SX1238 *tapas* Gene Insertion

Student William Wang
Advisor Nicole Lantz and Elizabeth Fox
Course Research in Molecular Genetics
Date February 28, 2024

Figures

Figure 1: **Adult SX1238 flies.** 1A (left) and 1B (right) show a male and female fly. The SX4 insertion is homozygous lethal. The flies have genotypes of *w-;SX4/Cyo;TM6B,Tb,Hu,e/e* and *w-/Y;SX4/Cyo;TM6B,Tb,Hu,e/e*. The flies have phenotypes of red eye, curly wings, humerals, and ebony body. Red eyes are a result of the *SX4 w+* dominant marker.

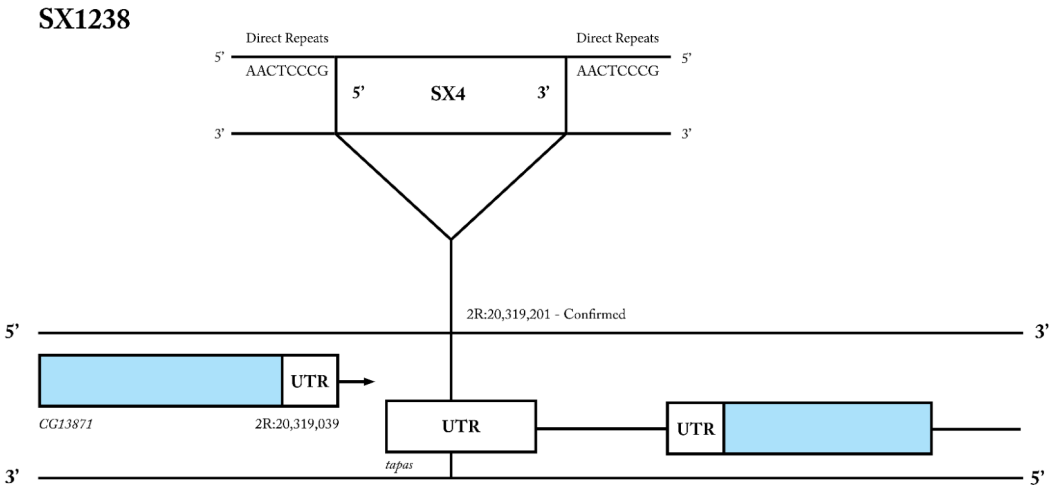

Figure 2: **SX1238 *SX4* insertion site**. The P element is inserted 5' to 3' at the insertion site 2R:20,319,201. The insertion site is within the 1st UTR of the *tapas* gene and within 200 bases upstream from the *CG13871* gene (Grammates et.al, 2022).

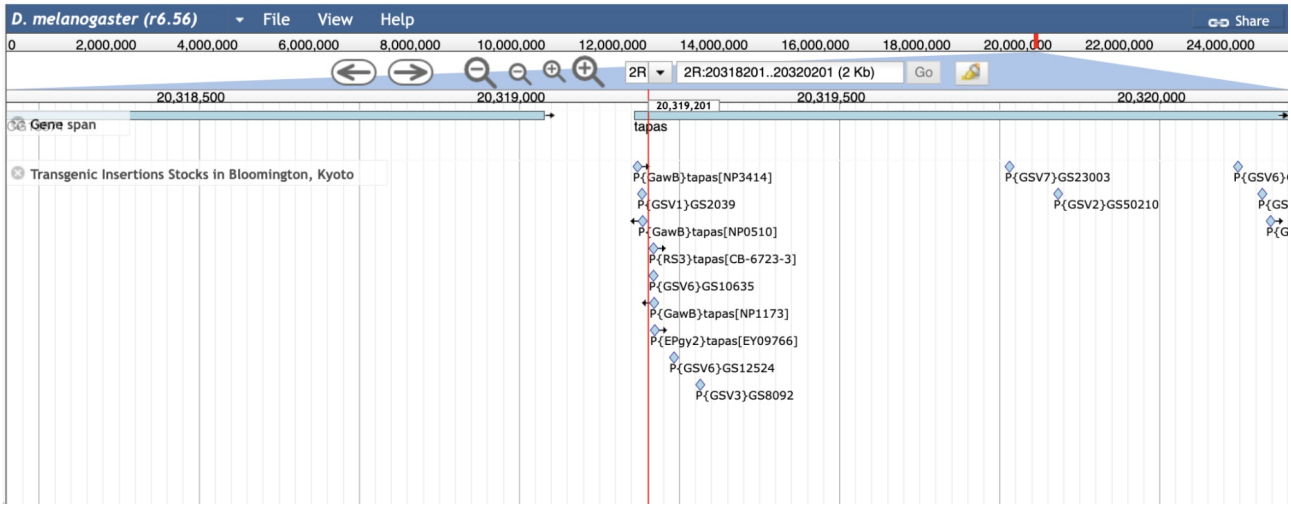

Figure 3: **Transgenic insertions close to SX1238 insertion site.** There are 9 transgenic insertions in close proximity to the SX1238 insertion site and 5 more within 1,000 bases upstream and downstream of the *SX4* insertion location (Grammates et.al, 2022). The thick red line represents the SX1238 insertion site of 2R: 20,319,201.

