



Module 6: Alternative Splicing

Leocadia Paliulis

Objectives

- Demonstrate how alternative splicing of a gene can lead to different mRNAs.
- Show how alternative splicing can lead to the production of different polypeptides and result in drastic changes in phenotype.

Prerequisites

- [Understanding Eukaryotic Genes](#) Modules 1–5

Class Instruction

- Introduce tra-RB
- Discuss differences between tra-RB and tra-RA. Reinforce concept of isoform.
- Investigation 1: How can there be different mRNAs encoded in the same gene?
- Investigation 2: Examine the *tra* 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.
- Discussion of gene models/wrap-up

Associated Videos and Resources

- [Genes and Isoforms Video](#)
- [Splicing and Phase Video](#)
- [RNA-Seq and TopHat Video](#)
- [Glossary for Understanding Eukaryotic Genes](#)

Table of Contents

Investigation 1: 2

Investigation 2. What are the consequences of alternative splicing on the polypeptides produced from each isoform? 7

Points for discussion:..... 8

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 GEP UCSC Genome Browser Mirror site at <https://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 to 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 “**chromosome range, or search terms, see examples**” text box: **contig1:9,700-11,000** and hit the “go” button to get a good view of the *tra* gene (Figure 1).

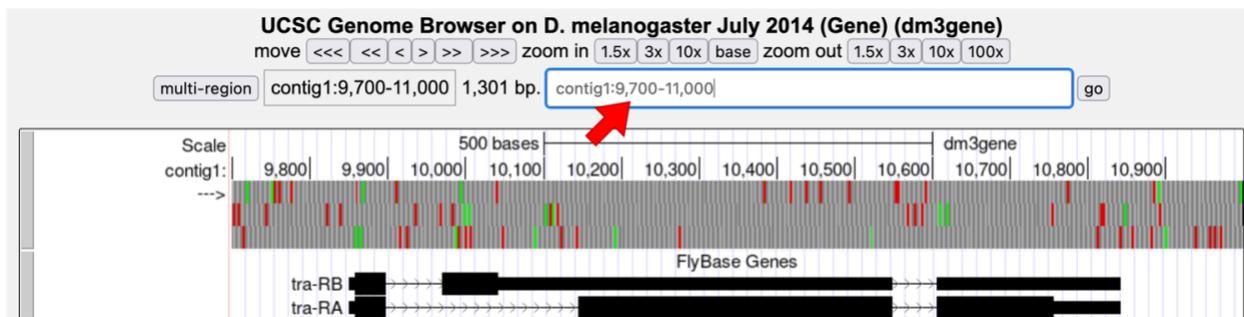


Figure 1 Center the Genome Browser on the *tra* gene.

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).

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

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

Figure 2 Enter settings for RNA-Seq tracks.

- Back on the Genome Browser main page, scroll down to “RNA Seq Tracks”, change the display settings for the “Exon Junctions” track to full, and then hit “refresh”. To review the use of RNA Seq data, watch the [RNA-Seq and TopHat video](#).

The screenshot shows the UCSC Genome Browser interface for *D. melanogaster*. The main track displays FlyBase genes (tra-RB, tra-RA) and RNA-Seq read coverage for Adult Females (blue) and Adult Males (red). A red box highlights a region of high coverage in the female track. Below the tracks, the 'RNA Seq Tracks' section is expanded, showing 'RNA-Seq Coverage' and 'Exon Junctions' both set to 'full' display mode. The interface includes navigation tools like 'move start', 'move end', and 'zoom in/out', as well as track management options like 'collapse all' and 'expand all'.

Figure 3 Genome Browser showing FlyBase gene annotations, RNA-Seq read coverage, 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). The splice sites are located in the intron just outside of the splice junction boxes (Figure 4). Notice that the diagrams for the first and second RNA-Seq Exon Junctions 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](#).

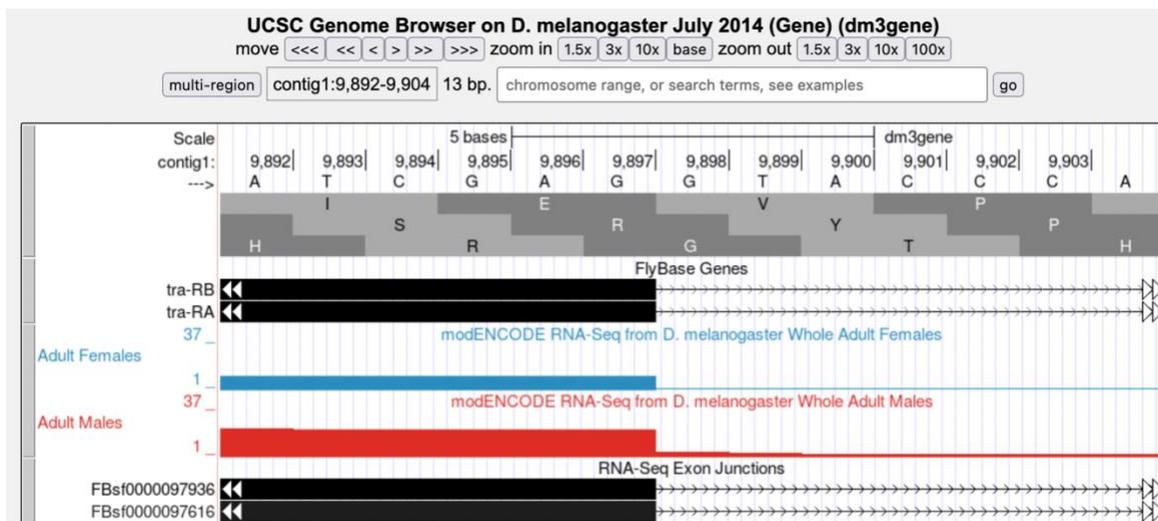


Figure 4 Zoom in on the first 5' splice site.

