Transcriptional regulation of *Drosophila melanogaster* courtship behavior by *fruitless* and *dissatisfaction*

Aron Gyuris

B.A. Brown University, 2004

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Division of Biology and Medicine at Brown University

Providence, Rhode Island

May 2014

This dissertation by Aron Gyuris is accepted in its present form by the Division of Biology and Medicine as satisfying the dissertation requirement for the degree of Doctor of Philosophy

Date_____ _____

Advisor, Dr. Michael McKeown

Recommended to the Graduate Council

Date_____ _____

Reader, Dr. Robert Reenan

Date_____ _____

Reader, Dr. Richard Freiman

Date_____ _____

Reader, Dr. Gilad Barnea

Date_____ _____

Reader, Dr. Michael Marr

Approved by the Graduate Council

Date_____ _____

Dr. Peter Weber
Dean of the Graduate School

Aron Gyuris

Brown University
Department Molecular Biology, Cell Biology, Biochemistry
Box G
Providence, RI 02912

(443) 676-5088
Aron_Gyuris@brown.edu

Education

Brown University, Providence, RI Fall 2007 – Present
Molecular Biology, Cell Biology, and Biochemistry
Advisor: Dr. Michael McKeown
Thesis: Transcriptional Regulation of *Drosophila melanogaster* Courtship Behavior by *fruitless* and *dissatisfaction*

Brown University, Providence, RI Fall 2000 – Spring 2004
B.A. Biology

Work

Brown University, Providence, RI Spring 2004 – Fall 2007
Research Assistant
Laboratory of Dr. Richard Freiman

- Maintained and genotyped transgenic mouse lines
- Conducted experiments examining the mutant phenotypes of MED26 (CRSP70), Brd2 and TAF4b

Teaching Experience

Brown University, Providence, RI Fall 2008
Teaching Assistant for Bio 0470 – Genetics

- Organized and implemented a weekly laboratory exercise for an undergraduate genetics course
- Graded lab reports and assisted in grading exams

Publications

Voronina E, Lovasco LA, **Gyuris A**, Baumgartner RA, Parlow AF, Freiman RN. (2007). Ovarian granulosa cell survival and proliferation requires the gonad-selective TFIID subunit TAF4b. Developmental Biology 303(2): 715-26.

Gyuris A, Donovan DJ, Seymour KA, Lovasco LA, Smilowitz NR, Halperin AL, Klysik JE, Freiman RN. (2009). The chromatin-targeting protein Brd2 is required for neural tube closure and embryogenesis. Biochim Biophys Acta. May;1789(5):413-21.

Table of Contents

Chapter 1: Introduction

Innate behaviors offer a unique area of study to examine the relationship between the input and output of these actions, the neural architecture that underlies them and the molecular mechanisms that establish the neurons and circuits from which they are effected. Courtship in *Drosophila melanogaster* is a robust, tractable behavior that is present in animals reared in isolation (Greenspan, 2000). This indicates that at least initially, this behavior is intrinsic to the properly formed male and female fly nervous systems. Furthermore, the neural basis for this behavior has been studied extensively (Yu, 2010). Finally, it is known that courtship is genetically determined by a well-defined developmental gene hierarchy (Billeter, 2006a). Together, the well-elucidated phenomenon of *Drosophila* courtship provides a solid foundation of genetic pathways and neuronal structures responsible for a distinct series of behaviors upon which functional studies of the path of genes to nervous system to behavior can be built.

Drosophila melanogaster **Courtship Behavior**

Courtship in *Drosophila melanogaster* is characterized by a conserved sequence of behaviors performed by the males and mate selection through receptivity by the females. When presented with an attractive courtship target (normally a *Drosophila melanogaster* female), males progress through a linear series of actions (Figure 1A) starting with orienting towards and following the target. Once sufficient proximity is established, the male taps the abdomen of the target with the forelegs, thereby sampling the cuticular hydrocarbon pheromone profile of the target. This contact chemoreception is used to determine the sex and species of the intended target (Johansson, 2007). The correct cues drive subsequent behaviors by the male, starting with unilateral wing

extension and the singing of the courtship song. Again, the purpose of this behavior is proper mate selection, this time with female receptivity affected by courtship song in a species-specific fashion (Tauber, 2003). If the male sings well enough, the female will stop and allow the male to lick her genitals. Here, gustatory cues mediate mate selection (Tompkins, 1983). Finally, at the end of this series of events is successful copulation with the male mounting the female, bending his abdomen to make genital contact and transferring a mix of sperm and seminal fluids, some of which will alter the behavior of the female post copulation (Billeter, 2006a). Interestingly, if courtship behavior is interrupted, the male will start at the beginning upon resumption of courtship (Hall, 1994), indicating that at least the linear progression of courtship behavior is a fixed action pattern (Manoli, 2006). As described, mate detection is mediated through the integration of multiple sensory modalities and no single sensory function is sufficient to confer the initiation or maintenance of *Drosophila melanogaster* courtship (Krstic, 2009).
Mate selection behavior by receptivity in females is also driven by sensory cues, predominantly olfactory (Gailey, 1986), visual (Markow, 1987) and auditory (Manning, 1967) stimuli.