Transgenic Insertions

While there are no LexA transgenic insertions within 1,000 bases upstream or downstream of the insertion site, there are 9 non-LexA insertions around the insertion site and 5 more within 1,000 bases upstream and downstream of the insertion site. This novel insertion allows researchers to study an entirely new portion of the fly genome using a LexA binary expression system which can be paired with a LexAop reporter and other binary expression systems to study tissue interaction in this location.

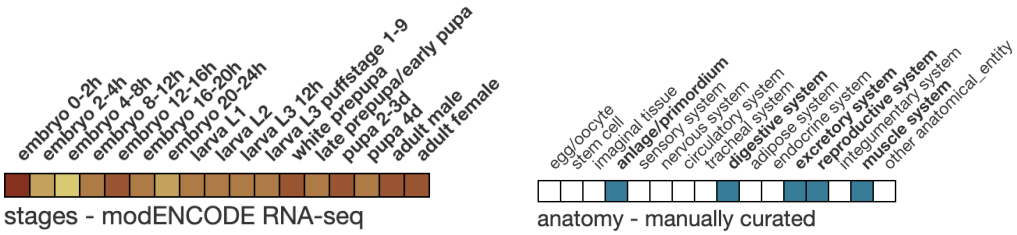

Figure 4: ***tapas* expression data.** Figure 4A (left) outlines the expression data for *tapas* within the Drosophila life cycle, with darker cells representing higher expression. In the L3 stage there is moderate expression of *tapas.* Figure 4B (right) indicates the expression data (RNA/protein) within each anatomical location. Colorless tiles represent when there is no data. RNA or protein expression of *tapas* occurs in the anlage/primordium, digestive system, excretory system, reproductive system, and the muscle system. There was no data in the FlyAtlas Anatomy Microarray (Grammates et.al, 2022).

Predicted Expression

tapas is expressed in the adult brain, embryonic/larval muscle system, female reproductive structures, germ layer derivative, and gut section. tapas and its role in preventing retrotransposition is likely not limited to only functioning in such areas but is likely most often occurring in such areas (all of which are critical to germline cells and their development.) When the *LexA* driver trapped by an enhancer associated with *tapas* is crossed with *LexAopGFP*, fluorescence will occur at the locations of highest expression at the larval stage. This expression will be centered around the adult brain, embryonic/larval muscle system, female reproductive structures, germ layer derivative, and gut section.

Gene

SX1238 was inserted within the first UTR of *tapas*, a protein coding gene. *tapas* encodes for proteins responsible for silencing retrotransposons, a transposon utilizing a copy-paste mechanism as opposed to a cut and paste mechanism for

duplication (Grammates et.al, 2022). *tapas* interacts with piwi-interacting (piRNAs), a special type of RNA which defends against transposable elements (Patil et al., 2014). tapas proteins are located in the P granule and perinuclear region of the cytoplasm, closely located to the nucleus of the cell, which interact with proteins involved in the piRNA retrotransposon silencing (Patil et al., 2014). The *tapas* gene is a paralog with the Tudor domain protein tejas which work together to maintain necessary piRNA components and production within Drosophila germline cells (Patil et al., 2014). tapas proteins work against the proliferation of retrotransposon by destabilizing its mRNA, making the RNA more susceptible to degradation (MRNA *Destabilization*, n.d.) Additionally *tapas* is involved in secondary processing, an amplification loop which generates additional piRNA (Fu & Wang, 2014).

The gene tapas is an ortholog to ten human genes including *TDRD7, TDRD6, TDRD1, TDRD15* and a number of other genes (Grammates et.al, 2022). *tapas* is not known to have any human disease association.

Genetics

- Genotype: *w-;SX4/Cyo;TM6B,Tb,Hu,e/e* and *w-/Y;SX4/Cyo;TM6B,Tb,Hu,e/e*
- Phenotype: red eyes, curly wings, tubby, humeral, ebony
- No non-curly flies were observed. This implies that the *SX4* insertion location is homozygous lethal.

