PCR amplification and DNA sequence analysis of conserved homeobox genes

January 26-27, 2026
Extract DNA from animal tissue:

We have a diverse collection of invertebrates that we obtained from the pond on the Bryn Mawr Campus. We will isolate genomic DNA from these samples and amplify a conserved homeobox gene from these animals using degenerate PCR primers. Students will work individually.

On Monday, January 26, 2026:

1. Put on gloves.
2. Rinse your work area with 70% EtOH (in squirt bottles) and wipe down the area with a couple of KimWipes.
3. Thaw your choice of tissue by removing it from the -20°C freezer and placing it at room temperature for a few minutes. Each student should choose a different organism.
4. Examine the size of your thawed organism.
   A. If the amount of tissue in your sample is small (~2 mm cubed or less), proceed to step 5.
   B. If the amount of tissue is large, place the thawed tissue on a fresh piece of parafilm and use a fresh razor blade to cut a 2 mm cube of tissue off of the larger piece. If the tissue is soft, place it directly into a fresh, labeled 1.5 ml eppendorf tube. If the tissue is firm, cut the 2 mm cube into several smaller pieces, and place the pieces into a fresh, labeled 1.5 ml eppendorf tube. Unused tissue can be refrozen in the original tube.
5. Add 180 µl buffer ATL to the tissue in the tube.
6. Add 20 µl proteinase K to the tissue/ATL mixture. Cap the tube tightly and mix by vortexing for 15 seconds.
7. Place the tube in the 56°C shaking incubator overnight.

On Tuesday, January 27, 2026:

8. Put on gloves and wipe down your work area with 70% EtOH.
9. Vortex your digested sample for 15 seconds.
10. Add 200 µl buffer AL to the sample. Vortex 15-30 seconds to mix.
11. Add 200 µl 100% ethanol to the sample. Vortex 15-30 seconds to mix.
12. Place a DNeasy Mini spin column into a 2 ml round bottom collection tube.
13. Pipet the sample (~600 µl) onto the DNeasy Mini spin column.
14. Spin sample in the microfuge at 8000 rpm for 1 minute. Don't forget to balance your tube when using the microfuge!
15. Pour flow-through into the sink.
16. Transfer DNeasy Mini spin column to a **fresh** 2 ml round bottom collection tube. Dispose of the old collection tube.
17. Add 500 µl buffer AW1 to the DNeasy Minispin column.

18. Spin the sample in the microfuge at 8000 rpm for 1 minute.
19. Transfer DNeasy Mini spin column to a **fresh** 2 ml round bottom collection tube. Dispose of the old collection tube.
20. Add 500 µl buffer AW2 to the DNeasy Mini spin column.
21. Spin the sample in the microfuge at 13,000 rpm for 3 minutes.
22. Carefully remove the DNeasy Mini spin column from the collection tube, being careful to avoid transferring any of the flow-through liquid at the bottom of the collection tube. If your DNeasy Mini spin column contacts the liquid flow-through, discard the liquid by pouring it into the sink, and spin the column/collection tube for one additional minute.
23. Place the DNeasy Mini spin column into a fresh, **labeled** 1.5 ml eppendorf tube. Be sure to label the tube with your initials, your organism, the word "DNA" and the date.
24. Add 200 µl buffer AE to the DNeasy Mini spin column. Incubate at room temperature, 1 minute.
25. Spin sample in microfuge at 8000 rpm for 1 minute.
26. Your DNA is now in the 1.5 ml eppendorf tube and is ready to be used for PCR.

Setting up the PCR:

You will set up a total of three separate PCR reactions. One reaction will serve as a negative control: it will have all of the reagents found in your experimental samples, except it will lack a DNA template. The other two reactions will be experimental samples; the only difference between these two samples will be the amount of template DNA in each of them.