Interestingly, the linear progression of courtship behavior maps to the *Drosophila* nervous system in an anterior to posterior fashion (Figure 1B) (Hall, 1979), in that earlier behaviors are localized more anteriorly and as the behaviors come in progression (except for licking), so do they map progressively more posteriorly, first in the brain and then in the ventral nerve cord (VNC). Briefly, following, tapping and wing extension behaviors are controlled by structures in the dorsal, posterior cortex of the supraoesophageal ganglion. Courtship song originates from the ventral thoracic ganglia. Licking is the

outlier in this progression, in that the focus of this behavior is localized to the dorsal brain. Though gustatory stimuli thought to underlie licking behavior are processed initially in the subesophageal ganglion (SOG), relaying these inputs to higher order processing centers (Scott, 2001; Vosshall, 2007) could explain the localization of this behavior. Lastly, attempted copulation and copulation behaviors originate in the abdominal ganglion (Billeter, 2006a), as this is the location of the motor neurons that innervate the abdominal muscles (Gailey, 1991; Finley, 1997) and the internal reproductive organs (Lee, 2001a; Lee, 2001b; Billeter, 2006b).

Female receptivity is much less well understood neuroanatomically, but research (Tompkins, 1983) indicates that this behavior depends on anterior dorso-medial protocerebrum, as one of likely many other structures being female.

Drosophila melanogaster **Sex Determination**

Sex specific differentiation in flies is determined by the X chromosome to autosome ratio (Schütt, 2000; Yamamoto, 2007). In females (2X:2A), the splicing factor *sex-lethal* (*Sxl*) is activated through alternative splicing in embryogenesis, just preceding the formation of the syncytial blastoderm. Alternatively, in males (1X:2A) *Sxl* is inactive owing to the inclusion of an exon containing multiple in-frame stop codons (Villella, 2008). In addition to directing dosage compensation, *Sxl* controls alternative splicing of *transformer* (*tra*) (Boggs, 1987; Bopp, 1991). Itself a splicing regulator, *tra*, in concert with the constitutive *tra2*, drives the female specific alternative splicing of the two major effectors of sexual differentiation, *fruitless* (*fru*; see section below) and *doublesex* (*dsx*). As *tra* is inactive in males, *fru* and *dsx* exhibit the male specific default splice forms.

Manipulating *tra* results in male body morphology and behavior in chromosomally XX animals as a result of splicing to the male forms of *fru* (Ryner, 1996) and *dsx* (Ryner, 1991), cementing the role of *tra* at the top of the *Drosophila* sex determination hierarchy (Figure 2A).

doublesex

doublesex (dsx) functions to pattern the soma in a sex specific fashion. It functions as a sexually dimorphic DNA binding transcription factor (Erdman, 1993; Erdman, 1996) that, under the control of *tra* alternative splicing, manifests as a male (*dsx*-M) or female (*dsx*-F) form in the respective sex (Burtis, 1989). Morphologically, reversal of the isotypes of *dsx* relative to chromosomal sex leads to a complete reversal of cuticular sex characteristics in both sexes (Rideout, 2010). Animals lacking *dsx* altogether exhibit intersexual body morphology (Baker, 1980) confirming that *dsx* is required to specify divergence in the bi-potential pathway of somatic sex determination. In the nervous system, the state of *dsx* drives sexual dimorphism in both neuronal cell number and projection patterns (Sanders, 2008; Rideout, 2010). Accordingly, *dsx* mutants display abnormal courtship behaviors (Villella, 1996), specifically in the courtship song (Taylor, 1994). Transcriptionally, *dsx* belongs to the Dmrt (*doublesex* and *mab-3*-related transcription factor) family, members of which are conserved in their function in sex determination in animals (Zarkower, 2002). The sex splicing of *dsx* directs alternative termini, responsible for mediating dimerization, to the carboxy end of the protein conferring distinct functional differences to these proteins (Figure 2B) (An, 1996). At the level of gene regulation, *dsx*-M and *dsx*-F bind the same consensus

sequence (Erdman, 1996), notably at the *yolk-protein* promoter where they either repress (M) or activate (F; in tandem with its cofactor *intersex*) transcription (Garabedian, 1985; An, 1995). Despite its prominent role in the sex determination of *Drosophila melanogaster*, *dsx* was not studied in this work, but drawing parallels to its function as a transcriptional regulator of sex determination could aid in understanding these functions in other proteins with similar functions.

