

# MODULE 6: ALTERNATIVE SPLICING

LEOCADIA PALIULIS

**Lesson Plan:**

<b>Title</b>	<b>Alternative Splicing</b>
<b>Objectives</b>	<ul style="list-style-type: none"> <li>• Demonstrate how alternative splicing of a gene can lead to different mRNAs.</li> <li>• Show how alternative splicing can lead to the production of different polypeptides and result in drastic changes in phenotype.</li> </ul>
<b>Pre-requisites</b>	Modules 1-5
<b>Order</b>	<ul style="list-style-type: none"> <li>• Introduce students to tra-RB, a second isoform of <i>tra</i>.</li> <li>• Investigation 1</li> <li>• Investigation 2</li> <li>• Discussion</li> </ul>
<b>Homework</b>	None included. Students could analyze a second gene on the browser using the work they have done on <i>tra</i> as a template.
<b>Class Instruction</b>	<ul style="list-style-type: none"> <li>• Introduce tra-RB</li> <li>• Discuss differences between tra-RB and tra-RA. Reinforce concept of isoform.</li> <li>• Investigation 1: How can there be different mRNAs encoded in the same gene?</li> <li>• Investigation 2: Examine the <i>tra</i> polypeptides by looking at the three possible reading frames. Review concept of reading frame and introduce phase if not previously introduced. Students will construct a gene model for tra-RB, using sequence information, and RNA-Seq data as evidence.</li> <li>• Discussion of gene models/wrap-up</li> </ul>
<b>Associated videos</b>	<ul style="list-style-type: none"> <li>• Genes and Isoforms Video: <a href="https://youtu.be/8jtTp_6vN4M">https://youtu.be/8jtTp_6vN4M</a></li> <li>• Splicing and Phase Video: <a href="https://youtu.be/JsvUfHy3eHE">https://youtu.be/JsvUfHy3eHE</a></li> <li>• RNA Seq and TopHat Video: <a href="https://youtu.be/qepVXEsfLMM">https://youtu.be/qepVXEsfLMM</a></li> </ul>

## INVESTIGATION 1.

In this investigation, we will focus on tra-RB, the second isoform of the *tra* gene, and will explore how multiple different mRNAs and polypeptides can be encoded by the same gene. The story of tra-RB is an exciting story of sex, alternative splicing, and poison exons!

1. To begin, open a web browser.
2. Go to the UCSC Genome Browser Mirror site at <http://gander.wustl.edu/> and follow the instructions given in Module 1 to open contig1 of *Drosophila melanogaster*, using the **July 2014 (Gene)** assembly rather than the **Aug. 2014 (BDGP Release 6 + ISO1 MT/dm6)** assembly. Once you are on the Genome Browser page, set “Base Position” (under the “Mapping and Sequencing Tracks” bar) to “full” so that you will be able see the three possible reading frames (remember that you will not see individual bases or amino acids until you zoom in, though). Also set “FlyBase Genes” (under the “Genes and Gene Predictions” bar) to **full**. Don’t forget to click on one of the “refresh” buttons to see your changes.
3. Enter the following coordinates into the ‘enter position or search terms’ text box: **contig1:9,700-11,000** and hit the ‘go’ button to get a good view of the *tra* gene (Figure 1).

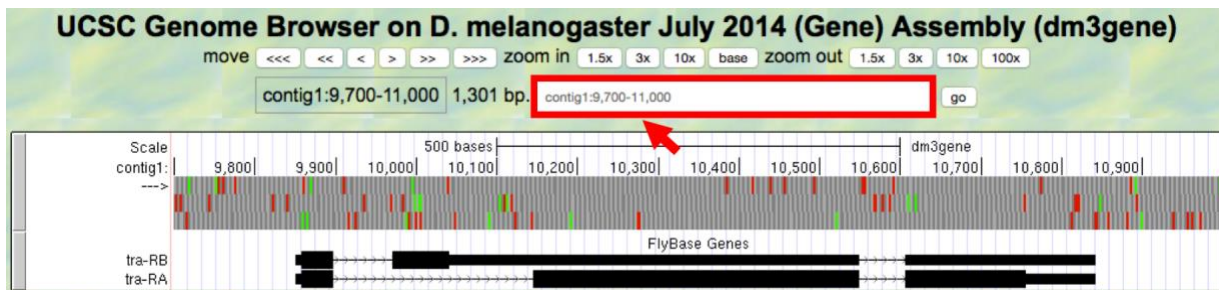


FIGURE 1: CENTER THE BROWSER ON THE *TRA* GENE.

