

# GEP Digital Laboratory Notebook Faculty Guide

Fall 2014

## I. Introduction

This GEP Digital Laboratory Notebook (DLN) was created for implementation of GEP annotation projects in introductory level biology courses. Students in these courses have a limited background in molecular genetics (alternative splicing, etc.), bioinformatics algorithms (BLAST and gene prediction), and modern molecular technologies (next-gen DNA sequencing and RNA-seq). To assist these entry-level students, the DLN instructions scaffold their learning and the DLN itself organizes their observations and analysis. Implementation of the DLN will help students make the most of the time they spend on GEP annotation projects.

## II. Background preparation

Before implementing the GEP DLN students should

- Be familiar with basic molecular genetics and gene expression from instruction in the course that is implementing the GEP annotation or in a pre-requisite course. Instruction in genomics and DNA sequencing is recommended as well.
- Watch the Next Generation Sequencing Video Tour on the GEP website at [http://gеп.wustl.edu/curriculum/course\\_materials\\_WU/introduction\\_to\\_genomics/nextgen\\_video\\_tour](http://gеп.wustl.edu/curriculum/course_materials_WU/introduction_to_genomics/nextgen_video_tour)
- Complete and/or reference the Basics of BLAST assignment or other Annotation Tutorial or Walkthrough from the GEP website - [http://gеп.wustl.edu/curriculum/course\\_materials\\_WU/annotation/tutorials\\_and\\_walkthroughs](http://gеп.wustl.edu/curriculum/course_materials_WU/annotation/tutorials_and_walkthroughs)

## III. Files and Implementation

### I. Implementation Tips

- Since the instructions are very explicit and there are videos to support student understanding (*please note that the videos are only visual aids and have no narration or sound*), it should be possible for students to complete the RpS3A example annotation exercise fairly independently. I would try assigning this annotation exercise as homework and have your students turn in completed example Contig2 DLN pages (see more about the various types of files below).
- During class time I would initially address questions about the RpS3A example annotation exercise that was completed as homework and then let the students work on the annotation of their claimed contig projects with supervision. One nice feature of the DLN is that students can document their work and move on. If a student runs into a difficult section of the gene and they need your help they can get assistance later if they have the question well documented in the DLN. Also, it is possible with the DLN to quickly look through all of a student's work and assess whether they have been making a consistent error and determine if they have or have not put in enough effort to produce a quality annotation.
- Once a student has completed their annotation work using the DLN they should use the instructions in the new "Annotation of a *Drosophila* Gene" exercise that Wilson Leung has created and posted on the GEP website to verify their annotation using the Gene Model Checker web application (page 26 to 32). The coordinates to run the Gene Model Checker Analysis are in the highlighted boxes of the DLN page 3 making the process of assembling the CDS coordinates simpler. Click the link below to download a copy of the "Annotation of a *Drosophila* Gene" exercise:

## II. Student files

On the GEP website there are several files for students to download. The DLN is organized as “pages” (1, 2, and 3) that students will work through sequentially. Each DLN “page” has three types of files. Below is a list of the files and a short explanation of the purpose of each type of file:

1. An instructions file that walks the students through an example annotation of the *RpS3A* gene in the *Drosophila biarmipes* 4<sup>th</sup> chromosome Contig2.
  - a. DLN Page 1 Synteny Instructions.docx
  - b. DLN Page 2 Gene Structure Instructions.docx
  - c. DLN Page 3 Coding DNA Sequence Instructions Part 1.docx
  - d. DLN Page 3 Coding DNA Sequence Instructions Part 2.docx
2. A DLN file for the *RpS3A* Contig2 example annotation
  - a. DLN Page 1 Synteny Contig 2 Example.docx
  - b. DLN Page 2 Gene Structure Contig 2 Example.docx
  - c. DLN Page 3 Coding DNA Sequence Contig 2 Example.docx
3. A DLN file for the student’s claimed contig annotation project
  - a. DLN Page 1 Synteny.docx
  - b. DLN Page 2 Gene Structure.docx
  - c. DLN Page 3 Coding DNA Sequence.docx

Also, each DLN “page” focuses on different learning objectives (LO) that start at low levels of Bloom’s taxonomy and then progress to higher levels of the taxonomy of cognition. Below are the LO’s each DLN “page” focuses on:

1. DLN Page 1 Synteny
  - a. LO1: Develop a basic level of skill using the GEP genome browser and Flybase website.
  - b. LO2: Describe the synteny or lack of synteny between the contig they have claimed for annotation and the orthologous chromosomal region in *Drosophila melanogaster*.
  - c. LO3: Decide whether the BLASTX evidence track in the GEP genome browser has revealed all of the genes in the claimed contig.
2. DLN Page 2 Gene Structure
  - a. LO1: Apply the GEP gene browser skills learned from DLN page 1 to analyze the structure of the gene isoforms in a contig annotation project.
  - b. LO2: Compare the gene structure of the query genes predicted by BLASTX to the gene structure of the reference *Drosophila melanogaster* genes.
  - c. LO3: Analyze the open reading frames for each predicted query exon to determine whether the predicted gene structure is supported or not.
3. DLN Page 3 Coding DNA Sequence
  - a. LO1: Investigate the coding DNA sequence of a gene.
  - b. LO2: Assess whether the independent evidence provided by the BLASTX, Genscan, and RNA-seq evidence tracks support or refute hypotheses about the location of the start codon, splice junctions, and stop codons in the coding DNA sequence.
  - c. LO3: Evaluate whether the coding DNA sequence coordinates supported by the evidence tracks produce an amino acid sequence that is similar to the reference *Drosophila melanogaster* amino acid sequence.
  - d. LO4: Document the results of this intensive coding DNA sequence analysis in a Digital Laboratory Notebook for reference later.

### III. Faculty files

For faculty there are a number of additional files beside this instructions file. Below is a brief list of the types of files and their purpose:

1. For each DLN “page” there is an answer key for the RpS3A annotation example. You can use these files to grade student work and to determine whether they completed the example annotation exercise.
  - a. DLN Page 1 Synteny Contig 2 Example Key.docx
  - b. DLN Page 2 Gene Structure Contig 2 Example Key.docx
  - c. DLN Page 3 Coding DNA Sequence Contig 2 Example Key.docx
2. For each DLN “page” there is also an example of student work from my class showing how students would use the DLN to document their annotation work for their claimed contig. You can use this as a model for the type of documentation your students should be able to produce.
  - a. DLN Page 1 Synteny Contig 2 Student Work Example.pdf
  - b. DLN Page 2 Gene Structure Student Work Example.pdf
  - c. DLN Page 3 Coding DNA Sequence Student Work Example.pdf
3. For the introduction to DLN Page 3, I created a figure to explain what a coding DNA sequence. I have included a Word file that contains the full size image. This image is editable if you would like to use it for teaching purposes.

Good Luck and Happy Gene Annotating!

Nick Reeves, Ph.D.

[nreeves@msjc.edu](mailto:nreeves@msjc.edu)

Associate Professor

Chair, Department of Biology

Mt. San Jacinto College

Menifee, CA