fruitless

fruitless (*fru*) is a complex gene. Transcription of *fru* is initiated from one of four promoters, named P1-P4, with P2-P4 producing only sexually monomorphic transcripts (Goodwin, 2000). These RNAs then splice to a 5 exon cassette that comprises the common body (Comm) shared by all *fru* transcripts. Alternative splicing from the *fru*-Comm exons to one of four 3' terminal ends (A, B, C or D) completes the variation seen in this gene (Figure 3A, Figure 4A). Transcripts of all four alternate 3' isoforms are produced (Ryner, 1996; Billeter, 2006b).

Ablation of the entire locus leads to partially penetrant lethality characterized by neurodevelopmental defects (Anand, 2001; Song, 2002), while disruption of P1 alone results in viable animals with overt courtship phenotypes. These findings describe a poorly studied, sex-independent function for the gene products of P2-P4. Substantially more attention has been focused on the expression, splicing and function of the sex-specific P1 promoter. Transcripts originating from P1 are spliced under the control of *tra* to either a female (*fru*-F) form in its presence or a default male (*fru*-M) in its absence (Ryner, 1996; Ito, 1996). In females, inclusion of an additional exon that introduces a

premature stop codon represses *fru-F* protein at the level of translation (Usui-Aoki, 2000). With respect to the role of *fru* in establishing function in the neural circuit responsible for courtship behavior, the female circuit (*fru* OFF) can be thought of as the default state. Conversely, in males, *fru* P1 protein is expressed with the inclusion of a male specific 101 amino acids N-terminal to the Comm region (Lee, 2000). Therefore, a combination of the four P1 transcribed *fru* isoforms is largely responsible for establishing male courtship behavior.

fru was initially observed in a mutagenesis screen for courtship behavior (Gill, 1963). *fru*-null males are largely absent of courtship behavior, with the exception of rare homosexual chaining behavior, where in a group of males each male is courting and being courted in a line (Ito, 1996). Copulation success in *fru*-null males is limited by an abdominal bending defect (Hall, 1978). Hypomorphic *fru* mutant males exhibit courtship at sub-optimal levels and disrupted at different steps according to the allelic combination tested (Anand, 2001; Lee, 2001b). Furthermore, *fru* mutant males that court do so indiscriminately with respect to the sex of the target. Acute functional disruption of synaptic transmission in the P1 circuit fully abrogates courtship behavior (Manoli, 2005; Stockinger, 2005), firmly establishing not only *fru-M* function, but also the neuronal circuitry it establishes as necessary for courtship behavior. Conversely, introduction of *fru-M* into females is sufficient to confer at least partial courtship behavior (Demir, 2005) and with this ectopic function established, it becomes clear that *fru* acts as a switch for male courtship behavior. As expected from the absence of proteins from the P1 locus in the female, no overt sexual phenotypes are seen in *fru*-null females (Hall, 1979; Villella, 1997).

Predicted by their vital role, P2-P4 transcripts are expressed broadly in the CNS and some non-neuronal tissues in both sexes throughout development (Ryner, 1996; Goodwin, 2000). *fru*-M expression, however, is more restricted and superimposes onto the known courtship controlling centers of the fly brain (Billeter, 2006a) from anterior to posterior (Figure 1C, Figure 3B). Within this broad domain, *fru*-M is localized to about 1700 neurons (~2% of the nervous system) in 20 distinct clusters (Lee, 2000). In the anterior sensory inputs, *fru*-M appears in olfactory, gustatory and auditory receptor neurons (Stockinger, 2005). *fru* positive olfactory receptor neurons (ORN) mediate sensation of cis-vaccenyl acetate (cVA), a male cuticular pheromone that promotes male-male repulsion (Datta, 2008). Molecularly swapping Or67d, the olfactory receptor that identifies this circuit, for BmOR1, an olfactory receptor specific for the silkmoth, *Bombyx mori*, pheromone bombykol, is sufficient to render these neurons responsive to bombykol and, more importantly, effect repression of courtship behavior through the *fru* circuit (Kurtovic, 2007). ORNs expressing *fru-M* project to sexually-dimorphic (DA1) glomeruli (Kondoh, 2003) in the antennal lobe (AL). Furthermore, *fru*-positive projection neurons from DA1 themselves innervate the protocerebrum in a sexually dimorphic fashion (Datta, 2008). Gustatory receptor neurons (GRN) expressing *fru* are found in taste sensilla in the dorsal labellum and the foreleg (Stockinger, 2005; Goldman, 2007) and project to the SOG where gustatory inputs are organized (Yamamoto, 2008). Elucidation of the mechanism by which *fru* regulates sex-specific gustatory sensory information remains to be seen, but it is known that *fru*-M is required for a male-specific contralateral projection pattern of GRNs in the VNC (Mellert, 2010). Auditory neurons

