Supplemental Materials for Teachers

Investigating the Molecular Mechanism of Evolution: Mutation and Natural Selection
Tina Hubler, Patti Adams and Jonathan Scammell

I. PCR and Electrophoresis (background information and student activity)  pg. 2

   PCR
   Electrophoresis

II. Sequences  pg. 6

   Squirrel monkey variable region
   Owl monkey variable region
   Human CG
   Squirrel monkey CG
   Owl monkey CG
   Human growth hormone
   Human oxytocin
   BLAST results: sm vs. om variable regions
   BLAST results: identification of an Alu sequence in sm variable region

III. Mobile genetic elements  pg. 12

   Transposon
   Retrotransposon
   Alu element

IV. Icefish materials  pg. 14

   Making of the fittest
   Into the Jungle

V. Assessment  pg. 15

VI. Vendors for supplies  pg. 16
I. PCR and Electrophoresis (background information and student activity)

PCR

PCR is a powerful tool that allows researchers to produce millions of copies of selected regions of DNA. This quantity of DNA is required for downstream applications such as DNA fingerprinting and DNA sequencing. The *in vitro* copying of DNA in the laboratory follows the same basic steps that occur *in vivo* in the cell each time DNA is replicated prior to cell division. However, there are some important differences. First, denaturing (separating) DNA is accomplished by heating the sample rather than by enzymatic activity (helicase), as it is accomplished in the cell. Second, DNA primers are used to initiate new DNA synthesis, rather than RNA primers that are synthesized by primase in cells. Third, Okazaki fragments are not created because the DNA is completely denatured by heat. Fourth, as a consequence of these differences between *in vitro* and *in vivo* DNA synthesis, a thermostable DNA polymerase is the only enzyme required for synthesis of new DNA strands. The most commonly used thermostable polymerase originated from the thermophilic bacterium *Thermus aquaticus* and is abbreviated as *Taq* DNA polymerase.

The reagents needed for PCR are: (1) a template, the DNA to be copied, (2) nucleotide triphosphates, the nucleotide subunits of DNA, (3) thermostable DNA polymerase, (4) DNA primers complementary to the ends of the DNA region to be copied, (5) buffer, to maintain pH and ionic strength of the reaction mixture, and (6) nuclease-free water. Supplies and equipment needed for PCR are: (1) a PCR cycler with a programmable heat block that allows one to change the temperature of the reaction mixture, (2) PCR tubes to hold the reaction mixture, and (3) pipettors and sterile pipet tips.

After the reagents are added to the PCR tube, reaction mixtures are placed into a PCR cycler and the stages of PCR are initiated. The stages are: (1) denaturation at 92-95 °C, (2) annealing of primers to the template DNA at approximately 45-65 °C, and (3) synthesis of new DNA at approximately 68-72 °C. Denaturation separates the DNA strands so that each may be used as a template for new DNA synthesis. Annealing requires lower temperatures to allow nucleic acids to associate by complementary base pairing. The primers are usually 15-25 nucleotides in length and are complementary to the ends of the region of DNA to be copied. The primer length and composition (percentage of A, T, C, and G) largely determine the annealing temperature. For example, longer primers with higher percentages of C and G are used with higher annealing temperatures. Synthesis of DNA requires the enzyme DNA Polymerase. As is the case for all enzymes, its optimal function depends on pH, ionic strength and temperature. The manufacturer of the polymerase will recommend the optimal conditions and the time needed for the synthesis.

The above three steps represent one PCR cycle that produces two DNA molecules from each template molecule. In the next cycle, the two DNA molecules serve as template and four DNA molecules are produced. The amount of DNA doubles with each cycle (i.e. 1, 2, 4, 8, 16, 32, etc.). This exponential amplification can produce more than 1 million copies of DNA after 20 to 30 cycles.

