

# Supplemental Material

*CBE—Life Sciences Education*

Harvey *et al.*

**Supplemental Figure 1.** Student self-assessment administered at the beginning and end of each semester.

**Supplemental Figure 2.** Molecular biology knowledge assessment administered at the beginning and end of each semester.

**Supplemental Figure 3.** Python Project Survey administered March 2014.

**Supplemental Figure 4. A.** *In silico* assembly of python transcript assembly. Students used the transcript from a reference species to search the python WGS for similar sequences. The alignments were assembled into the putative python transcript for their gene of interest. **B.** The putative python transcript was validated by translating into an amino acid sequence using ExPasy. Students selected the correct reading frame by determining whether STOP codons had been introduced. The amino acid sequence was then entered into Protein BLAST to determine whether the gene of interest was encoded by the translated transcript. **C.** Primer design and *in silico* validation. Students entered the assembled python transcript into Primer3 to obtain a list of possible primer sets for their gene of interest. Primer BLAST was then used to determine whether the intended product was amplified by the primer sets in a closely related species.

**Supplemental Figure 5.** *In vitro* primer validation. Students performed conventional PCR and cDNA synthesized from the ventricles of 3 day post-fed (dpf) pythons to visualize the size of the product amplified by their primers. Expected product size is calculated during the primer design phase. All products for qPCR are intended to be between 100 (red dashed line) and 200 (blue dashed line) bases long. Free primers and primer dimers run below the 100 base line. Students make an initial assessment of the quality of their primers using these gels. In this example, students 1, 2, 4, 6, 7, and 10 successfully amplified a PCR product of the correct size using their PCR primers.

**Supplemental Figure 6.** qPCR primer validation. Amplification curves (left columns) of primers designed for two genes show the cycle threshold (horizontal green bar) and

provide an estimate of the abundance of the gene in python ventricle obtained from snakes that were 6 days post fed (6dpf) ; students designed two sets of primers for each gene so each gene is shown twice. A single peak in the melt curves (right columns) of each primer set indicates that a single product is amplified in the PCR protocol. Each student in this example demonstrated successful amplified of a single product between cycles 20 and 30, reflecting the variable but adequate abundance of each of the genes.

**Supplemental Figure 7.** Midterm examination questions (Fall 2013) with Bloom's taxonomy. LOCS, lower order cognitive skill; HOCS, higher order cognitive skill.

**Supplemental Figure 8.** Example of a final research report.

**Supplemental Figure 9.** Faculty Course Questionnaire (FCQ) Summary. Results of the standard assessment administered by the University of Colorado at Boulder from 2010 to 2013. n = 96 student respondents.

---

# MCDB 4202

## The Python Project

---

Spring 2014  
M W 3-5:50pm, Gold A1B18

Instructor: Pamela Harvey, PhD  
E-Mail: [pamela.harvey@colorado.edu](mailto:pamela.harvey@colorado.edu)  
Phone: 617-501-4175 (emergencies)  
Office: 303-492-7191  
Office Hours: by appointment

### Overview

The Python Project is a three-credit laboratory course designed to help upper division students obtain laboratory experience. During the class, students will design experiments to examine the molecular mechanisms of organ growth in the Burmese python. To this end, students will:

- Use **modern molecular biology and bioinformatic techniques** to isolate RNA, synthesize cDNA, design primers, measure expression of candidate molecules of the python genome, and present data in the context of the research project,
- **Generate novel data** that will contribute to an ongoing research project in the Leinwand lab.

### Course objectives

The overriding goal of The Python Project is to provide students with sufficient training & guidance to become proficient in a number of molecular biology techniques including but not limited to gel electrophoresis, isolation of RNA from tissue, cDNA synthesis, PCR, and real time PCR. Unlike laboratory exercises that are designed to reinforce concepts that may accompany lecture topics, there is no certainty that any one particular project will succeed, which somewhat mirrors the risks inherent when undertaking novel research. The linear, goal-oriented nature of this research effort means that repetition of some steps will be required to get things to work optimally.

1. Understand how your data contributes to the research being performed in the Leinwand lab,
2. Obtain expertise in real time PCR experiments from beginning to end,
3. Design experiments that address specific scientific questions,
4. Successfully present a poster describing your data in a public poster session to be held during the final exam period,
5. Understand and be able to describe previous research on your gene of interest.

### Prerequisites

MCDB 3120 and 3500, or MCDB 3135 and 3145, and CHEM 4711 and 4731.

### Evaluation

#### Quizzes and worksheets:

Quizzes and worksheets will be completed approximately weekly. Late submissions will lose one grade level per day.

#### Midterm Exam:

The midterm exam for the Spring semester is scheduled for Wednesday, March 12 at 3:00pm. This date is provided beforehand so students can plan their schedules accordingly. In an effort to be fair to all the students taking the course, every effort should be

### Materials

There is **no** textbook for this course. Materials will be posted on D2L on Thursday for the following week.

### Course Milestones

---

RNA Isolation

---

RNA Integrity & Purity

---

Primer Design

---

cDNA Synthesis

---

PCR Validation of Primers

---

Production of a Standard Curve

---

Quantitative PCR

---

Data Analysis

---

Data Presentation

---



made to attend this exam. The exam will be a review of laboratory techniques and information covered in the first half of the semester. It should take about two hours to complete.

### Final Exam:

The day and time of the final exam for the Spring 2014 semester is to be determined. The final exam will be the Python Project Poster Session/Pot Luck. The day and time may not be during the scheduled exam period. Students will present their data in a poster format to the public and MCDB department. Students will also be required to turn in a completed lab notebook and a final written report during the scheduled exam period.



### Point Distribution:

	Weight
Quizzes & worksheets	20%
Review article	10%
Oral presentation	10%
Midterm written exam	20%
Abbreviated summary	5%
Lab notebook	10%
Final report	15%
Poster presentation	10%

### Attendance policy

**Attendance is mandatory.** Because lab courses are participatory, your physical presence is required. You will be allowed one unexcused absence without adversely affecting your grade. Each additional unexcused absence will result in the dropping of a full letter grade. An unexcused absence will be defined as failure to notify the course instructor prior to your absence. Notification can be in the form of personal communication, email or contact by cell phone (text or voice mail). However, the onus will be on the student to inform the instructor that he or she will be absent. This includes potential conflicts with other courses that schedule exams when during the time our class meets.

### Make-up Exam Policy

If you anticipate an excused absence will conflict with an exam, please contact me **before** the scheduled exam. If you miss an exam, it is your responsibility to contact me to arrange a make-up. The student is responsible for providing satisfactory evidence within one week of the end of the absence to document the necessity of the absence.

### Laboratory Conduct

Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, color, culture, religion, creed, politics, veteran's status, sexual orientation, gender, gender identity and gender expression, age, disability, and nationalities. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records. See policies at <http://www.colorado.edu/policies/classbehavior.html> and at [http://www.colorado.edu/studentaffairs/judicialaffairs/code.html#student\\_code](http://www.colorado.edu/studentaffairs/judicialaffairs/code.html#student_code)

### Students with Disabilities

If you qualify for accommodations because of a disability, please submit to me a letter from Disability Services in a timely manner so that your needs may be addressed. Disability Services determines accommodations based on documented disabilities. Contact: 303-492-

8671, Willard 322, and <http://www.Colorado.edu/disabilityservices>.

Disability Services' letters for students with disabilities indicate legally mandated reasonable accommodations. The syllabus statements and answers to Frequently Asked Questions can be found at <http://www.colorado.edu/disabilityservices>.

## Religious Observances

Campus policy regarding religious observances requires that faculty make every effort to reasonably and fairly deal with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. See full details at: [http://www.colorado.edu/policies/fac\\_relig.html](http://www.colorado.edu/policies/fac_relig.html)



## Discrimination and Harassment

The University of Colorado Boulder (CU-Boulder) is committed to maintaining a positive learning, working, and living environment. The University of Colorado does not discriminate on the basis of race, color, national origin, sex, age, disability, creed, religion, sexual orientation, or veteran status in admission and access to, and treatment and employment in, its educational programs and activities. (Regent Law, Article 10, amended 11/8/2001). CU-Boulder will not tolerate acts of discrimination or harassment based upon Protected Classes or related retaliation against or by any employee or student. For purposes of this CU-Boulder policy, "Protected Classes" refers to race, color, national origin, sex, pregnancy, age, disability, creed, religion, sexual orientation, gender identity, gender expression, or veteran status. Individuals who believe they have been discriminated against should contact the Office of Discrimination and Harassment (ODH) at 303-492-2127 or the Office of Student Conduct (OSC) at 303-492-5550. Information about the ODH, the above referenced policies, and the campus resources available to assist individuals regarding discrimination or harassment can be obtained at <http://hr.colorado.edu/dh/>

## Honor Code

All students of the University of Colorado at Boulder are responsible for knowing and adhering to the academic integrity policy of this institution. Violations of this policy may include: cheating, plagiarism, aid of academic dishonesty, fabrication, lying, bribery, and threatening behavior. All incidents of academic misconduct shall be reported to the Honor Code Council ([honor@colorado.edu](mailto:honor@colorado.edu); 303-735-2273). Students who are found to be in violation of the academic integrity policy will be subject to both academic sanctions from the faculty member and non-academic sanctions (including but not limited to university probation, suspension, or expulsion). Other information on the Honor Code can be found at <http://www.colorado.edu/policies/honor.html> and at <http://honorcode.colorado.edu>

## Plagiarism and Copyrights

As commonly defined, plagiarism consists of passing off as one's own, the ideas, words, or writings that belong to another. In accordance with this definition, you are committing plagiarism if you copy the work of another person and turn it in as your own, even if you have the permission of that person. Plagiarism is one of the most serious forms of academic misconduct.

All lectures, exams, handouts and other materials used in this course (including those provided in D2L) are copyrighted. Because these materials are copyrighted, you do not have the right to reproduce, transmit, provide or receive these materials without explicit permission of the instructor/authors. Any other use of these materials is considered "unauthorized" and is thus a form of academic dishonesty and an honor code violation.

**The Python Project**  
**MCDB 4202**  
**Real Time PCR Primer Design Protocol Part I**  
**Spring 2014**

**Procedure Overview:**

- 1. Obtain the mRNA sequence of your gene from a species that is closely related to the Burmese python and color code the alternating exons.**
- 2. Color-code exon-exon boundaries in the mRNA sequence.**
- 3. Find the Python contig sequences that align with pieces of the mRNA sequence obtained from lizard or chicken (or other closely related species like another reptile or bird).**
- 4. Map the exon alignment into both the cDNA (mRNA) sequence and the python contigs.**
- 5. Assemble the python transcript using the order of the exons from the reference species.**
- 6. Validate the assembly by translating the transcript into an amino acid sequence and performing a Protein BLAST search.**
- 7. Design primers using the assembled transcript.**
- 8. Validate the primer set in Primer BLAST.**
- 9. Refine the python transcript if necessary and validate the primer set in Primer BLAST.**

# 1. Obtain the mRNA sequence of your gene from a species that is closely related to the Burmese python and color code the alternating exons.

Go to PubMed: [www.ncbi.nlm.nih.gov/pubmed](http://www.ncbi.nlm.nih.gov/pubmed) (alternatively, Google 'Pubmed' and it's the first option: Home – PubMed – NCBI)

Scroll to the bottom of the page and click 'Gene' under 'POPULAR'

In the search box, type in the name of your gene (either the full name or the appropriate gene abbreviation). For this exercise, type 'SIRT3'.

If the expected gene does not appear, check Gene Card ([www.genecards.org](http://www.genecards.org)) for an alias or alternative full name that could be used for the search. In the case of SIRT3, several other names could be used for the search.

The gene abbreviation will appear, followed by the full name. In parenthesis, the species is indicated. The closest annotated relatives to the Burmese python are *anolis carolinensis* (Anole lizard) and *gallus gallus* (chicken). Human, mouse, or rat may also be used but the relationship is more distant so finding the transcript in the python transcriptome will be more challenging.

Scroll down until you find the species that is most closely related to the Burmese python. If there are many listed, checked the 'Top Organisms' list at the right of the page. The 'Tree' link will list the taxonomic groups. From this list, you can select 'Birds' and *gallus gallus* will usually be present.

Click on 'SIRT3 sirtuin (silent mating type information regulation 2 homolog) 3 (*S. cerevisiae*) [ *Gallus gallus* (chicken) ]'.

You should now see the **Full Report** for the gene, which may include the following information:

1. Summary: provides official naming scheme and basic information like whether the gene has been shown to encode a protein,
2. Genomic context: gives the location of the gene in the genomes (chromosome number),
3. Genomic regions, transcripts, and products: depicts the location of the gene along the chromosome relative to other genes. Single nucleotide polymorphisms are shown as well as coding regions and Exons,
4. Bibliography: lists published articles about the gene in PubMed and links to their abstracts,
5. Variation: links to single nucleotide polymorphism and genotype reports,
6. Pathways from BioSystems: links to specific pathways that involve the gene,
7. General gene information: homology to genes in other species with links to their nucleotide and animal acid sequences,
8. General protein information: nomenclature for the protein the gene encodes,
9. **NCBI Reference Sequences (RefSeq): links to mRNA and Protein sequences,**
10. Additional Links: links to Genome Browser and UniGene.

Under **NCBI Reference Sequences (RefSeq): links to mRNA and Protein sequences**, click on '[NM\\_001199493.1](#)' This link takes you to the mRNA sequence and translation. The link for the mRNA sequence (beginning with 'XM' or 'NM') is usually followed by the link for the protein sequence (beginning with 'XP' or 'NP').

The page will show the amino acid and mRNA sequence. Note that the format of the mRNA sequence includes numbers and spaces. To switch the format to eliminate the numbers and spaces, click 'FASTA' under the gene name. FASTA format removes numbers and spaces, making the sequence easier to navigate.

Copy the text from gene name ‘**Gallus gallus sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae) (SIRT3), mRNA**’ to the end of the sequence into a blank text document. Change the font to Courier New 8. Your text should appear as follows:

**Gallus gallus sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae) (SIRT3), mRNA**

NCBI Reference Sequence: NM\_001199493.1

[GenBank Graphics](#)

>gi|313747483|ref|NM\_001199493.1| Gallus gallus sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae) (SIRT3), mRNA

```
ATGGAGCGGGGGTTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAGCGGGTGGCC
TGGCTCTGTTCGCCCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTACAAGGGACCAGGCCCTTCTC
TCTGTCTGTGCCCGCAGTGTCTTCTAGGACTGGGCAGCTGGGGAGGTGACAGTGGGAAGCAGAAGCTC
ACCCTGCAGGATGTGGCAGAGCTCATTCGGAAGAAGGAGTGTCTGTCGAGTAGTGGTGTGGCCGGTGTG
GGATTAGCACCCCCAGCGGCATCCAGACTTCAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCA
GTACAACATCCCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTTACCCTGGCAATTATAGACCCAACCTACGCCCACTATTTCTGAGACTCCTGC
ATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAAATATTGATGGGCTGGAGCGAGTTGCTGGGATCCC
TCCTGATAGACTGGTGGAGCCACGGCACCTTTGCTACTGCCACGTGCACAGTCTGTTCGGAGGAAATC
CCAGGAGAGGACTTCAGGGGGACGTTATGGCAGACAAGGTCCTCACTGTCTGTCTGCACCGGAATCG
TCAAGCCTGACATCGTGTCTTTTGGCGAGGAGCTCCCGCAGCGCTTCTTCCTGCACATGACAGACTTCCC
CATGGCAGACCTGCTTTTTGTTCATCGGAACGTCCCTGGAGGTGGAGCCCTTTGCCAGCCTGGCAGGAGCT
GTTTCGCAACTCCGTTCCCGGGTCCCTCATCAACCGAGATCTTGTAGGACCGTTTGCCTGGCAACAACGCT
ACAATGACATAGCCAGCTGGGGGATGTGGTCACTGGGGTTGAGAAGATGGTAGAACTGCTGGACTGGAA
TGAAGAGATGCAAACACTAATTCAGAAAAGAAAAGAAAAGCTGGATGCAAAGACAAATAGGACAGCTGG
CTCCCTGTGCTGGAAAAGCAATGTTACCCCTAAAGATGAGAAGATGCTGCCATACCCAGCGTGAGCAGAG
AAAGAGACAATGGCAGGGAGCTGTGGAAGAGCCCATCCAGCTCTGCAGCAGCGAATCTACACAGGACTGG
CACTTGGGGAAAAGGAGGATGCTCATGGCTCAACCCAGCCAGCCTGGAAGC
```

Note that the sequence does not indicate the junctions between introns and exons. We need this information to design appropriate primers for real time PCR.

## 2. Color-code exon-exon boundaries in the mRNA sequence.

To obtain the color-coded mRNA sequence that indicates the locations of introns and exons, scroll down the right side of the page and select 'Ensembl' under 'LinkOut to external resources.' This will direct you to your gene in Ensembl. If there is no LinkOut, perform a Google search for Ensembl Genome Browser and follow the link. In this case, you will need to type in SIRT3. If hundreds of species are returned, limit the search to a closely related species. Often, a gene will not be found in PubMed but will be present in Ensembl. In the case of Sirt3, anole lizard gene sequence is available. Search for 'SIRT3' and click on 'Anole lizard' under 'Restrict species to'.

On the left side of the page, click **cDNA**, a DNA copy of the mRNA transcript.

You will now see the cDNA sequence for your gene, but there is more information than you need to find this gene in the python transcriptome. The format will need to be changed to include only the cDNA sequence with color-coded introns and exons.

On the left side of the page, click 'Configure this page'. Choose the following configuration:

Show exons:	yes
Show codons:	no
Show UTR:	no
Show coding sequence:	no
Show protein sequence:	no
Show RNA features:	no
Show variations:	no
High variations longer than 10bp:	no
Filter variations by consequence type:	no filter
Line numbering:	no

To save the configuration settings, click the ✓ at the upper right corner of the window.

Copy the text from the transcript name to the end of the cDNA sequence. Change the font to Courier New 8. It is also helpful to remove extra lines and hyperlinks. Your text should appear as follows:

**Transcript:** SIRT3-201 ENSGALT00000006685  
**Description**  
sirtuin [Source:RefSeq peptide;Acc:NP\_001186422]  
**Location**  
[Chromosome 5: 1,625,769-1,629,803](#) forward strand.  
**Gene**  
This transcript is a product of gene [ENSGALG0000004201](#) This gene has 1 transcript (splice variant)

**cDNA sequence**  
ATGGAGCGGGGGTTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG  
CGCGGTGGCCTGGCTCTGTTCGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTA  
CAAGGGACCAGGCCCTTCTCTGTCTGCTGCCGCCAGTGCTGTTCTAGGACTGGGCAGC  
TGGGGAGGTGACAGTGGGAAGCAGAAGCTCACCTGCAGGATGTGGCAGAGCTCATTCCG  
AAGAAGGAGTGTCTCGAGTAGTGGTGATGGCCGGTGTGGGATTAGCACCCACAGCGGC  
ATCCCAGACTTCAGGTCTCCGGGAGCGGCCTATAGTAACCTTGAGCAGTACAACATC  
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC  
ACTTTGGCCAAGGAGCTCTACCTGGCAATTATAGACCAACTACGCCCACTATTTCTCG  
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG  
GAGCGAGTTGCTGGGATCCCTCCTGATAGACTGGTGAAGCCACGGCACCTTTGCTACT  
GCCACGTGCACAGTCTGTCTGGAGGAAATTTCCAGGAGAGGACTTCAGGGGGGACGTTATG  
GCAGACAAGGTCCCTCACTGTCTGTCTGCACCGGAATCGTCAAGCCTGACATCGTGTTC  
TTTGGCGAGGAGCTCCCGCAGCGCTTCTTCTGACATGACAGACTTCCCCATGGCAGAC  
CTGCTTTTGTCAATCGGAACGTCCTGGAGGTGGAGCCCTTTGCCAGCCTGGCAGGAGCT  
GTTCCGCAACTCCGTTCCCGGGTCCCTCATCAACCGAGATCTTGTAGGACCGTTTGCCTGG  
CAACAACGCTACAATGACATAGCCAGTGGGGGATGTGGTCACTGGGGTTGAGAAGATG  
GTAGAAGTCTGGACTGGAATGAAGAGATGCAAACTAATTCAGAAAGAAAAGAAAAG  
CTGGATGCAAAAGACAAATAG



**\*\*\*When you begin your primer design document, include as many notes as possible. A new document should be started for each gene. This will be enormously helpful later on. Your document should be in the following format at this point (below). These documents, one for each of your genes, will be uploaded to D2L for review by 3pm on Monday, January 27, 2014.**

mRNA sequence obtained from PubMed:

**Gallus gallus sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae) (SIRT3), mRNA**  
NCBI Reference Sequence: NM\_001199493.1

[GenBank Graphics](#)

>gi|313747483|ref|NM\_001199493.1| Gallus gallus sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae) (SIRT3), mRNA

```
ATGGAGCGGGGGTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAGCGCGGTGGCC
TGGCTCTGTTCCGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTACAAGGGACCAGGCCCTTCTC
TCTGTCTGCTGCCGCCAGTCTTCTAGGACTGGGCAGCTGGGGAGGTGACAGTGGGAAGCAGAAGCTC
ACCCTGCAGGATGTGGCAGAGCTCATTCGGAAGAAGGAGTGTCTGCGAGTAGTGGTGTGGCCGGTGTCTG
GGATTAGCACCCCGAGGCATCCAGACTTCAGGTCTCCGGGGAGCGGCCTTATAGTAACCTTGAGCA
GTACAACATCCCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTTACCCTGGCAATTATAGACCCAACACTACGCCCACTATTTTCCTGAGACTCCCTG
ATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTGGAGCGAGTTGCTGGGATCCC
TCCTGATAGACTGGTGGAAAGCCACGGCACCTTTGCTACTGCCACGTGCACAGTCTGTCGGAGGAAATTC
CCAGGAGAGGACTTCAGGGGGGACGTTATGGCAGACAAGGTCCTCACTGTCTGTCTGCACCGGAATCG
TCAAGCCTGCACATCGTGTCTTTGGCGAGGAGCTCCCGCAGCGCTTCTTCCTGCACATGACAGACTTCCC
CATGGCAGACCTGCTTTTTGTCTATCGGAACGTCCTGGAGGTGGAGCCCTTTGCCAGCCTGGCAGGAGCT
GTTCGCAACTCCGTTCCCGGGTCCCTCATCAACCGAGATCTTGTAGGACCGTTTGCCGGCAACAACGCT
ACAATGACATAGCCAGCTGGGGGATGTGGTCACTGGGGTTGAGAAGATGGTGAAGTGTGGACTGGAA
TGAAGAGATGCAAAACCTAATTAGAAAAGAAAAGAAAAGCTGGATGCAAAAAGACAAATAGGACAGCTGG
CTCCCTGCTGCTGAAAGCAATGTTACCCATAAAGATGAGAAGATGCTGCCATACCCAGCGTGAGCAGAG
AAAGAGACAATGGCAGGAGCTGTGGAAGAGCCATCCAGCTCTGCAGCAGCGAATCTACACAGGACTGG
CACTTGGGGAAAAGGAGGATGCTCATGGCTCAACCCAGCCAGCCTGGAAGC
```

Color-coded mRNA sequence showing exon-exon boundaries. This sequence was obtained from Ensembl Genome Browser:

**Transcript: SIRT3-201 ENSGALT00000006685**

**Description**

sirtuin [Source:RefSeq peptide;Acc:NP\_001186422]

**Location**

[Chromosome 5: 1,625,769-1,629,803](#) forward strand.

**Gene**

This transcript is a product of gene [ENSGALG00000004201](#) This gene has 1 transcript (splice variant)

**cdNA sequence**

```
ATGGAGCGGGGGTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCCTGGCTCTGTTCCGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTCTGTCTGTCTGCTGCCGCCAGTGTCTGTTCTAGGACTGGGCAGC
TGGGGAGGTGACAGTGGGAAGCAGAAGCTCACCCCTGCAGGATGTGGCAGAGCTCATTCGG
AAGAAGGAGTGTCTGTCGAGTAGTGGTGTATGGCCGGTGTGGGATTAGCACCCCGAGCGGC
ATCCCAGACTTCAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTTACCCTGGCAATTATAGACCCAACACTACGCCCACTATTTTCCTG
AGACTCCTGCATGACAAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG
GAGCGAGTTGCTGGGATCCCTCCTGATAGACTGGTGGAAAGCCACGGCACCTTTGCTACT
GCCACGTGCACAGTCTGTGCGAGGAAATCCCAGGAGAGGACTTCAGGGGGGACGTTATG
GCAGACAAGGTCCTCACTGTCTGTCTGCACCGGAATCGTCAAGCCTGCACATCGTGTTC
TTTGGCGAGGAGTCCCGCAGCGCTTCTTCCTGCACATGACAGACTTCCCATGGCAGAC
CTGCTTTTTGTCTATCGGAACGTCCTGGAGGTGGAGCCCTTTGCCAGCCTGGCAGGAGCT
GTTCGCAACTCCGTTCCCGGGTCCCTCATCAACCGAGATCTTGTAGGACCGTTTGCCTGG
CAACAACGCTACAATGACATAGCCAGCTGGGGGATGTGGTCACTGGGGTTGAGAAGATG
GTAGAAGTGTGGACTGGAATGAAGAGATGCAAAACACTAATTCAGAAAAGAAAAGAAAAG
CTGGATGCAAAAAGACAAATAG
```

# The Python Project

## MCDB 4202

### Real Time PCR Primer Design Protocol Part II

#### Spring 2014

### 3. Find the Python contig sequences that align with the mRNA sequence obtained from lizard or chicken.

Go to [blast.ncbi.nlm.nih.gov](http://blast.ncbi.nlm.nih.gov). You can access this link easily by performing a Google search for 'Blast NCBI.' This will take you to the Basic Local Alignment Search Tool (BLAST) bioinformatics algorithm, which will allow you to compare your sequence from lizard or chicken to the nucleotide sequences of other genomes. In this case, we are interested in the unannotated whole genome shotgun (WGS) for Burmese python (*python molurus*).

In the BLAST window, under 'Choose a BLAST program to run', select 'nucleotide blast.'

Copy the entire color-coded cDNA sequence from Ensembl and paste it into the 'Enter Query Sequence' box. Note that each blue or black segment represents an exon; the sequence is a DNA copy of the mRNA.

**TIPS:** If your sequence is extremely long (over 750 bases), limit your search to a portion of the transcript rather than the entire transcript. You must have at least 200 bases in your search to design primers of an appropriate length for real time PCR.

Be sure that you are searching in the 'blastn suite.'