expressing *fru* are found in the Johnston's organ in the antennae (Stockinger, 2005), a structure responsible for collecting auditory stimuli and mediating courtship song (von Schilcher, 1979) and other courtship auditory inputs (Fabre, 2012). In the higher order processing centers, *fru* expression features prominently in the mushroom bodies (MB) and superior medial and lateral protocerebrum, a region dubbed the lateral protocerebral complex (Yu, 2010) and postulated to integrate multiple sensory inputs and direct corresponding motor outputs.

Motor output of the *fru* circuit maps clusters in the thoracic ganglia (Rideout, 2007; Clyne, 2008) responsible for generation of the courtship song and abdominal ganglia (Lee, 2001a; Lee, 2001b; Billeter, 2006b) where the neurons that innervate the internal reproductive organs reside. The presence of the *fru* circuit throughout the neuronal regions previously implicated in courtship behavior strengthens the role of *fru* as a major factor in patterning the male nervous system for courtship.

Of particular interest in the *fru* neuronal circuitry is the sexually dimorphic muscle of Lawrence (mOL) found in the midline the 5th abdominal segment (A5) in males (Figure 3C) (Gailey, 1991). This muscle is formed through coalescence of small longitudinal muscles at A5 and is dependent on *fru-M* expression in the neuron that innervates them (Ito, 1996). Females normally manifest only the small intersegmental muscles, but expression of *fru-M* was sufficient to direct ectopic mOL formation in chromosomally XX animal (Usui-Aoki, 2000). This phenomenon further describes the requirement of *fru* in establishing the male-specific circuitry required for courtship behavior.

Temporally, *fru* is expressed at all of the stages of the life cycle and appears to peak throughout the midpupal restructuring of the nervous system (Lee, 2000). This is a phase in development where the structures and functions of the adult nervous system are patterned (Kimura, 2011). In support of this developmental role for *fru*, stage specific disruption of *fru* in the desert locust, *Schistocerca gregaria*, affected courtship behavior when altered in the 3rd and 4th nymphal stages (Boerjan, 2012). This affect was attenuated when *fru* was reduced in later stages.

The effects seen in *S. gregaria* highlight the broad evolutionary conservation of *fru* as a master regulator of courtship behavior with both sequence and functional homology observed (Figure 4B) across the *Drosophila* genus through to species as divergent as the German cockroach, *Blattella germanica* (Clynen, 2011), mosquitos, *Anopheles gambiae*, the honey bee, *Apis mellifera* and the red flour beetle, *Tribolium castaneum* (Gailey, 2006).

It was postulated that *fru* functioned as a DNA binding transcription factor through domain analysis placing it within the BTB-Zinc Finger (BTB-ZnF) family (Ito, 1996; Ryner, 1996; Yamamoto, 1996). The BTB is a highly conserved protein-protein dimerization domain found in zinc-finger containing DNA binding proteins (Albagli, 1995; Stogios, 2005) named for the archetypical members of the family: *Broad Complex* (*BR-C*), *tramtrack* (*ttk*) and *bric-á-brac* (*bab*). The BTB-Zn family is large gene family conserved across *C. elegans*, *Drosophila melanogaster* and humans (Zollman, 1994;). At least 58 predicted BTB-ZnF family members can be found in the *D. melanogaster* genome alone. Many of the members of this family have been implicated in effecting gene expression programs controlling various aspects of development. Additionally,

