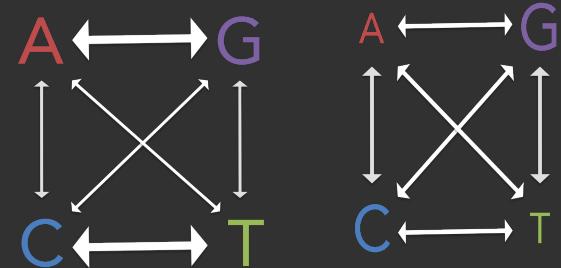
- Evolution
- Ecology and conservation
- Agriculture
- Human development and disease
 - HIV
- Squamates
 - Gecko feet
 - Venom evolution



Phylogenetic methods

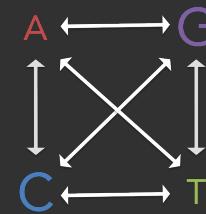
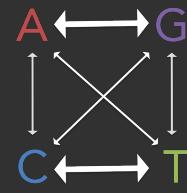
- Data
- Model
- Computer
- Tree with 20 tips
 - # of tree topologies > stars in the known universe

ATCCACACAAACAATACGAGCCGGAAACCAAAAGTGAAA
CGAACCGCGGGGAGAGGGGGTTTCCGATTCGGGCGC
TGAGCTGTATTAGACCTTGGCTTAATCAGGGCAAGC
TCACACACACATAACGACTGGAAACCAAAAGTGAAA
TCCACACACATAACGACTGGAAACCAAAAGTGAAA
CCAACCGCGGGGAGAGGGGGTTTCCGATTCGGGCGC
TGAGCTGTATTAGACCTTGGCTTAATCAGGGCAAGC
TCACACACACATAACGACTGGAAACCAAAAGTGAAA
TCCACACACATAACGACTGGAAACCAAAAGTGAAA
CCAACCGCGGGGAGAGGGGGTTTCCGATTCGGGCGC
TGAGCTGTATTAGACCTTGGCTTAATCAGGGCAAGC



What I use computers for:

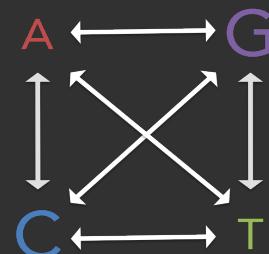
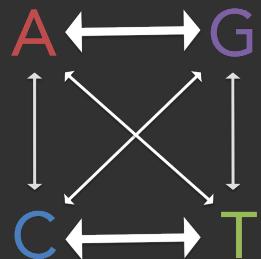
Study the behavior and variation of models and data



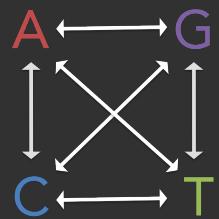
What I use computers for: Behavior and variation of models and data

- All models are wrong but some are useful

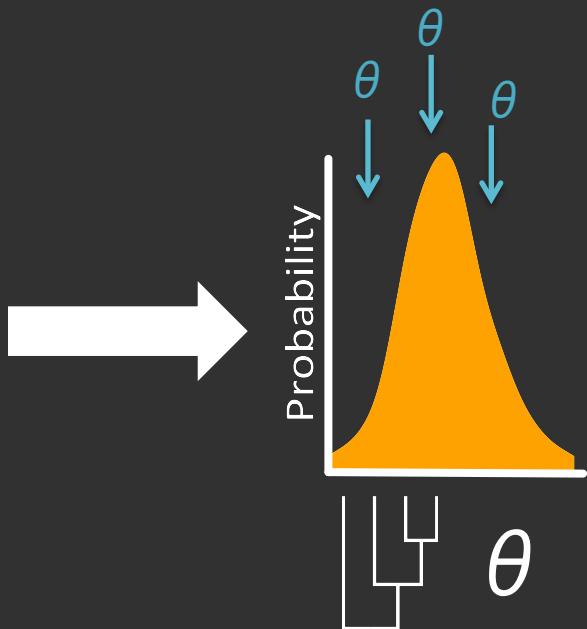
- Are our current models working the way we think they are?
- How wrong are they? And does it matter?