For this example, use the first four exons from the SIRT3 transcript:

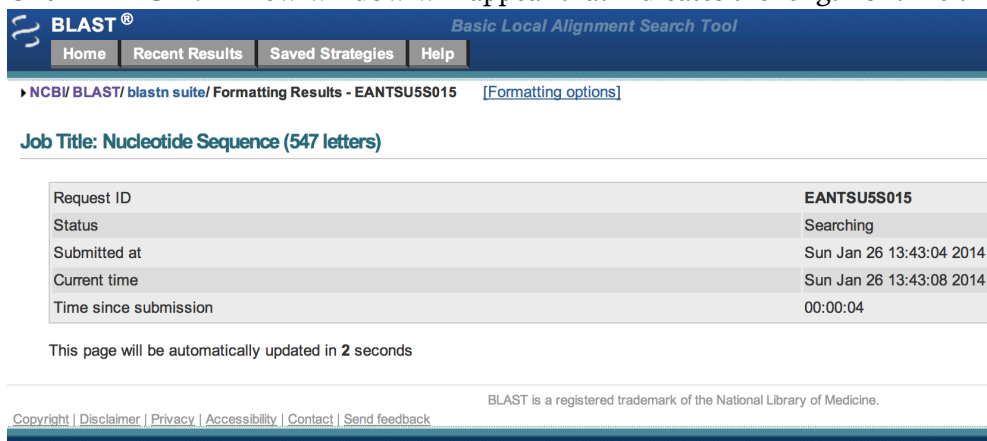
```
ATGGAGCGGGGGTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCCTGGCTCTGTTCGCCCTCAGTGCAGGACTGGCTGGGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTGTCTGTCTGCTGCCGCCAGTGCTGTTCTAGGACTGGGCAGC
TGGGGAGGTGACAGTGGGAAGCAGAAGCTCACCTGCAGGATGTGGCAGAGCTCATTCGG
AAGAAGGAGTGTCTCGTTCGAGTAGTGGTGATGGCCGGTGTGGGATTAGCACCCCCAGCGGC
ATCCCAGACTTCAGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTCTACCTGGCAATTATAGACCCAACCTACGCCCACTATTTCTG
AGACTCCTGCATGACAAAAGGGCTCCTTCTGCGTCTCTATACTCAGAAATTGATGGGCTG
GAGCGAG
```

Under 'Choose Search Set', select 'whole-genome shotgun contigs' in the pull-down menu.

Type 'python' in the 'Organism' box. A pull-down menu will appear. Select 'python molurus (taxid:51750).'

Verify that 'Highly similar sequences (megablast)' is selected in 'Program Selection'.

Click 'BLAST'. A new window will appear that indicates the length of time the search has been performed:



BLAST® Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

NCBI/BLAST/blastn suite/ Formatting Results - EANTSU5S015 [Formatting options]

Job Title: Nucleotide Sequence (547 letters)

Request ID	EANTSU5S015
Status	Searching
Submitted at	Sun Jan 26 13:43:04 2014
Current time	Sun Jan 26 13:43:08 2014
Time since submission	00:00:04

This page will be automatically updated in 2 seconds

BLAST is a registered trademark of the National Library of Medicine.

Copyright | Disclaimer | Privacy | Accessibility | Contact | Send feedback



When the search is complete, the results will be shown. In the example, where the *gallus gallus* SIRT3 Ensembl sequence is pasted into BLAST, no significant similarity will be found.

NCBI/ BLAST/ blastn suite/ Formatting Results - EANTSU5S015

[Edit and Resubmit](#) [Save Search Strategies](#) [Formatting options](#) [Download](#)

### Nucleotide Sequence (547 letters)

<b>RID</b>	<a href="#">EANTSU5S015</a> (Expires on 01-28 01:43 am)	<b>Database Name</b>	wgs
<b>Query ID</b>	lcl 18979	<b>Description</b>	<a href="#">See details</a>
<b>Description</b>	None	<b>Program</b>	BLASTN 2.2.29+ <a href="#">Citation</a>
<b>Molecule type</b>	nucleic acid		
<b>Query Length</b>	547		

**No significant similarity found. For reasons why, [click here](#)**

Other reports: [Search Summary](#)

If this message is returned, it does not necessarily mean that the sequence does not exist in the python. It is likely that the search criteria were too stringent. In most cases, changing the Program Selection to ‘**Somewhat similar sequences (blastn)**’ on the search page will solve the problem, especially when lizard sequence is not available and *gallus gallus* is used.

Return to the Query screen and change the Program Selection to ‘**Somewhat similar sequences (blastn)**’.

Click ‘**BLAST**’.

A graphic summary of the BLAST hits will appear:

BLAST® Basic Local Alignment Search Tool

[Home](#) [Recent Results](#) [Saved Strategies](#) [Help](#)

NCBI/ BLAST/ blastn suite/ Formatting Results - EAP1T9C801R

[Edit and Resubmit](#) [Save Search Strategies](#) [Formatting options](#) [Download](#)

### Nucleotide Sequence (547 letters)

<b>RID</b>	<a href="#">EAP1T9C801R</a> (Expires on 01-28 01:47 am)	<b>Database Name</b>	wgs
<b>Query ID</b>	lcl 214231	<b>Description</b>	<a href="#">See details</a>
<b>Description</b>	None	<b>Program</b>	BLASTN 2.2.29+ <a href="#">Citation</a>
<b>Molecule type</b>	nucleic acid		
<b>Query Length</b>	547		

Other reports: [Search Summary](#) [Taxonomy reports](#) [Distance tree of results](#)

#### Graphic Summary

Distribution of 19 Blast Hits on the Query Sequence

Mouse over to see the define, click to show alignments

Score Range	Color
<40	Black
40-50	Blue
50-80	Green
80-200	Pink
>=200	Red

In this case, 19 pieces of python transcript (contigs) were similar to the sequence that was queried (SIRT3 from *gallus gallus*).

If, for example, you scroll over the red line between 300 and 550, you will highlight ‘**AEQU02054582 python bivittatus**’. A red contig indicates that the match is fairly good between a part of the sequence you queried and the Python contig. Scores represent the length of the contig and the percentage of matches. **Remember that this sequence was derived from the whole python genome so includes both introns and exons.**

Click on the red line to see the alignment between the sequence you queried and the Python contig.

Download ▾ GenBank Graphics

Python bivittatus Python\_molurus\_bivittatus-5.0.2-723.9, whole genome shotgun sequence  
Sequence ID: [gb|AEQU02054582.1](#) Length: 11858 Number of Matches: 1

Range 1: 4905 to 5140 GenBank Graphics ▾ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
228 bits(252)	1e-57	192/236(81%)	0/236(0%)	Plus/Plus

Query	312	CAGGTCTCCGGGAGCGGCTCTATAGTAACCTTGAGCAGTACAACATCCCTTACCCCGA	371
Sbjct	4905	CAGGTCTCCTGGAGTGGACTATAACAATAATCTTCAGCAGTATAATATCCATACCCTGA	4964
Query	372	AGCCATCTTTGAAGTGGCTACTTCTTCATCAACCCCAAGCCATTCTTCACTTTGGCCAA	431
Sbjct	4965	AGCCATATTTGAGCTTAGTACTTTTTCAGAAATCCAAGCCTTCTTCAGTTAGCTAA	5024
Query	432	GGAGCTTACCCCTGGCAATTATAGACCAACTACGCCCACTATTTCCTGAGACTCCTGCA	491
Sbjct	5025	GGAGTTGTACCCTGGCAATTACAGACCAAACTATGCCCACTATTTCCTGACTTCTGTT	5084
Query	492	TGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTGGAGCGAG	547
Sbjct	5085	TGACAAAGGGCTCCTTCTGCGCTCTACACACAAAATATTGATGGGCTGGAGAGAG	5140

The **Query** line is the sequence you put into BLAST. The **Sbjct** line is the python contig. The sequences are 81% similar, a decent match especially with no gaps in the sequence. Plus/Plus indicates that the orientation of the *gallus* transcript is the same as the orientation of the python contig (5' to 3').

Click on Sequence ID 'gb|AEQU02054582.1|' to retrieve the corresponding Python contig sequence. The contig is a contiguous piece of genomic DNA that is 11,858 bases long. It contains one of the exons from the *gallus gallus* sequence.

Examine the exon boundaries from the *gallus gallus* sequence and the region that aligns with the python contig. Highlight the location of the python contig.

*Gallus gallus*:

```
ATGGAGCGGGGGTTCCGGCGGGAGCGGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCTGGCTCTGTTCGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTGTCTGCTGCCGCCAGTGCTGTTCTAGGACTGGGCAGC
TGGGGAGGTGACAGTGGGAAGCAGAAGCTCACCTGCAGGATGTGGCAGAGCTCATTCCG
AAGAAGGAGTGTGTCGAGTAGTGGTGTATGGCCGGTGTGGGATTAGCACCCAGCGGC
ATCCAGACTTCAGGTCTCCGGGGAGCGGCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAAGTGGCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTTACCCCTGGCAATTATAGACCAACTACGCCCACTATTTCCTG
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG
GAGCGAG
```

Highlighted in yellow is the sequence that aligned with **Python molurus contig AEQU02054582.1**.

Note that the python contig aligned mostly within a single exon.

Continue with this process until you have at least 200 continuous bases representing *at least* two exons mapped.

- TIPS:** For overlapping sequence, highlight in gray.  
Choose different colors for each section mapped.  
Avoid mapping very small lengths of sequence. The likelihood of these sequences being a true match is low.  
Choose alignments that have 0 gaps whenever possible.  
Choose alignments that have greater than 80% identity whenever possible.  
If the graphical depiction of the alignments has multiple contiguous sections, map these first.  
Make as many notes as possible. This will be enormously helpful later on if you have to redesign primers.

Continue mapping the alignment of your gene. Your text should be similar to that shown below. Make as many notations as possible to make clear what the identity of each sequence is.

The second best alignment has two regions within the same contig. This is indicated in the graphical representation by two bars separated by a vertical line. It is also indicated in the alignments as Range 1, Range 2, etc.

Download ▾ GenBank Graphics Sort by: E value ▾

Python bivittatus Python\_molurus\_bivittatus-5.0.2-3128.19, whole genome shotgun sequence  
Sequence ID: [gb|AEQU02140972.1](#) Length: 1628 Number of Matches: 2

Range 1: 1149 to 1192 [GenBank](#) [Graphics](#) ▾ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
44.6 bits(48)	0.027	36/44(82%)	0/44(0%)	Plus/Minus

```

Query 403 AACCCCAAGCCATTCTTCACTTTGGCCAAGGAGCTCTACCCCTGG 446
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct 1192 AACCCGAGAACCATTCTTCACTTTGGCTCGGGAGCTCTACCCAGG 1149
  
```

Range 2: 1305 to 1403 [GenBank](#) [Graphics](#) ▾ Next Match ▲ Previous Match ▲ First Match

Score	Expect	Identities	Gaps	Strand
39.2 bits(42)	1.2	68/99(69%)	0/99(0%)	Plus/Minus

```

Query 298 GGCATCCCAGACTTCAGGTCTCCGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAAC 357
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct 1403 GGGATCCCCGATTTCCGCTCACCTGGCACGGGGCTATATGCCAATCTACAGCAATACAGT 1344
  
```

```

Query 358 ATCCCTTACCCGAAGCCATCTTTGAACTGGCCTACTTC 396
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct 1343 CTACCATATCCTGAAGCCATCTTTGAGATTGGTTACTTC 1305
  
```

Range 1 is highlighted in the gallus gallus transcript in green. Range 2 is highlighted in purple.

```

ATGGAGCGGGGGTTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCTGGCTCTGTTCCGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTGTCTGCTGCCGCCAGTGCTGTTCTAGGACTGGGCAGC
TGGGAGGTGACAGTGGGAAGCAGAAGCTCACCTGCAGGATGTGGCAGAGCTCATTCCG
AAGAAGGAGTGTCTCGAGTAGTGGTGATGGCCGGTGTGGGATTAGCACCCCAAGCCG
ATCCAGACTTCAGGTCTCCGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
CAATTAAGCCCAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
CAATTAAGCCCAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG
GAGCGAG
  
```

Mapping these alignments did not help us achieve better coverage of more than one exon. It is likely that Range 2 is not a 'real' alignment because the genomic sequence would have intron between exons 4 and 5.

The third best alignment has one region mapped within contig AEQU02078587.1.

Download ▾ GenBank Graphics

Python bivittatus Python\_molurus\_bivittatus-5.0.2-1219.34, whole genome shotgun sequence  
Sequence ID: [gb|AEQU02078587.1](#) Length: 19849 Number of Matches: 1

Range 1: 716 to 747 [GenBank](#) [Graphics](#) ▾ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
41.0 bits(44)	0.33	28/32(88%)	0/32(0%)	Plus/Plus

```

Query 206 AGCTCACCTGCAGGATGTGGCAGAGCTCATT 237
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct 716 AGCTGACCTGCTGGTTTGGCAGAGCTCATT 747
  
```

This alignment is short but has 88% identity and no gaps. This alignment is mapped in turquoise.

```

ATGGAGCGGGGGTTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCTGGCTCTGTTCCGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTGTCTGCTGCCGCCAGTGCTGTTCTAGGACTGGGCAGC
TGGGAGGTGACAGTGGGAAGCAGAAGCTCACCTGCAGGATGTGGCAGAGCTCATTCCG
AAGAAGGAGTGTCTCGAGTAGTGGTGATGGCCGGTGTGGGATTAGCACCCCAAGCCG
ATCCAGACTTCAGGTCTCCGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTCTACCTTGGCAATTATAGACCCCAACTACGCCCACTATTTCCTG
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG
GAGCGAG
  
```



Python bivittatus Python\_molurus\_bivittatus-5.0.2-7712.8, whole genome shotgun sequence

Sequence ID: [gb|AEQU02215050.1](#) Length: 2436 Number of Matches: 1

Range 1: 914 to 941 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
37.4 bits(40)	4.0	25/28(89%)	0/28(0%)	Plus/Plus

```

Query 260 TAGTGGTGGTGGCCGGTCTGGGATTAG 287
          |||||
Sbjct 914 TAGTGGTGGTGGCAGCTGCTGGGATTAG 941
  
```

The sixth alignment is mapped in teal.

```

ATGGAGCGGGGGGTTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCCTGGCTCTGTTCCGCCCTCAGTGCAGGACTGGCTGCGGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTCTGCTGGCCAGTGTGTTCTAGGACTGGGCAGC
TGGGAGGTGACAGTGGGAAGCAGAACTCACCTGCAGGATGTGGCAGAGCTCATTGG
AAGAAGGAGTGTCTGTCGAGTAGTGGTGGTGGCCGGTCTGGGATTAGCACCCCAGCGGC
ATCCCAGACTTCAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTTACCCTGGCAATTATAGACCAACTACGCCCACTATTTCCTG
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAAATTGATGGGCTG
GAGCGAG
  
```

The sixth alignment is also in exon 3, but additional coverage is provided.

Seventh alignment:

Python bivittatus Python\_molurus\_bivittatus-5.0.2-5454.18, whole genome shotgun sequence

Sequence ID: [gb|AEQU02187563.1](#) Length: 5102 Number of Matches: 1

Range 1: 4842 to 4869 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
37.4 bits(40)	4.0	25/28(89%)	0/28(0%)	Plus/Minus

```

Query 196 GGGAAAGCAGAAGCTCACCCCTGCAGGATG 223
          |||||
Sbjct 4869 GGGAAAGCTGAGGCTCACCCCTGCTGGATG 4842
  
```

The first nucleotide of seventh alignment is mapped in purple.

```

ATGGAGCGGGGGGTTTCGGCGCGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCCTGGCTCTGTTCCGCCCTCAGTGCAGGACTGGCTGCGGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTCTGCTGGCCAGTGTGTTCTAGGACTGGGCAGC
TGGGAGGTGACAGTGGGAAGCAGAACTCACCTGCAGGATGTGGCAGAGCTCATTGG
AAGAAGGAGTGTCTGTCGAGTAGTGGTGGTGGCCGGTCTGGGATTAGCACCCCAGCGGC
ATCCCAGACTTCAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTTACCCTGGCAATTATAGACCAACTACGCCCACTATTTCCTG
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAAATTGATGGGCTG
GAGCGAG
  
```

The seventh alignment is across the gray and turquoise alignments. This is redundant and is therefore not mapped.

Eighth alignment:

Python bivittatus Python\_molurus\_bivittatus-5.0.2-4689.2, whole genome shotgun sequence

Sequence ID: [gb|AEQU02174810.1](#) Length: 6781 Number of Matches: 1

Range 1: 3917 to 3939 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
37.4 bits(40)	4.0	22/23(96%)	0/23(0%)	Plus/Minus

```

Query 177 CAGCTGGGGAGGTGACAGTGGGA 199
          |||||
Sbjct 3939 CAGCTGGGGAGGAGACAGTGGGA 3917
  
```

The first nucleotide of the eighth alignment is highlighted in bright green. The eighth alignment mostly overlaps with the gray alignment so it is not needed.

```
ATGGAGCGGGGGGTTCCGGCGGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCTGGCTCTGTTCCGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTCTGCTGCTGGCCCAATGCTGTTCTAGGACTGGGAGC
TGGGAGCTGACAGTGGGAAGCAGAACTCACCCCTGCAGGATGTGGCAGAGCTCATTCGG
AAGAAGGAGTGTCTGTCGAGTAGTGGTGATGGCCGGTGTGGGATTAGCACCCCCAGCGGC
ATCCCAGACTTCAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTCTACCCCTGGCAATTATAGACCCAACTACGCCCACTATTTCTCG
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG
GAGCGAG
```

Ninth alignment:

[Download](#) [GenBank](#) [Graphics](#)

Python bivittatus Python\_molurus\_bivittatus-5.0.2-4269.1, whole genome shotgun sequence  
Sequence ID: [gb|AEQU02166785.1](#) Length: 6376 Number of Matches: 1

Range 1: 286 to 310 [GenBank](#) [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
37.4 bits(40)	4.0	23/25(92%)	0/25(0%)	Plus/Minus

```
Query 120 ACAAGGGACCAGGCCCTTCTCTCTG 144
      |||
Sbjct 310 ACAAGGGACCAGGCCCTTCTCTCTG 286
```

The ninth alignment is highlighted in purple. The part that overlaps with the bright blue alignment is highlighted in gray.

```
ATGGAGCGGGGGGTTCCGGCGGGAGCGGCGCTGGTGGCGGCATGGAGAAGCCTGTGGGAG
CGCGGTGGCCTGGCTCTGTTCCGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGTA
CAAGGGACCAGGCCCTTCTCTCTGCTGCTGGCCCAATGCTGTTCTAGGACTGGGAGC
TGGGAGCTGACAGTGGGAAGCAGAACTCACCCCTGCAGGATGTGGCAGAGCTCATTCGG
AAGAAGGAGTGTCTGTCGAGTAGTGGTGATGGCCGGTGTGGGATTAGCACCCCCAGCGGC
ATCCCAGACTTCAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC
ACTTTGGCCAAGGAGCTCTACCCCTGGCAATTATAGACCCAACTACGCCCACTATTTCTCG
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG
GAGCGAG
```

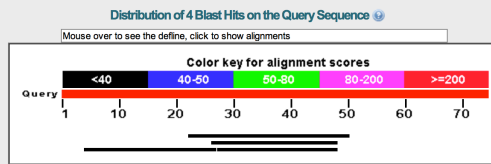
## NEW BLAST SEARCH:

At this point, the scores are becoming lower so a new BLAST is performed to limit the search to the unhighlighted region between the teal and yellow. To provide more sequence, which will improve the specificity of the alignment, the sequence between the turquoise and yellow (including the teal) is used:

```
CGGAAGAAGGAGTGTCTGTCGAGTAGTGGTGATGGCCGGTGTGGGATTAGCACCCCCAGCGGC
ATCCCAGACTT
```

Four alignments are returned:

## Graphic Summary



## Descriptions

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
<a href="#">Python bivittatus Python_molurus_bivittatus-5.0.2-7712.8, whole genome shotgun sequence</a>	37.4	37.4	37%	0.36	89%	<a href="#">AEQU02215050.1</a>
<a href="#">Python bivittatus Python_molurus_bivittatus-5.0.2-1648.17, whole genome shotgun sequence</a>	35.6	35.6	29%	1.3	95%	<a href="#">AEQU02096051.1</a>
<a href="#">Python bivittatus Python_molurus_bivittatus-5.0.2-1885.28, whole genome shotgun sequence</a>	33.7	33.7	28%	4.4	95%	<a href="#">AEQU02104492.1</a>
<a href="#">Python bivittatus Python_molurus_bivittatus-5.0.2-13.56, whole genome shotgun sequence</a>	33.7	33.7	31%	4.4	91%	<a href="#">AEQU02001995.1</a>

The first alignment is identical to the teal:

Download GenBank Graphics

Python bivittatus Python\_molurus\_bivittatus-5.0.2-7712.8, whole genome shotgun sequence  
Sequence ID: [gb|AEQU02215050.1](#) Length: 2436 Number of Matches: 1

Range 1: 914 to 941 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
37.4 bits(40)	0.36	25/28(89%)	0/28(0%)	Plus/Plus

```
Query 23 TAGTGGTGATGGCCGGTCTGGGATTAG 50
      |||
Sbjct 914 TAGTGGTGGTGCAGCTGCTGGGATTAG 941
```

```
CGGAAGAAGGAGTGTCTCGAC TAGTGGTGATGGCCGGTCTGGGATTAG CACCCCCAGCGGC
ATCCCAGACTT
```

Second alignment:

Download GenBank Graphics

Python bivittatus Python\_molurus\_bivittatus-5.0.2-1648.17, whole genome shotgun sequence  
Sequence ID: [gb|AEQU02096051.1](#) Length: 32243 Number of Matches: 1

Range 1: 30007 to 30028 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
35.6 bits(38)	1.3	21/22(95%)	0/22(0%)	Plus/Plus

```
Query 27 GGTGATGGCCGGTCTGGGATT 48
      |||
Sbjct 30007 GGTGATGGCTGCTGGGATT 30028
```

```
CGGAAGAAGGAGTGTCTCGAC TAGTGGTGTGATGGCCGGTCTGGGATTAG CACCCCCAGCGGC
ATCCCAGACTT
```

Third alignment:

Download GenBank Graphics

Python bivittatus Python\_molurus\_bivittatus-5.0.2-1885.28, whole genome shotgun sequence  
Sequence ID: [gb|AEQU02104492.1](#) Length: 47224 Number of Matches: 1

Range 1: 44174 to 44194 GenBank Graphics Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
33.7 bits(36)	4.4	20/21(95%)	0/21(0%)	Plus/Minus

```
Query 28 GTGATGGCCGGTCTGGGATT 48
      |||
Sbjct 44194 GTGTTGGCCGGTCTGGGATT 44174
```

The first nucleotide is highlighted in yellow. The third alignment mostly overlaps with the teal alignment.







CGCGGTGGCCTGGCTCTGTTCGGCCCTCAGTGCAGGACTGGCTGCGGGGCGTGCAGGGA  
CAAGGGACCAGGCCCTTCTCTCTGCTGGCTGGCCAGTG TGTCTAGGACTGGGCAGC  
TGGGGAGGTGACAGTGGGAAGCAGAA GCTCACCCCTGCAGGATGTGGCAGAGCTCATT CGG  
AAGAAGGAGTGTGCTCGAGTAGTGGTGATGGCCGGTGTGGGATTAG CACC DCCAGCGGC  
ATCCAGACT CAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC  
CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC  
ACTTTGGCCAAGGAGCTTACCCTGGCAATTATAGACCAACTACGCCACTATTTCTTG  
AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG  
GAGCGAG

From the olive alignment to the end of the yellow alignment, there are only 8 unmapped nucleotides. This is acceptable to proceed with assembling the python transcript.

# The Python Project

## MCDB 4202

### Real Time PCR Primer Design - Controls

#### Spring 2014

#### 1. Using the assembled python transcript, map the designed primers.

The assembled python transcript for SIRT3 is below.

CATT CGGA AGAAGGAGTGT CATGGAGTAGTGGTGGTGGCAGCTGCTGGGATTAG CACC CCCAGCTGCATCCCAG  
 ACTT CAGGTCTCCTGGGAGTGGACTATAACAATAATCTTCAGCAGTATAATATTCATACCCTGAAGCCATATTT  
 GAGCTTAGCTACTTTTTCCAGAATCCCAAGCCTTCTTCAGGTTAGCTAAGGAG

Unhighlight the assembled python transcript and color-code the exons in the python assembled transcript using the highlighted chicken transcript (below)

CATT CGGA AGAAGGAGTGT CGTCGAGTAGTGGTGTATGGCCGGTGGTGGGATTAG CACC CCCAGCGGCATCCCAG  
 ACTT CAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATCCCTTACCCCGAAGCCATCTTT  
 GAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTCACTTTGGCCAAGGAG

CATT CGGA AGAAGGAGTGT CATGGAGTAGTGTAGTGGTGGTGGCAGCTGCTGGGATTAG CACCCCGAGCTGCAT  
 CCCAGACTTCAGGTCTCCTGGGAGTGGACTATAACAATAATCTTCAGCAGTATAATATTCATACCCTGAAGCCA  
 TATTTGAGCTTAGCTACTTTTTCCAGAATCCCAAGCCTTCTTCAGGTTAGCTAAGGAG

Because Exons 1 and 2 are not mapped, focus on the most mapped regions of Exons 3 and 4:

GGGAGGTGACAGTGGGAAGCAGAAGCTCACCTGCAGGATGTGGCAGAGCTCATT CGG  
 A AGAAGGAGTGT CGTCGAGTAGTGGTGTATGGCCGGTGGTGGGATTAG CACC CCCAGCGGC  
 ATCCAGACTT CAGGTCTCCGGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATC  
 CCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTC  
 ACTTTGGCCAAGGAGCTCTACCCTGGCAATTATAGACCCAAC TACGCCACTATTTCTCTG  
 AGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTG  
 GAGCGAG

- Region1/Exon3
- Region2/Exon3
- Region3/Exon3
- Region4/Exon3
- Region5/Exon3
- Exon4

Identify the corresponding python sequence from the alignments. Copy and paste them in order. Verify that you are copying the SUBJECT sequence (python) and not the QUERY sequence (reference species).

GGGAAGGTGACAGTGGGAAGAAGAA

AGCTGACCCTGCTGGTTTTGGCAGAGCTCATT

AGAAGGAGTGT CATGGAGTAGTG

CCCAGCTGCATCCCAGACTT

TAGTGGTGGTGGCAGCTGCTGGGATTAG

CAGGTCTCCTGGGAGTGGACTATAACAATAATCTTCAGCAGTATAATATTCCATACCCTGA  
AGCCATATTTGAGCTTAGCTACTTTTTCCAGAATCCCAAGCCTTTCTTCAGGTTAGCTAA  
GGAGTTGTACCCTGGCAATTACAGACCAAACCTATGCCCACTATTTTCTTCGACTTCTGTT  
TGACAAAGGGCTCCTTCTGCGCCTCTACACACAAAATATTGATGGGCTGGAGAGAG

Place the highlighted alignments from the python contigs in the order of the reference species sequence. This is the preliminary assembled python transcript for the gene.