1. Put on gloves and wipe down your work area with 70% EtOH.
2. Using a sharpie, label three 0.5 ml PCR tubes with your initials and 0, 1 or 2. Be sure to label the tubes on the side near the top of the tube.
3. Add the following reagents, in the order shown in the table, to your PCR tubes:

	tube 0	**tube 1**	**tube 2**
ddH_2O	10 µl	9 µl	8 µl
2x PCR master mix	12.5 µl	12.5 µl	12.5 µl
20 µM HoxE primer	1.25 µl	1.25 µl	1.25 µl
20 µM HoxF primer	1.25 µl	1.25 µl	1.25 µl
DNA template	0 µl	1 µl	2 µl

4. Mix the PCR reactions by vortexing for 2-3 seconds. Place on ice until entire class is ready.
5. Place tubes in PCR machine (located in the T. Davis lab, Park 222).
6. The PCR program will run as follows:

94°C 1 minute
37°C 1 minute } repeat for 35 cycles
72°C 1 minute
72°C 10 minutes
4°C indefinitely (until program is cancelled and tubes are removed)

When the program is complete, the PCR tubes will be transferred to the -20°C freezer for storage until next week.

February 3, 2026
Analysis of PCR products using agarose gel electrophoresis:

1. Put on gloves. **Note:** It is **very** important that you wear gloves whenever you are handling any of the equipment for gel electrophoresis, because it almost certainly contains traces of Ethidium Bromide (EtBr). EtBr is a DNA intercalating agent that we will use to stain the DNA in our agarose gels. It is a mutagen, and you want to avoid direct contact with it.

2. (**NOTE**: step 2 will be done in advance for you) Preparation of a 2% agarose gel:
   -select a 250 ml Erlenmeyer flask, marked with an EtBr label
   -add 2 g agarose to the flask
   -add 100 ml 1x TAE buffer to the flask
   -gently stuff a folded KimWipe into the mouth of the flask
   -gently swirl the agarose/buffer mixture
   -microwave the mixture for 1 minute at 70% power
   -carefully remove the flask, remove the KimWipe, and gently swirl
   -continue to microwave in 1 minute (or less) intervals until the agarose is dissolved
   -allow the gel mixture to cool on the benchtop for 5-10 minutes;
   the mixture is cool enough when the flask can be held comfortably in your gloved hand
   -add 2 µl 10 mg/ml EtBr to the molten agarose and swirl gently
   -pour the agarose into the electrophoresis apparatus; **make sure that a comb has already been inserted!**
   -immediately rinse the empty Erlenmeyer flask several times with water and invert to drain

3. Place three 1.5 µl drops of 10x loading dye onto a piece of parafilm. The dye has 3 uses: (1) it lets you see the sample as you pipet it into a well of the gel, which itself is submerged in buffer; (2) its density is quite high due to a high concentration of sucrose – this ensures that your sample will be well-behaved and fall straight into the well as you pipet it; and (3) the dye is of particular size that's about the same as a 300-bp DNA fragment – so although the DNA itself is not visible while the gel is running, its progress can be tracked by watching the dye move through the gel.

4. Add 15 µl of PCR reaction from each of your samples to ONE of the three drops of dye so that you have one drop with 15 µl from tube 0, one drop with 15µl from tube 1, and one drop with 15 µl from tube 2. Keep track of which sample each drop corresponds to.

5. Turn the tray containing the solidified agarose gel 90° so the wells are adjacent to the black (negative) cathode. Add enough 1x TAE buffer to completely cover the gel. Carefully pull out the comb and rinse it immediately in water. Add 6 µl 10 mg/ml EtBr to the buffer that is located **opposite** the wells (at the "bottom" of the gel).