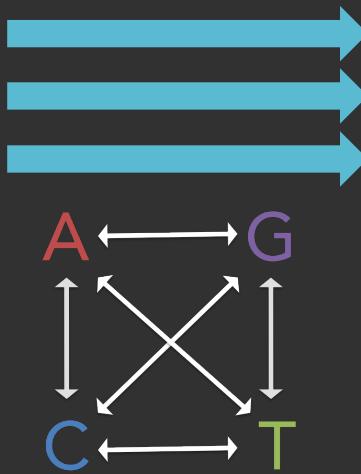
Model Behavior



Empirical Data & Model



Distribution of phylogenies and parameters



Expected distribution of trees

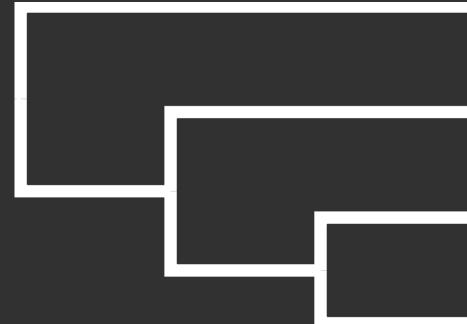
What I use computers for: Behavior and variation of models and data

- Variation within and across datasets

- Do all genes support the final phylogeny?
- Why do some genes result in different relationships between taxa?

```
TTCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
CCAACGGCGGGGAGAGGGGGTTCGTAATTGGCGCT  
TGAGTCGTATAGAGCTTGGCGTAATCAAGGCATAGC  
TTCCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
TTCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
CCAACGGCGGGGAGAGGGGGTTCGTAATTGGCGCT  
TGAGTCGTATAGAGCTTGGCGTAATCAAGGCATAGC  
TTCCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
TTCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
CCAACGGCGGGGAGAGGGGGTTCGTAATTGGCGCT  
TGAGTCGTATAGAGCTTGGCGTAATCAAGGCATAGC  
TTCCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA
```

```
TTCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
CCAACGGCGGGGAGAGGGGGTTCGTAATTGGCGCT  
TGAGTCGTATAGAGCTTGGCGTAATCAAGGCATAGC  
TTCCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
TTCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
CCAACGGCGGGGAGAGGGGGTTCGTAATTGGCGCT  
TGAGTCGTATAGAGCTTGGCGTAATCAAGGCATAGC  
TTCCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
TTCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA  
CCAACGGCGGGGAGAGGGGGTTCGTAATTGGCGCT  
TGAGTCGTATAGAGCTTGGCGTAATCAAGGCATAGC  
TTCCACACAAACATACGAGCCGGAAAGCAAAAGCTAAA
```



Variation in phylogenomic trees

A

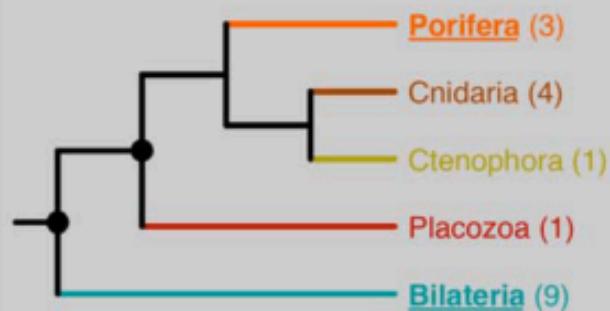
B

C

Schierwater et al. (2009)

15 mitochondrial and 34 nuclear genes

GTR and WAG models



Porifera



Placozoa



Cnidaria



Ctenophora

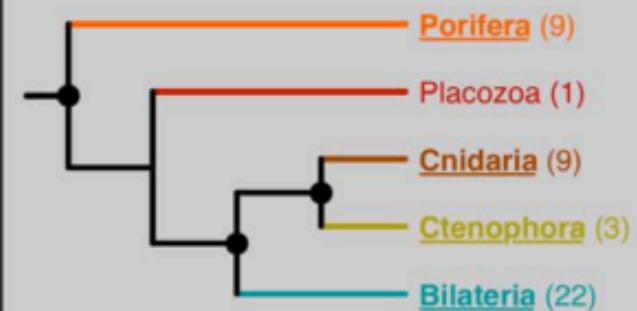


Bilateria

Philippe et al. (2009)

128 nuclear genes

CAT model



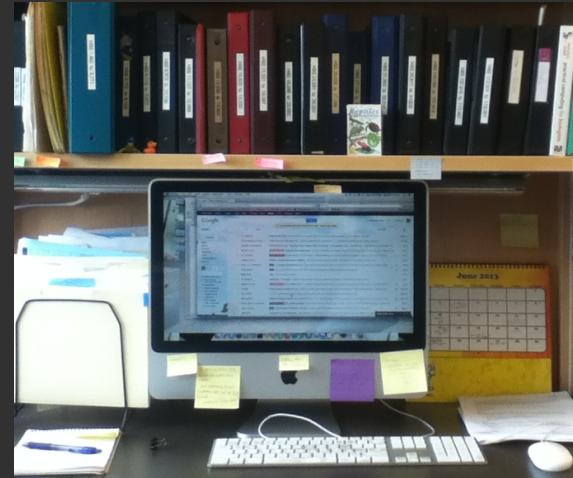
What I use computers for:

Behavior and variation of models and data and evolutionary biologists

- Communicate the importance of thorough data analysis
 - Make these methods easier to use

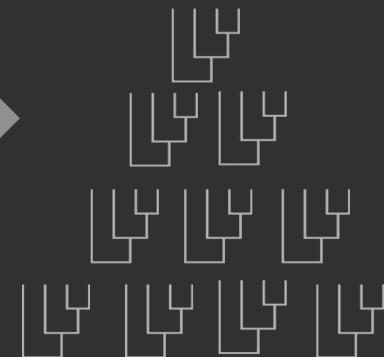
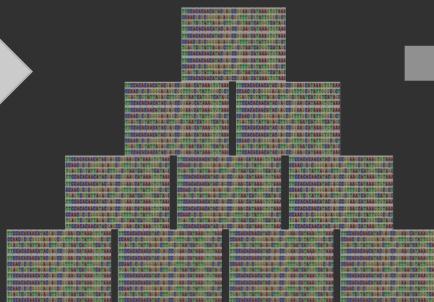
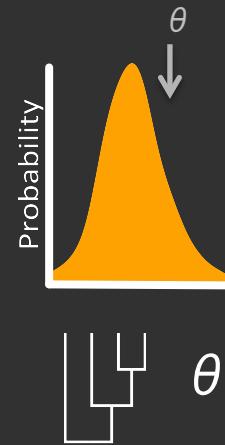
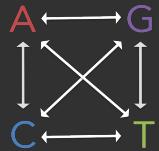
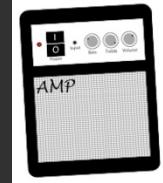
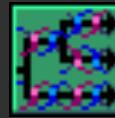


Happy evolutionary biologist

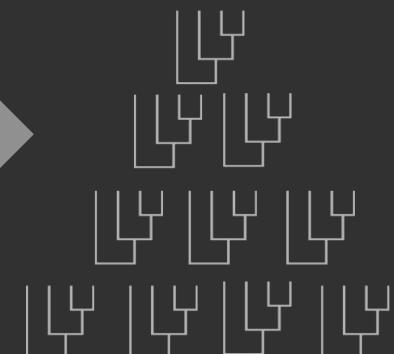
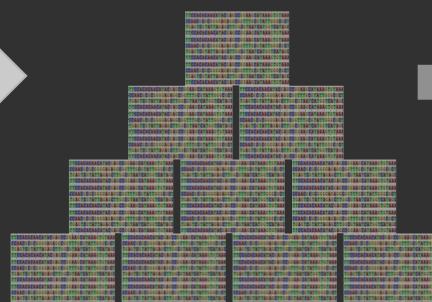
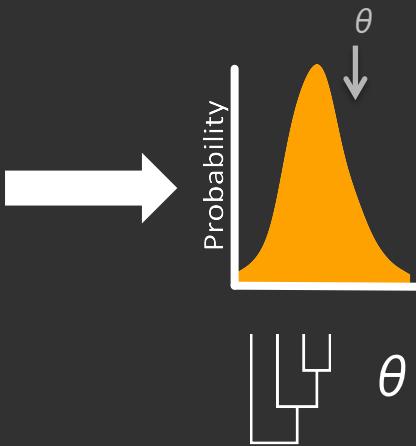
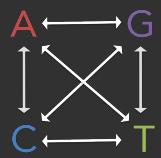


