Conch Phylogeny

This virtual laboratory explores the phylogenetic history of a group of gastropod mollusks, the conchs. The exercise is based on a live, in-class laboratory that was created by Dr. James Krupa and Dr. Andrew Bouwma-Gearhart, from the UK Biology Department.

We will examine 18 species of conchs (Family: Strombidae; genera: *Strombus* and *Lambis*) and one species of pelican's foot (Family: Aporrhaidae; genus: *Aporrhais*), the outgroup.

You will compare and contrast two phylogenetic trees for 18 species of conchs, and their outgroup, the pelican's foot: 1. a phylogenetic tree based on 11 morphological characters (or traits); and, 2. a phylogenetic tree based on the sequence of the Histone H3 (his3) gene. You will construct your phylogenetic trees with a software package known as PHYLIP (PHYLogeny Inference Package), created by Joseph Felsenstein at the University of Washington. Although Prof. Felsenstein provides this software package free of charge, for this lab we find it easier to access it through on online bioinformatics portal known as Mobyle, at the Pasteur Institute in France. This lab is divided into two exercises, one for each phylogenetic tree.

Exercise 1: A tree based on morphological data. (2+ hours)

Backround: Geerat Vermeij is perhaps the leading authority on morphological evolution of mollusks in general, and of gastropod mollusks in particular. What's particularly interesting about Prof. Vermeij is that he's been blind since the age of 3. He was introduced to mollusk shells as child, and fell in love with them immediately. Even as a child, he described their amazing beauty, purely with his sense of touch. Dr. Krupa, who created the live version of this lab, brings his collection of conchs to the teaching laboratory, and students become intimately familiar with these shells, both tactilely and visually. We will attempt to recreate this experience visually with 2D photos and interactive 3D models (see below).

For background information, Prof. Vermeij was a featured scientist in the PBS series, The Shape of Life, The Survival Game. Here's a link to the 54-minute video on youtube's Evolution Documentary Channel, and here's a link to Vermeij's biography on the PBS Shape of Life website.

Online Analyses and Software: For this first exercise, we will use 4 websites for access to data and software (the fourth one, ImageJ is optional).

1. The specimens: In the live version of this class, students hold one or more specimens of each species in their hands and take a variety of measurements. For this online virtual lab, we have provided you 2D photographs, and interactive 3D models here: http://darwin.uky.edu/~sargent/bio303/conchlab/models/. On the left side of the screen select a species, and you will see a 3D model. You can use your mouse to rotate and zoom in and zoom out for each 3D model. Also under each species (lower left of your screen) is a pair of links to dorsal and ventral photographs. From these images, you will record 10 characters/traits for each species (see below).
   Note: this website will NOT work in Internet Explorer, but works fine in Firefox, Chrome and Safari. Also, the interactive 3D models will not work on touch screens (e.g. iPads, other tablets, smart phones). You will need a PC or a Mac, with a mouse.
2. pars: pars is a program for constructing phylogenetic trees from discrete data (i.e. counts and data coded categorically as numbers), and is based on the assumption of parsimony. We will access this program through Mobyle's online PHYLIP portal. Documentation for how this program works is available on Felsensteins website: pars documentation.
3. FigTree: FigTree is a free stand-alone program that takes the output from your PHYLIP analysis and plots it as a tree.
4. ImageJ: ImageJ is a free image-analysis program that you may find helpful for taking measurements off the 2D conch photos (for traits 5 and 6 below). ImageJ is available free from the NIH, and you may find it broadly useful for all kinds of image analyses that are done routinely in the biological and biomedical sciences.

Morphological Traits: Below is a list of descriptions for each trait, illustrated with photographs.