6. Load your samples as follows:
   -4 µl pBR322-*Msp*I ladder (MW marker) – **note**: load ONE lane with MW marker per gel
   -15 µl tube 0 drop
   -15 µl tube 1 drop
   -15 µl tube 2 drop

7. When all students have loaded the gel, attach the lid to the electrophoresis apparatus. Insert the leads into the power supply. Turn on the power supply and set it to 150 volts.
8. The gel should run for ~45 minutes. During this time, the DNA, which is negatively charged, will migrate towards the positive (red) anode. Agarose gels are always set up to "run to red." Fragments of DNA will separate according to size, with the smallest fragments migrating the fastest, and therefore moving the furthest.
9. When the dark blue dye has migrated ~2/3 the length of the gel, turn off the power supply and disconnect the leads from the power supply. Then, carefully remove the lid. Carefully lift up the gel tray (**make sure you are wearing gloves!**) and tilt it slightly to drain off the buffer. Place the tray and the gel into a Tupperware.
10. You will visualize your gel using the GelDocXR, a piece of equipment that will expose your gel to ultraviolet (UV) light. EtBr fluoresces under UV light; therefore, you can visualize the position of DNA within the gel based on fluorescing "bands".
11. Analyze your gel. The minimum size expected is 146 bp; this is the size expected in the absence of introns within the genomic DNA you amplified. It is possible that you will see fragments larger than 146 bp. Smaller bands may represent artifacts or residual primers.

Ligation of putative homeobox gene fragment into a plasmid vector:

In order to generate enough product for DNA sequencing, you will further amplify the product by ligating it to a plasmid vector and allowing the recombinant plasmid to replicate in bacteria. The first step in this process is attaching the PCR fragment to the vector via ligation.

The ligation reaction you will perform takes advantage of the fact that *Taq* polymerase, the enzyme used in your PCR amplification, has a tendency to attach a single-stranded A on the end of otherwise double-stranded DNA fragments. This overhanging A is available to hydrogen bond with a complementary T. The vector you will use has been engineered as a linear molecule with an overhanging T. Therefore, the T on the vector and the A on your PCR fragment will hydrogen bond, generating a circular molecule, and DNA ligase will covalently connect the sugar-phosphate backbone. This technique is called "T/A cloning". Many different biotech companies sell T/A cloning kits; the one we will use is from Promega.

1. Put on fresh gloves and wipe down your work area with 70% EtOH.
2. Set up a ligation reaction as follows:
   5 µl 2x ligation buffer (gray capped tube)
   3 µl PCR product (either from tube 1 or tube 2)
   1 µl pGEM-T Easy vector (yellow capped tube)
   1 µl T4 DNA ligase (green capped tube)
   10 µl total volume
3. Mix by vortexing, 5-10 seconds.
4. Incubate the ligation in the 4°C refrigerator until next week.

February 10, 2026
Transformation of ligation products into bacterial cells:

Competent cells are very fragile and should be treated delicately. Follow all instructions carefully in order to maximize transformation efficiency.

1. Put on fresh gloves and wipe down your work area with 70% EtOH.
2. Thaw competent DH5α cells on ice.
3. Aliquot 100 µl DH5α cells into a 1.5 ml eppendorf tube and place the tube in your ice bucket. DO NOT pipet up and down repeatedly as this may lyse the cells.
4. Add 5 µl ligation mix to the competent cells.
5. Mix by flicking the tube gently and return it to the ice bucket for 30 minutes.
6. Heat shock the cell/ligation mixture by incubating the eppendorf tube in the 42°C heat block for exactly 45 seconds.
7. Return the tube to the ice bucket immediately. Incubate for 2 minutes on ice.
8. Add 900 µl LB broth to the cell/ligation mixture.
9. Incubate your sample at 37°C with shaking, 45 minutes - 1 hour.
10. Using a sharpie, label two LB-amp^{100} plates with the following information: DH5α, your sample type, your name or initials and the date. On one of the plates, write 10%; on the other plate, write 90%.
11. Carefully pour ~5 glass beads onto the agar surface of each of the plates.
12. Pipet 100 µl of your transformation onto the 10% plate. Put the lid on the plate. Spread the cells around by rapidly moving the plate up and down several times (12 o'clock to 6 o'clock), then side to side (3 o'clock to 9 o'clock) several times on the benchtop.
13. Microfuge the remaining 900 µl of cells at 13,000 rpm for 15 seconds. Remember to use a balance!
14. Uncap your eppendorf tube and invert it over the sink so that <u>most</u> of the liquid drains out. You want ~100-200 µl of liquid to remain in the tube, so do not shake the tube too hard!
15. Using your P200 set to ~150 µl, resuspend the cells in the residual liquid by pipetting up and down.
16. Pipet <u>all</u> of the cell suspension onto the 90% plate. Put the lid on the plate and spread the cells around as you did in step 12.
17. Tap the glass beads to one side of the Petri dishes. Dump the glass beads into the glass bead recycling container.
18. Invert your plates such that the side containing the agar is on the top. Place the plates on the shelf in the 37°C incubator. Plates may be stacked on top of each other. Incubate the plates overnight at 37°C to allow the transformed bacteria to grow.
19. The instructor will remove the plates from the 37°C incubator the next day and will place them in the 4°C refrigerator. You may come in to look at them at any time.