4. Let’s consider what we know about tra-RA and learn more about tra-RB.

**Q1.** Given that exons are shown by the black boxes, and introns are shown by thin lines with arrowheads in the FlyBase Genes track, what does this tell us about the first intron of *tra-RB* compared to that of *tra-RA*?

5. Now let’s look at the patterns of transcription. Scroll down to ‘**RNA Seq Tracks,**’ click on the “RNA-Seq Coverage” link. Change the track display settings as we did in Module 2 (“Display mode” field on **full**; “Data view scaling” field set to “**use vertical viewing range setting**”; “max” field under “Vertical viewing range” to **37**), but this time check both “Adult Females” and “Adult Males” under “List subtracks.” Hit the “Submit” button (Figure 2). Back on the browser main page, Under ‘**RNA seq tracks**’ → ‘**Exon Junctions**’ use the dropdown menu to select **full**,

then hit **refresh**. To review the use of RNA Seq data, watch the RNA Seq and TopHat video: <https://youtu.be/qepVXEsfLMM>

**RNA-Seq Coverage Track Settings**

## RNA-Seq Read Coverage [\(All RNA Seq Tracks\)](#)

Display mode:    [Reset to defaults](#)

Type of graph:

Track height:  pixels (range: 11 to 110)

Data view scaling:  Always include zero:

Vertical viewing range: min:  max:  (range: 1 to 250)

Transform function: Transform data points by:

Windowing function:  Smoothing window:  pixels

Negate values:

Draw y indicator lines: at y = 0.0:  at y =

[Graph configuration help](#)

List subtracks:  only selected/visible  all

- full  Adult Females modENCODE RNA-Seq from D. melanogaster Whole Adult Females [schema](#)
- full  Adult Males modENCODE RNA-Seq from D. melanogaster Whole Adult Males [schema](#)

FIGURE 2: ENTER SETTINGS FOR RNA SEQ TRACKS.

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

contig1:9,700-11,000 1,301 bp. enter position or search terms go

Scale: 500 bases dm3gene

contig1: 9,800| 9,900| 10,000| 10,100| 10,200| 10,300| 10,400| 10,500| 10,600| 10,700| 10,800| 10,900

tra-RB tra-RA

Adult Females 37

Adult Males 37

modENCODE RNA-Seq from D. melanogaster Whole Adult Females

modENCODE RNA-Seq from D. melanogaster Whole Adult Males

RNA-Seq Exon Junctions

FBs0000097936

FBs0000097616

FBs0000146317

move start  move end

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.

default tracks default order hide all add custom tracks configure reverse resize refresh

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.

collapse all expand all

**Mapping and Sequencing Tracks** refresh

Base Position GC Percent Restr Enzymes Short Match

full hide hide hide

**Genes and Gene Prediction Tracks** refresh

FlyBase Genes Genscan Genes N-SCAN Genes D. mel. cDNAs TSS Annotations tra Isoform