A typical PCR setup and conditions for a 1 kb template are listed below:

Setup:

- **10X reaction buffer (with MgCl2)**: 5 ul
- **dNTP mix (10 mM of each dNTP)**: 1 ul
- **Taq DNA Polymerase (5 U/ul)**: 0.25 ul
- **downstream primer (50 pmol/ul)**: 1 ul
- **upstream primer (50 pmol/ul)**: 1 ul
- **template DNA (10-100 ng/ul)**: 1-2 ul
nuclease-free water (adjust to a final 50 ul)

**Cycle conditions:**

95 °C  5 min  initial denature step
25-35 cycles:
  95 °C  30 sec  denature step
  55 °C - 65 °C  30 - 60 sec  annealing primer
  72 °C  90 - 120 sec  extension step
72 °C  7 min  final extension step

Additional technical information may be found at websites maintained by Promega Corp. and Integrated DNA Technologies, Inc.


Textbook references for PCR figures:

iGenetics, A Molecular Approach, Third Edition, Russell, ISBN 0321569768, Pearson Education Inc. (Figure 9.3)

Essentials of Genetics, Seventh Edition, Klug, ISBN 0321618696, Pearson Education Inc. (Figure 17.10)

Biology Today and Tomorrow, 4th Edition, Starr, ISBN 9781133364450, Brookes/Cole Cengage Learning (Figure 10.5)

Biology, 9th Edition, Campbell, ISBN 97803215558237, Pearson Education Inc. (Figure 20.8)

**PCR student activity**

As an introductory activity, students will graph the exponential growth of the number of DNA molecules accumulated during PCR cycling. Given the data below, ask students to prepare a graphical representation of the number of DNA molecules that would accumulate during PCR cycling if you start with 1 molecule. This activity may also be used as a review of graphical representation of data. Important points to discuss are: choosing and labeling axes, plotting points and drawing conclusions.

<table>
<thead>
<tr>
<th>cycles</th>
<th>number of molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>10</td>
<td>1024</td>
</tr>
<tr>
<td>15</td>
<td>32768</td>
</tr>
<tr>
<td>20</td>
<td>1048576</td>
</tr>
</tbody>
</table>
Electrophoresis

Electrophoresis is a common technique used in molecular laboratories to separate mixtures of charged molecules. In our activities the molecules are separated according to size. An agarose gel may be used as the supporting medium to separate nucleic acids. The gels are made by dissolving solid agarose in boiling buffer, pouring the partially cooled solution into a casting tray, and allowing it to solidify by cooling to room temperature. The gel casting tray includes a comb that forms sample wells as the gel solidifies. Agarose gels are a porous matrix, with the pore size depending on the concentration of agarose used. Higher concentrations of agarose yield gels with smaller pores. 1.2 - 1.5% agarose gels are used to separate DNA molecules from 200 to 10000 base pairs in size.

The gel is submerged in an electrophoretic chamber containing buffer (usually Tris Acetate-EDTA, pH=8). Samples are loaded into the sample wells and current is applied to the chamber. At pH=8, nucleic acids are negatively charged and travel under the influence of the electric field toward the opposite end of the gel (the positive electrode). Smaller molecules travel through the pores more quickly and migrate further in the gel.

Electrophoresis buffer is used to control pH and ionic strength of the gel and chamber. In contrast, sample buffer containing 10-20% glycerol is sometimes added to samples to cause samples to settle to the bottom of the sample well before electrophoresis begins.

After electrophoresis, sample components are usually not visible unless stained. Agarose gels are commonly stained with either methylene blue or ethidium bromide. Both stains have affinity for nucleic acids. Ethidium bromide is more sensitive and allows one to visualize smaller quantities of DNA. However, it is a mutagen and requires ultraviolet light for visualization. If sufficient quantities of nucleic acids are present, methylene blue is a safer alternative for visualization.