1. **Posterior projection of the outer lip: 0** absent, **1** present and rounded, **2** long and pointed. This is defined as portion of outer lip that extends behind the aperture (see illustration).
2. **Thickening of the outer lip: 0** none, **1** slight, **2** distinct. This is referring to how thick the outer lip is with respect to the adjacent areas of the shell.
3. **Digits on the outer lip: 0** none, **1** small, **2** moderate, **3** large. We will define digits as the extensions on the outer lip that occur below the shoulder, thus digits on posterior project do not count.
4. **Flaring on the outer lip: 0** none, **1** slight, **2** extreme. When viewed from above (the dorsal view), outer lip extends out not down.
5. **Length of adult shell: 0** <3cm, **1** 3-6cm, **2** 6-12cm, **3** >12cm. This is measured from the tip of the spire to the base of the aperture (see illustration), and is easiest in the dorsal view. You may stretch a piece of string over this distance (up and down on your computer monitor), and then calibrate your length of string to the ruler in the photo. Alternatively, you may cut and paste your photo into ImageJ, and use that to measure your length.
6. **Spire height as % of shell length**: **0** <20%, **1** 20-40%, **2** >40%. Spire height is measured from the shoulder where it flows into aperature and above this point.
7. **Shoulder spines: 0** absent, **1** short and knobby, **2** long and somewhat pointed. At least one shoulder spine has to be long and somewhat pointed to score the entire shell as a "2."
8. **Body whorls: 0** no knobs, **1** knobs. This is the body of the shell below the shoulder.
9. **Shell texture: 0** smooth, **1** rough. A shell and have knobs on body and still be smooth, just as a shell can have no knobs and still feel rough.
10. **Striations in aperture: 0** none, **1** inner edge, **2** outer edge, **3** both. Striations are about texture (like a washboard) rather than color (some striations have light ridges and dark valleys; others are light all over). These can be seen in the 2D photos, ventral view. It may be easiest to see these striations, or their lack, on the original 14.2Mp photos on this striations page. Right click on a photo, click on view image, and use the zoom tool to enlarge.
11. **Stromboid notch: 0** absent**, 1** present. A trait that distinguishes the Strombidae, the true conchs, from the rest of the gastropods is the Stromboid notch. These are easily seen in the ventral photographs, but unfortunately, do not show up well in the 3D models.

Open up the ConchMorph.xls file. It contains two worksheets: “Raw Data” and “PHYLIP Format.” You will enter your data into the “Raw Data” worksheet, which looks like this.

		Traits										
Species	Common name	1. Projection (0 to 2)	2. Outer lip (0 to 2)	3. Digits (0 to 3)	4. Flaring (0 to 2)	5. Length (0 to 3)	6. Spire (0 to 2)	7. Spines (0 to 2)	8. Whorl (0 to 1)	9. Texture (0 to 1)	10. Striations (0 to 3)	11. Notch (0 ti 1)
1. *Aporrhais pespelecani*	Common pelican's foot	x	x	x	x	x	x	x	x	x	x	x
2. *Lambis chiragra*	Chiragra Spider conch	x	x	x	x	x	x	x	x	x	x	x
3. *Lambis lambis*	Common Spider conch	x	x	x	x	x	x	x	x	x	x	x
4. *Strombus gigas*	Queen/pink conch	x	x	x	x	x	x	x	x	x	x	x
5. Strombus *gallus*	Rooster-tail conch	x	x	x	x	x	x	x	x	x	x	x
6. *Strombus alatus*	Florida fighting conch	x	x	x	x	x	x	x	x	x	x	x
7. Strombus *pugilis*	West Indian fighting conch	x	x	x	x	x	x	x	x	x	x	x
8. Strombus *bulla*	Bubble conch	x	x	x	x	x	x	x	x	x	x	x
9. *Strombus vomer*	Vomer conch	x	x	x	x	x	x	x	x	x	x	x
10. *Strombus microurceus*	Micro conch	x	x	x	x	x	x	x	x	x	x	x
11.Strombus *urceus*	Little bear conch	x	x	x	x	x	x	x	x	x	x	x
12. *Strombus vittatus*	Vittate conch	x	x	x	x	x	x	x	x	x	x	x
13. *Strombus costatus*	Milk conch	x	x	x	x	x	x	x	x	x	x	x
14. *Strombus. aurisdianae*	Diana conch	x	x	x	x	x	x	x	x	x	x	x
15. *Strombus canarium*	Dog conch	x	x	x	x	x	x	x	x	x	x	x
16. *Strombus peruvianus*	Peruvian conch	x	x	x	x	x	x	x	x	x	x	x
17. *Strombus epidromus*	Swan conch	x	x	x	x	x	x	x	x	x	x	x
18. *Strombus gibberulus*	Humpback conch	x	x	x	x	x	x	x	x	x	x	x
19. *Strombus fragilis*	Fragile conch	x	x	x	x	x	x	x	x	x	x	x

The first two columns contain species names, the Latin binomial followed by the common name. The next 11 columns are where you put your numerical value for each of the ten traits for each species (i.e. replace each x with a number).

The “PHYLIP Format” worksheet concatenates the data for trait values into a form that PHYLIP “wants” for input. It looks like this.

```
          19 11
Apespelec xxxxxxxxxxx
Lchiragra xxxxxxxxxxx
Llambis   xxxxxxxxxxx
Sgigis    xxxxxxxxxxx
Sgallus   xxxxxxxxxxx
Salatus   xxxxxxxxxxx
Spugilis  xxxxxxxxxxx
Sbulla    xxxxxxxxxxx
Svomer    xxxxxxxxxxx
Smicrourc xxxxxxxxxxx
Surceus   xxxxxxxxxxx
Svittatus xxxxxxxxxxx
Scostatus xxxxxxxxxxx
Saurisdia xxxxxxxxxxx
Scanarium xxxxxxxxxxx
Speruvian xxxxxxxxxxx
Sepidromu xxxxxxxxxxx
Sgibberul xxxxxxxxxxx
Sfragilis xxxxxxxxxxx
```

The numbers, 19 11, in the first line tells PHYLIP that your data set consists of 19 species, with 11 traits measured for each species. Each row contains a PHYLIP formatted, abbreviated species name, followed by a string of 11 x’s, all run together (concatenated). In the “Raw Data” worksheet, you will replace x for each trait for each species with its numerical value (0, 1, 2, whatever). The “PHYLIP Format” worksheet contains a formula that will concatenate those trait numbers for you into strings of ten numbers with no spaces between them.