Next week, you must come in <u>the day before your scheduled lab</u> section to count colonies and inoculate miniprep cultures. This should be done <u>after 4 pm</u> to prevent the cultures from overgrowing.

February 16-17, 2026
Isolating plasmid DNA from transformed bacterial cells:

On Monday, February 16, 2026: Inoculation of overnight cultures

1. Put on gloves and wipe down your work area with 70% EtOH.
2. Remove your transformation plates from the refrigerator. Look at the colonies and estimate the fraction of colonies that are white vs. blue, and record this estimate in your lab notebook. Keep in mind that white colonies are the best indicator of success in cloning your PCR product into the plasmid vector.
3. From the refrigerator, obtain 6 tubes, each containing 2 ml LB-amp. Label each tube with your initials.
4. Inoculate each culture by touching a **<u>white</u>** colony with a sterile toothpick, and then dropping the toothpick into a tube.
5. Place your cultures in the 37ºC incubator with shaking.

On Tuesday, February 17, 2026: Isolate plasmid DNA from bacterial cells

The overnight growth of your bacteria was in selective medium (nutrient broth plus the antibiotic ampicillin), ensuring that bacterial cells would grow only if they maintained the plasmid taken up during the transformation. Each liquid culture now has enough bacterial cells to provide the raw material to purify the plasmid. Cells will be lysed in a basic solution with detergent, then cellular debris and the bacterial chromosome will be precipitated when this solution is neutralized. The plasmid remains in the supernatant, binds to the matrix of a small column, and is recovered in a small volume for analysis.

1. Put on gloves and wipe down your work area with 70% EtOH.
2. Collect your miniprep cultures. Label the microfuge tubes with your initials and pour 1.5 ml of each culture into the appropriate tube. Centrifuge for 1 minute to pellet the cells.
3. Pour the supernatant culture medium into the sink. Tap the inverted tubes on a paper towel to remove any residual liquid. Use a Kimwipe to remove excess liquid from the tube if necessary.
4. Resuspend the cells in 250 µl buffer P1 by vortexing or by pipetting up and down. Use a fresh tip for each tube. (**<u>Hint</u>**: after you pipet the P1 into the tube, use your P200, set at 100 µl, to resuspend the cells - this will prevent the introduction of a lot of bubbles into the suspension.) It is important that there are not any clumps of cells left at the end of this step. If you are not certain you have resuspended the cells properly, please check with the instructor.
5. Add 250 µl buffer P2 to each tube. Mix by gently inverting 4-6 times. **<u>Do not</u>** pipet up and down. This step lyses the cells: the liquid should become less turbid and more clear. Do not let cells lyse for longer than 5 minutes.