Electrophoresis student activity

To ensure that students understand the principle of electrophoretic separation according to size, ask students to predict the migration of molecules from a mixture of DNA. Assume a DNA ladder is
added to one well in a gel and the following mixture of molecules is added to the second well: 1500 bp, 1000 bp, and 500 bp of DNA. Students can predict how far the individual molecules from the mixture will migrate and draw the positions on the gel below.
II. Sequences

**Squirrel monkey variable region (Alu element bold and underlined)**

```
agttcctctc taccttgtaa ctctccaga cccccggct aggccctgga
actaagaa aa ttctcacta aacaatgctt gcctcagctca tccgttccctc
actctctttc acctctcacc ttgatttccc agaggaggg gggagttgat
aagagaactg tcgagaacag cgtcattta cccgggaacct gcctatgagcc
agggacgtta cagacagctat cttgcttaag tttgacatca tcccatgag
ctgatcttc ttatttccc atattaaca aa tggagaatct gaggaatggg
aaagtttagt gactttgcca agctcacaata atgaagtagt gttaccaggc
agaaagtggc atatactcg tgggaccagc tgcataatga aatagtgggg
cctctgttaa aaaactatta atcgccgggg cgccgtggggt cagacacaggt
cqgaccac ac tcttgcaaca tcggtgaaacccggtctctac taaaaataaca
aaaaggtacc tggccctgggt gggtgccatgc tgtaatccca gctactcagg
aggtggaggg agggaatttg cctgaggcca ggaggcggag gtggcggggtg
ogradgatctcgccaggttatac ctctcacaatg tgggacctt ccagagagtata
gtaatcctgcta ccagagagtc tgcattttagc cgagacttgc tatggaaccag ggaacgggtgct
atggagatg ttcctagcgc aagtcgtagc tgggacgtcc ctcattcactg
```  

**Owl monkey variable region**

```
tagttctctc ataccttgta cctgtccag acacccggccc taggccttga cactgaggag
attctactaa aacaaatgct gcccgagctc atctcctctc atacccctacac ctgatgcc cagagggagg gggaggggg gaagttgann gggagttgac
aagagaactg tcgagaacag cgtcattta cccgggaacct gcctatgagcc
agggacgtta cagacagctat cttgcttaag tttgacatca tcccatgag
ctgatctttc tattttccc atattaaca aa tggagaatct gaggaatggg
aaagtttagt gactttgcca agctcacaata atgaagtagt gttaccaggc
agaaagtggc atatactcg tgggaccagc tgcataatga aatagtgggg
cctctgttaa aaaactatta atcgccgggg cgccgtggggt cagacacaggt
cqgaccac ac tcttgcaaca tcggtgaaacccggtctctac taaaaataaca
aaaaggtacc tggccctgggt gggtgccatgc tgtaatccca gctactcagg
aggtggaggg agggaatttg cctgaggcca ggaggcggag gtggcggggtg
ogradgatctcgccaggttatac ctctcacaatg tgggacctt ccagagagtata
gtaatcctgcta ccagagagtc tgcattttagc cgagacttgc tatggaaccag ggaacgggtgct
```  

**Human CG**  Accession no. NC_000019.9  GI:224589810

```
atgggagatg ttccaggtta gactgccgag ccctctgggc ccctctccac
cccctcgcag aatctcgtgc atgagaatgg gcagacaggct gtagcgtgtg gaagaggtccc
tctctctgag gagaagtgac ccctcagtaa cttcagcttg gccagttctg aaggggtgggg
atgtaatagc tggggcttc ctcgttcttc ttgcttcgtc gttgactctg gaagggcagg
tgtccgggttg gttggttctcg atagagagat gccgaggaaggt ctctcttggt cttggtgggt
```
Squirrel monkey CG  Accession no. GU117708.1

```
   atg gagatgtcct
   aggtaagact gcagggcccc tgggtacctt ccacgcctcc ccaggccatc actggcatga
   agagggcagc agtcgctgtcct gtctggcgtctt gctgggcgtgg tctggccgtc
   cagtaagcct tcatgggaga agttccctgag ggtggagaac gctgtgctgct
   aggctcgggt cagacccagc tctgggcgtct ttctggcgtgct ctctgggagtt
   gcggccgtgc agtcacatgg ctggggtcgtct ccctgttgct gtaagccttc tta
```