family members display gene complexity characterized by multiple promoters and alternative splicing of zinc-finger containing exons conferring differential DNA binding specificity in an isoform specific fashion. *BR-C* is a primary response gene under the control of ecdysone signaling (Emery, 1994) that mediates broad developmental gene programs in prepupal and pupal stages (Karim, 1993; von Kalm, 1994). Two *ttk* alternative zinc-finger isoforms control distinct gene expression programs in embryonic development (Read, 1992). Similarly, *bab* isoforms are broad regulators of developmental genes responsible for ovarian development, pigmentation of the cuticle and formation of the appendages (Couderc, 2002; Lours, 2003). A non-eponymous member of the BTB family, *abrubt* regulates genes in muscles involved in forming proper neuronal innervations (Hu, 1995). The homology of transcriptional function based on isoform specific gene regulation driving distinct developmental programs seen in the BTB-ZnF family holds great predictive power over the mechanism by which *fru* enacts establishment of the neuronal circuitry underlying courtship behavior.

Strangely, the hypothesis that *fru* is a DNA binding transcriptional regulator was not validated for 17 years. In the meantime, correlative evidence from the *yellow* locus showed that this gene required for male courtship behavior was downregulated in the absence of *fru* (Drapeau, 2003), but did not place *fru* at the *yellow* promoter. Biochemical interaction studies linked *bonus* (*bon*), a mammalian TIF1 homolog, as partners involved in differential recruitment of Histone deacetylase 1 or Heterochromatin 1 to effect sexually dimorphic neuronal arborization in the AL, foreleg sensilla and the midline crossing of GRNs in the VNC (Ito, 2012). Through bringing *fru* function closer to the promoter, these findings also did not establish DNA binding. Finally, work

published this month described direct *in vitro* DNA binding in an isoform-specific fashion (Dalton, 2013). *fru*-A, -B and -C each bind a distinct consensus sequence, present at an incompletely overlapping set of promoters. As such, they appear to regulate transcription of these downstream effector genes in partially overlapping pattern (Dalton, 2013). However, as expression from these genes was assayed by isoform-specific overexpression, not observed by direct promoter occupancy, it remains to be seen whether these are the genes directly under the control of *fru* or identified through artifact from overexpression.

In support of the argument that P1 specifies a spatial and temporal expression domain, not a novel biochemical function through inclusion of the M exon, is the observation that aberrant splicing around a gene trap Gal4 between the promoter and the common exons was sufficient to rescue some aspects of courtship behavior in a supposed null animal (Ferri, 2008). These "M-less" *fru* P1 transcripts were absent of the 5' male exon and were translated from the common exons on. Furthermore, while "M-less" females did not court, they did develop a mOL, normally only specified by the gene products of P1. That these "M-less" proteins could function somewhat normally is consistent with the functions predicted for the know domains and suggests that wild-type *fru-M* isoforms bind the same DNA sequences as all other *fru* proteins of that respective isotype, only specified to P1 target genes in space and time by their promoter. As no single *fru-M* isoform is sufficient to full rescue courtship behavior, although *fru*-C drives courtship to the highest degree (Billeter, 2006b), it is almost certain that individual isotypes regulate distinct sets of genes, all of which are needed to properly establish courtship behavior.

All together, *fru* is involved in an extensive neural circuit that governs all aspects of courtship behavior from initiation to completion.

dissatisfaction

dissatisfaction (*dsf*) was also identified in a genetic screen, this one for female sterile mutations (Schüpbach, 1991). Genetically, *dsf* functions as a regulator of both male and female courtship behavior.

dsf is an orphan nuclear hormone receptor (Finley, 1998) of the NR2e class (Figure 5A). Nuclear hormone receptors are a highly conserved superfamily of transcriptional regulators marked by a characteristic domain organization (Laudet, 1995; King-Jones, 2005). Within this broad superfamily, *dsf* exhibits significant homology to vertebrate Tailless (Tlx), especially in the DNA binding domain (DBD; 81% identity). Predictions based on the DNA binding activity of Tlx led to the elucidation of the *dsf* consensus binding sequence (Pittman, 2002). Similarity in the ligand binding domains (LBD) of Tlx and *dsf* (44% identity) is lower than that seen in the DBD, but is higher in this domain between these two proteins than between Tlx and any other protein outside of the Tailless family (Finley, 1998).