Less happy evolutionary biologist

Unhappy Old Workflow



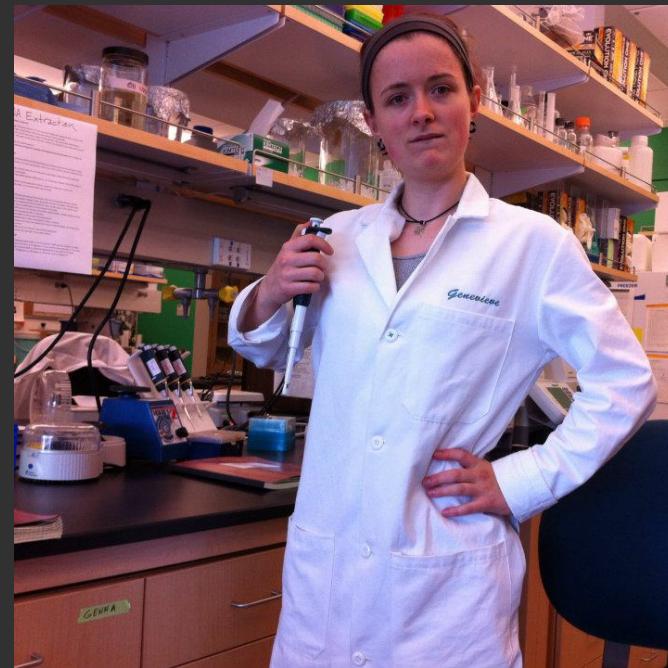
Happier New Workflow



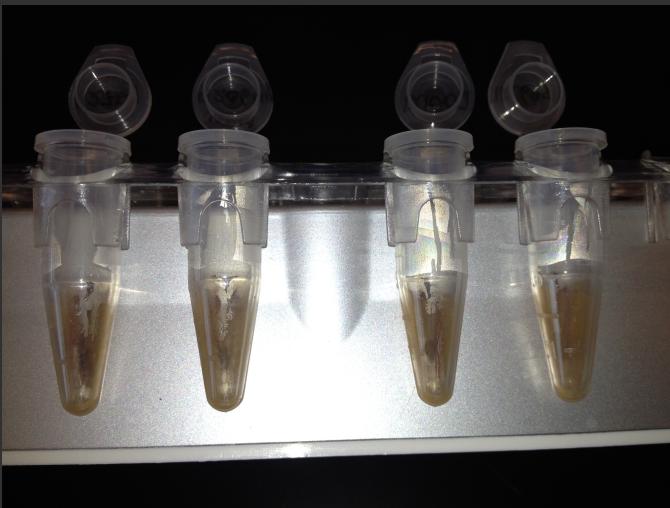
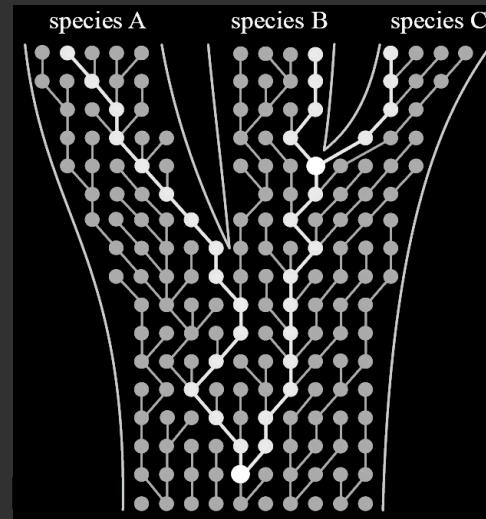
Why me, why computers?



Why me, why computers?



Why me, why computers?



Computers!!!

My first coding challenge.... Ever...

Computers!!!

NySpace.com | Home The Web MySpace Search Help | SignOut

Skin Therapy
Alpha Lipoic Acid, Q10 & MOE Complex Products. Free Shipping.

Alpha Lipoic Acid On Sale
Great Selection At 81% Off
Retail Shop and Save - Buy Now!

vitamin e Sale
vitamin e: Get 20% off or more on vitamin e. Plus Free Shipping \$50.

1 Vilie is 5 Xango, 7 Goli
Highest ORAC score in industry. Listed in PDR at 98% absorbability.

Ads by Google

Home | Browse | Search | Invite | Film | Mail | Blog | Favorites | Forum | Groups | Events | Videos | Music | Classifieds

Tom



Male
30 years old
Santa Monica, CALIFORNIA
United States

Last Login:
5/16/2006

[View My: Pics](#) | [Videos](#)

Contacting Tom

- [Send Message](#)
- [Forward to Friend](#)
- [Add to Friends](#)
- [Add to Favorites](#)
- [Instant Message](#)
- [Block User](#)
- [Add to Group](#)
- [Rank User](#)

MySpace URL:

Tom is in your extended network

Tom's Latest Blog Entry [Subscribe to this Blog]

MySpace Concert & Parties -Georgia, Orlando, Miami! (view more)

In Stores Today - MySpace Records Vol. 1! (view more)