For example, the data for *Strombus microurceus* might change from…,

```
Smicrourc    xxxxxxxxxxx
```

…to…

```
Smicrourc    02010000031
```

This happens for you automatically.

To run your data in pars in PHYLIP, do the following.

1. Open pars on the Mobyle website.
2. Copy and paste your "PHYLIP Format" data (only the first two columns and first twenty rows that contain text) into the Input File box at the top of the pars program.
3. Under "Bootstrap and Jumble options" (3 fields down), click on "choose a method" and select "randomize input order of sequences (Jumble)." Enter any odd random number in the next box, and enter the number,10, in "Number of times to jumble."
4. In your last box, you indicate which species' row contains your outgroup species. The default is row 1, which is why we put the outgroup, *A. pespelecani* in row 1. So, make sure that box has the number, 1, in it.
5. Click the "Run" button at the top of the page. To make sure you're not a "bot," the program will ask you for your email address, and to type in the word you see into a box. Do that and wait for your output.
6. You should get a crude tree in the first box. You can see it better by clicking the "full screen view" button under the top box.
7. In the second box are data for your output tree. Click the "save" link above the second box, which will save a file called, pars.outtree on your computer in the "downloads" folder.
8. On your computer, open the program, FigTree.
9. In FigTree's File menu, select open, and open pars.outtree from your "downloads" folder. FigTree should now display a tree for you.
10. Within FigTree's tree, click on the branch and label, Apespelec, which should highlight that branch.
11. Now click on the "Reroot" button at the top of FigTree, which will change your tree's appearance slightly. Your tree in FigTree should have the same layout as the output box for the pars program (step 6 above).
12. In FigTree's File menu, click on "Export to pdf..." and save your file as something you'll remember with the .pdf extension, like morphologytree.pdf. Keep this file handy, because you will answer questions about this tree below, and you will turn in this figure with your lab report.

Questions:

1. Were there any monophlyletic groups in your tree? If so, what were they?
2. Were there any paraphyletic groups in your tree?
    a. Is paraphyly a problem?
    b. If so, what might you do with Strombid taxonomy to resolve the problem of paraphyly?
3. What synapomorphy is shared by all of the Strombidae?
4. Which traits appear to be synapomorphies and which appear to be due to homoplasy?
5. Using the criterion of parsimony, map on your tree where you thing the following traits arose:
    a. Stromboid notch
    b. Projections (for this exercise treat all projections as 1 and their lack as 0)
    c. Digits
    d. Striations (for this exercise treat all striations as 1 and their lack as 0)

6. **Exercise 2: A tree based on DNA sequence data.** (2+ hours)

Background: This exercise is based on a recently published molecular phylogeny of the conchs (Latiolais et al, 2006), which came out of the labs of Michael Hellberg of LSU and Kaustuv Roy of UCSD. The phylogeny is based on a nuclear gene, histone H3 (his3) and a mitochondrial gene, cytochrome oxidase I (COI). For this exercise, we will focus only on his3.

Online Analyses and Software:

1. GenBank: GenBank is a publically accessible, NIH funded repository for published DNA and genome sequences. You will obtain the 327bp his3 sequence for each species from GenBank.
2. Clustal: Clustal is an online resource for analyzing sequence alignments for DNA sequences. We will access ClustalW2 through the EBI server.
3. dnapars: dnapars is a program within PHYLIP for constructing phylogenetic trees based on DNA sequence data. We will access this program through the Mobyle portal.
4. FigTree: FigTree is a free stand-alone program that takes the output from your PHYLIP analysis and plots it as a tree.

Part I: Gathering the genetic data.