6. Add 350 µl buffer N3 and mix immediately by inverting 4-6 times. You should see a cloudy white precipitate form from the detergent in buffer P2, cellular debris, & the bacterial DNA.
7. Centrifuge the mixture for 10 minutes.
8. Obtain 6 columns, each with a collection tube, and label them uniquely for each sample with your initials and a number (for example, I would label mine TD1, TD2, etc.).
9. After pelleting the precipitate, pour each supernatant into the appropriately labeled column. With each column in its collection tube, centrifuge for 1 minute to bind the plasmid DNAs to the column matrix. Discard the flow-through by pouring it into the sink and replace the columns in the same collection tubes.
10. Add 750 µl buffer PE to each column and centrifuge for 1 minute. This washes out any remaining impurities and the salts from the lysis step, but leaves the plasmids bound to the columns. Discard the flow-through by pouring it into the sink and replace the columns in same collection tubes.
11. Re-centrifuge for an additional 1 minute to remove traces of buffer PE from the columns.
12. Label 6 fresh 1.5 ml microfuge tubes with your initials, an identifying number, and the date. Place the columns in the appropriate fresh microfuge tubes.
13. To elute the plasmids from the columns, pipet 50 µl buffer EB into each column. Don't pipet the EB buffer onto the walls of the columns; center the pipet tip over the white matrix and watch the buffer as it is dispensed onto the surface of the matrix. Incubate the columns at room temperature for 1 minute.
14. Centrifuge for 1 minute. Discard the columns. The 1.5 ml microfuge tubes contain your plasmid DNA.

Analysis of plasmid DNA using restriction digestion

Next, you will confirm that you have isolated plasmid DNA by performing a restriction digestion followed by gel electrophoresis. Occasionally, minipreps yield plasmid DNA that fails to contain the desired insert or is not sufficiently concentrated or pure enough for sequencing. By running a sample of digested miniprep DNA on a gel, you will be able to assess which of your samples will be the best for sequencing. You will digest your plasmids with the restriction enzyme *Eco*RI, which should cut the plasmid twice, resulting in two linear products: the plasmid vector (~3000 bp) and the insert (~145 bp).

1. Put on gloves and wipe down your work area with 70% EtOH.
2. Label 6 microfuge tubes, one for each of the 6 digestion reactions; make sure the labels are unique and correspond to each of your miniprep samples (ie: dig TD1, dig TD2, etc; labeling these with "dig" will help to distinguish your digested miniprep DNA samples from your undigested miniprep DNA). Label a seventh microfuge tube "MM" for "master mix".
3. In the master mix microfuge tube, mix the following reagents <u>in the order shown below</u>:
   - 70 µl distilled water
   - 10.5 µl 10x *Eco*RI buffer
   - 7 µl *Eco*RI enzyme

   Vortex for 2-3 seconds to mix the reagents together.

4. Pipet 12.5 µl of the master mix from step 3 into each of the 6 **empty** restriction digestion tubes ("dig"). **DO NOT** pipet this mixture into the tubes containing your miniprep DNA!
5. Add 2.5 µl of each miniprep DNA to the appropriate restriction digestion tube. Vortex 2-3 seconds to mix.
6. Incubate the reactions for 45-60 minutes at 37ºC.

Analysis of restriction digestion products using agarose gel electrophoresis:

A 2% agarose gel will be prepared for you in advance of this lab session. To review the procedure for preparing an agarose gel, see the protocol from February 6th, step 2. **You must wear gloves any time you are handling the gel, including when you are loading your samples.** The gel contains ethidium bromide, which binds to DNA and fluoresces pink under UV light. Unfortunately, it's also a mutagen, and you should avoid direct contact with it.

1. After the restriction digestion reaction is complete, add 2 µl of 10x gel loading dye to each digested sample. Review the purpose of the loading dye in step 3 of the protocol from February 6th.
2. Turn the tray containing the solidified agarose gel 90° so the wells are adjacent to the black (negative) cathode. Add enough 1x TAE buffer to completely cover the gel. Carefully pull out the comb and rinse it immediately in water. Add 6 µl 10 mg/ml EtBr to the buffer that is located **opposite** the wells.
3. Load 4 µl of the 2-log ladder into the first well of the gel; only one lane of 2-log ladder is needed per gel. This is a molecular weight marker and will allow you to estimate the size of the bands produced as a result of the restriction digestion.
7. Load 15 µl of each of your 6 digested miniprep samples onto the gel.
8. When the gel is completely loaded with all student samples, place the lid on, plug the leads into the power supply, and turn the power supply on. The gel should be run at ~150 volts. The gel will run for about an hour, after which it will be photographed under UV light. Each student will get a copy of the image for analysis of her restriction fragments.
9. Each student will identify two miniprep samples that contain insert for DNA sequencing and should give the tubes containing these miniprep DNAs to the instructor. Samples will be sent out for sequencing, and the data will be available in lab next week.