Significantly, there is greater conservation between the LBD of *dsf* and vertebrate Tlx than between the LBD of *dsf* and *Drosophila tailless* (*tll*) (33%). Importantly, divergence between *dsf* and *tll* is found in the functional elements of the DBD important for sequence recognition suggesting disparate roles for these two proteins. Conserved homology in the LBD could provide clues as to the ligand of *dsf*, but as Tailless is also an orphan nuclear hormone receptor, the lack of knowledge about its cognate ligand

precludes this approach. The hinge region of *dsf* shares little homology with Tlx (Finley, 1998), but does appear to be similar to the hinge of *dsf* in other species (see below). Functionally, *tll* has been implicated in embryonic patterning and nervous system development in *Drosophila* (Pignoni, 1990). Tlx also plays a role in the formation of neural structures, most significantly involving olfactory sensory processing, in the mouse (Monaghan, 1997; Roy, 2004) positing an interesting hypothesis for analogous roles for *dsf* in neural specification.

An interesting cofactor involving a putative *dsf* ligand is *takeout*, a gene involved in transport of lipophilic ligands from the fat body surrounding the brain (Dauwalder, 2002; Fujii, 2002). *takeout* is itself under the control of the sex determination hierarchy and is induced by the male forms of *dsx* and *fru*. If *takeout* is the carrier for a *dsf* ligand, this observation offers a potential mechanism for a sexually dimorphic switch of *dsf* transcriptional activity. In support of this, feminization of the fat body results in the reduction of courtship behavior (Lazareva, 2007), potentially placing *takeout* mediated transport of a yet to be determined ligand upstream of *dsf* function.

Evolutionarily, *dsf* is highly conserved across 60 million years of insect evolution. Full homology can be seen in the DBD across *Drosophila melanogaster*, *D. virilis* (100% identity) and *Manduca sexta* (10 of 82 amino acid substitutions, with none in the DNA binding elements). The LBD shares 97% conservation across these species. Less conservation is seen in the hinge region, although the length (356 amino acids) is conserved and 59% of residues are homologous. Higher divergence is expected in the hinge as this largely unstructured region serves predominantly in spacing between the

DBD and LBD in nuclear hormone receptors (King-Jones, 2005). Functional domains in the hinge could also persist through the regions of homology observed.

Unlike *fru*, *dsf* does not undergo sex-specific splicing (Finley, 1998). The large intron separating exons 3 and 4 of *dsf* contains a potential *tra* splice site, but Northern blotting reveals only a single splice isoform. Analysis of splicing between the sexes confirms the monomorphic nature of *dsf*. Furthermore, a low level of expression of *dsf* appears to be similar in both males and females.

dsf is present throughout the life cycle from embryo to adult in both sexes, with a slight increase throughout the pupal stage. Anatomically, the expression of *dsf* is limited to a small number of neurons at any point in the life cycle in a pattern identical in both sexes. In larvae, a bilateral pair of clusters in the anterior brain adjacent to the esophagus and two additional clusters found more dorsally show *dsf* expression. Pupal brains showed limited expression in the AL, MB, SOG, anterior protocerebrum, retinal neurons (Finley, 1998) and in a very limited number of individual neuronal nuclei in the VNC cord and abdominal ganglia (Finley, 1998; Katheen Yan, unpublished results (Figure 5B). In the adult, a small number of cells positive for *dsf* are found in proximity to the AL and lateral protocerebrum (Figure 5B). The function of *dsf* expressing cells in the CNS is not known, but their proximity to the chemosensory and photosensory centers is suggestive that they are involved in processing sensory input from external stimuli.

Drawing parallels to *fru* where a change in specificity of a chemosensory receptor can confer courtship inducing properties to alternate sensory stimuli (Kurtovic, 2007), it is probable that *dsf* expressing neurons also mediate mate selection by sensory discrimination. Along these lines, feminization of neurons in the AL (Ferveur, 1995) or

MB (O'Dell, 1995) by *tra* results in aberrant courtship behavior further highlighting the importance of sensory input in mate selection choice. Further posteriorly, the cells labeled in the abdominal ganglia are in proximity to the neurons that innervate the uterus in females indicating a likely function for these cells. (Taylor, 1994; Ferveur, 1995). Finally, the anterior protocerebrum has been identified in mediating female receptivity (Tompkins, 1983) correlating this neuronal area to the phenotype observed in females.

dsf loss-of-function mutants are normal in general behavior including inactivity/sluggishness, ataxia in walking or flight and grooming behaviors. However, mutants of both sexes show abnormal courtship behavior. *dsf* -/- male courtship is indiscriminate relative to the sex of the target, directed at both sexes in mixed populations and directed at both males and females in single fly assays. As in *fru*, multiple mutant males will also display chaining behavior. Additionally, time to copulation in these animals is delayed due to a defect in abdominal bending. Whereas wild-type males readily bend to 180°, mutant males rarely (<10%) achieve an abdominal angle beyond 45°. The consequence of this is an effective block to copulation, as *dsf*-null males cannot make the genital-to-genital contact required. The neurological basis of this phenotype can be explained by abnormal innervation of the ventral longitudinal muscles of the abdominal segment 5 (vA5). The nature of the innervation defect is unknown, but morphologically presents as fewer and larger presynaptic termini or boutons. Muscle of Lawrence formation is unaffected by the loss of *dsf*, indicating that *dsf* is not required to masculinize this muscle in the fly abdomen (Finley, 1997).

