|  |  |
| --- | --- |
|  | F Element Project: Quick Start Guide |

# Types of Curriculum Materials

* **Lecture:** provide an annotated slide set;
* **Walkthrough:** highly scripted “click the button” tours that guide a student through a protocol;
* **Exercises:** require the student to go through the process answering questions (can be done for a grade);
* **Workflows:** one-page handouts to remind the student of the steps for a protocol, including the logic for decision making.
* **Resource:** project-specific materials (e.g., annotation protocols, strategy guides, and reports)

# Overview of the Concepts/Tools

All the GEP curriculum materials can be accessed through the [GEP Curriculum Search](http://thegep.org/curriculumsearch) page. Curriculum specific to the F Element Project are available under the “[Project Curriculum](https://thegep.org/projects/felement/#project-curriculum-header)” section.

The Quick Start Guide is organized around five major Concepts/Tools:

## Concepts/Tools 1, 2, and 3

* Beginning curriculum common to all GEP projects.

## Concepts/Tools 4

* Prepares students to work on their own annotation projects; submitting completed projects to GEP allows them to become co-authors (if they read/critique/approve the resulting group manuscript or participate in micropublication).

## Concepts/Tools 5

* If time permits, the project will benefit by students checking the Transcription Start Sites annotation for each gene as well.

## Optional Concepts/Tools

* Exploring motifs and analyzing repetitious elements are follow-on explorations that will contribute to student understanding of eukaryotic genomes, but these topics are not part of the current research protocol.

# Concepts/Tools Table

|  |  |  |
| --- | --- | --- |
| **Concepts/Tools** | **Minimum Recommended Materials** | **Optional/Additional Materials** |
| **1. Gene structure, Introduction to a genome browser, Introduction to gene model construction**  The recommended materials provide an introduction to gene structure by covering transcription, mRNA processing, and translation. The introduction to gene model covers topics including a codon consists of three nucleotides, that a coding exon can begin or end in the middle of a codon (i.e., phases of the splice acceptor and donor sites), and the need to maintain an open reading frame after splicing. | * Exercises: [UEG modules 1–6](https://thegep.org/ueg/) * Package includes [glossary](https://thegep.org/lessons/mlaakso-reference-ueg_glossary_beginning_students/) and [videos](https://youtube.com/playlist?list=PLjMi88MMUqsJ5xHbmVU9p-BJU44Wukgyn) | * Any of the following combinations of UEG depending on student background preparation: * [1](https://thegep.org/lessons/jstamm-lesson_with_exercises-ueg_module1_gene/), [2](https://thegep.org/lessons/msantisteban-lesson_with_exercises-ueg_module2_transcription_i/), [5](https://thegep.org/lessons/chowell-lesson_with_exercises-ueg_module5_translation/), & [6](https://thegep.org/lessons/lpaliulis-lesson_with_exercises-ueg_module6_alternative_splicing/) * [1](https://thegep.org/lessons/jstamm-lesson_with_exercises-ueg_module1_gene/), [4](https://thegep.org/lessons/mlaakso-lesson_with_exercises-ueg_module4_splicing/), & [6](https://thegep.org/lessons/lpaliulis-lesson_with_exercises-ueg_module6_alternative_splicing/) * [Condensed versions](https://wustl.box.com/s/j7y8ulnjk9tw7l1pv2oqjhdc3ek4hc7y) of UEG modules |
| **2. Homology/BLAST**  BLAST reports regions of sequence similarity; what is it looking for and how do we interpret the report? | * Walkthrough: [An Introduction to NCBI BLAST](https://thegep.org/lessons/wleung-lesson-simple_introduction_blast/) * Exercise: [Detecting and Interpreting Genetic Homology](https://thegep.org/lessons/wleung-lesson_with_exercises-detecting_interpreting_homology/) * Alternative (simpler) Exercise: [Introduction to BLAST using Human Leptin](https://thegep.org/lessons/jdiangelo-lesson_with_exercises-leptin_blast/) | **Background Lesson:**   * Lecture: [Using BLAST for Genomic Sequence Annotation](https://thegep.org/lessons/jbuhler-lecture-blast_genomic_annotation/)   (PowerPoint presentation by Dr. Buhler)   * Lecture Notes: [Detecting and Interpreting Genetic Homology: Lecture Notes on Alignment](https://thegep.org/lessons/jbuhler-lesson-blast_notes/) (accompanying notes for the PowerPoint presentation) |
| **3. RNA-Seq (Optional)**  We make extensive use of RNA-seq data; this is primarily derived from processed transcripts, so we can Identify splice sites based on spliced RNA-Seq reads. | * Lecture: [RNA-Seq: A Closer Look at Read Mapping](https://thegep.org/lessons/jbuhler-lecture-rnaseq_mapping/)   (PowerPoint presentation by Dr. Buhler)   * Exercise: [Browser-based annotation and RNA-seq data](https://thegep.org/lessons/jbuhler-lesson_with_exercises-browser_based_rnaseq/) |  |

