

## **PSC MARC Internship Summer 2015 Workplan**

**Project Title:** Detection of differential expression from de-novo RNA-seq data

**Student:** Iván L. Jiménez Ruiz

**Primary Mentors:** Alexander Ropelewski, Humberto Ortiz-Zuazaga, José Agosto Rivera

**Secondary Mentors:** Pallavi Ishwad, Selena Rodriguez Rivera, Antonio Camacho Flores

### **Abstract:**

- Detect differential expression in RNA sequenced data reads in the absence of a reference genome for *Drosophila Melanogaster*
- Generate an algorithm or pipeline to be used in de-novo assembly and/or transcription

**Scientific Objectives:** [Bulleted list of specific objectives or goals that you intend to get out of the internship project from a scientific basis; for example:

- Develop the project to be presentable in a scientific paper.
- Analyze Drosophila and foreign RNA-seq reads by using a variety of bioinformatics tools
- Apply bioinformatics tools for the analysis of spike-in data for the detection of differential expression at the gene level of *D. Melanogaster*
- Assemble the complete transcriptome of *D. Melanogaster*

### **Personal Objectives:**

- Learn about RNA sequencing in both biological and computational terms
- Gather important information that will be used in your thesis
- Improve algorithmic, problem solving, computational and general programming skills
  - UNIX
  - Trinity
  - Galaxy
  - SRA Toolkit
  - Databases
- Gain more experience and confidence as a programmer and a presenter

**Prior Work:** The project is based on the analysis of SRA data files of Drosophila published at the NCBI website, which can be downloaded from: <http://www.ncbi.nlm.nih.gov/sra>

### **Deliverables:**

- Possibly an algorithm used in the detection of differential expression of reads in a de-novo assembly project (using *D. Melanogaster* data).
- A rough draft of a PLOS one paper
- Quality multiple sequence alignment of *D. Melanogaster*

**Milestones:** [This is a timeline of what you expect to complete. List at least one milestone for each week of the internship, plus milestones for up to four semesters once you return to campus - for example:

1st week (June 1): Gather additional information about *Drosophila melanogaster* and RNA sequencing, continue to read the journals and papers that were published on PubMed, prepare the first of many weekly presentations and complete a Workplan

2nd week (June 8): Attend the Workshop and apply the techniques that are learned to my project

3rd week (June 15): Attend the Workshop and apply the techniques that are learned to my project

4th week (June 22): Find and collect some known *Drosophila* SRA file sequences (from databases such as PIR, and PDB), and perform an initial multiple sequence alignment of the sequences (using Trinity?), compare findings with literature (maybe create an initial phylogenetic tree).

5th week (June 29): Write a draft methods section for a PLOS One paper for my project, identify key notes on differential expression (maybe create revised alignments and phylogenetic trees)

6th week (July 6): Integrate sequence-based bioinformatics work with structural based bioinformatics.

7th week (July 13): Prepare poster for Duquesne University 2015 Summer Research Symposium

8th week (July 20): Present poster at Duquesne University 2015 Summer Research Symposium

9th week (July 27): Complete draft paper of bioinformatics work.

Fall Semester 2015: Experimental verification of Bioinformatics work.

Spring Semester 2016: Submit paper to PLOS One. Submit paper on computational work to XSEDE 16 conference and SC 16 conference.

Summer Semester 2016: Present computational work at XSEDE 16 conference.

Fall Semester 2016: Present computational work at SC16 conference.