GGGAAGGTGACAGTGGGAAGAAGAAGCTGACCCTGCTGGTTTTGGCAGAGCTCATTAGAAGGAGTGTTCATGGA  
GTAGTGTAGTGGTGGTGGCAGCTGCTGGGATTAGCAGGTCTCCTGGGAGTGGACTATAACAATAATCTTCAGCAG  
TATAATATTCCATACCCTGAAGCCATATTTGAGCTTAGCTACTTTTTCCAGAATCCCAAGCCTTTCTTCAGGTT  
AGCTAAGGAGTTGTACCCTGGCAATTACAGACCAAACCTATGCCCACTATTTTCTTCGACTTCTGTTTGACAAAG  
GGCTCCTTCTGCGCCTCTACACACAAAATATTGATGGGCTGGAGAGAG

Compare the highlighted reference species transcript from Ensembl to the assembled python transcript. Change the format of the reference species so it is exactly the same as the assembled python transcript; this will make it easier to identify potential problems in the assembly.

Chicken SIRT3 (mapped):

GGGGAGGTGACAGTGGGAAGCAGAAGCTCACCCCTGCAGGATGTGGCAGAGCTCATTTCGGAAGAAGGAGTGTTCGT  
CGAGTAGTGGTGTATGGCCGGTGTCTGGGATTAGCACCACCAGCGGCATCCCAGACTTCAGGTCTCCGGGGAGCGG  
CCTCTATAGTAACCTTGAGCAGTACAACATCCCTTACCCCGAAGCCATCTTTGAACTGGCCTACTTCTTCATCA  
ACCCCAAGCCATTCTTCACTTTGGCCAAGGAGCTCTACCCTGGCAATTATAGACCCAACTACGCCCACTATTTTC  
CTGAGACTCCTGCATGACAAAGGGCTCCTTCTGCGTCTCTATACTCAGAATATTGATGGGCTGGAGCGAG

There is missing sequence between mapped regions in the python transcript. Use the nucleotides from the reference species to fill in these gaps.

GGGAAGGTGACAGTGGGAAGAAGAAGCTGACCCTGCTGGTTTTGGCAGAGCTCATTTCGGAAGAAGGAGTGTCA  
TGGAGTAGTGTAGTGGTGGTGGCAGCTGCTGGGATTAGCACCACCAGCTGCATCCCAGACTTCAGGTCTCCTGG  
GAGTGGACTATAACAATAATCTTCAGCAGTATAATATTCCATACCCTGAAGCCATATTTGAGCTTAGCTACTTTTT  
TCCAGAATCCCAAGCCTTTCTTCAGGTTAGCTAAGGAGTTGTACCCTGGCAATTACAGACCAAACCTATGCCCACT  
TATTTTCTTCGACTTCTGTTTGACAAAGGGCTCCTTCTGCGCCTCTACACACAAAATATTGATGGGCTGGAGAG  
AG

There are six extra nucleotides in the python sequence. The problem is likely between mapped regions.

Between olive and turquoise, there is an extra 'A'.

GGGAAGGTGACAGTGGGAAGAAGAAGCTGACCCTGCTGGTTTTGGCAGAGCTCATTTCGGAAGAAGGAGTGTCA  
GGAGTAGTGTAGTGGTGGTGGCAGCTGCTGGGATTAGCACCACCAGCTGCATCCCAGACTTCAGGTCTCCTGG  
AGTGGACTATAACAATAATCTTCAGCAGTATAATATTCCATACCCTGAAGCCATATTTGAGCTTAGCTACTTTTT  
CCAGAATCCCAAGCCTTTCTTCAGGTTAGCTAAGGAGTTGTACCCTGGCAATTACAGACCAAACCTATGCCCACT  
ATTTTCTTCGACTTCTGTTTGACAAAGGGCTCCTTCTGCGCCTCTACACACAAAATATTGATGGGCTGGAGAGA  
G

TAGTG is repeated between green and teal.

GGGAAGGTGACAGTGGGAAGAAGAAGCTGACCCTGCTGGTTTTGGCAGAGCTCATTTCGGAAGAAGGAGTGTCA  
GGAGTAGTGGTGGTGGCAGCTGCTGGGATTAGCACCACCAGCTGCATCCCAGACTTCAGGTCTCCTGGGAGTGG  
ACTATAACAATAATCTTCAGCAGTATAATATTCCATACCCTGAAGCCATATTTGAGCTTAGCTACTTTTTCCAGA

ATCCCAAGCCTTTCTTCAGGTTAGCTAAGGAGTTGTACCCTGGCAATTACAGACCAAACCTATGCCCACTATTTTCTTCGACTTCTGTTTGACAAAGGGCTCCTTCTGCGCTCTACACACAAAATATTGATGGGCTGGAGAGAG

**TIPS:** Be sure to eliminate repeated sequence in the overlap between mapped contigs.

If there was missing sequence in the python (indicated by a '-' in the python contig), fill in the space with the nucleotide from the reference species.

If there was missing sequence in the reference query sequence, delete the corresponding nucleotide from the python.

## 6. Validate the assembly by translating the transcript into an amino acid sequence.

Translate sequence in all 3 reading frames, looking for an intact open reading frame (ORF) without introduced STOP codons. To do this, access the online translation algorithm <http://web.expasy.org/translate/>. The site can be easily found by performing a Google search for 'translate mrna sequence.' It will return the translation for all three reading frames in the 5' to 3' direction as well as 3' to 5'. Ignore the 3' to 5' translation. ORFs that begin with methionine will be highlighted in red, however, your sequence may not begin at the start so consider ALL of the translation, regardless of whether there is a highlighted ORF. When the assembled python transcript for SIRT3 is entered, the following results are returned:

### Translate Tool - Results of translation

Open reading frames are highlighted in red. Please select one of the following frames - in the next page, you will be able to select your initiator and retrieve your amino acid sequence:

#### 5'3' Frame 1

G K V T V G R R S **Stop** P C W F W Q S S F G R R S V **Met** E **Stop** W W W Q L L G L A P P A A S Q T S G L L G V D Y T I I F S S I I F H T L K P Y L S L A T F S R I P S L S S G **Stop** L R S C T L A I T D Q T **Met** P T I F F D F C L T K G S F C A S T H K I L **Met** G W R E

#### 5'3' Frame 2

G R **Stop** Q W E E A D P A G F G R A H S E E G V S W S S G G G S C W D **Stop** H P Q L H P R L Q V S W E W T I Q **Stop** S S A V **Stop** Y S I P **Stop** S H I **Stop** A **Stop** L L F P E S Q A F L Q V S **Stop** G V V P W Q L Q T K L C P L F S S T S V **Stop** Q R A P S A P L H T K Y **Stop** W A G E

#### 5'3' Frame 3

E G D S G K K K L T L L V L A E L I R K K E C H G V V V V A A A G I S T P S C I P D F R S P G S G L Y N N L Q Q Y N I P Y P E A I F E L S Y F F Q N P K P F F R L A K E L Y P G N Y R P N Y A H Y F L R L L F D K G L L L R L Y T Q N I D G L E R

Frame 1 has several STOP codons so this frame is incorrect.

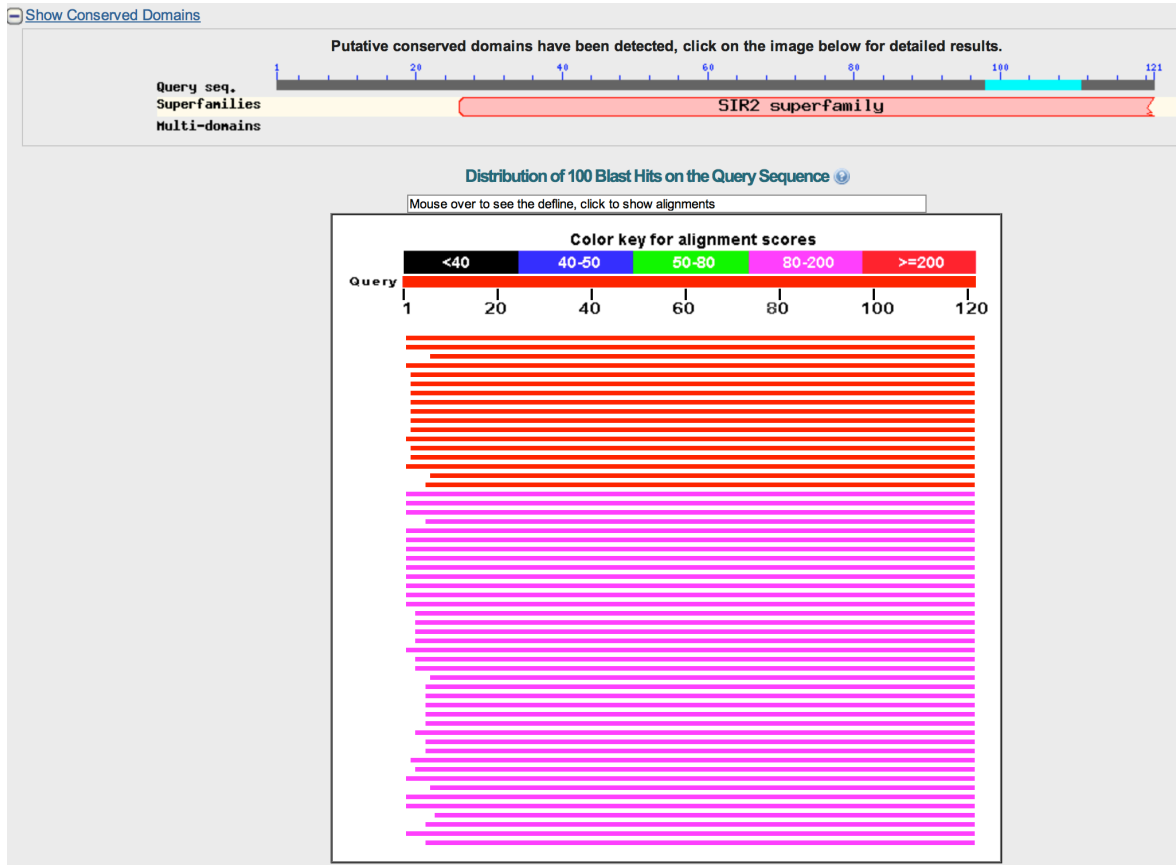
Frame 2 also has several STOP codons so this frame is incorrect.

Frame 3 has no introduced STOP codons. Proceed to Step 7 with this amino acid sequence.

## 7. Validate the amino acid sequence.

Copy the best predicted ORF. Paste it into Blastp (accessed by going to the Blast home page and choosing 'protein blast' under 'Choose a BLAST program to run'). Do not specify a species in this initial search. It should return chicken or lizard with a high degree of similarity and very few gaps, insertions or deletions. Move the exon boundaries as necessary.

The results for SIRT3, 5'3' Frame 3 are:



100 hits are identified. The format of BLASTp is the same as BLASTn. The identities of the hits are below the graphical depiction of the alignments.

Descriptions

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> PREDICTED: NAD-dependent deacetylase sirtuin-3, mitochondrial-like [Meleagris gallopavo]	219	219	99%	6e-68	86%	<a href="#">XP_003206163.1</a>
<input type="checkbox"/> sirtuin [Gallus gallus]	216	216	99%	9e-67	88%	<a href="#">NP_001186422.1</a>
<input type="checkbox"/> NAD-dependent deacetylase sirtuin-3, mitochondrial, partial [Ophiophagus hannah]	215	215	95%	3e-66	90%	<a href="#">ETE60955.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X1 [Zonotrichia albicollis]	215	215	99%	9e-66	83%	<a href="#">XP_005494853.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial [Falco cherrug]	207	207	98%	5e-64	85%	<a href="#">XP_005445118.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X1 [Anas platyrhynchos]	210	210	98%	6e-64	83%	<a href="#">XP_005030250.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X2 [Anas platyrhynchos]	210	210	98%	6e-64	83%	<a href="#">XP_005030251.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial [Falco peregrinus]	207	207	98%	6e-64	85%	<a href="#">XP_005239984.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent deacetylase sirtuin-3, mitochondrial-like [Anolis carolinensis]	210	210	98%	7e-64	83%	<a href="#">XP_003214844.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X3 [Pelodiscus sinensis]	209	209	98%	8e-64	84%	<a href="#">XP_006111794.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X1 [Pelodiscus sinensis] >ref XP_006111793.1  PREDIC	209	209	98%	1e-63	84%	<a href="#">XP_006111792.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X2 [Zonotrichia albicollis]	207	207	99%	2e-63	83%	<a href="#">XP_005494854.1</a>
<input type="checkbox"/> NAD-dependent deacetylase sirtuin-3, mitochondrial, partial [Anas platyrhynchos]	204	204	98%	6e-63	83%	<a href="#">EOA93867.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X3 [Anas platyrhynchos]	205	205	98%	8e-63	83%	<a href="#">XP_005030252.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial-like, partial [Melospittacus undulatus]	201	201	99%	1e-62	82%	<a href="#">XP_005140529.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial [Chrysemys picta bellii]	205	205	95%	3e-62	85%	<a href="#">XP_005313181.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial [Alligator sinensis]	202	202	95%	7e-61	83%	<a href="#">XP_006033400.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X2 [Ictidomys tridecemlineatus]	199	199	99%	2e-60	80%	<a href="#">XP_005341665.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X1 [Ictidomys tridecemlineatus]	199	199	99%	3e-60	80%	<a href="#">XP_005341664.1</a>
<input type="checkbox"/> PREDICTED: NAD-dependent protein deacetylase sirtuin-3, mitochondrial isoform X1 [Macaca fascicularis]	198	198	99%	3e-59	78%	<a href="#">XP_005576761.1</a>

*Gallus gallus* SIRT sequence appears as the second hit with 88% identity but it does not specify the isoform. However, SIRT3 is identified in 99 other species, including Anolis.

## 8. Design primers using the assembled transcript.

Perform a Google search for 'Primer3'. Click the first link to Primer3 – BioTools – the University of Massachusetts Medical School.

Use the presumptive transcript sequence to design primers in Primer3. Design two sets for each of TWO PCR products. The products will be 100-200 bases long, spanning at least one exon-exon boundary. Enter at least 200 bases from the assembled python transcript with the exon-exon boundary in the middle.

```
CATTTCGGAAAGAAGGAGTGTTCATGGAGTAGTGGTGGTGGCAGCTGCTGGGATTAGCACC  
CCCAGCTGCATCCCAG  
ACTTCAGGTCTCCTGGGAGTGGACTATACAATAATCTTCAGCAGTATAATATTCCATACCCTGAAGCCATATTT  
GAGCTTAGCTACTTTTTCCAGAATCCCAAGCCTTCTTCAGGTTAGCTAAGGAG
```

This is 207 bases long. By examining the reference species transcript (chicken), the exon-exon boundary is between the red and yellow regions.

Copy the assembled transcript and paste it into Primer3 search window.

Check 'Pick left primer or use left primer below' and 'Pick right primer or use right primer below (5' → 3' on opposite strand).'

Verify that the following conditions are selected:

Product Size	Min: 100	Opt: 150	Max: 200	
Primer Size	Min: 18	Opt: 20	Max: 27	
Primer Tm	Min: 57.0	Opt: 60.0	Max: 63.0	Max Tm Difference:
Product Tm	Min:	Opt:	Max:	
Primer GC%	Min: 20.0	Opt:	Max: 80.0	

Max Self Complementarity: 8.00    Max 3' Self Complementarity 3.00

Salt Concentration: 50.0    Annealing Oligo Concentration: 50.0

Primer3 will produce the following output:



Scroll down to 'Primer Pair Specificity Checking Parameters' and change the default organism (Homo sapiens) to Anolis. Leave all other settings as the default.

Click 'Get Primers'. The algorithm will now search all organism databases for sequences that your primers will amplify if a PCR were to be performed. Click 'Check' periodically if the site does not update.

The search returned the following for primer set 1:

Input PCR template: none  
 Specificity of primers: Target templates were found in selected database: Refseq mRNA (Organism limited to Anolis carolinensis)  
 Other reports: [Search Summary](#)

▼ Detailed primer reports

**Primer pair 1**

	Sequence (5'->3')	Length	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CAGCTGCTGGGATTAGCAC	19	58.60	57.89	7.00	1.00
Reverse primer	AGAAAGGCTTGGGATTCTGG	20	57.48	50.00	5.00	0.00

**Products on target templates**

>XR\_118763.1 PREDICTED: Anolis carolinensis WD repeat and FYVE domain-containing protein 3-like (LOC100566554), miscRNA

product length = 4455  
 Forward primer 1 CAGCTGCTGGGATTAGCAC 19  
 Template 4691 ..AT.....CT.....A 4709

Forward primer 1 CAGCTGCTGGGATTAGCAC 19  
 Template 9145 AG...GA..A..... 9127

>XM\_003221160.1 PREDICTED: Anolis carolinensis neuron navigator 3, transcript variant 2 (nav3), mRNA

product length = 3022  
 Forward primer 1 CAGCTGCTGGGATTAGCAC 19  
 Template 764 .....A..C.GC..T.. 782

Forward primer 1 CAGCTGCTGGGATTAGCAC 19  
 Template 3785 TT.....C.CT..... 3767

>XM\_003221159.1 PREDICTED: Anolis carolinensis neuron navigator 3, transcript variant 1 (nav3), mRNA

product length = 3016  
 Forward primer 1 CAGCTGCTGGGATTAGCAC 19  
 Template 764 .....A..C.GC..T.. 782

Forward primer 1 CAGCTGCTGGGATTAGCAC 19  
 Template 3779 TT.....C.CT..... 3761

>XM\_003220732.1 PREDICTED: Anolis carolinensis inter-alpha (globulin) inhibitor H5 (tih5), mRNA

product length = 1762  
 Reverse primer 1 AGAAAGGCTTGGGATTCTGG 20  
 Template 2544 .....C.TT.C.G.... 2525

Reverse primer 1 AGAAAGGCTTGGGATTCTGG 20  
 Template 783 T.C....AA...C..... 802

None of the products is SIRT3. If this result is returned, either continue using Primer-BLAST to validate the other returned sets or try a different region of the assembled transcript.

The first three primer sets share the same first primer. The fourth primer set was Primer-BLASTed:

NCBI/Primer-BLAST : results: Job id=-AXP6Q1DRE2qTKIU4sA20imMspbuS\_P [more...](#)

Input PCR template: none  
 Specificity of primers: Target templates were found in selected database: Refseq mRNA (Organism limited to Anolis carolinensis)  
 Other reports: [Search Summary](#)

▼ Detailed primer reports

**Primer pair 1**

	Sequence (5'->3')	Length	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTGCATCCCAGACTTCAGGT	20	59.38	55.00	4.00	3.00
Reverse primer	TTGCCAGGGTACAACCTCCTT	20	58.85	50.00	4.00	1.00

**Products on target templates**

>XM\_003214796.1 PREDICTED: Anolis carolinensis NAD-dependent deacetylase sirtuin-3, mitochondrial-like (LOC100555836), mRNA

product length = 153  
 Forward primer 1 CTGCATCCCAGACTTCAGGT 20  
 Template 692 TG.....G..T..... 711

Reverse primer 1 TTGCCAGGGTACAACCTCCTT 20  
 Template 844 .....T.....T..... 825

## 10. Refine the python transcript and validate the primer set in Primer-Blast, if necessary.



# The Python Project

## MCDB 4202

### Real Time PCR Primer Design - Controls

#### Spring 2014

### Goal:

Design a reverse primer that will amplify intronic sequence. The reverse primer will be used with the forward primer that was already designed. It will amplify a larger product (approximately 500 bases).

### Purpose:

- as a positive control, to validate the primer design method in genomic DNA,
- as a negative control, to validate the real time PCR experiment.

### Procedure:

#### 2. Using the assembled python transcript, map the designed primers.

The assembled python transcript for SIRT3 is below.

```
CATTCGGAAGAAGGAGTGTTCATGGAGTAGTGGTGGTGGCAGCTGCTGGGATTAGCACC  

CCCAGCTGCATCCCAGACTTCAGGTCTCCTGGGAGTGGACTATAACAATAATCTTCAGCAGTATAATATTC  

CATAACCCTGAAGCCATATTTGAGCTTAGCTACTTTTTCCAGAATCCCAAGCCTTTCTTCAGGTTAGCTAAGGAG
```

**Unhighlight** the assembled python transcript and color-code the exons in the python assembled transcript using the highlighted chicken transcript (below)

```
CATTCGGAAGAAGGAGTGTTCGTCGAGTAGTGGTGGTGGCAGCTGCTGGGATTAGCACC  

CCCAGCGGCATCCCAGACTTCAGGTCTCCGGGAGCGGCCTCTATAGTAACCTTGAGCAGTACAACATCCCTTACCCCGAAGCCATCTTT  

GAACTGGCCTACTTCTTCATCAACCCCAAGCCATTCTTCACTTTGGCCAAGGAG
```

Below is the unhighlighted python transcript with color-coded exons:

```
CATTCGGAAGAAGGAGTGTTCATGGAGTAGTGTAGTGGTGGTGGCAGCTGCTGGGATTAGCACCC  

CCCAGCTGCATCCCAGACTTCAGGTCTCCTGGGAGTGGACTATAACAATAATCTTCAGCAGTATAATATTC  

CATAACCCTGAAGCCATATTTGAGCTTAGCTACTTTTTCCAGAATCCCAAGCCTTTCTTCAGGTTAGCTAAGGAG
```

Primers that validated for this assembled transcript are below. If you had to choose primers that did not validate using Primer BLAST, list the primers that were ordered.

#### Primer pair 1

	Sequence (5'→3')	Length	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTGCATCCCAGACTTCAGGT	20	59.38	55.00	4.00	3.00

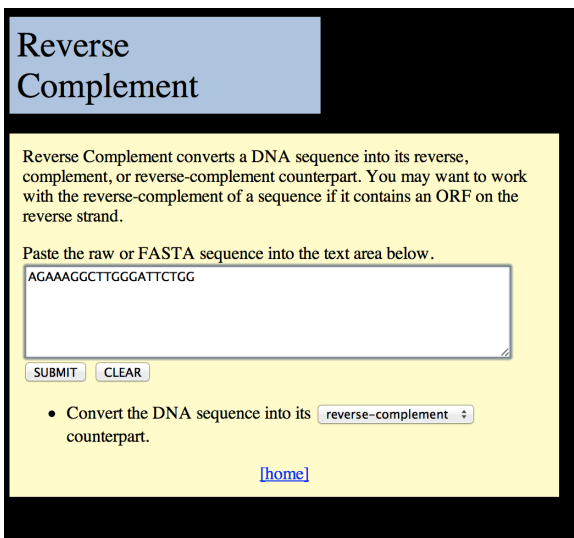
**Reverse primer** AGAAAGGCTTGGGATTCTGG 20 57.48 50.00 5.00 0.00  
Products on target templates

>XM 776345.3 PREDICTED: Strongylocentrotus purpuratus NAD-dependent protein deacetylase sirtuin-3, mitochondrial-like (LOC575990), mRNA  
product length = 122  
Forward primer 1 CTGCATCCCAGACTTCAGGT 20  
Template 624 .G..... 643  
  
Reverse primer 1 AGAAAGGCTTGGGATTCTGG 20  
Template 745 .....T..G....T. 726

Forward: CTGCATCCCAGACTTCAGGT  
Reverse: AGAAAGGCTTGGGATTCTGG

Create the *reverse complement* of the *reverse primer*. Use the website, [http://www.bioinformatics.org/sms/rev\\_comp.html](http://www.bioinformatics.org/sms/rev_comp.html), to design the reverse complement of the sequence. This site can also be found by performing a Google search for 'reverse complement bioinformatics.'

Enter the sequence of the reverse primer into the search box and click 'Submit,'



The website will calculate the reverse complement for you:

The Sequence Manipulation Suite: Reverse Complement  
Results for 20 residue sequence starting "AGAAAGGCTT".

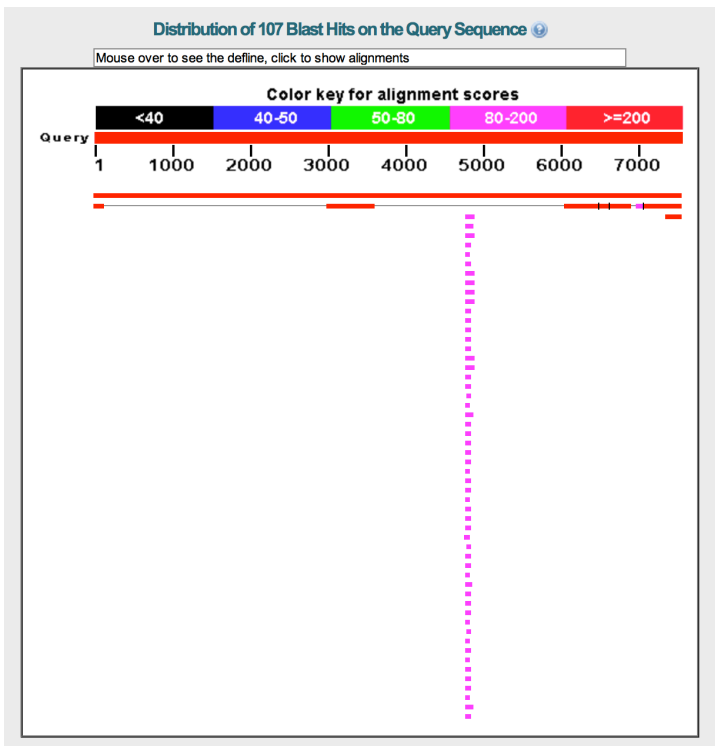
CCAGAATCCCAAGCCTTTCT

Reverse complement: CCAGAATCCCAAGCCTTTCT

Search for the forward primer and the reverse complement of the reverse primer in the assembled python transcript and highlight their locations. Verify that the ENTIRE sequence of the primer is correct and that the primers amplify a product that spans an intron.

CATTTCGGAAGAAGGAGTGTCATGGAGTAGTGTAGTGGTGGTGGCAGCTGCTGGGATTAGCACCCCCAGCTGCAT





The first hit is the correct contig; it should be in red and span the entire length of the Query. Click on the alignment and then select the Sequence ID in blue. This will link to the NCBI site for the contig, as it was deposited to the database.