Owl monkey CG  Accession no. JN613228  GI:372126675

```
   ctctcagctc ccctgtgtgc ctctgctgtgc gcttgctgtgc gctgtctgtgc
   gcacccagc cactgctgtgc gcgtgtgctgc gcgtgtgctgc gcgtgtgctgc
   gcacccagc cactgctgtgc gcgtgtgctgc gcgtgtgctgc gcgtgtgctgc
   gcacccagc cactgctgtgc gcgtgtgctgc gcgtgtgctgc gcgtgtgctgc
```
Human growth hormone  Accession no. E00140.1  GI:224589808
Human oxytocin  Accession no. NC_000020.10  GI: 224589812

tgtctgtcgtg gctctcctgcc gctgacccctc gctctgcata cacctcagctg gccagccctcg
ggcaagaggg cccgcgcggga cctcgacgtgc gcaaggtgac gccccgctcgc ggcctccgc

gagaactcaca ggcagcgcagc cctcgagcgtg cgacagctgtg gcgcgctcgc gggagccccgg

ttcgcagctcg gctccctggcg cctgacctcc gcctgctaca cccagaactcg ccccctggga

gcaagaggg ccgcgcggga cctcgacgtgc gcaaggtgac gccccgctcgc ggcctccgc
BLAST results: sm vs. om variable regions

Query = squirrel monkey variable region
Sbjct = owl monkey variable region

Query  1    AGTTCCCTCTCTACCTGTCTGCTACCTGTTCCCAGCTTCTGGACACTAGGGAAA  60
          |||||||||||||||||||||| |||||||||||||||||||||||| |  
Sbjct  2    AGTTCCCTCTCTACCTGCTCTGCTACCTGTTCCCAGCTTCTGGACACTGAGGAGA  61
Query  61   TTCTTACTAAACAAATGCTTGCCCAGCTCATCCGTCCCTCACTCTTCTCTACCTCACC  120
          |||||||||||||||||||||  |||||||||||||||||||||||||  
Sbjct  62   TTCTTACTAAACAAATGCTTGCCCAGCTCATCTCCCCTCACTCTTCTCTACCTCACC  121
Query  121  TTGATTCCCCAGAGGAGGAGGG-----------------GAAGTGATAAGAGAAGACTGTCG  163
          |||||||||||||||  |||||||||||||||  |||||||||||||  
Sbjct  122  TTGATTCCCCAGAGGAGGAGGGAAAGGGGGAGGGGAGGGGAAGTGGNNNGAGAATTGACG  181
Query  164  AGAACAGCTGCTATTTACCAGGAGCTTGATTTGGGCCAGGGACTTTAACAGACACATTCT  223
          |||||||||||||||||  |||||||||||||||  |||||||||||||  
Sbjct  182  AGAACAGCTGCTATTTAGCCGGGACTTTGCTATGGGCCAGGGACTTTACAGACACATTCT  241
Query  224  GTCTAAGTTTTGACATCATCCATCCATGAAGTTGATCTTTACTATTATCCCAATTAACAAATGA  283
          |||||||||||||||  |||||||||||||||  |||||||||||||  
Sbjct  242  GTCTAAGTTTTGACATCATCCATCCATGAAGTTGATCTTTACTATTATCCCAATTAACAAATGA  301
Query  284  GAAATCTGAGGCATGGGAAAGTTAAGTGACCTTGCAAGCTACATAATGAGATGTGGGT  343
          |||||||||||||||  |||||||||||||||  |||||||||||||  
Sbjct  302  GAAATCTGAGGCATGGGAAAGTTAAGTGACCTTGCAAGCTACATAATGAGATGTGGGT  360
Query  344  ACCAGGCCAGAACTTCGCTATATTATCTGCTTGGGACCAGGTGCAAAATAGGCAAAATGCTGGGGCCCT  403
          |||||||||||||||  |||||||||||||||  |||||||||||||  
Sbjct  361  ACCAGGCCAGAACTTCGCTATATTATCTGCTTGGGACCAGGTGCAAAATAGGCAAAATGCTGGGGCCCT  420
Query  404  CTGTTTTAATAACTTTAATC    423
          |||||||||||||||  |||||||||||||||  |||||||||||||  
Sbjct  421  CTGTTTTAATAACTTTAATC    440