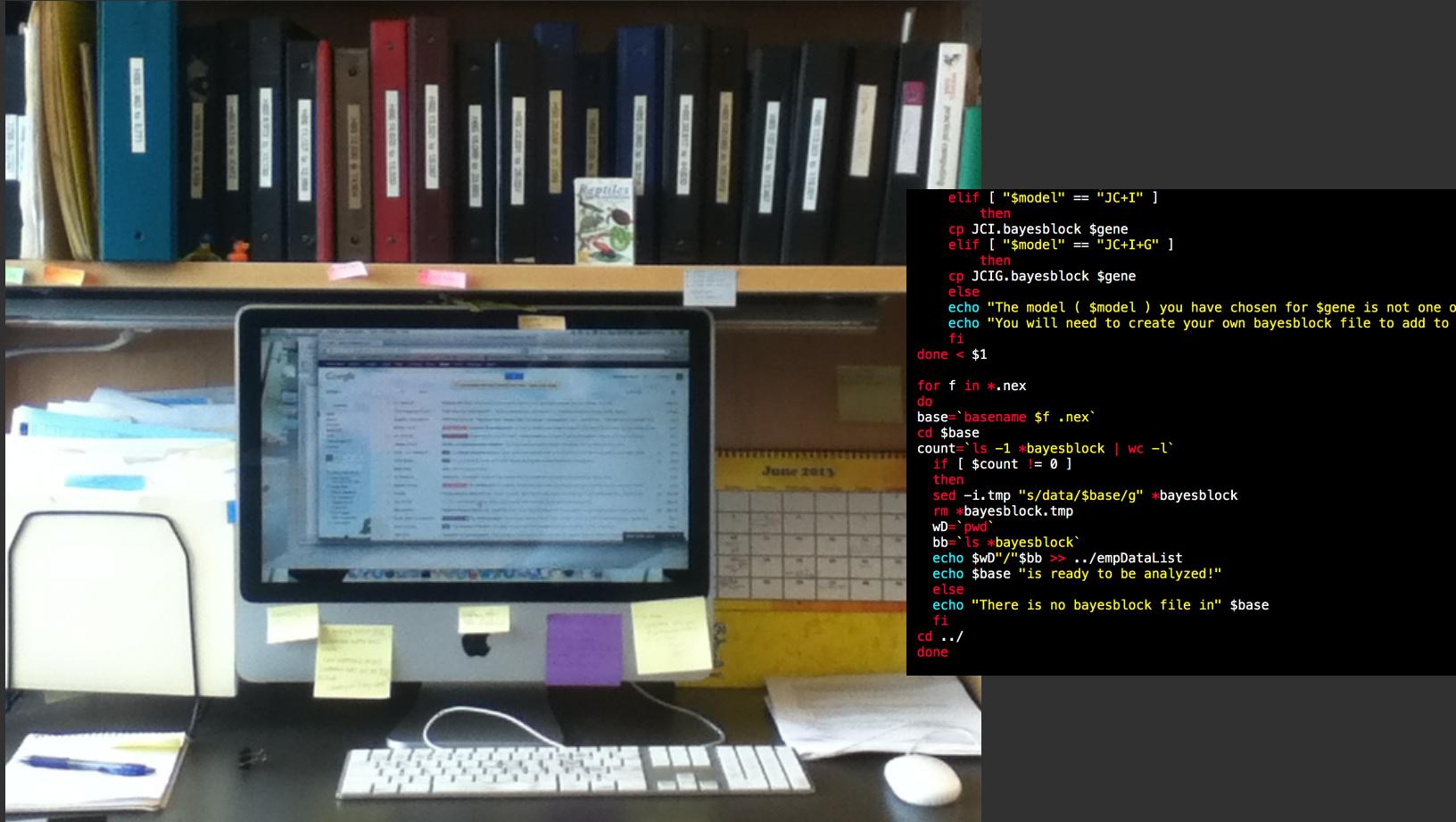
MySpace Records - get more photos for your profile! (view more)

October 29th - MySpace 2-Year Anniversary Concert! (view more)

NY, QOTSA, I

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<View All Blo<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en-AU">
<head>
<p>@Page_subIncludes @P<
<title>@P@Computed_name="dspTitle" useKeyword="AppName" @P>: <P@Computed_name="@h
</head>
<body>
<!-- For non-visual user agents: -->
<div id="top"><a href="#main-copy" class="doNotDisplay doNotPrint">Skip to main
<p>This is a test</p>
<!-- ##### Header ##### -->
<div id="header">
<div class="superHeader">
<div class="left">
<span class="doNotDisplay">Current Host:</span>
<P@UserName value="abbreviated" @P> @ <P@Computed_name="@Host" @P>
</div>
```

Computers!!!



Acknowledgements

Questions?

*Herpetology and other
friends:*

Chris Austin Brad Shaffer

Zach Rodriguez Ben Johnson

Sara Ruane Phil Spinks

Matt Brady Muge Gidis

Jessica Oswald Sarah Hendricks

Clare Brown Hilary Rollins

Computer friends:

Jeremy Brown

Sebastian Höhna

Bob Thomson

Lyndon Coghill

David Morris



HPC | **PCP** HIGH
PERFORMANCE
COMPUTING