NCBI Resources How To

Nucleotide  [Limits](#) [Advanced](#)

[Display Settings:](#) GenBank [Send to:](#)

### Python bivittatus Python\_molurus\_bivittatus-5.0.2-11215.3, whole genome shotgun sequence

GenBank: [AEQU02239036.1](#)  
[FASTA](#) [Graphics](#)

---

[Go to:](#)

LOCUS AEQU02239036 7483 bp DNA linear VRT 09-SEP-2013  
 DEFINITION Python bivittatus Python\_molurus\_bivittatus-5.0.2-11215.3, whole genome shotgun sequence.  
 ACCESSION AEQU02239036 [AEQU02000000](#)  
 VERSION AEQU02239036.1 GI:540702772  
 DBLINK BioProject: [PRJNA61243](#)  
 KEYWORDS WGS.  
 SOURCE Python bivittatus (Burmese python)  
 ORGANISM [Python bivittatus](#)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Lepidosauria; Squamata; Bifurcata; Unidentata; Episquamata; Toxicofera; Serpentes; Henophidia; Pythonidae; Python.  
 REFERENCE 1 (bases 1 to 7483)  
 AUTHORS Castoe,T.A., de Koning,J.A., Hall,K.T., Yokoyama,K.D., Gu,W., Smith,E.N., Feschotte,C., Uetz,P., Ray,D.A., Dobry,J., Bogden,R., Mackessy,S.P., Bronikowski,A.M., Warren,W.C., Secor,S.M. and Pollock,D.D.

Contigs can be extremely long, up to as many as 30,000 bases. Change the format of the NCBI page to FASTA so the sequence of the contig (shown at the bottom of the page) is without numbers or spaces.

Copy the ENTIRE contig sequence into a text document. Change the font to Courier, font size 8.

**Python bivittatus Python\_molurus\_bivittatus-5.0.2-11215.3, whole genome shotgun sequence**

GenBank: AEQU02239036.1

[GenBank Graphics](#)

>gi|540702772|gb|AEQU02239036.1| Python bivittatus Python\_molurus\_bivittatus-5.0.2-11215.3, whole genome shotgun sequence

TATGGGCAAAGACCAAAGCAGCTGCAGTATCACCAGAGAAAGGATAAGATGGCAGACATTCCTCCATAC  
AGTCCACAGGAGTCAGGGCTGACTCAGTGGCATCTAACCAACACAGTAAACACCTTGATTAAGGCTGAG  
CGCAACAAGTGAAGTACATAAATACTATTAGAGGAAAATGCTATATAAGCAAGCAATATCTGCC  
AAAGCCCCTCATTAATAAACCTCTCCCAAACTCAGCAAGTCTTGGAAATAATCCACAATGCACCTAA  
GTCTGCAATCAAGGCAGATGTAAGCAGAAGTTACTAAAATCACTGCATCATTTTTTCTTTATTTAACCA  
CCTCCATTTCTTCTCCTCCTCCTTCTTCTATGTGAGTAAGTAGGGGTGTTCAAAGTACTTACACACC  
ACTGAGTCAAATCTAGTAAGAGCTGAAAGAGCTATTAGAAAATTAAGCGTATTGCTTACAGGTACAGGT  
TTTATTTCTCCTCAGGGAGCTTTAGGCAATTTGGATTTGTCATGCAGCCCTCCATCTGTAGGATTAATCTGGG  
GAAAGCCATGACAGTTAAACTGGCATAAAAATGAGATAGGTATGCATAAAAATGTAATGATCTTCCC  
TTTCAAGGGCAAAAATATAGCCCTTCTGAAAGAGAAAATAACAAGATAACAGAGGCAGTCTCCATATACC  
TTTGGTGAAGTGTATAGGTCCAAAGTAGTTGGCATCCATGATCTTTTGTCAAGTTCAGTACATGTTG  
TGAAGTGTCCCTTGGCTGTTGCTAGCAGTGTGAATGAGTACATCAACGCATCCGTAACAATCCAGGA  
TTTCTTTAGCTACATCTTGGATACAGTTTACATCAGCAAGATCTAAAAGGATAAGCTTCGGGTGAAATGT  
CTACAAAAGAAGCAGGGTTTTTTTGTACCGCACTATGTCGCTGTAATACAGAGATATCTTAGTAGTTTT  
TTGAAGTACAATACTTAACAAAAAAGAGCTAATTTCCGTATGAGACAATAAAGCACCTTCCCAATTTG  
CATACTTTTGGATGTAGTGTCCAGAATTTCCAGTCAAGCTAGCTGACGGTAGAAGAAAATGGGGTTTAT  
CAGATGAGCACAAACATGTAGAAGGAAATCAGATCCAAAGATCAACATAAAGTAGAATCAGTACATTTAA  
GGTATAATTAGTAGCAGTTCCTTTGATAGAAGTGTCACTTTAGAAGATTTGGTAGTCTCCTGCTGATTA  
AATCCAGGATCACCTATGTCTCTTTTATATAAATTTTCTGACCTCATATTTGGAATAAAGTGTGTGAT  
TATATGGGAAAAAATGTGTATCATGTACACTACCTTTTACTATAAAGCAGCATCATTCATTCATCAA  
TTAAAAACAACCAACTAGACTGGCTCTAATTTTGATATTTTCCAGTAAATGCATAAATAATATAAATAAA  
CAGTGTATCATCTCACTAAGATAAATACTGAGAGACAACAAAATCAAAAATGCAAAGAGAGGCAA  
TAAAACTAATTTTAAATGTCCTTTGGGAAATAAGTTTGTCTTGATCGAATGCTAAAAATATTGCGAGGTT  
ATACTGTATGCACTGCATACTGTTTGAATGTTGAGTGTATCTATATTTTTCTGAAACTACATCTTAC  
AGCCTGTGGTCTATTGGCCTTTTATCAAACTAATTTACACAGGTTGAGAATGAAATAAATTAACCTGTA  
GATAATTAGTTTGTGGCAAGCAATTTAATTTAGACATACTAAAAAACAACAAAACCTGGCCATCT  
TATATCAAGTGTCTCTTTTGTTCCTTTTATATAGTTCTTTAAAGATATGCAATACAAGGGACTTACTGCA  
GTAGGATCAGCCACACTGTTTAAAGCATATAGAGAGCTTCTAATTTGTCAGGTGTTTCCCACACAAA  
TCAGCCTGTCTCCTCTGATGAATACTCTGTGAACATTTCTGGAAGATAAACCAAGGGATAATTCAGTT  
AATTTGGAATAGTAATACTGTATCTCTCTTCTTGTCTCATCAACCTAACACAGAAAAAATATCTTCTC  
ACCTCAGCGTTACCACACAGATGGACCCCCACACACACACTCAACTGTTTATGCTTCAAATACAGA  
TCTTATGGATGGTTAAACAAATGCGCTGTACCTCAAGCTATAATCCGTACATTTCCAAACAAACATATT  
TGTACCTCCGTTTGTACAAATGGATTCTATGAAGTCTGTGCAAGCATTTGGGTGCAACAAGTAGAAC  
CCAGGACCACTGCAGTTGGGCTACAGTACTGAAAACGAAAGACGTCATGATGTTTCCAGTGCAGATC  
CTCTGCAGTGCAGGCGTATGCAAGAAGTAAAGTGAATCATCAGAAATCATCAACCAAGAAATACCGT  
TCATGTAGGGGGCTGTGGACATAGGAGCTCAGGGTCCATTTCTCAGATCACCATAAATGCATCCATG  
TTGCTAGAGCAAGGTGAGGAGCTGTCATGCTATGCTGAGTGCAGGTAAGGAGCAGCTTATTTAATAC  
TCCTCATCAATGTAGAATAATGATGAGGATACAGAGCAGGATTTTTGCTTTGTCTGATTTCTTGGCGT  
CATCCACACAGGATGTGGTGAAGAGATGTATCTCTCTGTTACCAGCTTGGAAATGAATTGCTGGAAA  
TATCTTCCATAGAAAGTCAACCTGGGCGAGTTGTAGTGTGCAGTGGTGTATGCTTCCAAAGGTCTC  
CCCTCCCATCTGGAAGGTAAAGCTGGGAACCTGGGCAGCCAACACTTTCATGAGCTTCAACAACGTA  
CAGTATGTCAGTGGCAAGAGTGGCTACAGTACTGAAAACGAAAGACGTCATGATGTTTCCAGTGCAGAT  
ATAGAAAAAAGAAAGAAATCTACCCCTGGACATCTCAAGAAATAGCCTAGGATACACAGAGCATGCT  
CAGTAGCTGCAGCATGGAATGCCAGATGGAAGAGAAAACCAGAAAATACAGTGTAGCAGGAAAGAGAC  
AAAATGGCAGCTTGGAAAGGAGCAATGTTTGTGGCAAAACCCAGAGTAAAATCAATAACTGAATCAA  
ACTTTTAAATTAACCTGATTTCTTGGGTTTATACAACACAGAGTTAAAAATTAAGCCACCACATTTTA  
GATTAGTGAGATGCATGAACTCAACCAAGTATGCATCTCTCTATGTGAAAAAAGGACAAAATATTT  
AAAATAGGTGAACAGCATTTCTCTTTTCAAAAAACATACCACAATCTTCTTGTAGGATACCTATATAA  
GAGAAATCTGGCAATATGATTAAGCAAGCTTATAGAACAAGAAACACAGCTATCTACAATTTGGAATA  
TCTCTACCTATTTAAACAAGAGATTTCTTACATCTTCCACACTCTGTTCTGGTCCAGCTCCAGTAAACACA  
AGAGTTCTGCAAAATTCCTACTTGGGATGAAATGTATATCTGTATAATGTTGCTTGTCTGTCACCT  
GATGCAATTTTCTAGTATGTTTGGCAATGGTGCACATGACCCCAATGATTTCTCAGTTGATGTCCTCTG  
TTTCTCCCTCCTGACTGGCAGCAGCAGCAAGCTTCTGCAATCTTCTGAAAGGATCATTTTCTATT  
CCCCTGACATTTCTGTAGCTATGAGCTTCATTCATTCACCAAAACCATGGTGCAGGTTCAAGGTCTTTC  
TCCAAGGTCTGCAAAATAATTTTCTTCAAGATAACCCCTTGTATCTCAACTGCTTGTCTGACCTGACCTTAA  
AACTCAACCTGGTGCATAATTTAGGATACAGAAAATAAGCTATTTATATCACACTGCTGCCACTTCAT  
ACGTCCTGTATGCGCAAAATGGTTCTCACCACACTACAGTGCAGCCAGCCAACTATAAATAGATTTTTTTT  
TCCTAAGGCATACAGTTTGTAGAGGAGTTTTTCCAACATAATGTGGATCTAAGTGCCTTCTTCTTAAAC  
CCACATTTCAAGTATGCTTGAAGAGATTTGGTGGATGCAAAATCTTAATAAATAACAATTAAGTTT  
TTATGCTGCTTACAGTTTGTGCAAAATATTTATTTTAAACCATGAAATTTGAAATTTGGCGCCA  
AACAAAACAACAAAGCAATGCTAGATTCAGATTTCTTGGCAAAATCATACTACTGATGATCTATTGTTATA  
CATTTATACTGTCAATGCAGAGTTCAGTCCCTTTTTTACCCTTCCCGTCTTCAATACCCCATGGAT  
TATGTTTTAGTCTCTGATTTATGTAAGCTGCAATATTTGAAATTCGGGATTTCTGCAGACAAATGAAATTT  
GTTCCAAGGTGTTATAAATAGGTAGTACAGATGCATATGGACATTAAGAAGGCAATATTTAGATTTGGGACA  
GAACATAGATTTATTTCTGTAGGAAAGAAATGTATAGGTGAATAGGAAATAACAGAGTAAATTTGGAGAAC  
TCCTCTGCGAATGTGCTCACTTGAAGGGCCACTCTGCTCCAAATACACTATTTGCATGGCTTCCCTATT  
GATTTACGGGGCTGAAGCCTCGATCGTGGTCCAGTTTCTAGTTTTTGTAGTATTTAATGCTTCTGTATT  
TCTAGGTTTTCTTTTAAATCAAGGCTTACATTTCCCTATTTCAAGTGTCTCAATCTTAGCAACTTT  
AAGATGGCTGGAGTCCAGCTCCAGAATTTCCCTAGCCAACACTGAAGTCCAGATTTCTTAAAGTTGTTAA  
GGTTGAGAAACAGGGCCTATTTCTTTTAAATGAATGTGAAGGCTTGAATACTGATGTGTACGCA  
TCAAAATAGCTAGATTTCCAGGAGGCAAACTGCAGTATAGAAATGTTTACTTTTTGCAACCTGTGTACGTTT  
GTGTGGTTACACTATAGAAATATAAGTCACTGCTGCGCTGCTTACCCTTCCCAAGCCAGAGAGATCAT  
ACTGATTTATACAACCTTGTCCGCATAGCAGACTTCGACATCAGCTGAGTAACCATCCGATAAATGTAT

ACAATCCCAGTATTCCAACCACAAGCAGTGGGAGGATTAGTATCGCCAAAATACCCATCGTCACCTAGC  
GAAAAAGGTAATTTGAAGTCACAGAGAAGATCCATCAAATTTGCCACAAAAGATTCATTTTTCTTTGAAA  
ATATACTTAAGCAAAGATCAGTGACGAGGAAACAGGTAATCTATTGCATATCTTGACCCAAACCCATAA  
TTAAGATAGAACATTATTTATTCAGCATCTCTCAGTACATATGGAGATTTTCAAAGTGCACAAAAGCCTA  
TTGTTTTGTTTTGCTTTGGAAATGCAGAGCATTTTTTTCTCTTCAAATGAACCCATGCTATATATTT  
TTATCTTGTAATAAAGACTGTGACACAAATCCATTTAAATCCATTTCTTATTGCTACTATATATTA  
TAAACAGCTGTGTGCAGCAATCCATTTTATTGTCAGTGCAAAAAAAAAAAAAAAAAAGGAAAGGAGTAGAG  
GAAAAAAAAAACTTAAGTATTTCTCTTAAATTTATAAAGACCAAACAGCAAATTACGTCAGACATTAAT  
CCCATGTGACATCCAACCTACAGTTCTGCAATAAAGCTAATAACTTACTTGGAAAGGCAAGCGGTGAGTCTG  
CTGATAAGGTGATGTCTCTTGCCAACAGGTTTATCCAGAAGTTAATTAACCTCGCAGACATTTTTGGA  
AAGGAAAGCAGTCCACCAGCTGTAGACTTTGCTCTGGAGCAAAAAGCAGCTATGTTTGTAGTGCCAGCTGC  
ATCCAGACTTTAAAAAAGTAAATATAGTATCTGTGGTCTCAGCCACATACCTGACCTATTTTGTTCAC  
AGGGCTTCCACGCTCTTCTATAGCCTGATACAGAATGTAGAAAGAAAGTCTCTTCAATTTATTTGTTGTC  
TCTCAGAACTGTAGTTACACGAGGCAAGGAAATCATTTTATAGGACCCATCTGCATTCGCATGGTTTC  
TGGCCCCATTTGCAATAGTACACGATGGAAAGCTCACTTTCCACAATCCTTATGTGGTGGGAATGCCACT  
AGAACTGACTATAACATTTCTATATGTAGATAAGAATATTTTTTTAAGTATTAATATATTTCTAGTTTA  
CCACGTAGTGTTCCTCCAGTTTGGTGTGCCACTGTGTCTCTTTAAGTGTATCGCTTTGGGTGTTTCT  
AATTAGTTAGAAGAGGAATAGTACCAGTAAATATATAGGCTCTGTGTGAGGTTAAAGTTTGGCATGGCA  
GAGTTAGTTCCACACTTAATTTAGAGCTAGGCACAAATGACATTTTGAAATGATTAAGGCTGAGCCAG  
CGACCTCTGGGGAAGTAAATGGCAAACGAAGATGGCTTCACCTTCAGCGGAGCCGACAACCTGTGGCATA  
ATGAGGAACAGTGAAGGAGCAAAACCTAGAAATGTCTCAACTTTTCACTTTATAATGAGCTATGTCTGA  
TCCATGGGATGTGATGTGCATGACAGCTTGATCAGTCCATTTGTGCTGATACCTGAAAGATGAAGTA  
CCTCTGGCAGTGAATCTCTGGGTCTCATGAAGACTCTATACCTGGAAAACAAAATAATGCAAGTAAAT  
TCATACTTGGGCTCAGTATGTTTGTGTATTAATTTTTTTTGGCATGGTTGAACATTTGATAGTCTG  
GTTAATAGTTCAGACAATGGGTATTTTAGTGGACTCTGAGGATTTTGAATGACAGATACCTTTTAC  
TGTAATAGGACTGCGAGTCTTGGATCTTAATGAAAATATACTTTTTCTTTGCATAGCACATACAGTGTG  
ATTTAAACATGTTTTTAATAAGAATAAAGGAGTTGTTTTTACTAATCATAAAGCAAAGTCCAGAGCTTA  
AGTTTTCTGAAAACCTTGGTGTAGACTCTAAAGTCCCTTCAAATGTGCCAGTATTTTATGTGCCAGTA  
GAAGGAAAATATTTGATATGTGAGAGTAATCATATTAAGTGGATGTAATTAATTAAGACAAATTCATTT  
TGCTCTGGTATGCGATTAAAAGTTCTACTGTGATTTTTCTGAAAACAGTTGTGCTATAAAATGTATAC  
ATTCTCTGTCAAATGTGATTTCACTTATTTCTCCAGTCAATTTAAGAAAGTGAATATAACATTTGGGA  
AATGGAAAATAGAGCAGATAGTAAATTTTACAATCAGGCATCTCTGATTCATTCAGCGGCATTTTGT  
TCCCATCAACGTGAAGTCAATTTTCAATATGTGATCTTTTAGTAGTAGCAGACTGCAATTTACAGCTGAG  
AGAGGATGGATGTAGTATACCTTACAAAAGACATCTAGACAAGGGGACTACAATCATTCT

#### 4. Map the forward primer in the python contig genomic DNA sequence.

Search for the forward primer in the sequence. If it cannot be found, search for the reverse complement of the sequence. This would be necessary if the reference species transcript aligned in the plus/minus orientation with the python contig.

Highlight the forward primer. In this case, the primer spans an exon boundary, highlight only the part that lies in this exon.

Select approximately 500 bases from the beginning of the primer, if available. If 500 bases are not available, select as many as possible but not less than 200.

For this example, 494 bases were selected.

**CTGCATCCAG**ACTTTAAAAAAGTAAATATAGTATCTGTGGTCTCAGCCACATACCTGACCTATTTTGTTCACAGGGCTTCCACGTCCTTCTATAGCCTGATACAGAATGT  
TAGAAAGAAAGTCTTTCATTTATTTGTTGCTCTCAGAATCTGTAGTTACACGAGGCAAGGAAATCATTTTTAGGAGCCCATCTGCATTTCGCATGGTTTCTGGCCCCATT  
TGCAATAGTACACGATGGAAAGCTCACTTTCCACAATCCTTATGTGGTGGGAATGCCACTAGAACTGACTATAACATTTCTATATGTAGATAAGAATATTTTTTTAAGTA  
TTAATATATTTCTAGTTTACCACGTAGTGTTCCTCCAGTTTGGTGTGCCACTGTGTCTCTTAAAGTGTATATCGCTTTGGGTGTTTCAATTAGTTAGAAGAGGAATAGTAC  
CAGTAAATAATATGGCTCTGTGCAGGTTAAAGTTTGGCATGGCA

#### 5. Design a reverse primer that amplifies a PCR product that is approximately 450 bases long.

Enter the section of the contig sequence that contains the forward primer. If the forward primer is incomplete as in SIRT3 (the primer spans an exon-exon boundary), design a new set of forward and reverse primers with the forward primer in the exon as close to the junction as possible and the reverse primer in the intron.

**CTGCATCCAG**ACTTTAAAAAAGTAAATATAGTATCTGTGGTCTCAGCCACATACCTGACCTATTTTGTTCACAGGGCTTCCACGTCCTTCTATAGCCTGATACAGAATGT  
TAGAAAGAAAGTCTTTCATTTATTTGTTGCTCTCAGAATCTGTAGTTACACGAGGCAAGGAAATCATTTTTAGGAGCCCATCTGCATTTCGCATGGTTTCTGGCCCCATT  
TGCAATAGTACACGATGGAAAGCTCACTTTCCACAATCCTTATGTGGTGGGAATGCCACTAGAACTGACTATAACATTTCTATATGTAGATAAGAATATTTTTTTAAGTA



```

RIGHT PRIMER      453  21  59.99  42.86  2.00  0.00  GAAACACCCAAAGCGATAACA
PRODUCT SIZE: 436, PAIR ANY COMPL: 4.00, PAIR 3' COMPL: 1.00

3 LEFT PRIMER     28  20  59.02  50.00  6.00  2.00  CAGCTGCATCCCAGACTTTA
RIGHT PRIMER     427  20  60.37  55.00  5.00  2.00  AGAGACACAGTGGCACACCA
PRODUCT SIZE: 400, PAIR ANY COMPL: 3.00, PAIR 3' COMPL: 0.00

4 LEFT PRIMER     28  20  59.02  50.00  6.00  2.00  CAGCTGCATCCCAGACTTTA
RIGHT PRIMER     452  20  58.15  40.00  2.00  0.00  AAACACCCAAAGCGATAACA
PRODUCT SIZE: 425, PAIR ANY COMPL: 4.00, PAIR 3' COMPL: 2.00

```

Statistics

	con	too	in	in	no	tm	tm	high	high	high			
	sid	many	tar	excl	bad	GC	too	too	any	3'	poly	end	
	ered	Ns	get	reg	GC%	clamp	low	high	compl	compl	X	stab	ok
Left	794	0	0	0	84	0	280	230	0	26	17	13	144
Right	872	0	0	0	1	0	499	157	0	10	0	14	191

Pair Stats:

considered 1452, unacceptable product size 1325, high end compl 35, ok 92  
primer3 release 1.1.4

Map the primer set with the best (lowest) *any* and 3' scores into the exon/intron sequences. Note the junction between the exon and intron. For SIRT3, the first set is used. Again, the reverse complement of the reverse primer must be made to map it.

```

GGAGCAAAAGCAGCTATGTTTAGTGCCCAGCTGCATCCCAG | ACTTTAAAAAAGTAAATATAGTATCTGTGGTCTCAGCCACATACCTGACCTATTTTGTTCACAGGGCTT
CCACGTCCTTCTATAGCCTGATACAGAATGTTAGAAAAGAAAGTCTCTTCATTTATTTGTTTGCTCTCAGAATCTGTAGTTACACGAGGCAAGGAAATCATTTTATAGGAGCCC
ATCCTGCATTTCGCATGGTTTCTGGCCCCATTTGCAATAGTACAGATGGAAAGCTCAGTTTCCACAATCCTTATGTGGTGGGAATGCCACTAGAAACTGACTATAACATTTT
TATATGTAGATAAGAATATTTTAAAGTATTAATATATTCAGTTTACCACGTAGTGTTCCTCCAGTTTGGTGTGCCACTGTGTCTCTTAAAGTGTATATCGCTTTGGG
TGTTTCTAATTAGTTAGAAGAGGAATAGTACCAGTAAATAATATGGC

```

The forward primer should be entirely in the exon. The second set is mapped as an alternative:

Reverse complement of the reverse primer:

TGTTATCGCTTTGGGTGTTTC

```

GGAGCAAAAGCAGCTATGTTTAGTGCCAGCTGCATCCCAG | ACTTTAAAAAGTAAATATAGTATCTGTGGTCTCAGCCACATACCTGACCTATTTTGTTCACAGGGCTT
CCACGTCCTTCTATAGCCTGATACAGAATGTTAGAAAAGAAAGTCTCTTCATTTATTTGTTTGCTCTCAGAATCTGTAGTTACACGAGGCAAGGAAATCATTTTATAGGAGCCC
ATCCTGCATTTCGCATGGTTTCTGGCCCCATTTGCAATAGTACAGATGGAAAGCTCAGTTTCCACAATCCTTATGTGGTGGGAATGCCACTAGAAACTGACTATAACATTTT
TATATGTAGATAAGAATATTTTAAAGTATTAATATATTCAGTTTACCACGTAGTGTTCCTCCAGTTTGGTGTGCCACTGTGTCTCTTAAAGTGTTATCGCTTTGGG
TGTTTCTAAATTAGTTAGAAGAGGAATAGTACCAGTAAATAATATGGC

```

The following product will be produced:

```

TGTTTAGTGCCAGCTGCATCCCAGACTTTAAAAAGTAAATATAGTATCTGTGGTCTCAGCCACATACCTGACCTATTTTGTTCACAGGGCTTCCACGTCCTTCTATAGC
CTGATACAGAATGTTAGAAAAGAAAGTCTCTTCATTTATTTGTTTGCTCTCAGAATCTGTAGTTACACGAGGCAAGGAAATCATTTTATAGGAGCCCATCCTGCATTCGCATGG
TTTCTGGCCCCATTTGCAATAGTACAGATGGAAAGCTCAGTTTCCACAATCCTTATGTGGTGGGAATGCCACTAGAAACTGACTATAACATTTCTATATGTAGATAAGAAT
ATTTTAAAGTATTAATATATATTCAGTTTACCACGTAGTGTTCCTCCAGTTTGGTGTGCCACTGTGTCTCTTAAAGTGTTATCGCTTTGGGTGTTTC

```

Perform a Word Count to determine the expected size of the product. This product should be 437 bases long.

Clearly note at the end of the document the primer(s) you have designed. If your forward primer for real time PCR mapped entirely to a single exon, you will note *one reverse primer*. If your forward primer spanned an exon-exon boundary, you will note *one forward primer and one reverse primer*.



# The Python Project

## Fall 2014

### Primer Test 1 – PCR

#### Purpose:

The purpose of the test is to determine how many products the primers for real time PCR amplify. This experiment will also indicate the size of the product produced.

#### Method Overview:

The primers will be added to 750ng cDNA synthesized from RNA isolated from hearts of pythons at a time point of your choice. A master mix containing dNTPs, *Taq* polymerase, and buffer containing magnesium will be added to the cDNA and primers. The number of reactions set up will equal the number of sets of real time PCR primers tested.

#### Method:

1. Prepare the master mix that is specific to your primers.

Master Mix	X 1	X n + 2
PCR Reaction Buffer containing magnesium	2.5 $\mu$ L	
10nM dNTPs	1 $\mu$ L	
100ng/uL cDNA	1.5 $\mu$ L	
Water	16 $\mu$ L	

2. Add 21 $\mu$ L of Master Mix to each PCR tube.
3. Add 1 $\mu$ L each of 12.5 $\mu$ M Forward and 12.5 $\mu$ M Reverse Primers so the total PCR reaction volume is 25 $\mu$ L.
4. Store on ice.

