Last update: 02/25/2020

MODULE 4 HOMEWORK: SPLICING

MEG LAAKSO

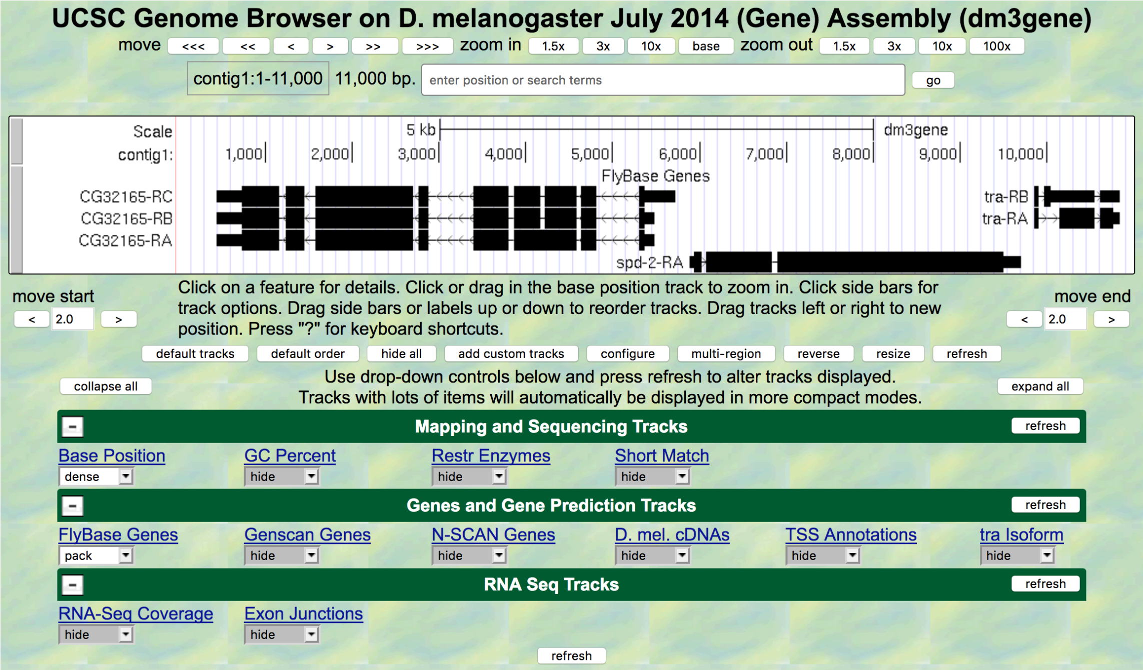
DETERMINING SPLICE SITES FOR THE *SPD-2* GENE

1. Open a web browser on your laptop. Internet Explorer, Mozilla Firefox, Safari, or Chrome will work for this investigation. Go to the GEP Mirror of the UCSC Genome Browser at http://gander.wustl.edu. Follow the instructions given in Module 1 to navigate to the contig1 project in the *D. melanogaster* 'July 2014 (Gene)' assembly.

Reminder: Configure the Genome Browser Gateway page as follows:

1. Select "***D. melanogaster***" under "REPRESENTED SPECIES"
2. Select "**July 2014 (Gene)**" under "*D. melanogaster* Assembly"
3. Enter "**contig1**" into the "Position/Search Term" text box
4. Click on the "**GO**" button

As you will remember, this section of DNA is 11,000 base pairs long and is part of the left arm of chromosome 3, which is about 28,100,000 bp long (Error! Reference source not found.). If your Browser window is showing other evidence tracks, reset by clicking on "default tracks."


UCSC Genome Browser on D. melanogaster July 2014 (Gene) Assembly (dm3gene)

contig1:1-11,000 11,000 bp.

Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. Press "?" for keyboard shortcuts.

Use drop-down controls below and press refresh to alter tracks displayed. Tracks with lots of items will automatically be displayed in more compact modes.

FIGURE 1 A SCREEN SHOT OF 'CONTIG1' PROJECT.

2. Let's start by setting the evidence tracks we want to see. Click on the "hide all" button. Then open only the tracks that will provide data for this investigation.