full hide hide hide hide hide

**RNA Seq Tracks** refresh

RNA-Seq Coverage Exon Junctions

full full

refresh

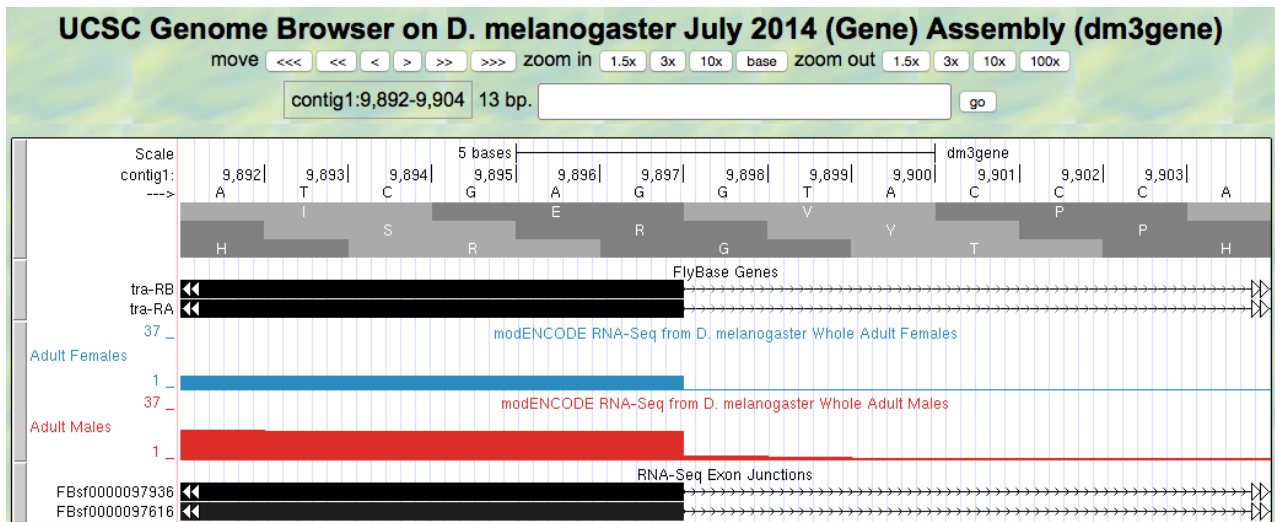
FIGURE 3: BROWSER SHOWING GENE PREDICTORS, RNA-SEQ TRACKS, AND RNA-SEQ EXON JUNCTIONS FOR MALE AND FEMALE *DROSOPHILA MELANOGASTER*.

Now we can see the RNA-Seq data for males (red) and females (blue). Recall that peaks in RNA-Seq Read Coverage tracks usually correspond to the regions of the genome that are being

transcribed. These two samples generally show similar RNA-Seq read coverage along the entire span of the *tra* gene. However, the adult female sample shows substantially lower RNA-Seq read coverage at around 9,971-10,145 (red box in Figure 3). We can also see the RNA-Seq Exon Junctions track, which shows the location of splice sites supported by the RNA-Seq data (as you saw in Module 4). Recall that the black boxes in the FlyBase Genes track are exons and the thin lines with arrowheads show the locations of the introns. Notice that the diagrams for the first and second RNA-Seq Exon Junctions tracks have the same 5' splice site but different 3' splice sites. Let's see what we can find out about these splice sites. First, we need to establish the reading frame for the first exon. Zoom in on the 5' end of the transcript around position 9850-9860.

**Q2.** Given what you know about the initiation of translation, which of the 3 possible reading frames is used for both the tra-RA and tra-RB products? \_\_\_\_\_

Now zoom in on the location of the 5' splice site at the end of the first exon in both tra-RA and tra-RB (Figure 4). We will also be thinking about the concept of phase here. To review splicing and phase, watch the Splicing and Phase video at: <https://youtu.be/JsvUfHy3eHE>



**FIGURE 4: ZOOM IN ON THE FIRST 5' SPLICE SITE.**

**Q3.** Give the coordinate for the last base of the first exon for tra-RA \_\_\_\_\_

**Q4.** Give the coordinate for the last base of the first exon for tra-RB \_\_\_\_\_

**Q5.** What is the consensus sequence for the 5' splice site (donor site) \_\_\_\_\_.

**Q6.** What are the coordinates for the 5' splice site in tra-RA? \_\_\_\_\_.

**Q7.** What are the coordinates for the 5' splice site in tra-RB? \_\_\_\_\_

**Q8.** What is the phase at this splice site? \_\_\_\_\_

6. Now zoom out and zoom in on the start of the second exon in *tra-RB*, just after the 3' splice site. We can identify the second exon by the RNA-Seq data, in particular using the RNA-Seq Exon Junctions data.