# The Python Project, Spring 2014

## RNeasy Protocol

### Introduction

The RNeasy Kits are designed to purify RNA from small amounts of starting material. They provide a fast and simple method for preparing up to 100 µg total RNA per sample. The purified RNA is ready for use in downstream applications such as:

- RT-PCR and real-time RT-PCR
- Differential display
- cDNA synthesis
- Northern, dot, and slot blot analyses
- Primer extension
- Poly A+ RNA selection
- RNase/S1 nuclease protection
- Microarrays

The RNeasy Kits allow the parallel processing of multiple samples in less than 30 minutes. Time-consuming and tedious methods, such as CsCl step-gradient ultracentrifugation and alcohol precipitation, or methods involving the use of toxic substances, such as phenol and/or chloroform, are replaced by the RNeasy procedure

### Principle and procedure

#### **RNA purification using RNeasy technology:**

The RNeasy procedure represents a well-established technology for RNA purification. This technology combines the selective binding properties of a silica-based membrane with the speed of microspin technology. A specialized high-salt buffer system allows up to 100 µg of RNA longer than 200 bases to bind to the RNeasy silica membrane.

Biological samples are first lysed and homogenized in the presence of a highly denaturing guanidine-thiocyanate-containing buffer, which immediately inactivates RNases to ensure purification of intact RNA. Ethanol is added to provide appropriate binding conditions, and the sample is then applied to an RNeasy Mini spin column, where the total RNA binds to the membrane and contaminants are efficiently washed away. High-quality RNA is then eluted in 30–100 µl water.

With the RNeasy procedure, all RNA molecules longer than 200 nucleotides are purified. The procedure provides an enrichment for mRNA since most RNAs <200 nucleotides (such as 5.8S rRNA, 5S rRNA, and tRNAs, which together comprise 15–20% of total RNA) are selectively excluded. The size distribution of the purified RNA is comparable to that obtained by centrifugation through a CsCl cushion, where small RNAs do not sediment efficiently.

#### **Isolation of Total RNA from Animal Cells and Tissues:**

Samples (maximum  $1 \times 10^7$  cells or 30 mg tissue, are disrupted in lysis buffer containing GITC (Buffer RLT) and homogenized. Ethanol is then added to the lysate, creating conditions that promote selective binding of RNA to the RNeasy membrane. The sample is then applied to the RNeasy mini spin column. Total RNA binds to the membrane, contaminants are efficiently washed away, and high-quality RNA is eluted in water.

#### **Equipment and Reagents to Be Supplied by User:**

When working with chemicals, always wear a suitable lab coat, disposable gloves, and protective goggles. For more information, consult the appropriate material safety data sheets (MSDSs), available from the product supplier.

**For all protocols:**

14.3 M  $\beta$ -mercaptoethanol ( $\beta$ -ME) (commercially available solutions are usually 14.3 M)  
Sterile, RNase-free pipet tips  
Microcentrifuge (with rotor for 2 ml tubes)  
96–100% ethanol  
Disposable gloves  
Equipment for sample disruption and homogenization  
Blunt needle and syringe  
Mortar and pestle

## Important Notes

**Determining the amount of starting material:**

It is essential to use the correct amount of starting material in order to obtain optimal RNA yield and purity. The maximum amount that can be used is determined by:

- The type of sample and its RNA content
- The volume of Buffer RLT required for efficient lysis
- The RNA binding capacity of the RNeasy spin column

When processing samples containing average or low amounts of RNA, the maximum amount of starting material can be used. However, even though the RNA binding capacity of the RNeasy spin column is not reached, the maximum amount of starting material must not be exceeded. Otherwise, lysis will be incomplete and cellular debris may interfere with the binding of RNA to the RNeasy spin column membrane, resulting in lower RNA yield and purity.

**Eliminating genomic DNA contamination:**

Generally, DNase digestion is not required with RNeasy Kits since RNeasy silica membrane technology efficiently removes most of the DNA without DNase treatment. However, further DNA removal may be necessary for certain RNA applications that are sensitive to very small amounts of DNA (e.g., TaqMan RT-PCR analysis with a low-abundance target). In these cases, residual DNA can be removed by optional oncolumn. The DNase is efficiently removed in subsequent wash steps. Alternatively, residual DNA can be removed by a DNase digestion after RNA purification.

# RNeasy Mini Protocol for Isolation of Total RNA from Animal Tissues

Important notes before starting:

- Use an appropriate amount of tissue.
- Some tissues, including heart, spleen, and brain are difficult to homogenize. The volume of lysis buffer may need to be increased to facilitate complete homogenization and to avoid reduced yields.
- Fresh or frozen tissue can be used. To freeze tissue for long-term storage, flash freeze in liquid nitrogen and transfer immediately to  $-70^{\circ}\text{C}$  for storage up to several months. To process, do not allow tissue to thaw (e.g. during weighing) prior to disruption in Buffer RLT. Tissue lysates (in Buffer RLT) can also be stored at  $-70^{\circ}\text{C}$  for several months. To process frozen lysates, thaw samples and incubate for 10 min at  $37^{\circ}\text{C}$  in a water bath to dissolve salt. Continue with step 2.
- Buffer RLT may form a precipitate upon storage. If necessary, warm to redissolve.
- $\beta$ -Mercaptoethanol ( $\beta$ -ME) must be added to Buffer RLT before use. Add  $10\ \mu\text{l}$   $\beta$ -ME per 1 ml of Buffer RLT. The solution is stable for 1 month.
- Buffer RPE is supplied as a concentrate. Before using for the first time add 4 volumes of ethanol (96–100%) as indicated on the bottle to obtain a working solution.
- All steps of the RNeasy protocol (including centrifugation) should be performed at 20 to  $25^{\circ}\text{C}$ . During the procedure, work quickly.

## 1. Disrupt tissue and homogenize lysate.

Note: Incomplete disruption and homogenization will lead to significantly reduced yields, and can cause clogging of the RNeasy mini spin column.

Simultaneously disrupt and homogenize the sample using Buffer RLT and a rotor–stator homogenizer.

Place fresh or frozen tissue in a suitably sized vessel for the homogenizer. Add the appropriate volume of Buffer RLT (see below), and homogenize immediately until a completely homogeneous lysate is obtained (typically 20–40 sec).

Starting material Volume of Buffer RLT

up to 20 mg:  $350\ \mu\text{l}$

20 to 30 mg or if tissue is difficult to lyse:  $600\ \mu\text{l}$

Note: Ensure  $\beta$ -ME is added to Buffer RLT before use starting.

Disruption and homogenization of starting materials

Efficient disruption and homogenization of the starting material is essential for all intracellular RNA isolation procedures. Disruption and homogenization are two distinct steps.

**Disruption:** Complete disruption of cells walls and plasma membranes of cells and organelles is absolutely required to release all the RNA contained in the sample. Different samples require different methods to achieve complete disruption. Incomplete disruption results in significantly reduced yields.

**Homogenization:** Homogenization is necessary to reduce the viscosity of the cell lysates produced by disruption. Homogenization shears the high-molecular weight genomic DNA and other high-molecular-weight cellular

components to create a homogeneous lysate. Incomplete homogenization results in significantly reduced yields.

**Disruption and homogenization using rotor–stator homogenizers:** Rotor–stator homogenizers thoroughly disrupt and simultaneously homogenize, in the presence of lysis buffer, animal tissues in 5–90 sec depending on the toughness of the sample. Rotor–stator homogenizers can also be used to homogenize cell lysates. The rotor turns at a very high speed causing the sample to be disrupted and homogenized by a combination of turbulence and mechanical shearing. Foaming of the sample should be kept to a minimum by using properly sized vessels, by keeping the tip of the homogenizer submerged and holding the immersed tip to one side of the tube. Rotor–stator homogenizers are available in different sizes and operate with differently sized probes. Probes with diameters of 5 mm and 7 mm are suitable for volumes up to 300  $\mu$ l and can be used for homogenization in microfuge tubes. Probes with a diameter of 10 mm or above require larger tubes.

**2. Centrifuge lysate for 3 min at maximum speed in a microcentrifuge and use only the supernatant in subsequent steps.**

For some samples very small amounts of insoluble material will be present, making the pellet invisible.

**3. Add 1 volume (600  $\mu$ l) of 70% ethanol to the cleared lysate, and mix well by pipetting. Do not centrifuge.**

If some lysate is lost during homogenization, reduce volume of ethanol accordingly. A precipitate may form after the addition of ethanol, but this will not affect the RNeasy procedure.

**4. Apply 700  $\mu$ l of the sample, including any precipitate that may have formed, to an RNeasy mini spin column sitting in a 2-ml collection tube. Centrifuge for 15 sec at 8,000 x g (10,000 rpm).**

If the volume of the mixture exceeds 700  $\mu$ l, load aliquots successively onto the RNeasy column and centrifuge as above. Reuse the collection tube but discard flow-through\* after each step. Reuse the collection tube in step 5.

**5. Pipet 700  $\mu$ l Buffer RW1 onto the RNeasy column, and centrifuge for 15 sec at 8,000 x g (10,000 rpm) to wash.**

Discard flow-through\* and collection tube.

**6. Transfer RNeasy column to a new 2-ml collection tube (supplied). Pipet 500  $\mu$ l Buffer RPE onto the RNeasy column, and centrifuge for 15 sec at 8000 x g (10,000 rpm) to wash.**

Discard flow-through and reuse the collection tube in step 7.

Note: Ensure ethanol is added to Buffer RPE before use.

**7. Pipet 500  $\mu$ l Buffer RPE onto RNeasy column, and centrifuge for 2 min at maximum speed to dry the RNeasy membrane. Continue directly with step 8, or to eliminate any chance of possible Buffer RPE carryover, continue first with step 7a.**

It is important to dry the RNeasy membrane since residual ethanol may interfere with subsequent reactions. This spin ensures that no ethanol is carried over during elution.

Note: Following the spin, remove the RNeasy column from the collection tube carefully so that the column does not contact the flow-through as this will result in carryover of ethanol.

**7a. Place the RNeasy spin column in a new 2-ml collection tube (not provided), and discard the old collection tube with the filtrate. Centrifuge at full speed for 1 min.**

**8. Transfer RNeasy column into a new 1.5-ml collection tube (supplied) and pipet 30–50  $\mu$ l of RNase-free water**

**directly onto the RNeasy membrane. Centrifuge for 1 min at 8,000 x g (10,000 rpm) to elute. Repeat if the expected RNA yield is >30 µg.**

If a second elution step is performed, elute into the same collection tube using another 30–50 µl RNase-free water.

# The Python Project

## Spring 2014

### cDNA Synthesis

#### RATIONALE:

cDNA or complementary DNA is a single stranded DNA copy of mRNA sequences. An RNA-dependent reverse transcriptase is used to elongate the cDNA. As with other polymerases, reverse transcriptase requires a double-stranded sequence at the 3' end. Hybridization of random hexamers is therefore required *before* addition of the reverse transcriptase. We will add random hexamers (primers) and dNTPs (nucleotides) to the RNA samples and then heat them to promote **hybridization**. We will then add the SuperScript III reverse transcriptase and heat again to promote **elongation**. Superscript III achieves full activity at 50°C. RNase OUT is also added in this step to degrade RNases that might be activated. This is a one stage PCR without repeating cycles.

**NOTE:** The non-template control (NTC) produced in this cDNA synthesis will contain no cDNA because there will be no Superscript III. DO include RNA in this step of the cDNA synthesis. The NTC will be used in the real time PCR reaction as an indication of whether the RNA contained contaminating genomic DNA that can be amplified with your real time PCR primers.

#### METHOD OVERVIEW:

Prepare tubes with each sample containing *RNA plus dNTPs, Random hexamers, and water* and place on ice

Heat samples containing *RNA plus dNTPs, Random hexamers, and water* at 65°C for 5 minutes in the thermocycler

Transfer the sample tubes to ice

Add the Master Mix containing *first strand buffer, DTT, RNase OUT, and SuperScript III or Water* AFTER the samples have been heated at 65°C

Transfer the sample tubes to the thermocycler for one-step elongation

#### METHOD DETAILS:

1. Calculate volume (based on spectrophotometer reading) of 2,000ng (2µg) RNA.
2. Set up a chart: Samples containing RNA from heart will be used for cDNA synthesis using random hexamers. One tube will also be created for a non-template control (NTC) that will contain no SuperScript III. In the last column of the chart, calculate the amount of water that will be required to make the final volume 13µL.

Experiment	Volume of Sample Containing 500ng RNA (µL)	dNTPs	Hexamers	Volume of Water Required to Make a Final Volume of 13µL
Fasted		1µL	1µL	
1dpf		1µL	1µL	
3dpf		1µL	1µL	
10dpf		1µL	1µL	
NTC – this can include RNA from any sample without SSIII		1µL	1µL	

3. Calculate volumes for the Master Mixes:

**NOTE: DO NOT ADD MASTER MIX UNTIL AFTER THE RNA/Hex/dNTP/water SAMPLES HAVE BEEN HEATED FOR 5 MINUTES AT 65°C**

Two Master Mixes are required: 1 for Random Hex cDNA synthesis and 1 for NTC (does not contain SSIII)

MMI (cDNA):	5X FSB	4 $\mu$ L	
	0.1M DTT	1 $\mu$ L	
	RNase OUT	1 $\mu$ L	
	<b>SSIII</b>	<b>1<math>\mu</math>L</b>	
		-----	
			7 $\mu$ L Total X number of samples (4) + a couple of extra

MM2 (NTC):	5X FSB		4 $\mu$ L
	0.1M DTT	1 $\mu$ L	
	RNase OUT	1 $\mu$ L	
	<b>SSIII</b>	<b>0<math>\mu</math>L</b>	
	<b>H<sub>2</sub>O</b>	<b>1<math>\mu</math>L</b>	
		-----	
			7 $\mu$ L Total X number of samples (1) + an extra

4. Set up tubes:

Tubes 1-4:	Samples containing RNA, dNTPs, Random Hexamers, Water
Tube 5:	NTC containing dNTPs, Random Hexamers, Water

**PCR reaction:**

STEP 1	HEAT AT 65°C	5 Minutes
STEP 2	PLACE AT 4C (on ice)	
STEP 3	ADD THE MASTER MIX (7 $\mu$ L) to each tube	
STEP 4	ELONGATION at 55°C	60 Minutes
STEP 5	INACTIVATION at 70°C	15 Minutes

\*\*\*Assuming that the reaction is 100% efficient (the amount of RNA put in compared to the cDNA that is produced is 1:1), 20 $\mu$ L total reaction volume containing 2,000ng RNA will produce cDNA at 100ng/ $\mu$ L.



# The Python Project

## Spring 2014

### cDNA Synthesis

#### RATIONALE:

cDNA or complementary DNA is a single stranded DNA copy of mRNA sequences. An RNA-dependent reverse transcriptase is used to elongate the cDNA. As with other polymerases, reverse transcriptase requires a double-stranded sequence at the 3' end. Hybridization of random hexamers is therefore required *before* addition of the reverse transcriptase. We will add random hexamers (primers) and dNTPs (nucleotides) to the RNA samples and then heat them to promote **hybridization**. We will then add the SuperScript III reverse transcriptase and heat again to promote **elongation**. Superscript III achieves full activity at 50°C. RNase OUT is also added in this step to degrade RNases that might be activated. This is a one stage PCR without repeating cycles.

**NOTE:** The non-template control (NTC) produced in this cDNA synthesis will contain no cDNA because there will be no Superscript III. DO include RNA in this step of the cDNA synthesis. The NTC will be used in the real time PCR reaction as an indication of whether the RNA contained contaminating genomic DNA that can be amplified with your real time PCR primers.

#### METHOD OVERVIEW:

Prepare tubes with each sample containing *RNA plus dNTPs, Random hexamers, and water* and place on ice

Heat samples containing *RNA plus dNTPs, Random hexamers, and water* at 65°C for 5 minutes in the thermocycler

Transfer the sample tubes to ice

Add the Master Mix containing *first strand buffer, DTT, RNase OUT, and SuperScript III or Water* AFTER the samples have been heated at 65°C

Transfer the sample tubes to the thermocycler for one-step elongation

#### METHOD DETAILS:

1. Calculate volume (based on spectrophotometer reading) of 2,000ng (2µg) RNA.
2. Set up a chart: Samples containing RNA from heart will be used for cDNA synthesis using random hexamers. One tube will also be created for a non-template control (NTC) that will contain no SuperScript III. In the last column of the chart, calculate the amount of water that will be required to make the final volume 13µL.

Experiment	Volume of Sample Containing 500ng RNA (µL)	dNTPs	Hexamers	Volume of Water Required to Make a Final Volume of 13µL
Fasted		1µL	1µL	
1dpf		1µL	1µL	
3dpf		1µL	1µL	
10dpf		1µL	1µL	
NTC – this can include RNA from any sample without SSIII		1µL	1µL	

3. Calculate volumes for the Master Mixes:

**NOTE: DO NOT ADD MASTER MIX UNTIL AFTER THE RNA/Hex/dNTP/water SAMPLES HAVE BEEN HEATED FOR 5 MINUTES AT 65°C**

Two Master Mixes are required: 1 for Random Hex cDNA synthesis and 1 for NTC (does not contain SSIII)

MMI (cDNA):	5X FSB	4 $\mu$ L	
	0.1M DTT	1 $\mu$ L	
	RNase OUT	1 $\mu$ L	
	<b>SSIII</b>	<b>1<math>\mu</math>L</b>	
		-----	
			7 $\mu$ L Total X number of samples (4) + a couple of extra

MM2 (NTC):	5X FSB		4 $\mu$ L
	0.1M DTT	1 $\mu$ L	
	RNase OUT	1 $\mu$ L	
	<b>SSIII</b>	<b>0<math>\mu</math>L</b>	
	<b>H<sub>2</sub>O</b>	<b>1<math>\mu</math>L</b>	
		-----	
			7 $\mu$ L Total X number of samples (1) + an extra

4. Set up tubes:

Tubes 1-4:	Samples containing RNA, dNTPs, Random Hexamers, Water
Tube 5:	NTC containing dNTPs, Random Hexamers, Water

**PCR reaction:**

STEP 1	HEAT AT 65°C	5 Minutes
STEP 2	PLACE AT 4C (on ice)	
STEP 3	ADD THE MASTER MIX (7 $\mu$ L) to each tube	
STEP 4	ELONGATION at 55°C	60 Minutes
STEP 5	INACTIVATION at 70°C	15 Minutes

\*\*\*Assuming that the reaction is 100% efficient (the amount of RNA put in compared to the cDNA that is produced is 1:1), 20 $\mu$ L total reaction volume containing 2,000ng RNA will produce cDNA at 100ng/ $\mu$ L.

# The Python Project - Fall 2013

## Standard Curve Production using a Protein Assay

### Materials:

2mg/mL Stock Bovine Serum Albumin (BSA)  
 ddH<sub>2</sub>O  
 Microtiter Plate  
 P200 Pipetman  
 P200 pipette tips  
 1.5mL Tubes

### Method Details:

1. Prepare 5 Standard, as follows:

Prepare six 1.5mL tubes by labeling them: 1, 2, 3, 4, 5, 6

Dilute the stock BSA (2mg/mL) to 500µg/mL, a 1:4 dilution, into the tube labeled '1'.

VORTEX!

Prepare the first dilution (250µg/mL), a 1:2 dilution of the 500µg/mL (Standard #1), by transferring 500µL of '1' into tube '2' and adding 500µL ddH<sub>2</sub>O

VORTEX!

Prepare the second dilution (125µg/mL), a 1:2 dilution of Standard #2, by transferring 500µL of '2' into tube '3' and adding 500µL ddH<sub>2</sub>O

VORTEX!

Continue making 1:2 dilutions into tube #4 & #5, vortexing after each tube is prepared

Add only water into tube #6 (the no protein control)

*NOTE: These standards may be frozen for later use, if you need to repeat your standard curve.*

Tube	Stock BSA	1	2	3	4	5	6
Concentration	2mg/mL	500µg/mL	250µg/mL	125µg/mL	62.5µg/mL	31.25µg/mL	0µg/mL
BSA	---	250µL (Stock)	500µL (1)	500µL (2)	500µL (3)	500µL (4)	0µL
ddH <sub>2</sub> O	---	750µL	500µL	500µL	500µL	500µL	1000µL
Total Volume	---	1000µL	1000µL	1000µL	1000µL	1000µL	1000µL

2. Add 50µL of each Standard in triplicate to the microtiter plate.

3. Add 40µL Dye Reagent to each of the Standards and Water

4. Incubate at room temperature for at least 5 minutes but not more than one hour.

# The Python Project

## Spring 2014

### Real Time PCR Experimental Setup

#### Background:

You will be performing a real time PCR experiment to examine whether the expression of your gene of interest changing in the hearts of Burmese pythons at various times after feeding.

**NOTE:** Real time PCR is also referred to as quantitative PCR and is sometimes abbreviated RT-PCR. This is confusing given that reverse transcriptase PCR used to produce cDNA from your RNA is *also* abbreviated RT-PCR. The newest convention recommends using qPCR as an abbreviation for real time PCR. Accordingly, in this class, we will use **RT-PCR** for reverse transcriptase PCR and **qPCR** for real time PCR.

#### You will be creating the following for your qPCR plate:

1. Dilutions of a pooled sample of all cDNA samples for production of the standard curve

\*\*This is the most important part of setting up your experiment. Be especially careful when creating these dilutions. If your standard curve is not 'good', you will not be able to interpret the expression data produced by the real time PCR machine.

2. Dilutions of all cDNA samples from 100ng/ $\mu$ L to 1ng/ $\mu$ L

\*\* 2 $\mu$ g of RNA were used to produce your cDNA. The volume used was 20 $\mu$ L, so your concentration of RNA in the RT-PCR was 100ng/ $\mu$ L. Assuming that the RT-PCR was 100% efficient, the concentration of your cDNA is 100ng/ $\mu$ L as well.

3. One Master Mix containing the fluorescent mix and primers for each primer set used (gene of interest and reference gene).

4. Primer dilutions.

\*\*Oligonucleotides are most stable at higher concentrations. We reconstitute our lyophilized primers to 100 $\mu$ M when we received them so they can be stored safely. The working concentration of these primers is 12.5 $\mu$ M. A dilution from 100 $\mu$ M to 12.5 $\mu$ M is therefore required for the Master Mixes.

#### Supplies needed for qPCR set up:

1. Ice
2. PCR Strips

One strip for pooled cDNA sample and serial dilutions for standards  
One strip for working dilution of cDNA (1ng/ $\mu$ L)

3. Six 1.5mL Tubes

Two tubes for dilution of the forward and reverse primers for the gene of interest to 12.5 $\mu$ M

Two tubes for dilution of the forward and reverse primers for the reference gene to 12.5 $\mu$ M

One tube for the gene of interest Master Mix

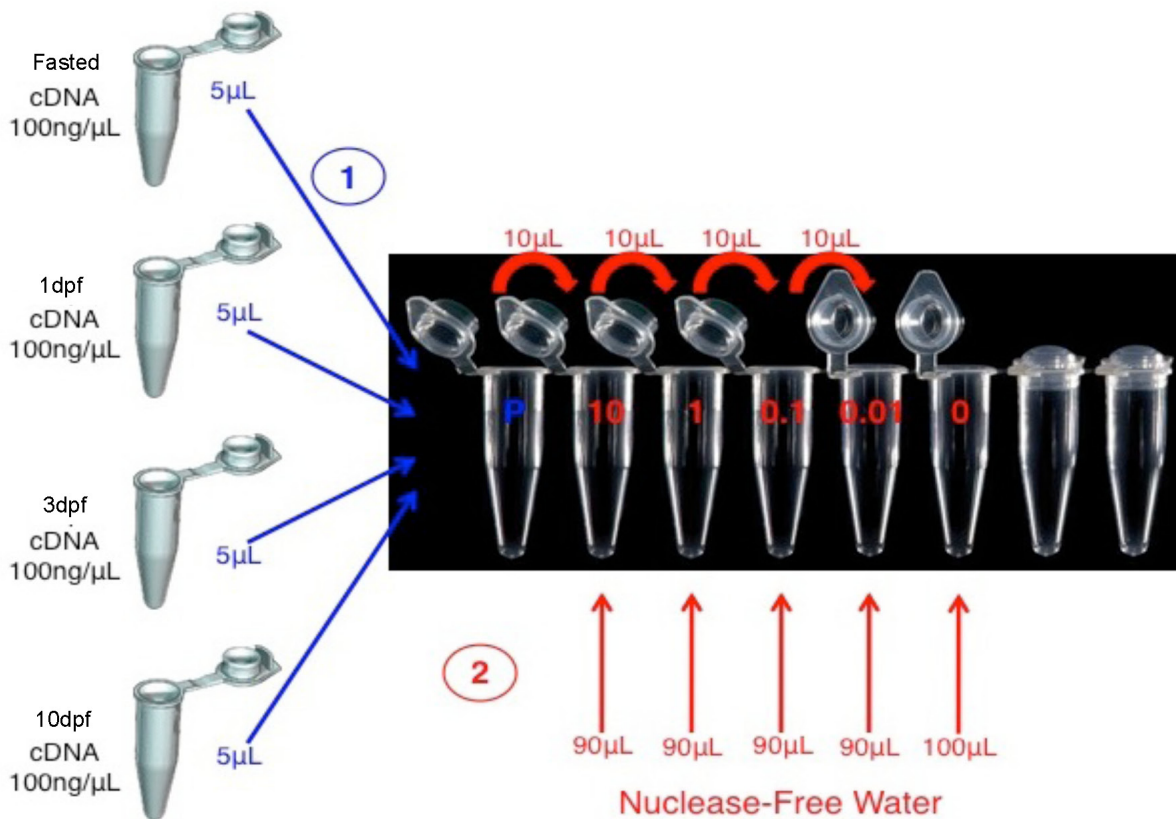
One tube for the reference gene Master Mix

4. One 96 well plate

5. 96 well plate holder

Getting started:

**1. Begin to dilute your pooled sample 1:10, creating a serial dilution.**



Prepare a fresh PCR strip by labeling the first 6 tubes: P, 1, 2, 3, 4, 5

Transfer 5 $\mu$ L of each cDNA sample into the tube labeled 'P' so a final volume of 20 $\mu$ L of cDNA is in the tube

**VORTEX!**

Prepare the first dilution (10ng/ $\mu$ L), a 1:10 dilution of the pooled cDNA, by transferring 10 $\mu$ L of 'P' into tube '1'

**VORTEX!**

Prepare the second dilution (1ng/μL), a 1:10 dilution of Standard 1, by transferring 10μL of '1' into tube '2'

### **VORTEX!**

Continue making 1:10 dilutions into tubes 3 and 4, vortexing after each tube is prepared

Add only water into tube 5 (a non-template control)

Tube	<b>P</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Concentration	<b>100ng/μL</b>	<b>10ng/μL</b>	<b>1ng/μL</b>	<b>0.1ng/μL</b>	<b>0.01ng/μL</b>	<b>0ng/μL</b>
cDNA	<b>5μL each</b>	<b>10μL (P)</b>	<b>10μL (1)</b>	<b>10μL (2)</b>	<b>10μL (3)</b>	<b>0μL</b>
Nuclease-Free H <sub>2</sub> O	<b>0</b>	<b>90μL</b>	<b>90μL</b>	<b>90μL</b>	<b>90μL</b>	<b>100μL</b>
Total Volume	<b>25μL</b>	<b>100μL</b>	<b>100μL</b>	<b>100μL</b>	<b>100μL</b>	<b>100μL</b>

**NOTE: Do not add your non-template control sample to the pool!**

### **2. Dilute 'Stock' cDNA to 'Working' cDNA**

Your stock cDNA is 100ng/μL. It needs to be diluted to 1ng/μL.