Our insertion site is 2R: 20,319,201. The insertion site is in the first UTR of *tapas* and 200 bases downstream from *CG13871*. The fact that the SX4 insertion location is homozygous lethal implies that the insertion location overlaps with a piece of the genome necessary for fly survival, which is currently unknown since the insertion location within the first UTR of *tapas* does not disrupt the expression of *tapas*. The disruption as a result of SX4's insertion influences a piece of genomic DNA, resulting in an inability for flies to survive. The insertion location could overlap with a gene not yet discovered, an enhancer, RNA, or another element important for biological function currently unknown.

Product

By studying the *Drosophila* genome, which shares many similar biological processes to the human genome, researchers can study human diseases. Specifically, by creating flies with identified *SX4* P element insertion locations and *LexA* drivers, we can manipulate the fly with corresponding reporters through binary expression systems that alter tissue function where the trapped enhancer is active. Since flies share similar biological processes and genes to humans through human orthologs, studying *Drosophila* informs understanding surrounding human genetics and the development of medicine. Specifically, dual binary expression systems such as LexA/LexAop and Gal4-UAS can be used together to study multiple tissues simultaneously or different effects of the same tissue simultaneously. Each binary expression system does not affect the other since the transactivator LexA and Gal4 don't affect each other's reporters.

For example, SX1238's insertion in the first intron of *tapas* could create a LexA driver for *tapas*. Researchers using reporters such as reaper could investigate the effect of the elimination of the cells in which *tapas* is expressed and its influences on fly biological function. A dual binary expression system could be used with Gal4-UAS expressing a reporter gene shRNA connected with *tejas* which reduces the expression of *tejas*. Thus, researchers can investigate the effects of the elimination of the cells which express *tapas* and the reduced expression of *tejas* and its influences on a fly's biological function.

Process

Once an insertion is created and maintained in a stable stock, inverse PCR and DNA sequencing is used to locate the insertion site. DNA extraction is the process used to purify the fly's genomic DNA, separating the genomic DNA from other macromolecules such as lipids, proteins, carbohydrates, and RNA through DNA purification. Digestion then takes place through the use of the restriction enzyme HpaII with the restriction site CCGG, creating 3 sticky cuts, one to the left, one to the right, and another within the P element DNA. Ligation occurs using the enzyme ligase to self-ligate (form of loop) out of the 2 sections of DNA after digestion, forming a circular piece of DNA. The circular piece of DNA contains the genomic DNA, a restriction site, and either a 5' or 3' end of the P element. Inverse PCR amplifies the genomic DNA near the insertion site through 2 PCR primers, one extending towards the restriction site and the other towards the end of either the 3' or 5' end of the P element. Chris is the 5' PCR primer with sequence GCA CAC AAC CTT TCC TCT CAA C. Dove is the 3' PCR primer with sequence CCA CGG ACA TGC TAA GGG TTA A. A sequencing primer (Sp1 for the 5' end and Ulf for the 3' end) is placed in front of the PCR primer towards the end of the P element. Through a 3-step process (denaturation, primer annealing, and extension), the genomic DNA surrounding the P element is amplified. Sanger Sequencing is used to determine the sequence of genomic DNA through PCR and gel electrophoresis; sequencing produces two FASTA files (one from Sp1 and Ulf) associated with the genomic DNA adjacent to *SX4*.

SX1238 had high-quality data for Ulf sequencing, including an identifiable TTTCATCATG and CCGG restriction site with a distinct confirmed insertion site 2R: 20,319,201. *SX1238* also had high-quality data for Sp1 sequencing, including an identifiable TTTCATCATG and CCGG restriction site with the same insertion site 2R: 20,319,201. Since both the Sp1 and Ulf sequencing results produced acceptable data with the same direct repeats and insertion sites, it is not necessary to design another PCR primer to confirm the insertion site.

References

- Fu, Q., & Wang, J. (2014, February 7). *Mammalian piRNAs*. National Library of Medicine: National Center for Biotechnology Information. Retrieved February 28, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4114582/#: :text=While
- *mRNA Destabilization*. (n.d.). Mouse Genome Informatics. Retrieved February 28, 2024, from https://www.informatics.jax.org/vocab/gene_ontology/GO:0061157#: :text=Definition
- Patil, V. S., Anand, A., Chakrabarti, A., & Kai, T. (2014, October 6). *The Tudor domain protein Tapas, a homolog of the vertebrate Tdrd7, functions in the piRNA pathway to regulate retrotransposons in germline of Drosophila melanogaster*. National Library of Medicine National Center for Biotechnology Information. Retrieved February 28, 2024, from https://pubmed.ncbi.nlm.nih.gov/25287931/