|  |  |  |
| --- | --- | --- |
| **4. Annotation**  How to use multiple lines of evidence to establish the presence of a gene, determine its *D. melanogaster* ortholog, and construct the best gene model, as defined by the available evidence. In most cases our evidence includes:   1. sequence homology = evolutionary conservation; 2. *ab initio* gene models = rules of ORF, start and stop codons; 3. RNA-seq data = local transcription, position of exons, splice site borders. | * Lecture: [Annotation of *Drosophila* Primer](https://thegep.org/lessons/wleung-lecture-drosophila_annotation_primer/)   (Describes annotation goals and strategies. Accompanies the walkthrough below.)   * Walkthrough: [Annotation of a *Drosophila* Gene](https://thegep.org/lessons/wleung-walkthrough-annotation_drosophila_gene/) * Exercise (optional): [Simple Annotation Problem](https://thegep.org/lessons/vsundaram-lesson_with_exercises-simple_annotation_problem/) * Resource: [F Element Project: Annotation Report](https://thegep.org/lessons/wleung-project_resource-annotation_report/)   (A detailed Word document; students must complete this form providing evidence for their annotation to contribute to the research and be eligible to be a co-author.)   * Resource: [GEP Annotation Workflow](https://thegep.org/lessons/wleung-workflow-gep_annotation/) * Resource: [Annotating Splice Sites Workflow](https://thegep.org/lessons/jwong-workflow-splice_sites/) * Resource: [Identify the *D. melanogaster* Ortholog](https://thegep.org/lessons/wleung-workflow-identify_ortholog/) | The following PowerPoint presentations provide alternatives to the “[Annotation of *Drosophila* Primer](https://thegep.org/lessons/wleung-lecture-drosophila_annotation_primer/).”  (Slides can be mix-match to fulfill your pedagogic needs)   * Lecture: [Annotation of *Drosophila*](https://thegep.org/lessons/wleung-lecture-drosophila_annotation/) * Lecture: [Annotation for *D. virilis*](https://thegep.org/lessons/cshaffer-lecture-annotation_dvirilis/)   **Other Optional Materials**   * Resource: [Annotation Instruction Sheet](https://thegep.org/lessons/cshaffer-project_resource-annotation_instruction_sheet/)   (A more in-depth description of the GEP approach; complements the “Annotation for *D. virilis*” PowerPoint.)   * Resource: [Annotation Strategy Guide](https://thegep.org/lessons/jwong-project_resource-annotation_strategy_guide) (How GEP strategies can be applied to more challenging cases.) * Exercise: [Browser-based Annotation and RNA seq-data](https://thegep.org/lessons/jbuhler-lesson_with_exercises-browser_based_rnaseq/) (This exercise provides more practice on comparative annotation, using BLAST and RNA-Seq.) * Resource: [GEP Digital Laboratory Notebook](https://thegep.org/lessons/nreeves-project_resource-gep_digital_lab_notebook/)   (One example of a digital notebook to guide students.) |

|  |  |  |
| --- | --- | --- |
| **5. Annotating the Transcription Start Sites (TSS) (Optional)​**  Introduces concepts of promoter architecture and the experimental techniques for characterizing promoters (e.g., CAGE and RAMPAGE; RNA-Seq; ChIP-Seq data for RNA Polymerase II and transcription factors; DNase I Hypersensitive Sites in chromatin; 9-state chromatin models based on histone modifications). | **Note:** We now have RAMPAGE data for the four species used here. The draft TSS annotation protocols and curriculum materials which use the new RAMPAGE data are available on [Box](https://wustl.box.com/s/bw2160o4ybgk3v7et82kvz7i8mwig0mx).  **Posted Materials:**  Lecture: [Searching for Transcription Start Sites in *Drosophila*](https://thegep.org/lessons/wleung-lecture-search_tss_drosophila/)  Walkthrough: [Annotation of Transcription Start Sites in *Drosophila*](https://thegep.org/lessons/wleung-lesson-annotation_tss_drosophila/)  Resource: [TSS Annotation Workflow](https://thegep.org/lessons/wleung-workflow-tss_annotation/) | Exercise: [TSS Module Primer: Review of Transcription, Promoter Structure, and Chromatin Packaging](https://thegep.org/lessons/jsiders-lesson-tss_modules_primer/)  (This exercise provides an overview of promoter structure in eukaryotic genomes that might be helpful to students before they work on the TSS Modules.)  Exercise: [TSS Modules (1-4)](https://thegep.org/tss/) |

# Notes

## Concept/Tools 1:

**Starting Points**. Freshmen/sophomores just learning about eukaryotic gene structure (exons/introns, etc.) may be best served by taking the time to walk through all six modules of “Understanding Eukaryotic Genes” (UEG). Upper level students with a good grounding in molecular genetics need only review this material, which also introduces the nomenclature and tools of a genome browser (UCSC Genome Browser Mirror). UEG is highly scripted, and can be given as an “at home” assignment before the start of the semester.

## Concept/Tools 4:

The three lectures are being used in different settings to launch students on the structural annotations of the F Element:

* The “[Annotation for *D. virilis*](https://thegep.org/lessons/cshaffer-lecture-annotation_dvirilis/)” uses the *mav* gene to illustrate the annotation protocol. The *mav* gene only has two coding exons, so it provides the quickest way to provide an overview of the annotation process.
* The “[Annotation of *Drosophila*](https://thegep.org/lessons/wleung-lecture-drosophila_annotation/)” presentation is a more detailed version of the “Annotation for *D. virilis*” presentation. It provides additional information regarding the RNA-Seq data, interpretation of the BLAST search results, and phases of splice donor and splice acceptor sites. Some faculty adapt the curriculum by combining a subset of slides from the “Annotation of *Drosophila*” presentation with the “Annotation for *D. virilis*” presentation.
* The “[Annotation of *Drosophila* Primer](https://thegep.org/lessons/wleung-lecture-drosophila_annotation_primer/)” presentation is typically used in workshops where the participants work on the “Annotation of a *Drosophila* Gene” walkthrough during the same training session. The presentation uses the same gene as the walkthrough (i.e. *CG31997*) to illustrate the key steps of the annotation protocol. For example, the group attending the Summer 2022 ABLE meeting used a modified version of the “Annotation of *Drosophila* Primer” and the walkthrough in their presentation.