Label a new strip of PCR tubes

Dilute each Stock cDNA sample into the new tubes so the final volume is 100μL

### **3. Create the Master Mixes:**

The SYBR mix for a real time PCR is similar to a conventional PCR. It contains the following:

SYBR Green I Dye

AmpliTaq Fast DNA Polymerase

Uracil-DNA glycosylase (UDG) (enzyme from E. Coli, removes uracil from DNA)

dNTPs

Optimized buffer components

For one sample, the master mix for each gene contains:

#### **MM X 1 sample:**

6μL SYBR

0.6μL Forward Primer

0.6μL Reverse Primer

Label two fresh 1.5mL tubes with 'MM GOI' (master mix, gene of interest) and 'MM RG' (master mix reference gene).

You have five samples and one NTC that each need to be run in triplicate. You also have five standards that each need to be run in triplicate. Calculate the total number of samples for which you

will need master mix (number of samples X 3 + 2 extra). Then calculate the volumes of each component of the master mix that need to be added to tube 1.

Repeat for your reference gene and add the appropriate volumes to tube 2.

**VORTEX and spin down you master mixes!**

**Reference Gene MM X (number of wells + 2):      GOI MM X (number of wells + 2):**

μL    SYBR

μL    SYBR

μL    Ref Forward Primer

μL    GOI Forward Primer

μL    Ref Reverse Primer

μL    GOI Reverse Primer

- 4. Create a 'map' of your plate layout using the Real Time PCR Template.**
- 5. Label your plate and put the samples/master mix in the wells. (6μLMM + 4μL cDNA)**
- 6. Place the sticker over the 96 well plate, cover in foil, label, and store at 4°C.**
- 7. Store your cDNA (100ng/μL), Working cDNA (1ng/μL) and standard dilutions at -20°C.**
- 8. Discard extra master mix by flipping open the cap and placing in solid waste.**

# The Python Project

## Fall 2013

### Primer Test 2 – Melt Curve

#### **Purpose:**

The purpose of the test is to determine how many products the primers for real time PCR amplify.

#### **Method Overview:**

The primers will be added to 8ng cDNA synthesized from RNA isolated from hearts of pythons that are fasted or 6dpf. A master mix containing dNTPs, a fluorescent molecule that associates with double stranded DNA (SYBR), polymerase, and buffer containing magnesium will be added to the cDNA and primers. Two reactions will be set up in a 96 well plate. The 96 well plate will be shared by all students who are testing their primers.

#### **Method:**

1. Prepare the master mix that is specific to your primers.

#### **Master Mix**

SYBR Green	20 $\mu$ L
Forward Primer (12 $\mu$ M)	2 $\mu$ L
Reverse Primer (12 $\mu$ M)	2 $\mu$ L

NOTE: the master mix above will provide enough volume for two reactions and are specific to a single set of primers. If more than one set of primers is tested, separate master mixes must be prepared.

2. To each of two wells in the 96 well plate, add 6uL master mix.
3. To each of two wells in the 96 well plate, add 4uL cDNA (100ng/uL).
4. Label the plate with your sample names.
5. When the plate is completed, cover with clear film.
6. Wrap the plate in foil and label.
7. Store at 4°C.



Name: \_\_\_\_\_

**The Python Project  
MCDB 4202  
Worksheet I  
Spring 2014**

1. Describe how to make the following:

150mL of a 1X TAE solution from 25X TAE stock

5mL of 1X Ethidium Bromide from 50X stock

750mL of 1.5% agarose using powdered agarose and 1X TAE

1L of 2% agarose using powdered agarose and 50X TAE (remember that you will need to use 1X TAE in your final solution)

500mL of 10% solution of sodium dodecylsulfate (SDS) from 20% SDS stock (liquid)

**Name:** \_\_\_\_\_

125mL of 25mM anhydrous magnesium chloride (solid, powder), molecular weight = 95.211

125mL of 70% ethanol from 100% ethanol (liquid)

500mL of 3.7% formaldehyde in 1X phosphate buffered saline (PBS) using the following stocks:  
37% formaldehyde (liquid) and 50X PBS (liquid)

2. After diluting your DNA sample 1:40, you obtain an A<sub>260</sub> of 0.14. What is the concentration of DNA in your original sample?

3. After diluting your RNA sample 1:20, you obtain an A<sub>260</sub> of 0.2. What is the concentration of RNA in your original sample?



NAME \_\_\_\_\_

4. How would you make 5L of 1X TBS from a 25X stock?

5. How would you make 100mL of 25mM NaCl from 1M NaCl stock?

6. You dilute 5uL of your DNA sample in 100uL of water. The measured absorbance of the diluted DNA at 260nm is 0.11. What is the concentration of your DNA sample?

**Bonus:**

Who runs the sponsor laboratory for this course?

NAME: \_\_\_\_\_

**The Python Project**  
**MCDB 4202**  
**Worksheet III Spring 2014**

1. The python WGS contains (circle all that apply):

- a. Introns
- b. Exons
- c. Promoter regions
- d. Poly (A) tails
- e. Stop codons

2. The forward primer you designed for real time PCR can be found in a 5' to 3' transcript without having to create the reverse complement of the sequence. Explain why the reverse primer will be found as the reverse complement.

3. You assemble your python transcript using WGS contigs that aligned with the exons of the chicken transcript for your gene. When you translate your assembled sequence, you notice that a Stop codon occurs about halfway into your translation. Which of the following is the likely cause?

- a. A stop codon was introduced because a chicken transcript was used.
- b. The exon-exon boundary has duplicate sequence from aligning with two different contigs that overlap.
- c. The snake transcript is slightly different from the chicken transcript; the transcript in python is shorter.
- d. It's not really a stop codon. It looks like a stop codon but unlike the chicken, the codon encodes a useful amino acid in the snake.
- e. All of the above.

NAME: \_\_\_\_\_

4. You designed a primer that anneals within an intron to serve as both a positive and negative control.

What template will you use when the primer is used in a reaction that will serve as a *positive* control?

- a. Genomic DNA
- b. Ribosomal RNA
- c. Messenger RNA
- d. cDNA
- e. Any of the above will work

Explain the purpose of the *positive* control. What does it demonstrate?

What template will you use when the primer is used in a reaction that will serve as a *negative* control?

- a. Genomic DNA
- b. Ribosomal RNA
- c. Messenger RNA
- d. cDNA
- e. Any of the above will work

Explain the purpose of this *negative* control. What does it demonstrate?

5. Why are primers for real time PCR designed so the product spans an intron?

**NAME:** \_\_\_\_\_

6. If your real time PCR primers were used with cDNA, approximately what size product do you expect to amplify?

7. If your real time PCR primers were used with genomic DNA, approximately what size product do you expect to amplify?

8. When you perform a Primer BLAST on your real time PCR primers, you amplify the correct gene. However, there are several mismatches between your primers and the targets listed. Is this a problem? Why or why not?

9. Primer BLAST returns several 'hits' when you search for your real time PCR primers. The first two are the correct gene with an expected product length of 153 bases. The third hit is a different gene with a product length of 95 bases. Is this a problem? Explain why or why not.

**PART I: Self-evaluation.**

**1. Rate your ability to accomplish the following tasks:**

- 1 = I am **not confident** in my ability to complete this task.  
2 = I am **somewhat confident** in my ability to complete this task.  
3 = I am **very confident** in my ability to complete this task.

- \_\_\_\_\_ Read and understand scientific literature---
- \_\_\_\_\_ Interpret data from techniques and draw logical conclusions
- \_\_\_\_\_ Troubleshoot problems during scientific experiments
- \_\_\_\_\_ Use resources (books, internet, databases, colleagues/peers) to gather information about new scientific techniques and concepts
- \_\_\_\_\_ Work effectively as a member of a group
- \_\_\_\_\_ Presenting data orally to a group of peers
- \_\_\_\_\_ Presenting data in written form

**2. Rate your understanding of the following topics:**

- 1= I have **never heard** of this topic.  
2= I have heard of this topic, but I am **unsure what it is** or how it relates to the python lab.  
3= I understand this topic, but I am **unsure how it relates** to the python lab.  
4= I **understand this topic** and how it applies to the python lab.  
5= I understand this topic, how it applies to the python lab, and I **am able to apply** knowledge of this topic to other contexts.

- \_\_\_\_\_ DNA structure (nucleotide bases and base pairing, bonds, double helix structure)
- \_\_\_\_\_ DNA directionality (5' to 3')
- \_\_\_\_\_ The steps in the polymerase chain reaction (PCR)
- \_\_\_\_\_ Sequence alignment programs such as BLAST
- \_\_\_\_\_ Gel electrophoresis
- \_\_\_\_\_ DNA primers
- \_\_\_\_\_ Real time PCR

**3. Choose the statement that best reflects your confidence in using online Bioinformatics tools (e.g., BLAST, NCBI database searches).**

- \_\_\_\_\_ I have only used these in a class.
- \_\_\_\_\_ I do not have a lot of experience with Bioinformatics tools, but can follow directions to learn a new tool.
- \_\_\_\_\_ I understand Bioinformatics tools well and am comfortable learning new tools.

**4. Choose the statement that best reflects your understanding of experimental design.**

- \_\_\_\_\_ I do not know the components of experimental design or how they are used in creating an experiment.
- \_\_\_\_\_ I understand the components of experimental design, but not how to use them to create an experiment.
- \_\_\_\_\_ I understand the components of experimental design and I am beginning to understand how to use them to create an experiment.
- \_\_\_\_\_ I understand the components of experimental design, their importance in experimental design, and know how to use them in creating an experiment.

**5. Choose the statement that best reflects your experience troubleshooting unexpected data from an experiment.**

- \_\_\_\_\_ I have no experience in recognizing unexpected data or suggesting ways of fixing the problems (troubleshooting).
- \_\_\_\_\_ I am able to recognize unexpected data, but I am not able to troubleshoot the problem.
- \_\_\_\_\_ I am able to recognize unexpected data and I am able to troubleshoot the problem.



**PART II: Basic molecular biology assessment.**

**6. What is wrong with this statement?**

*When using gel electrophoresis to separate DNA, RNA, or proteins by size, the largest molecules migrate faster in the agarose gel in response to the electric charge differential in the buffer.*

**7. Please answer *True or False* to the following questions and *explain* your answer.**

- a. Positive controls, when available, are essential to scientific experiments.
- b. Negative controls are not necessary because they will not show any results.

**8. What type of bonds holds the DNA double helix together?**

- a. Covalent bonds
- b. Hydrogen bonds
- c. Ionic bonds
- d. Bail bonds

**9. Below is a short sequence of DNA written in the 5' to 3' direction. Please write the sequence of the reverse complement of this sequence and indicate the directionality of the complementary strand.**

5' ATCGTAACGTCGT 3'

**10. The ability of primers to anneal to DNA can be dependent on many factors. Please choose all of the factors that would affect annealing.**

- a. Length of fragment
- b. Temperature
- c. G/C content
- d. Secondary structure
- e. Complementarity

**11. What is unique about the *Taq* polymerase used in PCR?**

**12. How many grams of magnesium chloride ( $\text{MgCl}_2$ ) are required to make a 1L solution of 100mM  $\text{MgCl}_2$ ? The molecular weight of  $\text{MgCl}_2$  is 203.3 g/mol.**

**13. How do you make a 1X concentration from a 10X stock solution?**

**14. How do you measure the concentration of a DNA sample?**

**15. Briefly explain what is means when gene expression changes in a cell. You may find it useful to describe this in terms of the Central Dogma of biology.**



The Python Project - University of Colorado at Boulder

Dear Python Project Alumni,

As you may already know, we are extremely interested in your opinion about your participation in The Python Project and in how the course impacted your education and career beyond your undergraduate education. You are invited to complete a survey about your experience. A link is provided below. The survey has only 19 questions and will take approximately 10 minutes to complete. Participation is completely voluntary, and answers are anonymously reported to us by Survey Monkey. Your answers will not in any way affect your current or future relationship with the University of Colorado. The results of this survey will be incorporated into a publication on the course. It is our hope that the publication will be beneficial in the development of similar courses and in obtaining funding to continue The Python Project in the future. If you choose to participate, please read the questions carefully and complete the survey within two weeks. Any questions can be directly to: pameia.harvey@colorado.edu.

Thank you!  
The Instructors of the Python Project

Completing this questionnaire online means you have read and understand the voluntary status of your participation in this survey. We thank you for your valuable participation.

**1. I had an overall positive experience in The Python Project laboratory course.**

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

**2. Compared to other laboratory courses I was enrolled in at University of Colorado, I prefer the structure of lab courses similar to The Python Project.**

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

**3. The Python Project taught me less about laboratory research compared to other laboratory courses I was enrolled in at University of Colorado.**

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

**4. The Python Project prepared me to engage in an independent research project.**

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

**5. My decision to engage in independent research was positively influenced by my participation in The Python Project.**

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

8. My participation in The Python Project increased my understanding of performing novel research.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

9. I liked that the data generated by the class contributed to an ongoing research project in a lab at University of Colorado.

- Strongly Disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

10. After taking The Python Project course, I worked in a research laboratory during my undergraduate education. (paid, internship, volunteer, or independent/honors credit)

- Yes
- No

11. Have you completed your undergraduate education?

- Yes
- No

12. If you answered 'No' for Question #11, are you currently enrolled in classes at University of Colorado?

- Yes
- No

13. Are you currently enrolled in postgraduate education?

- Yes
- No

14. Did you complete your postgraduate education?

- Yes
- No

15. If you did not engage in postgraduate education, what type of job do you currently hold?

16. If you did engage in postgraduate education, what type of postgraduate education are/were you enrolled in?

- Dental School
- Graduate School (Master's) in Biological Sciences
- Graduate School (PhD) in Biological Sciences
- Medical School
- Pharmacy School
- Other

17. If you have completed your postgraduate education, what type of job do you currently hold?

18. Gender

- Male
- Female
- Prefer not to identify

19. In what year were you enrolled in The Python Project?

- 2009
- 2010
- 2011
- 2012
- 2013

A.

Primer Design: OPA1

Parts 1&2

mRNA sequence obtained from PubMed on January 22, 2014. It has not been validated, but has the status of PROVISIONAL. The gene type is listed as "protein coding." It is suspected to be a mitochondrial precursor in the Dynammin family.

Gallus gallus optic atrophy 1 (autosomal dominant) (OPA1), mRNA

NCBI Reference Sequence: NM\_001039309.1

GenBank Graphics

>gi|86129525|ref|NM\_001039309.1| Gallus gallus optic atrophy 1 (autosomal dominant) (OPA1), mRNA

AGCGGCGCGCCGGGCGCCGGCGGGCGGTTCCCTCCCTGGCCCGGCTCCGGGGCCGAGGC
CGGGGGAGGAGCTTTGTGCGGGCTTTCCCTCCCTCCCTCCCGGCACTCCCTGCGCTCAGGAATG
GGCGGCGAAGCGGGCGGGCTTTGTGTCATCTGCAGGAGTCTAGCACAATAGCAACTAGGAATAAAAA
GAAATCACCACTTCAAAACTTGCACCTGGTTCCCGAAACAATCATCATCCCTACCAAGTCTGAAA
TTTCAAGACGTCCTAAGGATATCCCTCAACAAGTCTCTCTCTTAACTCGTCTCCCTACGCAAAA
CAAAACTTTAAATGTAAAAATATGGTTACCACTCAGTCCACAGGAACCTTTGGCTAGCTAGACTAGCATCAAG
GCTTCTCAAAAATCTGCATCTTATATGGGGTCTGCTGTAGGTGGTGTACACAGCTAAGAAAGACATAT
GATCATGTGGGAGGACATGATGCCTGATCTCGATGAATAAAGTGGATTTCTCTGACTTCACTGGGAGC
TTGATGAGCATATTGACTTGAAGAAAATTTAACAAGCCCTCTCCGATGCGGATGACCTTGCACAACTTCT
GCTGACTTCGAGAAGATTGGAGAGGCTTCACTCAGTAAAGGGATTTTCTCTGGTTACAAATTTG
GTTAGTGAAGTATAGGAGCTTCTGATCTACTTCTGTTGTTAGGTACTCCAGGAGAAACAGCATTCAGAG
CTACGGACCAAGGTTATGACAGCGACAAACAGTACAAAAGGTTTCTGACAAAAGAAAGATTGATCAACT
TCAGGAGAACTTCTGCCACTCAGTAAAATACCACAGGATGCTTGAGCGATTAGAGAAGGAAACAAA
GAATTAAGAAAATGGTACTGCAAAAGAGTGA CAAAGAAATTCACAGAGAAATTAAGAAATCTTTGA
TTGACATGTATTCTGAAATGATGATATCTGCTGATTTATGATGCCAGCTATAACACTCAAGAACAACCT
ACCTCAGTGTGATGCTGTTGAGATCAAAATGCGAGGAAAACAAGTGTATTAGAGATGATGTTCTCAAGC
CGAATATCCCTAGAGGGTCTGGAGAGATGATGACACGTTCCCTGTAAAAGTAAACCCCTTAGTGAAGTCC
CTCACACGTGGCTTTTCAAGAGACAGCTCTCGGGAGTTTGTATGACCAAGAAAGAGGATCTTGCAGC
TTTGAAGAAATGAGATAGAAAATCAAAATGAGGAATAAGTGAAGGAGGGCTGCACTGTTAGCACA GAGACC
ATCTCTTAAAGTGTAGAGGTTCTGGTTTACAGAGAATGGTATTTGTTGATTTACTCGGAGTCAATGATA
CTGTGACATCAGGATGCTCCAGATACAAAAGGAAACAATCTTTAGCATCAGCAAGGCTCATGACAGAA
TCAAAATGCCATCACTCTTTGATTTCAAGATGGATCAGTGGATGCAAGAACGCAATGTTGTCACAGACTTA
GTCAGCCAAATGGATCCCGAGGAAAAGGACAAATTTTGTGCTGACTAAAGTTGATCTTGCATGAGAAA
ATGGGCTAGCCCAAGCAGGATCAAAATAAATGAAAGGCAAACTCTTCCCAATGAAAGCTTTGGGTTA
CTTTGCAAGTTGTTA CTGAAAAGGAAAACGACAGTGAAGCAATGAATCTTAAAGAGTATGAAAGGAA
TTTTTTCAAAAATTCCAAAGCTGTGAAAACATCCATGCTAAAGGCAACCAAGTAAACAAGAACTTAA
GTCCTTGTCTCAATGATGCTTTTGGAAAATGGTGAAGAGTCTGTGAGCAGCAAGCAGATGCTTTTAA
AGCCACACGTTCAATCTTGAGACAGAAATGGAAGAACAAATTAACCCCGGTGCGAGAGCTTGCAGGAAT
GAATCTTTTGAAGAACAAAAGAAATGATTTCTTGAAGTCAAGTCTGACTCAAGTCAACCTCAACCTAAGC
ACTGGGAGGAGATTTCTCAGAAGACATTAAGGAGAGGGTACTACTCACGTGATGAGAAATATCTACT
TCCAGCAGCAGACAAATGAACCTCAGGACATTTAAACCACTGTTGACATCAAACTGAAGCAGTGGACT
GACAAAGAACTGCCAAA TAAAGCAGTAGAGGTGGCGTGGAGACTTTGCAAGAGGAGTTTTTCTGTTTCA
TGACAGAA CAAAAGGAAAAGAGCATGATGATATCTTTGATAACTGAAA CAAGCTGTCAAAGAGAAAG
TATTAAGCGCCATAAATGGAACGAGAGAGCAGAGGATA GTCGCTGTGATCCAGCAAACTGTTAGAA
GATCGGTCATAATGTA TAAACAGCAATGGATGCACTATTCAATTTTATGGAAGAGACTCTCCAAAAGT
GTCTCAAAGACACTGAA TCTGTTATTGAA GATATGGTGGTCCGGATTTGAAAAGAGGTTGGCTATACTG
GATAAGCCGCA CCAAAGAGCAGAAATATACGTAATGAAA CAAAATGAAC TTAGAAGTAAATCAAAATGC
AATGAAGAACATGCGACATACTGGCAAA TGAATGAGTGCACACTGTGAGAAA GAATCTGAAGCGAGAG
GAATAACAGTGGACCCATGCTGATAAAA GACAGTGGCCACAAATTTATAGAAGGATTTCTGAAAGC
TGCTTTGAACCACTGCAACCTTTGTCGAA GAGGTTCTATTAATCAACGACATTTCTGTA GACTCGGAG
CTTGAATGTAATGATATTGTCCTCTCTGCGGATACAGCGAATGCTGGCAATCAGCAGAAATACACTGA
GGCAGCAGCTTACTAATACAGAAATAGACGCTTGGAGAAAGTGAAGGAAAGTCTTGAAGATTTTGG
TGAGGCAATGAAAAGAGGTTAAGCTCC TAACTGGCAAAAGGGTCCAGCTTGCAGAGGATCTCAAAA
GTTCGAGAAATCCAGGAAAACCTTGAAGCTTCTATAGAAGCCCTCAACAAAGAAAATAGATTGTC AAGC
AGCAGCAACTTAAACACAGGAATGCGAGACTTGTGAGGATACCGTCCCAAAGAAAGCTGAA CAAAATGTC
TTCTGTGCTCCATCTGTGCTTACCTTTTGGGGAGCATTGTGAAATGTTGGAGAAAAGCAAAA TGA
TGCTCTCTGGGATTGAGAAAATCAATCAAAA CCGGCTGTTTTTTTAAATCTTGAAGCTACGTTGATATG
GATGAGATAGCTGGAAAGCTTGCAGAGTCTGCAATTTAAGAGTCTTTTTTCTGATGAGGAAA AAAATG
CTTTTTGTTGTTGATCTGCTCTGTA CAGTGGAGCTGGATCAGATCATAAAGGGATTTTACTCTGTGTC
AGTTTTCTGGCTGCAAGTGTCTC TAGAAGTTGACATAAAGTTTTTGTATTTGTTGTTGATCA TAAAAC
CCTTCGAAATGCAATGCCACTTCTCAGTATGCTCTGTTATTGAAA CAGTCTCTGATCACTCTGTTACAGAT
TTTAAAGCATGTTGATGATATTGAGGGGAAA TTGTTGAAATTTGTTAAACTCAAAA CTCTTGAAGAGCGT
CATTCAGGAAATTCAGTGTCTTCAAAGTGAAGTGCAGACTGCGCTTCACTGACACTTAAAGGTTTGC
CTCAGGAAATGATCAACAGGGGGCTTTTTTTTTACAGTCTTCTTAATACATGTTGTTCTTTGGAA
TCGTGGCCGATTACATTTCA TTTCTGCCATGACAGACTGATTTCTTTTAAATCTTCCCTTCCAAAT
AGGCA CAAAGTAACTGATGTTGTAATGTAAGAAAGTAGCCGTCACAA TTTCCAACTTCTCCAACTTT
TGATGTAAAAGAAA GCAAAAATAGAAATTA CTTTCACTGTTTCTAATAGCTACCGTAAGC CAAAATTA
TGATAGCAATGCTATGCAAAAGTCTATCATAACCATGTGATATCTGTAATGCAATACAGTAAAATAA
AACTAAGTATAATTTCTTCTCTGCA TTTTAAAGTTTGTCTAGGAAATGCTACTTCTGACAAATGTA
GTATAATA TAGGCTCTGATCATTAGGTGCTGGTGA TTAATCA CCACTCTATTTTCTGCCACACTTCC
GCAACAGATGATCTGGA TACACTTCTGAA TACTCTGTGGTATTAGATAAAGAA GATATTAGGCTTTAC
ACGATTTGTTTGAAGTTGCTCTATTATGAAAGTTGAAAATTTGATGATGTTTATTTCTTTTTGGAAA
AGGCTATGATTTGTCGACTCTGGCAAGTAAAGAAAGAGGAGCTGACTCTTTTATAACAGTACTGTGTC
TGTTTTGCTCCCAAGTGAACCAAATAGAAA TTTTGGTAAAGATGTTCCCAAGATA TTAGATAATTAAGTAC
TTAGGCTATGGATGAGAACTTAAATAGAGCTTATCA TAAAAGTTGAGAGGTTAAACATGCGAGGTG
ATCTACTGCTTGGGGA TTGTTTTGTTG TTGTTAAAAC TGTCTTTCTGAAA GTGAAAAGAAAGACTTC
TATCACTACGCTCTGACAA TAGCCTTAGAGATGATATTTGACAGTCTGTGAAAAGAGATTCCTGAT

A. (continued)

GCTGTACGCAAGTAAATAAGTCTGTACTTGTGTAAATTCCTCTTTGTATAAAAAAAAAATACACCAA  
ACTTTTGCATTACGTTAAGAGATCACTCAGAAAAATCTGTGTAAATGTGCAAAATATTACATC  
CAGTGGTCCAGTGTTCATCTCCAGTGTGTTCCTGTAGTCTAGTTCCTACCTCTGTGTCTCATC  
TAAGACTTAGCATGATGATTCATGGAACTAGGTATTTAATGCATAAAATGTAACTGAAGGAAAAAT  
AAAAATGTGGATTTTGAAACCGGAAA

Color-coded mRNA sequence showing exon-exon boundaries. This sequence is found in *Gallus gallus* and was obtained from Ensembl Genome Browser on January 22, 2014. This sequence has 29 exons.