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? _____

- 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).

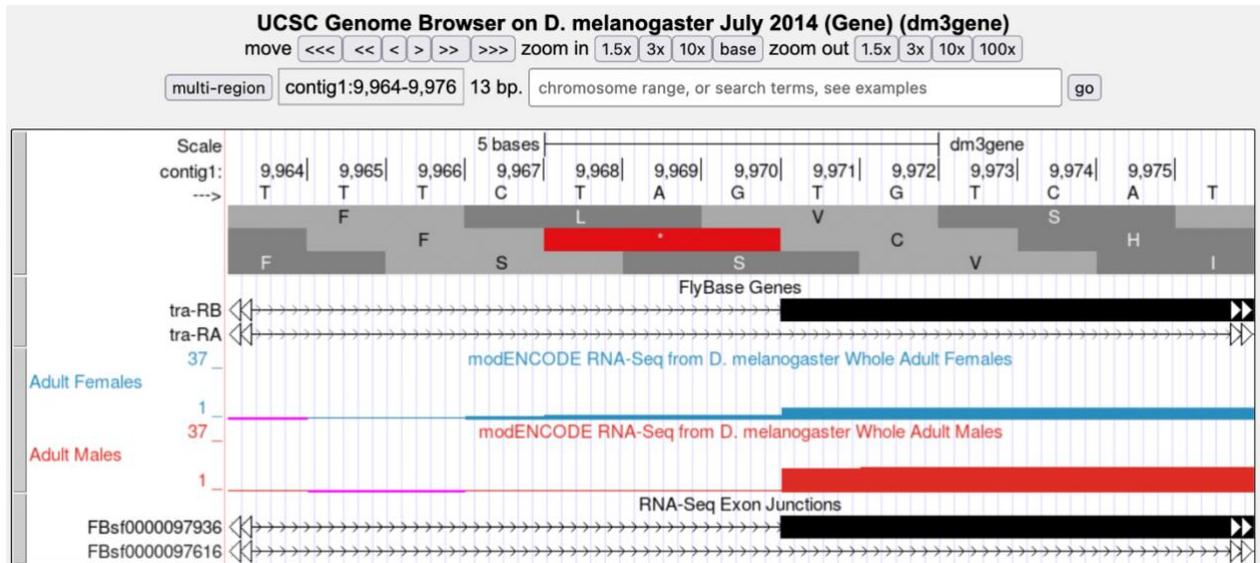


Figure 5 Zoom in on start of second exon in tra-RB.

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? _____

- 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).

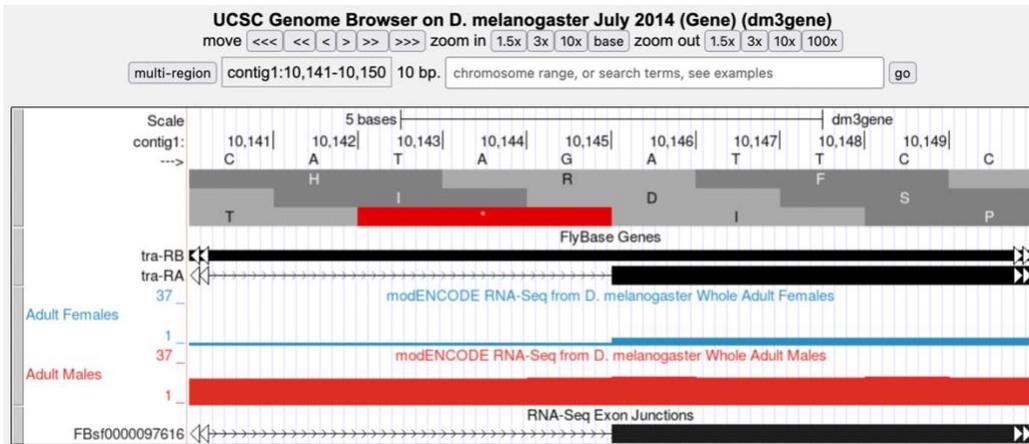


Figure 6 Zoom in on start of second exon for tra-RA.

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 first 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](#).

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

Q21. Look back at Module 5, Q17. Using the coordinates you’ve written in your answer to that question, calculate the number of amino acids in the protein translated from tra-RA. Do this by calculating the size of each exon, added the sizes of all exons together, and dividing by 3 (the number of bases in a codon). _____

Now look at the tra-RB isoform:

Q22. Write down the coordinates for exon 1. _____

Q23. Looking at exon 2 of tra-RB, is the entire second exon translated or is there a stop codon within the exon? _____

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.