February 24 & March 3, 2026
Sequence analysis of the cloned homeobox genes:

DNA sequencing reactions were performed by Eurofins Genomics. Today's lab session will start with a short lecture on DNA sequencing methodologies. The remainder of the lab time will be used for analysis of your sequence data, which will be available on the course Moodle site. Feel free to bring your laptop computer to lab, as you'll be using publicly available databases for analysis of your sequences.

Once you have downloaded your sequence data from the course Moodle site, the first step is to determine how much of the sequence data you have obtained is from your PCR clone, versus how much of the sequence is from the adjacent vector.

1. Open Blast (www.ncbi.nlm.nih.gov/blast). Scroll down to "Specialized searches" and click on "VecScreen"; this link will bring you to a program that will allow you to analyze your sequence for vector sequence.
2. Run your sequence data through this program. The output will tell you what portions of your sequence are likely to be from the cloning vector. In theory, vector sequence will flank your PCR fragment insert. Be sure to note where your insert sequence begins and ends.
3. Delete the vector sequences from your sequence data (simply delete the appropriate nucleotides). In theory, the sequence that remains should be the same size as your original PCR product. Save this new sequence to use for the remaining analyses. All sequences obtained from the class will be posted in a single document on Moodle for further analysis (below).

Next, you will want to determine whether the sequence you've obtained is what you expect… sequence of a homeobox gene!

4. Compare your product sequence to the sequence of the HoxE and HoxF primers (the primer sequences are available in the lab introduction PowerPoint, available on Moodle, and on the handout you received in lab at the beginning of the semester). Which end of your product corresponds to the HoxE primer? Which end corresponds to HoxF? Is the sequence you obtained consistent with the degeneracy in the primers?
5. If your sequence begins with HoxF, you'll want to "flip" it around so that it is in the correct forward orientation. To "flip" the sequence, you need to obtain the complementary sequence in the reverse orientation (the **antiparallel** sequence). You can do this using the sequence editing program found at https://www.bioinformatics.org/sms/rev_comp.html.
6. Once your sequence is in the forward orientation, with the HoxE sequence located at the beginning of the fragment, you can translate your sequence. One program for doing this can be found at: https://web.expasy.org/cgi-bin/translate/dna2aa.cgi. Keep in mind that since you sequenced an internal fragment of the Hox gene you amplified, you should NOT expect your translated sequence to start with Methionine. You can compare the results of your translation with the expected amino acids at the beginning of the sequenced region (available in the handout and in the Hox lab introduction PowerPoint on Moodle) to ensure that you are using the correct reading frame for subsequent analyses. Based on this information, does is look like you cloned a homeobox gene?
7. You can also determine if you amplified the sequence you expected by searching for conserved structural/functional domains. Navigate back to the BLAST homepage (www.ncbi.nlm.nih.gov/blast). Under Specialized searches, click on "CD-search", which will take you to a program that will identify conserved structural/functional domains in your sequence. You can enter either nucleotide or amino acid sequence information into the query box. Based on the information that is returned, did you amplify a homeobox gene?

Next, you will perform nucleotide-nucleotide and protein-protein Blast searches (blastn and blastp) to determine if your sequences are similar to any other sequences in the database.

8. Navigate back to the BLAST homepage (www.ncbi.nlm.nih.gov/blast) and click on Nucleotide BLAST. This program will compare your sequence will all of the nucleotide sequences found in the database. Make sure your search set is "Standard databases (nr etc.)",

and that you select "more dissimilar" or "somewhat similar" under program selection; "highly similar" will probably be too stringent.