Transcript: OPA1-202 ENSGALIT00000011583  
Description: dynamin-like 120 kDa protein, mitochondrial precursor [Source:RefSeq peptide;Acc:NP\_001034398]  
Location: Chromosome 9: 12,507,512-12,555,260 reverse strand.  
Gene: This transcript is a product of gene ENSGALG00000007150 This gene has 2 transcripts (splice variants)

cDNA sequence  
ATGTGGCGGACGAAAGCGCGCGGGCTTGTGTCTCTGAGGAGTCTAGCACATAGCAAC  
TATGGAATAAAAAAGAAATCACCATTCAAAACCTTGCACTGGTTCCTCCGAAAGCATACAT  
CATCCCTACCATCCAAAGTCTGAAATTCAAAGAGCTCCATTAAGGATATCCTTCAACAG  
TTCTCTTCTTAA TCGTCTTCCCTACGCAAAAACAAACTTTTAAATGTAATAATGGT  
TACCACTCTACAGGAACTTTGGCTAGCTAGACTAGCATCAAGGCTTCTCAAAATTCGC  
TATCTTATATGGGCTCTGCTGAGTGGTGGTTACACAGCTAAGAAACATATGATCAG  
TGAAAGGACATGATGCCCTGATCTCGATGAAATAAAGTGATATTTCTGACTTCATCTGG  
GAGCTTGAATGACATATGACTTAGAAAACCTTATCAAGCCCTTCTGATGCCGATGAC  
CTTGCCAAACCTTCTGCTGACTTCGAGAAATGGAGAGAGCTTCACTTCACTGAAAGGG  
ATTTTTCCTCTGGTTACAAATTTGGTGTAGTGAAGTCAATAGGAGCTTCTGATCTACTCTG  
TTGTTAGGTACTCCAGGAGAAACAGCAATTCAGAGCTACGGACCAAGGTTATGACAGCGAC  
AAACATCAAAAAAGTTTCTGACAAAGAAAGAAATGATCAACTTCAAGGAACTTCTG  
CGCACTCAGCTAAAAATCCAAACGGATGCTTGAAGGATGAGAAAGAAAGAAAGAAATTA  
AGAAAATGGTACTGCAAGAGATGACAAAGAAATTCATCAAGAAAGTTAAAGAAATCT  
TTGATGACATGATTTCTGAGTACTTGAATCTCTGATCTGATGATGATGCCAGCTATAAC  
ACTCAAGATCACTACCTCGAGTGGTACTGCTTGGAGATCAAAGTGCAGGAAAAACAAGT  
GTATTAGAGATGATGCTCAAGCCGAATATTTCCCTAGAGGGTCTGGAGAGATGATGAC  
CGTTCCTCTGTAAGGTAACCCCTTAGTGAAGGTCCTACCAAGTGGCTTTATCAAGAGC  
AGCTCTCGGAGTTGATCTGACCAAAGAGAGGATCTTGCAGCTTTGAGGAAATGAGATA  
GAAATCAGAATGAGGAA TAGCGTGAAGGAGGGCTGCACTGTTAGCACGGAGACTATCTCC  
TTAAGTGTGAAAGGTCCTGGTTTACAGAGAAATGGTATTTGTTGATTTTCACTGGAGTCAAT  
AGTACTGTGACATCAGGTATGGCTCCAGATACAAAGGAAACAACTTTAGCATCAGCAAG  
GCCTACATGCAAA TCCAAATGCCATCATCTCTTGTATCAAGATGGATCAGTGGATGCA  
GAACCGAGTATTTCTCACAGACTTAGTCAGCCAAATGGATCCCAAGGAAAAAGGACAAAT  
TTTGTGCTGACTAAAGTGTACTTGTGTAAGAAAAATGTGGCTAGCCCAAGCAGGATCCAA  
CAAAATATGAAAGCAACTCTTCCCAATGAAAGCTTTGGGTTACTTTGCAGTTGTTACT  
GGAAAAGGAAA CAGCAGTGAAGCATGATCTATTAAGAGATGAGAAAGAAATTTTT  
CAAAAATCCAAAGCTTTGAAACATCCATGCTAAAGGCACCAAGTAACAAACAAAGAAC  
TTAAGTCTTGTGCTCAGATGCTTTTGGAAAATGGTGAAGAGTCTGTGGACAGCAA  
CGAGATGCTTTTAAAGCCACACGTTTCAATCTTGAACAGAAATGGAA GAA CAATTA TCC  
CGTTCGGAGAGCTTGACAGGAA TGAATCTTTGAAAAGCAAAAGATGAGATCTTGTGAT  
GAAGTCAATAGTCTGACTCAAGTCAACCTTAA GCACTGCGAGGAGATTTCTCAGAAAGC  
TTATGGAGAGGGTATCTACTCACTGATGAGAAATCTACTCTCCAGCAGCACAGACA  
ATGAACTCAGGGACATTTAAACACACTGTGGACATCAAACTGAAGCAGTGGACTGACAAAG  
CAACTGCCAAAATAAGCAGTAGAGTGGCTGGAGACTTTGAAAGGAGTTTCTCTCG  
TTCATGACAGACAAAAGGAAAAGAGCATGATGATCTTTGATTAACCTGAAACAGCT  
TTCAAAAGAAAGAAATTTAAGCGCATAAATGGAACAGAGAGGAGGAAATCTCTCGT  
GTGATCCAGCAAA TCCCTTAGAAGATCGGCTAATACTGATAAAACAGCAATGGGATGCA  
GCTATTCATTTTATGGAAGAGACTTCCAAAGTCTCTCAAAGACCTGAATCTGTTATT  
GAAGA TATGGTGGGTCCTGGATTGAAAAAGAGGTGGCTATACTGGATAAGCCGCAACAAA  
GAGCA GAA TATACGTAATGAAACAAAAATGAACTTGAAGAA GTTAATCAATGCAATGAA  
GAACA TGCAGCATACTTGGCAAA TGA TGAAGTGAACA CTG CAGAA GAA TCTTGAAGCG  
AGAGGAA TAA CAGTGGACCCATGCTGATAAAAGATACGTGGCA CCAAATTTATAGAAAG  
TATTTCTGAA GACTGCTTTGAA CCACTGCAACCTTTGTCGAAAGGTTTCTA TTA CTAT  
CAACGACATTTCTGATAGCTCGGAGCTTGAATGTAATGATATTTCTCTCTCTGCGGATA  
CAGGAAATGCTGGCAATCAGCAAAA TACACTGAGGCAAGCACTTACTAATACAGAAAT  
AGACGCTTGGAGAA GAA TGTAAA GGAAGTCTTGAAGATTTTGTGAGGCAATGAAAAAG  
AAGGTTAAGTCTTAACTGGCAAAAGGTTCCAGCTTGCAGAGGATCTCAAAAAAGTTCCA  
GAAATCCAGGAAAAACTTGAAGCTTCTATGAAAGCCCTCCATCAAGAAAAATAG

Ensembl release 74 - December 2013 © WBI / EBI

Primer Design Parts 3 & 4

I took a portion (428 bp) of the cDNA sequence from the whole OPA1 sequence found in *Gallus gallus*. The portion (3 exons) has been copied below and highlighted above in bright green. Using BLAST to map the sequence I highlight the ranges of conservation below. This was completed on January 27, 2014.

GGAGGAGATTCCTCAGAAGACATTTATGGCAGAGGGTATCTACTCACGTGATGAGAAATATCTACCTTCAGGAC



A. (continued)

CACAGACAA TGAAC T CAGGGACATTTAACACCACTGTGGACATCAAAC TGRAGCAGTGGACTGACRAGCAACTG  
 CCAAATAAAGCAGTAGAGGTGGCGTGGGAGACTTTGCAAGAGGAGTTTCTCGTTTCATGACAGAACAAAAGG  
 AAAAGAGCATGATGATATCTTTGATAAACTGAAACAAGCTGTCAAAGAAGAAAGTATTAAGCGCCATAAATGGA  
 ACGAGAGAGCAGAGGATAGTCTGCGTGTGATCCAGCACAAATGCCCTAGAAAGATCGGTCAAATATCTGATAAACAG  
 CAATGGGATGCA  
 GCTATTTCATTTATGGAAGAGACTCTCCAAAGTCGTCTCAAAGACA

Python bivittatus Python\_molurus\_bivittatus-5.0.2-2085.10, whole genome shotgun sequence  
 Sequence ID: gb|AEQU02110921.1|Length: 15457Number of Matches: 3  
 Related Information  
 Range 1: 6753 to 6921GenBankGraphics **Range 1 is highlighted in grey.**  
 Next Match  
 Previous Match

Alignment statistics for match #1				
Score	Expect	Identities	Gaps	Strand
174 bits (192)	2e-41	140/169(83%)	0/169(0%)	Plus/Plus
Query 1	GGAGGAGATTCTTCAGAGACATTTATGGGAGGGTATCTACTCACGTGATTGAGATAT	60		
Sbjct 6753	GGAGGAAATTCTTCAGAAGTCTTTATGGGACAGAGTTTCGACACACCGTATAGAAAAAT	6812		
Query 61	CTACCTTCCAGCAGCAGACAATGAACTCAGGGACATTTAACACCCTGTGGACATCAA	120		
Sbjct 6813	CTACCTTCCAGCGGCGCAGTCTCTGAAATGCTGGGACTTTCAACACCACCGTAGACATCAG	6872		
Query 121	ACTGAAGCAGTGGACTGACAAGCAAAGTCCCAAATAAAGCAGTAGAGGTTG	169		
Sbjct 6873	GCTGAAGCAGTGGACTGACAAGCAGTCCCAACAAAGCTGAGAGGTTG	6921		

Range 2: 7982 to 8136GenBankGraphics **Range 2 is highlighted in turquoise.**  
 Next Match  
 Previous Match  
 First Match

Alignment statistics for match #2				
Score	Expect	Identities	Gaps	Strand
127 bits (140)	2e-27	121/155(78%)	0/155(0%)	Plus/Plus
Query 165	AGGTGGCGTGGGAGACTTTGCAAGAGGAGTTTTCGTTTCATGACAGAACAAAAGGAA	224		
Sbjct 7982	AGGTGGCCTGGGAGACCCCTCCAGGAGGATTTTCGCTTCATGACTGAGCAGAGGGCA	8041		
Query 225	AAGAGCATGATGATATCTTTGATAAACTGAAACAAGCTGTCAAAGAAGAAAGTATTAAGC	284		
Sbjct 8042	AGGAGCATGACGACATCTTCGACAACTGAAAGCAGCCGTAAGAGGAGGAGCATCAAGC	8101		
Query 285	GCCATAAATGGAAACGAGAGAGCAGAGGATAGTCTG	319		
Sbjct 8102	GGCACAAGTGGAAACGAGCGGGCGGAAGACAGCCTG	8136		

Overlapping portions of range 1 and 2 are highlighted in grey.

Primer Design Parts 5-10

I analyzed the aligned python contigs that were mapped to the *Gallus gallus* OPA1 transcript and then identified the corresponding python sequence based on the alignments given by BLAST (subject line). Overlapping regions are highlighted separately in grey. The following work was completed on January 29, 2014.

Mapped Region / Exon1-2

GAAGGAAATCTCAGAAAGCTTATATGGACAGAGTTTCAACACGTTATAGAAACCACTGACTTCCAGGGGCGGAGTTCGAAATCCGAGAACTTTAAACAGCAAGG  
 ACATCAG GCTGAAGCAGTGGACTGACAAGCAGCTGCCCAACAAGCTGTAAGGTG

Mapped Region/ Exon1-2

AGGTGGCGTGGGAGACCCTCCAGAGGAGTTTTCGTTTCATGACTGAGCAGAAAGGCAAGGAGCATGACGACATCTTCGACAACTGAAAGCAGCCGTAAGAGGAGAGAG  
 CATCAAGCGGCACAAGTGGAAACGAGCGGGCGGAAGACAGCCTG

Both PYTHON regions put together with the overlapping sequence only shown once:

## A. (continued)

```
GGAGGAAATTCCTTCAGAAAGTCTTTATGGGACAGAGTTTCGACACACGTGATAGAAAAATCTACCTTCCAGCGG
CGCAGTCTCTGAATGCTGGGACTTTCAAACACCAACCGTAGACATCAGGCTGAAGCAGTGGACTGACAAAGCAGCTG
CCAAACAAAGCTGTAGAGGTGGCCTGGGAGACCCCTCCAGGAGGAGTTTCTCGCTTCATGACTGAGCAGAAGGG
CAAGGAGCATGACGACATCTTCGACAAACTGAAGCAGGCCGTAAAGGAGGAGAGCATCAAGCGGCACAAGTGGA
ACGAGCGGGCGGAAGACAGCCTG
```

A comparison to the original exons mapped in GALLUS gallus:

```
GGAGGAGATTCCTTCAGAAAGACATTTATGGGAGAGGGTATCTACTCACGTGATGAGAAATCTACCTTCCAGCAG
CACAGACAATGAACTCAGGGACATTTAAACACCACCTGCGACATCAAACTGAAGCAGTGGACTGACAAAGCAACTG
CCAAATAAAGCAGTAGAGGTGGCCTGGGAGACTTTGCAAGAGGAGTTTCTCGCTTCATGACAGAACAAAAAGG
AAAAGAGCATGATGATATCTTTGATAAACTGAACAAGCTGTCAAAGAAGAAAGTATTAAGCGCCATAAATGGA
ACGAGAGAGCAGAGGATAGTCTG
```

Using ExPASy, I translated the python transcript. This validates the python assembly:

### Translate Tool - Results of translation

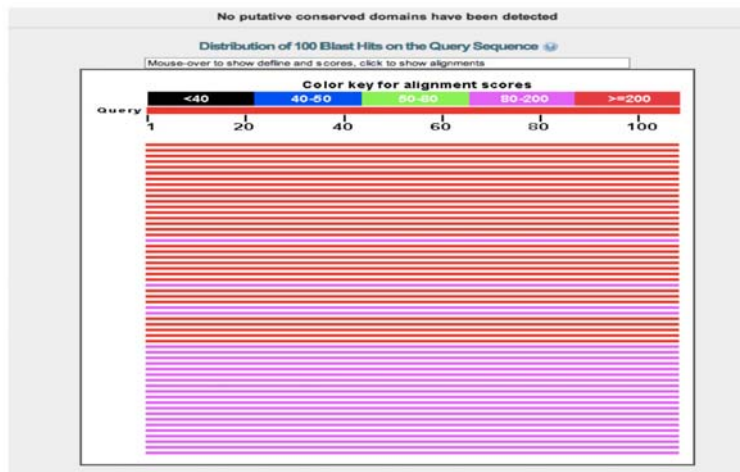
```
5'3' Frame 1
G G N S S E V F Met G Q S F D T R D R K H L F S S G A V S E C W D F Q H H R R H Q A E A V D Stop Q A A A Q Q S C
R G G L G D P P G G V F S L H D Stop A E G Q Q A Stop R H L R Q T E A G R K G G E H Q A A Q V E R A G G R Q P

5'3' Frame 2
E E I L Q K S L W D R V S T H V I E N I Y L P A A Q S L N A G T F N T T V D I R L K Q W T D K Q L P N K A V E V
A W E T L Q E E F S R F Met T E Q K G K E H D D I F D K L K Q A V K E E S I K R H K W N E R A E D S L

5'3' Frame 3
R K F F R S L Y G T E F R H T Stop Stop K T S T F Q R R S L Stop Met L G L S T P F Stop T S G Stop S S G L T S S
C P T K L Stop R W P G R P S R R S F L A S Stop L S R R A R S Met T T S S T N Stop S R P Stop R R R A S S G T S
G T S G R K T A
```

Frames 1 & 3 contained STOP codons, which makes them incorrect. However, Frame 2 contained no introduced STOP codons, so this is correct!

I continued to validate the amino acid sequence by using frame 2 and BLASTp:







## C. (continued)

The first primer shown above yielded the following results:

NCBI/Primer-BLAST : results: Job id=RI2tdd1\_dFtLaU9ILkt9GzVmTwomeVIP [more...](#)

Input PCR template none  
Specificity of primers Target templates were found in selected database: Refseq mRNA  
Other reports [Search Summary](#)

▼ Detailed primer reports

**Primer pair 1**

	Sequence (5'->3')	Length	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	AAGCAGTGGACTGACAAGCA	20	59.82	50.00	6.00	0.00
Reverse primer	TTTACGGCCTGCTTCAGTTT	20	58.02	45.00	4.00	0.00

**Products on target templates**

>[XM\\_005419092.1](#) PREDICTED: *Geospiza fortis* optic atrophy 1 (autosomal dominant) (OPA1), mRNA

product length = 143  
Forward primer 1 AAGCAGTGGACTGACAAGCA 20  
Template 2158 ..... 2177

Reverse primer 1 TTTACGGCCTGCTTCAGTTT 20  
Template 2300 ..G..A..... 2281

>[XM\\_005506637.1](#) PREDICTED: *Columba livia* optic atrophy 1 (autosomal dominant) (OPA1), mRNA

product length = 143  
Forward primer 1 AAGCAGTGGACTGACAAGCA 20  
Template 2140 ..... 2159

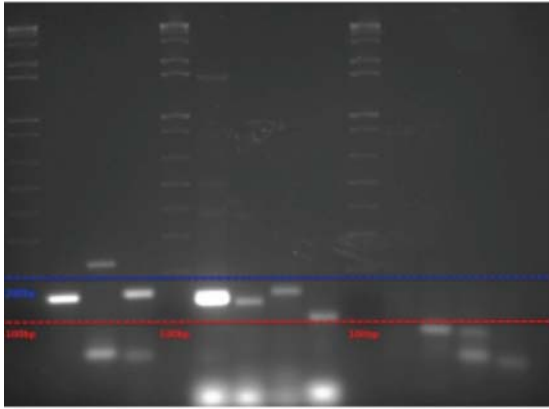
Reverse primer 1 TTTACGGCCTGCTTCAGTTT 20  
Template 2282 ..G.....T..T..... 2263

>[XM\\_005527101.1](#) PREDICTED: *Pseudopodoces humilis* optic atrophy 1 (autosomal dominant) (OPA1), mRNA

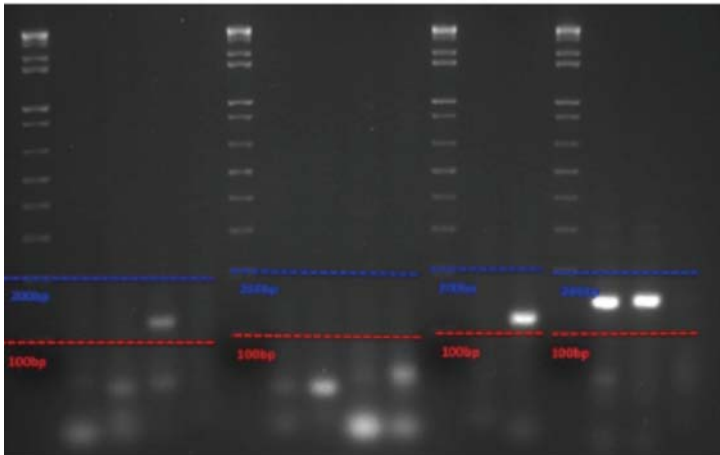
product length = 143  
Forward primer 1 AAGCAGTGGACTGACAAGCA 20  
Template 2122 ..... 2141

All products amplified are OPA1 and give a product length of 143. This strongly indicates that this will be an ideal primer that will anneal to the python transcript. I also looked at several of the other primers listed, but the results began to vary and amplify different genes.

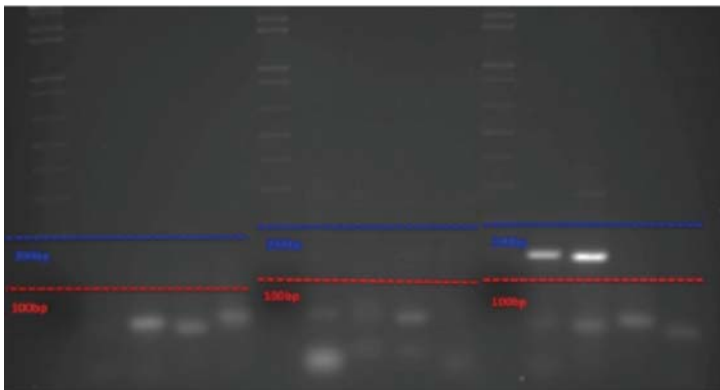
Student 1      Student 2      Student 3

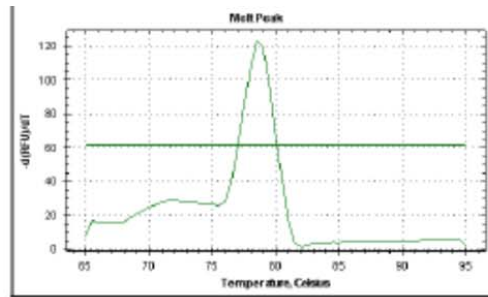
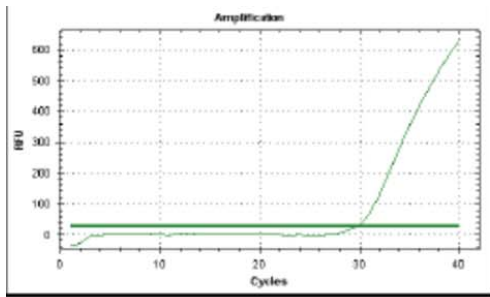


Student 4      Student 5      Student 6      Student 7

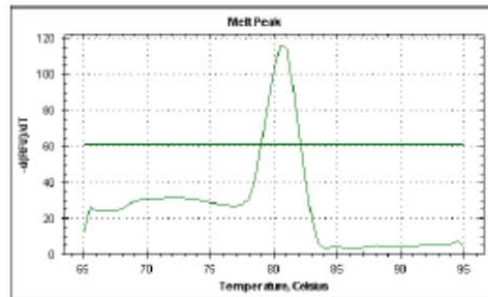
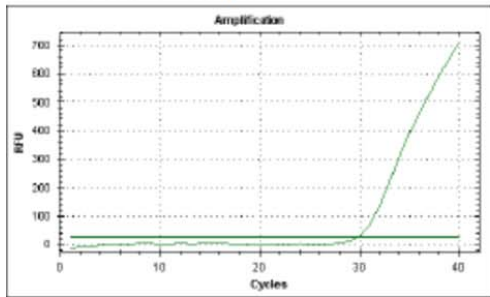


Student 8      Student 9      Student 10

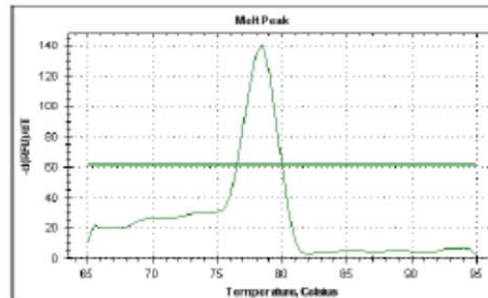
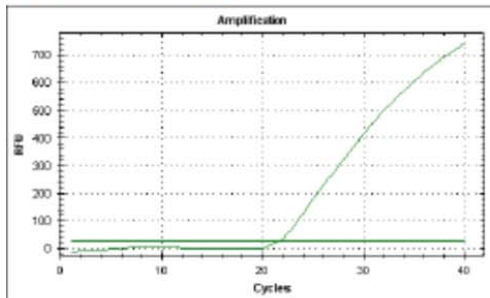




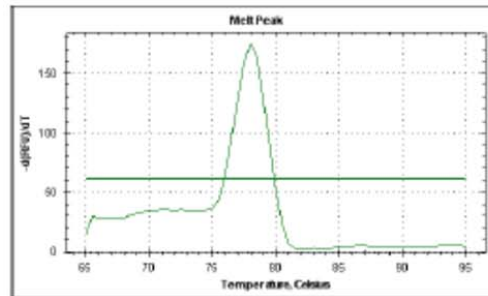
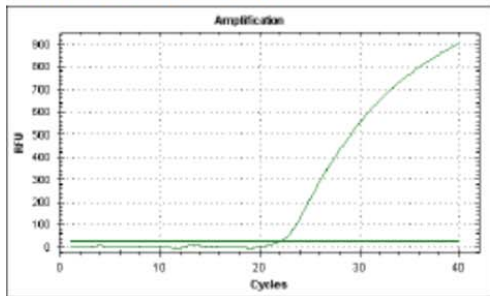
IGF2 F1R1 – 6dpf



IGF2 F2R1 – 6dpf



TIMP3 F1R1 – 6dpf



TIMP3 F1R2 – 6dpf

Question	Bloom's Taxonomy			
Although the heart of the Burmese python is quite similar to the human heart, what is the major structural difference between the hearts of pythons and humans?	Knowledge (LOCS)			
How might this structural difference affect oxygenation of peripheral tissues in the snakes?	Application (LOCS/HOCS)			
<i>Digestion of a large meal is extremely energetically demanding to the python system. On a cellular level, mitochondrial biogenesis is one of the consequences of this demand. What signaling protein is activated in response to low ATP to initiate the process of mitochondrial biogenesis?</i>	Knowledge (LOCS)			
What are three cellular processes that when increased require higher ATP production in a cell?	Knowledge (LOCS)			
What type of bond is broken when ATP is hydrolyzed to release free energy?	Knowledge (LOCS)			
<i>The authors of Riquelme, et al., 2011 examined the response of the Burmese python heart to infusion of several factors (see graph). Explain each of the results shown and their significance.</i>	Analysis (HOCS)			
<i>The authors of Riquelme, et al., also performed real time PCR to examine the expression of several mRNAs in the heart of the Burmese python. Given the graph below, explain the experiment performed and the significance of the results with regard to the type of growth that occurs in the heart after feeding. Be specific about the conditions tested and the purpose of each condition.</i>	Analysis (HOCS)			
<i>As you begin your research on T<sub>3</sub>Rα1, you learn that thyroid hormone (T<sub>3</sub>) induces mitochondrial biogenesis by binding to T<sub>3</sub>Rα1, a receptor that acts as a transcription factor. The T<sub>3</sub>-T<sub>3</sub>Rα1 complex is imported into the mitochondrial matrix through a TOMM channel and increases mitochondrial DNA transcription. The complex can also be imported into the nucleus to induce transcription of other mitochondrial transcription factors such as NRF1 and PGC1α, as seen below. How does binding of T<sub>3</sub> to the receptor (T<sub>3</sub>Rα1) allow the complex to move into the nucleus and mitochondrial matrix?</i>	Application (LOCS/HOCS)			
<i>To complement your studies, you decide to measure expression of the gene targets of the transcriptional complex. You design primers to measure expression of NRF-1 and PGC-1α, but you would also like to measure expression of gene targets in the mitochondrial DNA. What are two types of proteins that are transcribed from mitochondrial DNA?</i>	Knowledge (LOCS)			
<i>Mitochondrial DNA is different from nuclear DNA so will require slightly different techniques for measuring gene expression. Name two features that are unique to mitochondrial DNA.</i>	Knowledge (LOCS)			
<i>During cardiac hypertrophy, hypoxia may occur in the tissue because of the inability of the vasculature to keep pace with cellular growth. What is the effect of low oxygen conditions on the proton motive force generated in the mitochondria?</i>	Analysis (HOCS)			
How might this effect influence expression of T <sub>3</sub> Rα1?	Analysis (HOCS)			
<i>After obtaining the gene sequence for the T<sub>3</sub>Rα1 receptor in the Burmese python transcript for T<sub>3</sub>Rα1, you would like to validate your work before beginning primer design. Describe how would you determine whether the assembled python transcript is correct for T<sub>3</sub>Rα1 in silico.</i>	Knowledge (LOCS)			
<i>You begin designing primers using Primer3. You set the parameters in Primer3 and examine the sequence of the primers that you designed to determine whether the primers are suitable for PCR. What are three desirable qualities in your primer sequences that would make them 'good' primers?</i>	Knowledge (LOCS)			
<i>To ensure that the primers are amplifying specifically, you perform a Primer BLAST, which returns the following results:</i>				
Primer pair 1	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer GGAGGCTCTTCTGGAGAGGT	60.33	60	7	3
Reverse primer GGATGCTAGCTTGGTTCAG	58.33	55	6	1
Products on target templates				
>XM_003641189.2 PREDICTED: Gallus gallus receptor alpha 1 (T3Rα1), mRNA product length = 124 Forward primer 1 GGAGGCTCTTCTGGAGAGGT 20 Template 408 T..C.....G...C. 427 Forward primer 1 GGAGGCTCTTCTGGAGAGGT 20 Template 4863 A...AGA.....C.. 4844				
>XM_004948456.1 PREDICTED: Gallus gallus T3 receptor alpha 1 (T3Rα1), mRNA product length = 2325 Forward primer 1 GGAGGCTCTTCTGGAGAGGT 20 Template 117 A.....A.A..G....G 136 Forward primer 1 GGAGGCTCTTCTGGAGAGGT 20 Template 2441 .....G.GAAC..... 2422				
Utilizing the information provided by Primer BLAST, interpret the results in the context of your real time PCR experiment.	Analysis (HOCS)			
Why are there mismatches between the primers and the template?	Analysis (HOCS)			
<i>You receive your primers from Life Sciences and plan to validate them using conventional PCR. To test your primers, you will need to synthesize cDNA from RNA isolated from python ventricle. Why is this step necessary if you are interested in visualizing the ability of the primers to amplify mRNA transcripts?</i>	Knowledge (LOCS)			