Interestingly, *dsf* function in vA5 innervation appears to be dependent on the state of *tra*. In chromosomally female animals, disrupting both *tra* and *dsf* recapitulates the change in boutons seen in *dsf*-null males. Females lacking *tra* alone do not exhibit a defect in vA5 innervation. If *dsf* functions differently in the presence or absence of *tra*, the question of which downstream sex-determination factor is epistatic to *dsf* remains. Experiments with gain-of-function *dsxM* animals in the context of this *dsf* phenotype showed that *dsf* -/- females expressing *dsxM*, but otherwise normal for *tra* have normal innervation at vA5. As changing the sex of *dsx* does not affect *dsf* function, *fru* is left as the other likely candidate. (Finley, 1997)

dsf -/- females are courted actively by wild-type males, suggesting that their pheromonal hydrocarbon profile remains unchanged by loss of *dsf*. Despite this attractiveness, virgin *dsf* females display vigorous rejection of male courtship attempts through ovipositor extrusion, increased movement when being courted and actively resisting a male attempting to copulate by flicking the wings and kicking, behaviors normally only seen in recently mated females (Burnet, 1973). They are also unable to lay eggs due to a defect in motor neuron innervation of the circumferential muscles of the uterus. Phenotypically, this manifests as a distended abdomen from un-laid eggs collecting in the uterus. Though sperm is stored normally in the spermatheca of these females, no fertilized eggs can be observed in the uteri of these animals. Therefore, is not known whether the eggs that are produced are competent to be fertilized and produce viable progeny. Notably, mOL formation is absent in *dsf*-null females indicating that, at least in this location, *dsf* functions in a *fru*-independent manner (Finley, 1997). Previously, neither the male or female phenotypes were seen in heterozygous animals

(Finley, 1997), but more recently *dsf*^{7}/+ females were shown to exhibit mild courtship behaviors (Figure 6A) (Troy Shirangi, unpublished results).

Examining the epistatic relationship of *fru* and *dsf* indicated that *dsf* is responsive to the presence of *fru* in determining courtship behavior and sex-specific neural development (Figure 5A, B and C) (Shirangi, unpublished results). As a baseline, the classic *dsf*–null female phenotype is characterized by ectopic male courtship behavior (Figure 5A). The absence of *fru-M* in these animals establishes that without *fru*, *dsf* functions as a negative regulator of male behavior. At low intermediate levels of *fru*, the loss of *dsf* continues to increase courtship, but as the strength of the *fru* mutation decreases, a threshold is reached where a sufficient level of *fru* switches the effect of *dsf* on courtship from negative to positive (Figure 5B). At the other extreme of the continuum, females expressing *fru-M*, but otherwise female with respect to *tra*, exhibit male behavior and loss of *dsf* in this background leads to decreased courtship (Figure 5C). All together, it would seem that *dsf* function is dependent on the state of *fru* and switches from a pro-female/anti-male to a pro-male/anti-female factor as *fru* is introduced. Neurologically, loss of *dsf* only results in abnormal innervation at vA5 in males. Promisingly, the motor neurons that innervate these muscles are part of the *fru-M* circuit(Figure 5E). Disrupting *fru* alone in this context appears to have no effect on vA5 synaptic terminals, but the double mutant alleviates the aberrant phenotype observed in *dsf*-null males (Figure 5D). In the female, these neurons are unaffected in the absence of *dsf* and continue to appear normal with either loss of *fru* alone or in tandem with *dsf*. The question remains whether the synapses in the double mutant males are functionally wild-type or simply lack the overt bouton phenotype. As *fru* mutants themselves exhibit

abdominal bending defect, loss of both factors clearly does not restore normal function, but it would be interesting to see if these terminals reflect the full complement of pre- and post-synaptic proteins seen in the wild-type neuromuscular junction (Keshishian, 1996). Despite these questions, innervation at vA5 appears to show that *fru* and *dsf* interact genetically.