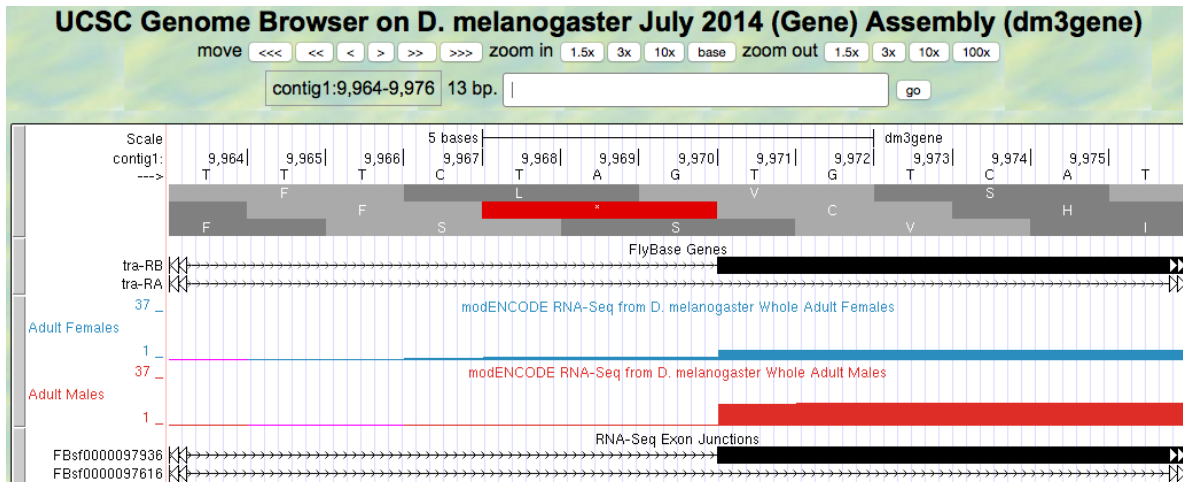


FIGURE 5: ZOOM IN ON START OF SECOND EXON IN *TRA-RB*.

**Q9.** What are the coordinates for the first base of the second exon in *tra-RB*? \_\_\_\_\_

**Q10.** What is the consensus sequence for the 3' splice site? \_\_\_\_\_

**Q11.** What are the coordinates for the 3' splice site in intron 1 of *tra-RB*? \_\_\_\_\_

**Q12.** What phase do we anticipate? \_\_\_\_\_

**Q13.** Given this, what is the reading frame for *tra-RB* exon2? \_\_\_\_\_

**Q14.** Does this make sense, given the location of stop codons? \_\_\_\_\_

7. Now zoom out and zoom in on the 3' splice site for *tra-RA*. This can be identified from the RNA-Seq data, particularly the RNA-Seq Exon Junctions.

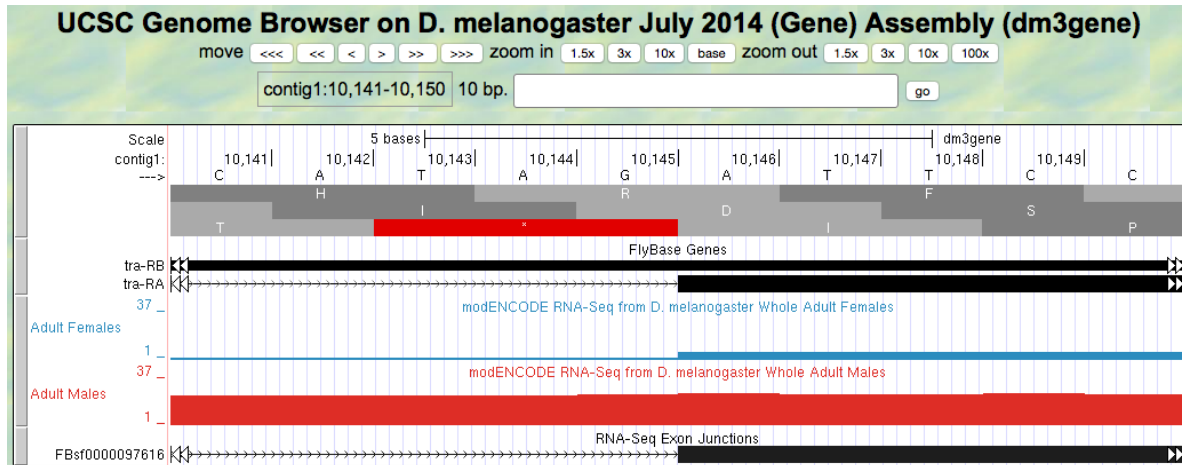


FIGURE 6: ZOOM IN ON START OF SECOND EXON FOR *TRA-RA*.