- Base position "**dense**"
  - Note that you will not be able to see the DNA sequence until you zoom in. If the base position is changed to "full", you can see the amino acid tracks also.

- RNA-Seq coverage "**full**"
  - You will see blue and red histograms representing RNA-Seq data generated using RNA samples from adult females and adult males, respectively.

- Exon Junctions "**full**"
  - Viewing the exon junctions will help you to find the splice donor and splice acceptor sites.

- D. mel. cDNAs "**full**"
  - This track was used extensively in previous modules, and is useful for confirming the sequence of the mature mRNA. Remember that a cDNA is a DNA copy of an mRNA.

3. Zoom in to view only the *spd-2* gene by entering "**contig1:5,750-9,800**" in the "enter position or search terms" text box.

We are now looking at the region of chromosome 3 where the *spd-2* gene is located.

Summary Question 1. How many exons does *spd-2* have? __________

Summary Question 2. How many introns does *spd-2* have? __________

HOMEWORK PART 1: IDENTIFYING SPLICE SITES FOR INTRON 1

You remember from class that short sequences are present at the beginning and end of each intron that allow the spliceosome to precisely remove each intron, leaving only the exon sequences in the mature mRNA. The first two nucleotides of the intron are "GT" and the last two nucleotides are "AG" (**Figure 2**).

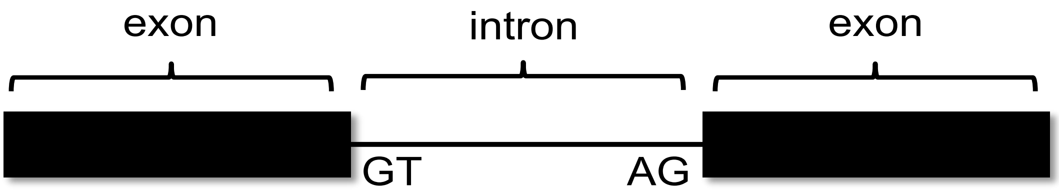


FIGURE 2 A BOWTIE DIAGRAM

4. Zoom in to the end of exon 1. Set the screen so that you can see about 15-20 nucleotides.

Summary Question 3. What are the last three nucleotides of exon 1? __________ What is the coordinate of the last nucleotide in exon 1? ____________________

Summary Question 4. What are the first two nucleotides of intron 1?

____

5. Zoom out so that you can see all of intron 1, and use TopHat to help find the end of the intron. Examine the Exon Junctions track. Then zoom in so that you can see the nucleotide sequence.

Summary Question 5. What are the last two nucleotides of intron 1?

____

Summary Question 6. What is the coordinate of the first nucleotide in exon 2?

____

HOMEWORK PART 2: IDENTIFYING THE 5' SPLICE DONOR AND 3' SPLICE ACCEPTOR SITES FOR INTRON 2

Let's use the same approach to map the exon-intron boundaries for intron 2.

Review the steps you used in Part 1, then repeat the process to answer the following questions about the 5' splice donor and 3' splice acceptor sites for intron 2.

6. Zoom in to the sequence surrounding the 5' splice donor for intron 2.

Summary Question 7. What are the last three nucleotides of exon 2?

____

Summary Question 8. What is the coordinate of the last nucleotide in exon 2?

____

Summary Question 9. What are the first two nucleotides of intron 2?

____

7. Zoom out as needed so that you can see all of intron 2. Use TopHat to find the end of intron 2.
8. Click and drag so that the end of intron 2 is centered in the viewer. Then zoom in so that you can see the nucleotide sequence.

Summary Question 10. What are the last two nucleotides of intron 2?

____

Summary Question 11. What is the coordinate of the first nucleotide in exon 3?

____

Summary Question 12. Using the information you've gathered so far, make a graphical picture of the *spd-2* gene with 3 exons and 2 introns. Number each exon and intron. Add the coordinates for first and last nucleotide of the exons that you have found so far. Add the sequences of the splice donor and splice acceptor sites at the appropriate locations. Finally, add a bent arrow for the transcription start site.