Note gap in sequence from Alu insertion in sm sequence

Query  722  aaaaaaCTATTTAATCTATTTCAAGACCAAGACCAAGACCAAGACATTAATGCAAGAGTAGGGCC  779
          |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || || |  
Sbjct  426  AAAAAACCATTTAATCTATTTCAAGACCAAGACCAAGACCAAGACATTAATGCAAGAGTAGGGCC  482
BLAST results: identification of an Alu sequence in sm variable region

Color key for alignment scores

Query = squirrel monkey variable region
Sbjct = Alu sequence

Query  424  GGCCGGGCGCGGTGCTCAAGCCTGTAATCCCAAGCACTTTTGGGAGGCCGAGGTGGGTGGA  483
Sbjct  1   GGCCGGGCGCGGTGCTCACGCCTGTAATCCCAAGCACTTTTGGGAGGCCGAGGCGGGCGGA  60
Query  484  TCACAGGGTGCTGAGATCGAGACCATCCTGGTCAACATGGTGAAACCCCGTCTCTACTAA  543
Sbjct  61   TCACAGGGTGCTGAGATCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAA  120
Query  544  AAATACAAAAAGTTAGCTGGGCGTGGTGGCAGTGCTAATCCCAAGCTACTCAGGAGG  603
Sbjct  121  AAATACAAAAA-TTAGCTGGGCGTGGTGGCAGTGCTAATCCCAAGCTACTCAGGAGG  179
Query  604  CTGAGGCAGGAGAATTGCCTGAGCCCAAGCGGAGGCGGAGGTTGCGGTGAGCCGAGATCGCGC  663
Sbjct  180  CTGAGGCAGGAGAATTGCCTGAGCCCAAGCGGAGGCGGAGGTTGCGGTGAGCCGAGATCGCGC  239
Query  664  CATTGCACTCCAGCTGGGTAACAAGAGCGAAACTCCGTCTCaaaaaaaa  713
Sbjct  240  CACTGCACCTCCAGCTGGGCA-C-AGAGCGAGACTCCGTCTCaaaaaaa  287
III. Mobile genetic elements

Transposons

Over half of the human genome can be traced to mobile sequences. Transposons are mobile genetic elements whose DNA sequence can be copied and inserted into new locations within the genome [1, 2]. The DNA sequence included in transposons often encodes a unique set of genes, some of which are involved in transposition and some of which may provide a benefit to the organism (e.g. antibiotic resistance genes) [3]. If an inserted gene confers a biological advantage in the organism (e.g. positive selection), the frequency of the gene may increase in the population. The insertion of mobile genetic elements into various sites of a genome may also have deleterious effects on the organism [4].

Retrotransposons

Retrotransposons are transposons that are copied and inserted into new locations in the genome via RNA intermediates. This involves transcribing a transposon DNA segment into RNA, followed by reverse transcribing the RNA into DNA. The reverse-transcribed DNA is then inserted into a new location within the genome by the action of enzymes (transposases) that are encoded by sequences contained in the the same or another transposon [5].