- Q15.** What are the coordinates for the first base of the second exon in *tra-RA*? \_\_\_\_\_
- Q16.** What is the consensus sequence for the 3' splice site? \_\_\_\_\_
- Q17.** What are the coordinates for that sequence in intron 1 of *tra-RA*? \_\_\_\_\_
- Q18.** Given the phase at the donor site, what phase are we looking for here? \_\_\_\_\_
- Q19.** Given this, what is the reading frame for *tra-RA* exon 2? \_\_\_\_\_
- Q20.** Does this make sense, given the location of stop codons? \_\_\_\_\_

The 3' acceptor site for the second intron in *tra-RA* is found inside the second exon of *tra-RB*. This intron is *alternatively spliced*. Alternative splicing is one way eukaryotes produce different proteins from the same coding regions of DNA. Here the alternative decision is made in a sex-specific manner; male fruit flies have targeted the spliceosome to use the first 3' acceptor site identified by the RNA-Seq Exon Junction data, while female fruit flies have targeted the spliceosome to use the second 3' acceptor site identified. This change in splicing has profound effects – in fact, it drives the programming of male and female characteristics in the developing fly. To review alternative splicing, watch the Genes and Isoforms video:

[https://youtu.be/8jtTp\\_6vN4M](https://youtu.be/8jtTp_6vN4M).

8. Reset your browser by entering **contig1:9,700-11,000** into the “enter position or search terms” text box and hit go. Let’s analyze the consequences of this alternative splicing on production of a protein product.

**Q21.** From your analysis of the RA isoform of *tra* in Module 5, how many amino acids does the *tra-RA* protein product have? \_\_\_\_\_

Now look at the tra-RB isoform:

**Q22.** Write down the coordinates for exon 1. \_\_\_\_\_

**Q23.** Given the reading frame that you established for tra-RB, does translation continue through exon 2, or is it terminated by a stop codon? \_\_\_\_\_

**Q24.** Write down the coordinates for the translated portion of exon 2. \_\_\_\_\_

**Q25.** How many amino acids does the protein translated from the tra-RB isoform have?  
\_\_\_\_\_

**Q26.** Is it likely that the protein translated from tra-RB could play the same functional role played by the protein translated from tra-RA? \_\_\_\_\_

The Tra protein has an important function in female *Drosophila*, and is itself a splicing factor that regulates splicing. Careful annotation of genes, as we have done here, can provide many insights into biological control mechanisms.

## INVESTIGATION 2. WHAT ARE THE CONSEQUENCES OF ALTERNATIVE SPLICING ON THE POLYPEPTIDES PRODUCED FROM EACH ISOFORM?

Now that we know that *tra* is alternatively spliced to make two isoforms, tra-RA and tra-RB, and that males express one isoform while females express the other, let's try to figure out how alternative splicing affects the polypeptides produced from translating these mRNAs. To do this, we need to produce a gene model for tra-RB and compare it to the gene model for tra-RA that you constructed in Module 5, showing where the start codons and stop codons appear in each isoform.

Use what you learned in Module 5 to construct a gene model for tra-RB. Locate the start codon, splice sites, and the stop codon. Construct the gene model below.

**Q27.** Coordinate for start of translation: \_\_\_\_\_

**Q28.** Coordinate for last base of exon 1: \_\_\_\_\_

**Q29.** Coordinate for first base of exon 2: \_\_\_\_\_

**Q30.** Coordinate for last base of exon 2: \_\_\_\_\_

**Q31.** Coordinate for first base of exon 3: \_\_\_\_\_

**Q32.** Stop codon coordinates: \_\_\_\_\_

### ***Points for discussion:***

How does the polypeptide translated from the tra-RB isoform differ from the polypeptide translated from the tra-RA isoform? What are the consequences of these differences on protein function?

Discuss how the bigger mRNA leads to creation of a smaller polypeptide!!

Consider how alternative splicing could allow many different proteins to be encoded by the same gene.

Based on the gene structure of the two isoforms of *tra* shown in the "FlyBase Genes" track, provide a hypothesis that could explain this difference in RNA-Seq read coverage between the adult males sample and adult females sample.