<i>You perform spectrophotometry on your isolated RNA. You obtain an A260 of 4.345.</i> Calculate the concentration of your isolated RNA sample.	Application (LOCS/HOCS)
<i>You calculate the ratio of A260:A280 from your spectrophotometer readings and obtain a value of 1.3.</i> What does this ratio tell you?	Application (LOCS/HOCS)
<i>To check the integrity of your RNA, you perform gel electrophoresis.</i> Calculate the volume of your RNA sample required to load 2,500ng RNA on a gel, based on the concentration you calculated above.	Application (LOCS/HOCS)
<i>The RNA gel yields two bands.</i> What specific species of RNA do these bands represent, and why are only two bands present when there are many different types of RNA present in the sample?	Knowledge (LOCS)
<i>Because you feel unsure about the quality of your RNA, you obtain samples from the Leinwand lab. You can assume that the cDNA synthesis reaction is 100% efficient.</i> If your reaction volume is 20uL in your cDNA synthesis reaction, what amount (mass) of RNA must be added to obtain 100ng/uL cDNA?	Application (LOCS/HOCS)
The synthesized cDNA sequences contain the following (circle all that apply):	Knowledge (LOCS)

- A U C G nucleotides
- A T C G nucleotides
- Random primers
- cDNA copies of genomic DNA
- Exonic sequence

<i>Your primers arrive from Invitrogen and you suspend them in RNase-free water at 100µM.</i> Explain how to dilute these primers to 125µL of 12.5µM, your stock concentration? Show your work.	Application (LOCS/HOCS)
<i>You are now ready to set up your PCR reaction.</i> Fill in the following chart to plan for your PCR: (7 points, 1 point for each calculation) Round values up to the nearest 1/10.	Application (LOCS/HOCS)

	Stock	Volume in 1X Reaction
25ng Python heart cDNA	50 ng/µL	
1X PCR reaction buffer (with MgCl <sub>2</sub> )	10X	
200µM dNTPS	10mM	
1µM Forward Primer	12.5µM	
1µM Reverse Primer	12.5µM	
0.15 Units Taq polymerase	5 Units	
MQ H <sub>2</sub> O (DI/RO)	----	

25µL

In which direction is DNA <i>made</i> by Taq polymerase?	Knowledge (LOCS)
<i>You program your thermocycler to use three standard temperatures.</i> Label each temperature with the appropriate PCR event that you expect to occur.	Comprehension (LOCS)
72°C	
94°C	
60°C	
<i>You decide to examine your PCR products on a 2.2% agarose gel to select a set of primers for real time PCR.</i> How do you make 200mL of 2.2% agarose using powdered agarose and 50X TAE stock? Show your calculations and be specific about the procedure you would use.	Application (LOCS/HOCS)
<i>You are given 10mg/mL ethidium bromide, a 20,000X concentrated stock for your agarose gel.</i> What volume of ethidium bromide would you add to 50mL of liquid agarose so the final concentration of ethidium bromide in the gel is 1X? Show your calculations and be specific about the procedure you would use.	Application (LOCS/HOCS)
<i>The agarose gel results are shown below.</i> Describe each of the bands present in the reactions and interpret the results.	Analysis (HOCS)
Describe three experimental approaches that could be taken to improve the results:	Synthesis (HOCS)

LOCS - lower order cognitive skills  
HOCS - higher order cognitive skills

## Introduction

In its 2013 annual report, the American Heart Association emphasized how pivotal a healthy lifestyle alone is to avoiding cardiovascular disease (CVD), calling for the collective need to focus on prevention rather than treatment of established CVD (Go et al., 2013). As of 2009, CVD accounted for 1 of every 3 deaths in the United States, and with 40.6% of those deaths attributable to high blood pressure, 13.7% to smoking, and 13.2% to poor diet, the need for a mass lifestyle makeover seems dire (Go et al., 2013). But such changes are slow, and millions of Americans, including those with conditions unrelated to diet or lifestyle, suffer and will continue to suffer from CVD. It is in this light that researchers are engaged in expansive efforts to search for new ways to better understand and treat heart disease.

At the University of Colorado Boulder, the Leinwand lab has found a unique model to study cardiovascular disease in the Burmese python. Dr. Leslie Leinwand and her team believe that what they continue to discover about the python and the changes that occur after it eats may provide insight into the processes involved in pathological conditions of the human heart. The python is not only a relatively inexpensive and easy subject of study, in its minimal care requirements; it is also an incredibly unique masterpiece of nature's craftsmanship and evolution's fine-tuning abilities.

After eating, the python experiences an unparalleled increase in metabolic activity (as much as a 44-fold increase in metabolic rate), which is accompanied by widespread organ growth (at least 8 of the python's organs are known to grow in size after a meal), all within 48-72 hours of consumption of a large meal (Riquelme et al., 2011; Secor, 2008). This is an apparent adaptation of the python's sit-and-wait hunting strategy, which allows these snakes to go weeks or months between meals (Secor, 2008). The python heart grows in mass by 40% within three days of feeding. After 10 days, however, the heart (and other organs) returns to its normal, fasted size (Riquelme et al., 2011).

Physiologic hypertrophy of the heart does occur in humans, such as in pregnancy or athletic training, and this growth is reversible. But it is when the heart grows in response to a pathological stimulus like elevated blood pressure, that the growth is irreversible. The heart muscle becomes fibrous, such that the heart no longer functions properly, and this is a nearly certain road to heart failure. What Dr. Leinwand and her team of researchers found was that the hypertrophy of the postprandial python heart is most likely physiologic, rather than pathologic—at least, it does not appear to be harmful for the snake. Evidence for this conclusion includes: an increase in SOD2 (which mediates the reduction of reactive oxygen species), increased expression of alpha-myosin heavy chain, and a lack of increased fibrosis (Riquelme et al., 2011). Leinwand's team has spent the past few years investigating the mechanisms of hypertrophy in the python heart, making a novel discovery that a specific trio of fatty acids identified in postprandial python plasma induces physiological hypertrophy in mammalian cardiomyocytes (Riquelme et al., 2011). This research may provide a new mechanism for treating human heart disease. Understanding how the heart regresses, however, demands further investigation.

The work of the Python Project is used to extend and supplement the research done in the Leinwand lab. This semester the class is investigating the regression of the post-fed python heart, attempting to address a novel question: what molecular mechanisms allow the python's heart to again return to its normal, fasted size? Although a generous amount of research is being done on the mechanisms of heart regression, there is still much to be understood. A model commonly used in the study of regression is transverse aortic constriction (TAC) in mice (Hariharan et al., 2013). Microarray results from a TAC/DeTAC study (in which mice hearts were subject to aortic constriction for one week, after which the constriction was removed) were used to select candidate genes to be studied by this semester's Python class. Looking at these genes in python heart tissue may provide valuable insight into mechanisms of physiologic regression (whereas TAC is a pathologic model).

### LC3 and Autophagy

Organs are limited in the mechanisms by which regression occurs: their cells can undergo apoptosis, necrosis, or autophagy. Often it is a delicate balance, dependent on a tightly regulated network, which directs the cell one way or the other. Although some cells in the heart, such as endothelial cells and fibroblasts, may be undergoing apoptosis, it is unlikely that the cardiomyocytes are regressing in this way. Cardiomyocytes are essentially unable to divide, and if apoptosis were clearing out these cells, there would be a decrease in the python's heart size over time. This particular investigation will be utilizing whole heart tissue, not only cardiomyocytes, and thus there may be some detection of increased apoptosis in this tissue. But with that in mind, there is stronger reason to hypothesize that the heart returns to its fasted size via autophagy or mitophagy,

the degradation of the mitochondria (a process that is not currently being looked at but is one that could be a fruitful investigation in the future of this study).

Autophagy (or macroautophagy, specifically) is a tightly regulated process used by cells to degrade proteins and organelles, helping to maintain homeostasis and a healthy cell environment by removing waste and damaged cell parts. The general steps through which autophagy occurs are: initiation, formation of the autophagosome, fusion of the autophagosome with the lysosome to form the autolysosome, and degradation and recycling of engulfed cargo (Przyklenk et al., 2012).

A well-known marker of autophagy is LC3-II, which plays a crucial role in the formation of the autophagosomal membrane. Studies have shown that the inhibition of LC3-II and its upstream effectors has deleterious effects on cell viability and does not allow for the formation of the autophagosome (Sou et al., 2008; Nishida, Yomaguchi, and Otsu, 2008). Microtubule-associated protein light chain 3 (LC3 or MAP1LC3) is an Atg8-like protein (Atg8 is the autophagic marker in yeast). In mammals it has two other known homologues, GABARAP and GATE-16, also members of this Atg8-like family and also localized to the autophagic membrane; though, while much is now known about the role of LC3, little is known about its homologues (Sou et al., 2008). LC3 has two alternative splicing variants, LC3A and LC3B. Both variants are shown to be autophagosomal markers, as both are known to undergo the same activating cleavage and modification as LC3 that allows it to localize to the autophagic membrane. And while it has been shown that there is a divergence in tissue distribution among these variants, all three are known to be highly abundant in rat heart tissue (Wu et al., 2006). For simplicity, reference will only be made to LC3 and its post-translationally modified forms, LC3-I and LC3-II.

LC3 is immediately converted to its active cytosolic form, LC3-I, upon translation, via Atg4 (an autophagy related protein). LC3-I is further activated by Atg7, and then transferred to Atg3, and with the facilitation of an Atg5-12 complex, is modified to its final, membrane-bound form, LC3-II. This final modification involves the conjugation of LC3-I to a lipid group, phosphatidylethanolamine (PE). Once lipidated, LC3-II localizes to the membrane of the pre-autophagosome, called the phagophore, allowing for its elongation and synthesis into a mature autophagosome (Hanada et al., 2007; Wu et al., 2006). In one study, a pull-down assay was used to determine the necessity of these Atg proteins upstream of LC3-II. It was shown that Atg7, Atg3, and the Atg5-12 complex are all essential for the conjugation of the PE group to form LC3-II (Hanada et al., 2007).

Because of its crucial role in the formation of the autophagosome, LC3-II serves as an effective marker for autophagy. Thus, determining its levels in this tissue at different time points may provide valuable insight into the molecular mechanisms the python heart tissue undergoes as it regresses.

#### **Predicted Outcomes**

The expression levels of both LC3A and LC3B were tested via quantitative PCR (qPCR), using cDNA synthesized from RNA isolated from fasted, 3 days post fed (DPF), 6 DPF, and 10 DPF python ventricle. Following the hypothesis that autophagy is the means by which the python heart regresses, it was expected that there would be an increased level of LC3B and LC3A expression in 6 DPF tissue. Although post-translational modifications dictate LC3 activity—LC3 is not active until it is cleaved by Atg4 in the cytosol, and further, does not serve its purpose in the elongation of the autophagosomic membrane until lipidated to form LC3-II—there is still expected to be an increase in its expression levels (Hariharan et al., 2013; Tamida et al., 2005). It is well understood, though, that these mRNA expression levels, as represented by the cDNA from the respective samples, are alone not sufficient to make any significant conclusions.

One way to generate a more comprehensive understanding of the molecular activity taking place at these time points would be to also obtain data from a number of other critical genes upstream of LC3-II in this autophagic pathway, such as Atg3, Atg7, Atg5, Atg12, and Atg16 (Hanada et al., 2007). The qPCR data from these genes, in conjunction with that from LC3A/B, may provide strong enough evidence to indicate autophagy in these cells. Another way would be to determine protein levels of LC3-I and LC3-II via immunoblotting; however, there are two immediately known problems with this method: 1) that LC3-II is itself degraded as the autophagosome matures into the autolysosome, and 2) LC3-II levels are often much more sensitive to detection than LC3-I (Mizushima and Yoshimori, 2007).

This study was seen, then, as a preliminary investigation into the mechanics of a complex network of

interactions, and the information obtained from it only serve as a means, albeit a solid one, to continue examining the role of autophagy and other pathways in the process of regression in the future.

## **Methods**

### ***Gene selection***

Genes were selected primarily via theoretical understandings of the pathways involved in autophagy and through findings presented in the literature surrounding autophagy and heart regression (Gustafsson and Gottlieb, 2009).

### ***Primer design and validation in silico***

Primers were designed for both LC3A and LC3B. First, a nucleotide BLAST (basic local alignment search tool) was performed using the cDNA sequence, with color-coded exons, of the gene from *Gallus gallus* (chicken) against the python whole genome shotgun sequence (WGS). Python contig sequences that conferred a high similarity to the cDNA sequence were mapped to the cDNA (the sequences from the chicken that were shown to be similar to the python were highlighted in the cDNA). The python transcript was assembled using the order of the exons denoted in the highlighted chicken sequence. This assembly was validated by translating the transcript into an amino acid sequence and performing a protein BLAST. Primers were then designed with Primer3 using the assembled transcript and were validated in a primer BLAST. Search parameters included both *Gallus gallus* and *Anolis carolinensis* (lizard) as acceptable matches. This was done separately for each gene.

### ***RNA isolation***

Tissue obtained from the python heart ventricle was disrupted in a lysis buffer and homogenized. (Note: tissue from each time point—fasted, 3 DPF, 6 DPF, and 10 DPF—was collected and pooled from two snakes.) Ethanol was then added to the lysate, after which the sample was applied to a Qiagen RNeasy mini spin column and eluted in water. The procedure followed the protocol provided by the Qiagen RNeasy kit. Note: the RNA that was ultimately used for the experiment was isolated using a phenol-chloroform extraction.

The integrity of the RNA was tested with a denaturing gel, and the concentration and purity of the RNA was determined using a spectrophotometer. The gel was 2.2% agarose in Tris EDTA with 6X Orange G loading dye, and 5,000X Ethidium Bromide was used to visualize the products on the gel.

### ***cDNA synthesis***

cDNA was synthesized from the RNA using standard protocol. Superscript III was the reverse transcriptase (RT) used in the reaction. The appropriate amount of RNA (as determined by the concentration estimation) to yield 50 ng/uL of cDNA was added first along with random hexamers, dH<sub>2</sub>O, and deoxyribonucleotides (dNTPs). Then the RNase OUT, first strand buffer (FSB), RT, and dithiothreitol (DTT) were added to initiate the PCR (concentration of each component was determined in the protocol). The PCR program included one cycle consisting of two steps, one at 65 degrees Celsius (deg C) and one at 50 deg C.

### ***Primer validation***

Primers for LC3A and LC3B were first tested using a PCR, the products of which were run on a 2.2% agarose gel with Tris EDTA and 6X Orange G loading dye and visualized with the addition 5,000X Ethidium Bromide. This was used to validate product size. To confirm that only each primer set only yielded one product, the primers were additionally subjected to a qPCR melt curve analysis.

### ***Quantitative real-time PCR (qPCR)***

Gene expression levels were determined using qPCR with SYBR Green. The reaction mix consisted of 2 ng/uL of cDNA, 12.5 uM of validated forward and reverse primers for the gene, and complete SYBR reaction mix. Serial dilutions for the standard curves were made using pooled cDNA from each of the time points (F, 3 DPF, 6 DPF, and 10 DPF) and followed a 1:10 dilution. 18S was used as the reference gene, and both standard curves and samples were loaded in triplicate. Analysis of the data generated by the qPCR was conducted using the [delta] [delta] Ct method, with 18S as the reference gene and fasted as the control condition. The data were then subjected to a Student's *t*-test, and statistical significance between the time points was determined. Error bars represent the standard error between each test, and data that also did not fit within two standard deviations from the mean of each test per time point were not included.

### ***Primers used in qPCR***

LC3A: GGTACAGCTGAATCCAACC (forward, 5' to 3')  
CCCAAAGTTTCCTGTGAAG (reverse, 5' to 3')



LC3B: TACTGGTGAACGGACACAGC (forward, 5' to 3')  
AAGAATCCGTCCTCGTCCTT (reverse, 5' to 3')

## Results

The protein BLAST results of the amino acid sequences from both genes confirmed their identity, i.e. that the python transcript was correctly aligned for each gene and that its protein product agreed (sequences from both LC3A and B generated 97-100% matches for MAP1LC3A and B, respectively, in both *Gallus gallus* and *Anolis carolinensis*). This confirmation then allowed for those transcripts to be used in primer design, the products of which were validated using primer BLAST. Similarly, these also yielded positive results, assuring *in silico* that these primers effectively generated the products of interest.

The denaturing gel used to examine the isolated RNA confirmed its integrity, as two clear bands (one representing the 28S rRNA and the other the 18S rRNA) were observed. Intactness and purity thus established, the RNA was then used to synthesize the cDNA utilized in the remainder of the experiments. The PCR showed a single product of the correct size for each primer set (figure 1). The product for LC3A was 152 base pairs and for LC3B, 85, and both gels showed products to be about these sizes. The melt curves provided additional assurance of the primers' effectiveness in amplifying a single product, as both primer sets generated a single, clean melt peak (figures 2a and 2b).

The qPCR experiment showed that both genes were relatively high in abundance in the python tissue. Sufficient standard curves were generated that showed amplification at about 22-25 cycles (figures 3a and 3b). The data for LC3A did not indicate any statistically significant changes in gene expression between fasted and any of the other time points. There did seem to be a slight decrease at 3 DPF and slightly greater increase at 6 DPF, but as error was high, these results could not be validated (figure 4a). The data for LC3B did, however, suggest a decrease at 3 DPF from the fasted level, which translated to about a 0.6 fold-change (figure 4b). Although a slight increase at 6 DPF, rising closer to the fasted level, was observed, this, again, was not determined to be a statistically significant change in expression levels.

## Discussion

The purpose of this study was to quantify changes in gene expression in postprandial python ventricle tissue, a preliminary means to investigate the molecular mechanisms involved in the regression of the heart. As a well-known marker of autophagy, LC3 was expected to increase at 6 DPF, when regression is understood to occur in the organs of the python. Thus if autophagy were indeed the means by which the heart regressed, it was thought that this gene would be upregulated and an increased expression would be observed. Another hypothesis, though, was that no significant change would occur between the expression levels at the various time points, as the LC3 protein is only involved in elongating the autophagosomal membrane when it is post-translationally modified to its active, lipid-linked form, LC3-II.

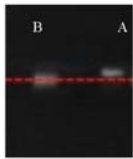
The data did not indicate any clear increase at 6 DPF; neither did they necessarily show that expression levels were remaining constant. The only statistically significant result was a decrease in LC3B from fasted to 3 DPF. This result, in conjunction with data collected by others about the expression patterns of other genes upstream of LC3 in this autophagic pathway, such as FOXO3 and Atg7, may indicate that autophagy is *not* taking place as expected and that the heart may indeed be regressing via other mechanisms. No such ruling can be established, however, without looking at protein levels, especially of LC3-II and GABARAP1, which is downstream of LC3 and is thought to be involved in the completion of the autophagosomic membrane.

The observed decrease at 3 DPF is congruent with the understanding that hypertrophic processes are most active around that time point, and as the cell is not wired to run two opposing pathways concurrently, it makes sense that the cell would be depressing the activity of autophagy-related genes as it undergoes a process of growth. This result may simply provide further validation that LC3 is abundant in the python heart and that its known function is most likely conserved in the python (as would be expected and as could primarily be insinuated from the BLAST results).

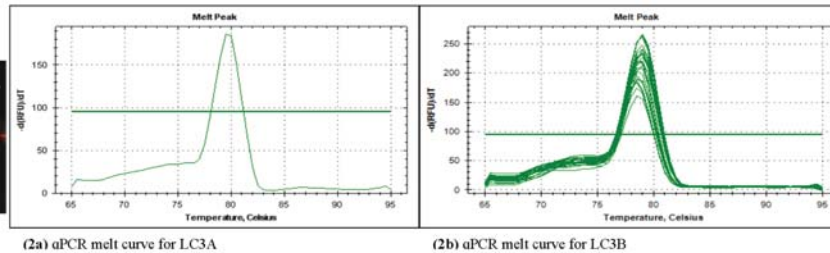
A couple of immediate limitations with this study include: the limited information that can be provided by examining mRNA expression alone (represented by the cDNA), the sample size used (two snakes per time point), the inability to isolate cardiomyocytes from the ventricle tissue (thus making conclusions more difficult as various cell types might be undergoing different pathways to regress), and the lack of information known about the difference, if any, between splice variants of LC3.

Autophagy may still be the mechanism used by the python heart to regress, though it would be prudent to first attempt to rule out this possibility by looking at protein levels of LC3-II and others in the pathway and then to consider other possible mechanisms such as general proteolysis or mitophagy. This study, though preliminary, serves as an important springboard from which to continue investigations about the python heart, its other organs, regression in general, and the implications it might have in treating human heart disease. As a profound issue in this country today, human CVD demands attention both in the action taken to prevent its prevalence and the research done to mediate its effects. Understanding the mechanisms of heart regression is especially important, as many conditions could be reversed if regression were either induced to correct pathologic hypertrophy or prevented from progressing to irreversible atrophy and heart failure. The unique model of the Burmese python may indeed be the means to unravel these issues and unlock an effective treatment for human CVD.

## Figures



(1) PCR primer test of LC3B (left) and LC3A (right)

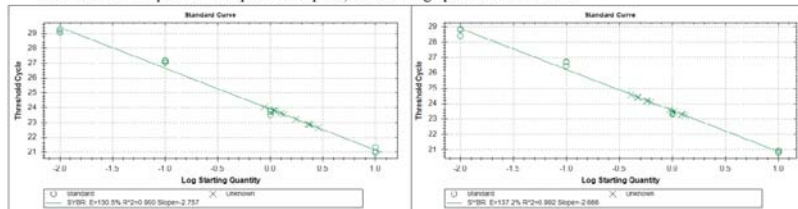


(2a) qPCR melt curve for LC3A

(2b) qPCR melt curve for LC3B

Figure 1 illustrates the amplification of a single product by the PCR, as well as products of the correct size. The LC3B product was designed to be 85 bp and LC3A 152 bp. The red line represents the 100 bp mark of the ladder.

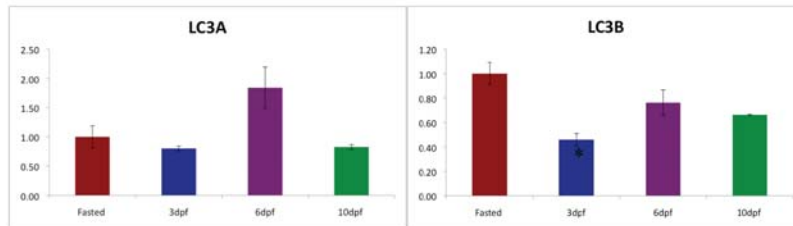
Figures 2a and 2b demonstrate the amplification of a single product, as suggested by the single melt peak. Figure 2a was obtained from melt tests run prior to the experimental qPCR, as the other graphs were unattainable.



(3a) qPCR standard curve for LC3A

(3b) qPCR standard curve for LC3B

Figures 3a and 3b were obtained from the qPCR experiments and demonstrate primer efficiency, relative product abundance, accurate pipetting, and sound dilutions. Ct values for both LC3A and LC3B were seen to be about 22-25.



(4a) LC3A gene expression

(4b) LC3B gene expression

Figures 4a and 4b represent the gene expression levels from the two qPCR plates used in this experiment. No significant changes were determined for LC3A. LC3B shows a 0.6 fold-change decrease from fasted to 3 DPf, with a p-value > 0.05.

## Works Cited

- Go A, et al. (2013) Heart Disease and Stroke Statistics—2013 Update, A Report From the American Heart Association. *Circulation* 127: e6-e245.
- Gustafsson, Å. B. and R. A. Gottlieb. (2009) Autophagy in Ischemic Heart Disease. *Circulation* 104:150-158.
- Hanada T, et al. (2007) The Atg12-Atg5 Conjugate Has a Novel E3-like Activity for Protein Lipidation in Autophagy. *J. Biol. Chem.* 282: 37298-37302.
- Hariharan N, et al. (2013) Autophagy Plays an Essential Role in Mediating Regression of Hypertrophy during Unloading of the Heart. *PLoS ONE* 8(1): e51632. doi:10.1371/journal.pone.0051632
- Mizushima N and Yoshimori T. (2007) How to Interpret LC3 Immunoblotting. *Autophagy* 3:6, 542-545.
- Nishida K, Yamaguchi O, and Otsu K. (2008) Crosstalk Between Autophagy and Apoptosis in Heart Disease. *Circulation Research* 103: 343-351.
- Przyklenk K, et al. (2012) Autophagy as a therapeutic target for ischaemia /reperfusion injury? Concepts, controversies, and challenges. *Cardiovascular Research* 94: 197–205.
- Riquelme C, et al. (2011) Fatty Acids Identified in the Burmese Python Promote Beneficial Cardiac Growth. *Science* 334 (6055): 528-531.
- Secor S. (2008) Digestive physiology of the Burmese python: broad regulation of integrated performance. *The Journal of Experimental Biology* 211: 3767-3774.
- Sou Y, et al. (2008) The Atg8 Conjugation System Is Indispensable for Proper Development of Autophagic Isolation Membranes in Mice. *Molecular Biology of the Cell* 19: 4762– 4775.
- Tanida I, et al. (2005) Lysosomal Turnover, but Not a Cellular Level, of Endogenous LC3 is a Marker for Autophagy. *Autophagy* 1:2, 84-91.
- Wu J, et al. (2006) Molecular cloning and characterization of rat LC3A and LC3B—Two novel markers of autophagosome. *Biochemical and Biophysical Research Communications* 339: 437–442.