Retrotransposons are classified as either autonomous or nonautonomous. Autonomous retrotransposons consist of DNA sequences that encode the necessary enzymes for replication and insertion [6]. LINES, such as L1 are examples of autonomous retrotransposons [7, 8]. Nonautonomous retrotransposons are able to recruit the necessary enzymes, but these “trans” factors have been produced from other DNA sequences. SINES, such as Alu, are examples of nonautonomous retrotransposons that require other DNA sequences to achieve their transposition [8].

Alu elements

Alu elements are the largest family of retrotransposons in the human genome [9]. Alu elements are approximately 300 bp in length and are dispersed throughout the human genome over a million times [1]. The exact sequence of an Alu element varies as a consequence of changes incurred during copying and insertion [10]. Sequence variation has produced a number of Alu subfamilies that appear to be primate lineage-specific. For example, the Y subfamily is found primarily in old world monkeys and apes [10, 11].

Insertion of Alu sequences has several impacts on genomes. Not only do Alu insertions increase the size of genomes, they also disrupt regulatory regions and coding regions [12]. Alu insertions have been identified in a number of human disorders including hemophilia, immunodeficiency, cholinesterase deficiency, optic atrophy, hyperparathyroidism, hemolytic anemia, lipoprotein lipase deficiency, complement deficiency, breast cancer and neurofibromatosis [9, 10, 13]. Whether beneficial, neutral or harmful, the movement of mobile genetic elements inevitably increases genome size and produces genetic variation in organisms.
References

IV. Icefish materials


V. Assessment

We suggest short quizzes on the material covered in each lab. Instructors should tailor this to their instruction. Additionally, we suggest a multiple choice quiz on the icefish article, after allowing students to complete the worksheet and ask questions. Finally, a comprehensive assessment should include all topics from lecture and lab. We suggest asking the students to contemplate and prepare answers to the essay question:

In your own words, explain how mutations and selection may result in the evolution of new species. Advise students that their answers should address: 1) where mutations occur in the genome  2) how selective pressures act on a) neutral, b) harmful and c) beneficial mutations and  3) how the genome serves as a historical record of selection. Then the instructor may employ a rubric to score each section of the essay.
**VI. Vendors for supplies (catalog numbers in parentheses)**

Note: The authors have discussed with vendors the possibility of offering the reagents as a commercially available kit. We will update this information if that occurs. Until then, the following vendors are suggested.

<table>
<thead>
<tr>
<th>Materials</th>
<th>Vendor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmids containing h, om or sm CG genes or variable regions</td>
<td>Author Tina Hubler, 256-765-4761, <a href="mailto:trhubler@una.edu">trhubler@una.edu</a></td>
</tr>
<tr>
<td>PCR tubes (TFI-0201 or TBI-0501)</td>
<td>Bio-Rad 1-800-424-6723</td>
</tr>
<tr>
<td>Gel electrophoresis apparatus : Mini-Sub Cell</td>
<td></td>
</tr>
<tr>
<td>Gel System (164-0300 includes power supply, electrophoresis cell, gel tray, comb and gel caster)</td>
<td></td>
</tr>
<tr>
<td>Thermal cycler (PTC-1148C)</td>
<td></td>
</tr>
<tr>
<td>Agarose (161-3103)</td>
<td></td>
</tr>
<tr>
<td>Electrophoresis buffer, 50X Tris/Acetate/EDTA (161-0743)</td>
<td></td>
</tr>
<tr>
<td>Primers for PCR</td>
<td>Eurofins MWG Operon 1-800-688-2248 or</td>
</tr>
<tr>
<td></td>
<td>Integrated DNA Technologies 1-800-328-2661</td>
</tr>
<tr>
<td>DNA ladder (G5711)</td>
<td>Promega 1-800-356-9526</td>
</tr>
<tr>
<td>GoTaq PCR Master Mix (M7112)</td>
<td></td>
</tr>
<tr>
<td>Nuclease free water (DW0991)</td>
<td></td>
</tr>
<tr>
<td>Methylene blue stain (875911)</td>
<td>Carolina Biologicals (1-800-334-5551)</td>
</tr>
</tbody>
</table>