9. Analyze the data you obtain. The "E value" given for possible matches indicates the likelihood that a similar sequence would be detected by chance. Therefore, the smaller the "E value", the more significant the hit. And by small, I mean SMALL… 1 x 10^{-7}, for example. Did you obtain any "hits" identical to your sequence? What "hits" did you obtain? Do they make sense based on what you know about your sample? What type of homeobox gene did you obtain? If you had more than one high-quality sequence from your organism, do they both represent the same class of homeobox gene?
10. Go back to the BLAST homepage and perform a protein blast (under Protein BLAST) with the amino acid sequence(s) you generated in step 6. Compare the results you obtain from this search with the results you obtained from the nucleotide blast (steps 8&9). Do you expect them to yield the same results? Do they yield the same results? If not, why might that be the case?

If you have more than one sequence to analyze, you will now compare your two sequences to each other. If you only have one sequence, you can perform this step with sequence data obtained from another student (or two, or three!) or data obtained by students who took this class in a previous year: a single document containing all of this year's class sequences is posted on the class Moodle site and a separate document containing data from previous years is also available on Moodle.

11. Any two sequences suspected of having sequence similarity can be aligned. For example, you can compare your two sequence samples to each other. You can also compare other sequences to yours: pick an invertebrate you expect to be close to yours (two flies for example), and one you expect might be more distant (a fly and a snail, for example). To do this, go back to the BLAST homepage (www.ncbi.nlm.nih.gov/blast) and click on "Nucleotide BLAST". At the bottom of the first box, click on "Align two or more sequences". You will want to select "more dissimilar" or "somewhat similar" under program selection for this search; "highly similar" will probably be too stringent.
12. Align the two sequences you've chosen to analyze; this will allow you to determine how similar they are to each other. Do you expect your sequences to be similar or different? Are they similar across their entire lengths, or do they have regions that are dissimilar (the ends, for example, as compared to the middle)? Are the data consistent with your expectations?
13. Alignments can also be performed for multiple sequences, and you can align all of the sequences obtained by students in our class to generate a cladogram. Navigate to Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo), which will allow you to get both the alignment data and a phylogenetic tree based on the samples you submitted. Make sure you toggle the search tool to "DNA" before you submit your query. When the results are available, click on "Tool Output" or "Results Files" to access the alignment. You can also click on "Guide Tree" and "Phylogenetic Tree" to access those images. Alternatively (or also!), you can try http://www.phylogeny.fr/ (they claim "robust phylogenetic analysis for the non-specialist"… that's us!); scroll down a little and click on "One Click" under Phylogeny Analysis. When running these alignments, make sure that all of your input sequences are in the same orientation – if they're not, the results you get may be misleading because the "forwards" vs. "backwards" sequences will group differently. Once you have executed the

alignment, look carefully at the data that are returned. Does anything strike you as odd? Can you "clean up" the output by eliminating problematic samples (ie: samples with poor sequence quality or samples that are significantly longer than expected)? Does the cladogram make sense? When thinking about whether your cladogram makes sense based on what we know about the phylogeny of the organisms analyzed, you should consider whether your cladogram results will differ if the sequences you align are all from the same class of homeobox gene *vs*. if the sequences you align are derived from multiple different types of homeobox genes.

14. You can also compare your protein sequence with the translated proteins derived from other students' sequences. What results do you get from this analysis? Does the cladogram look the same? Why might the results look different?

Next, you will extend the analysis of the sequences you obtained by examining full-length Hox gene sequences based on their similarity to the sequences you amplified. This part of the lab is still under development and may be modified as the semester progresses; final information will be provided in class.

15. Using the short homeobox sequence you amplified, use blastn to identify a full-length gene that most closely corresponds to your sequence.
16. With this full-length gene sequence, use Blast to identify 5 other homeobox genes from the same species. Why are you able to find multiple homeobox genes within a single species? Align these sequences to determine how similar they are to each other. Are there specific regions of the sequence that have higher levels of similarity than others? If so, how might you explain this? Generate a cladogram with these sequences. How do you interpret these data?
17. With the full-length gene sequence you identified in step 15, use Blast to identify homeobox genes from 5 **different** species. Align these sequences to determine how similar they are to each other. Generate a cladogram with these sequences. How do you interpret these data?