1. Go to GenBank. In the first field, make sure it says, Nucleotide (this field contains all the GenBank databases).
2. You will search the Nucleotide database for his3 sequences for all 19 species above, in the same order as above. In the GenBank search field type, Genus species his3, so for the outgroup, type Aporrhais pespelecani his3.
3. You will see a page that provides information on this gene for this species. Make sure its for 327bp (line 1), and that the reference is the same paper cited above.
4. In the "Display Settings" at the top of the page, select FASTA (text), which is a standard format for later analysis. For this species, it should look like…

```
>gi|106896191|gb|DQ525277.1| Aporrhais pespelecani histone
H3 (his3) gene, partial cds
CGCAAGTCCACCGGAGGAAAAGCTCCTCGCAAGCAGCTGGCCACCAAGGCCGCACGTAAAAGTGCTCCCG
CCACCGGCGGTGTCAAGAAGCCCCATCGTTACAGGCCCGGAACCGTGGCTCTCCGTGAGATCCGTCGTTA
CCAGAAGAGCACCGAGCTGCTGATCCGCAAGCTGCCCTTCCAGCGTCTGGTGCGTGAGATCGCCCAGGAC
TTCAAGACGGACCTGCGCTTCCAGAGCTCGGCCGTCATGGCTCTGCAGGAGGCCAGCGAGGCCTACCTGG
TGGGTCTCTTCGAGGACACCAACCTGTGCGCCATCCACGCCAAGCGT
```

5. Open Notepad and paste the contents there. In Notepad, you will want to truncate the first line so that it contains only, >Apespelicani, so the above sequence now looks like this…

```
>Apespelecani
CGCAAGTCCACCGGAGGAAAAGCTCCTCGCAAGCAGCTGGCCACCAAGGCCGCACGTAAAAGTGCTCCCG
CCACCGGCGGTGTCAAGAAGCCCCATCGTTACAGGCCCGGAACCGTGGCTCTCCGTGAGATCCGTCGTTA
CCAGAAGAGCACCGAGCTGCTGATCCGCAAGCTGCCCTTCCAGCGTCTGGTGCGTGAGATCGCCCAGGAC
TTCAAGACGGACCTGCGCTTCCAGAGCTCGGCCGTCATGGCTCTGCAGGAGGCCAGCGAGGCCTACCTGG
TGGGTCTCTTCGAGGACACCAACCTGTGCGCCATCCACGCCAAGCGT
```

   Make sure there are no spaces.

6. Repeat steps 2-5 for the other 18 species, with a space between each species. Once you have all 19 species in Notepad save the .txt file.
7. Open the ClustalW2 program, and paste your FASTA formatted .txt file from Notepad into the box. Press the Submit button.
8. When your results are ready, scroll down to see the alignments. The aligned sequences will be broken up into blocks of 60 nucleotides. You will want to make sure the sequences are aligned. When you have sites along the sequence where all the bases are identical (called invariant sites), an asterisk will appear below that site. Save the aligned file to answer the questions at the end of this activity.

Part II: Generating a genetic-based phylogenetic tree.

1. Return to the ClustalW2 program above.
2. Again, paste in your FASTA formatted sequences from Notepad, but this time we will set an option to put those data into a form for PHYLIP.
3. Under step 3 in the ClustalW2 program, click on More Options.
4. That box expands and at the bottom of the box where it says Format, click on PHYLIP.
5. Submit the job.
6. When it's done copy and save the PHYLIP formatted sequences as a .txt file in Notepad.
7. Open dnapars. In the top window, paste your PHYLIP formatted DNA sequence data.
8. In the fourth box down, Randomization Options, under Jumble check yes, enter an odd random number, and 10 as the number of times to Jumble.
9. In the last box, Other Options, set the outgroup species number to wherever *Aporrhais pespelecani* ended up in your sequence data list. For me, it was species number 19.
10. Hit run, and answer the email and bot questions. When it's done, you will have a tree in your top box. If you hit full screen, you'll see what it looks like.
11. Save your dnapar.outtree file on your computer in your downloads folder.
12. Open FigTree, open up your dnapars.outtree file, and reroot it to Apespele. Save this tree as a pdf, and print in out to compare it with your previous tree.

Questions:

1. Calculate the percentage of the 327 sites that are variable among the19 taxa.
2. For each cluster of species that appear on your phylogenetic tree, determine which base pairs are different from the other species to function as synapomorphies.
3. Do you have examples of monophyletic groups or paraphyletic groups? What does the presence of paraphyletic groups tell you? If Strombus turns out to be paraphyletic, what does this tell you about the genus? What should a systematist do about a paraphyletic genus?

4. What differences, if any, do you find between your morphological and genetic phylogenetic tree. Which one better produces the more realistic phylogeny? Why?
5. Why might a phylogeny based on a single gene not necessarily reflect the actual phylogeny of the group of species in question?

References:
Strombidae: http://seashellsofnsw.org.au/Strombidae/Pages/Strombidae_intro.htm

Aporrhaidae:

PBS Video – The Shape of Life: The Survival Game

PBS Website – Geerat Vermeij's Biography

Latiolais, J. M., M. S. Taylor, K.Roy and M. E. Hellberg. 2006. A molecular phylogenetic analysis of strombid gastropod morphological diversity. Molecular Phylogenetics and Evolution, 41: 436-444. (pdf)