However, the relationship between *fru* and *dsf* is not completely penetrant. Although the two share similarities in courtship phenotypes, significant differences exist. Fully null *fru* males are largely absent of courtship behavior, whereas *dsf*-null males court indiscriminately. In addition, the abdominal bending defect resides in different groups of muscles, as mOL formation is not affected by loss of *dsf*. In the female, *dsf* must function independent of *fru*, not only because of the lack of *fru* protein, but also because *dsf*-null females exhibit behavioral phenotypes whereas *fru* mutant females do not. Importantly, the expression of *dsf* is not dependent on *fru* (Finley, 1998). Most simply, *dsf* is expressed in females where *fru-F* is translationally off, but even in males, loss of *fru* does not appear to disrupt the normal expression or neural distribution patterns of *dsf*. And finally, *fru* and *dsf* are expressed in a largely non-overlapping subset of neurons in both sexes. Even where the expression of *fru* and *dsf* localization appears to be in proximity, it has not been shown that these proteins are found in a cell autonomous fashion. As an example, loss of *fru* in a cluster of cells in the anterior protocerebrum does not affect the presence of *dsf* in cells in that region (Finley, 1998). Nevertheless, the behavioral changes and vA5 innervation effects seen in the double mutants establish that while *fru* and *dsf* have functions independent of each other, it is plausible that they act coordinately to regulate events underlying these phenomena in a restricted region of the

nervous system. Based of the predicted functions for *fru* and *dsf,* coordinated regulation of gene expression by these two factors would be an exciting mechanism by which these phenotypes could be accounted for. Again, as described for the neuronal function and the output of behavior, *fru* and *dsf* undoubtedly regulate disparate genes in the absence of one another, but could potentially co-regulate genes in the areas of were overlap is seen.

Transcriptionally, *dsf* functions as a sequence specific DNA binding repressor. Homology to Tlx predicted that *dsf* could bind a similar recognition sequence. This was confirmed experimentally and *dsf* was shown to preferentially bind as a monomer to the AAGTCA half site *in vitro.* In line with nuclear hormone receptor DNA binding activity, dimeric binding to a direct repeat separated by a variant dinucleotide (AAGTCAnnAAGTCA) was also observed (Pittman, 2002). Functionally, it was determined that *dsf* repressed the activity of a reporter construct containing the cognate recognition sequence in tissue culture. This repression was reversed by a recombinant *dsf* coupled to a VP16 activation domain (*dsf*-VP16). Although these results indicated similar levels of transcriptional regulation in both the half site and the direct repeat, this does not preclude *dsf* homodimerization or heterodimerization with a yet to be discovered cofactor *in vivo.* Domain analysis revealed that the repression by *dsf* functions predominantly through the LBD. A second repressive domain was found in the amino-terminal region of the hinge. DNA binding was also attenuated by regions in the hinge. The *in vivo* transcriptional behavior of *dsf* was tested through various enhancer-Gal4 lines driving UAS-*dsf*. These experiments showed ectopically expressed *dsf* to be a potent transcriptional repressor. Broadly driven (*actin, dll, elav*) *dsf* was invariably lethal. Expression of *dsf* in a more limited domain (*ALK, dpp*, *GMR*) resulted in viable animals

with domain-specific anatomic abnormalities. More importantly expression of UAS-*dsf*-VP16 alleviated the observed phenotypes and the *in vivo* findings required the presence of the DBD, confirming the role of *dsf* as a transcriptional repressor. Restricting *dsf* to its own domain with a *dsf* enhancer-Gal4 driving a chimeric protein could provide a clearer answer to the *in vivo* consequences of *dsf* overexpression.

Though together these findings establish *dsf* as a repressor of transcription, open questions remain regarding the *in vivo* transcriptional activity of *dsf*. First, what is the role of the putative ligand for *dsf*? Nuclear hormone receptors are known to reverse transcriptional activity from repression to activation upon ligand binding (King-Jones, 2005). Given the sexual monomorphic expression of *dsf*, a ligand or even sexually dimorphic ligands could be the mechanism through which *dsf* regulates differential gene expression to produce the different neural outcomes between the sexes.

Open Questions and Hypothesis

Given the essential roles *fru* and *dsf* play in establishing the neuronal circuit that drives courtship behavior in *Drosphila melanogaster* and their known and predicted biochemical functions, we hypothesized the *fru* and *dsf* are DNA binding transcriptional regulators that control distinct gene expression programs that underlie development of the courtship circuit and set out to uncover the mechanisms by which they enact those functions.

Chapter 2: Results

Section 2.1: *in vitro* Sample Preparation

Cloning *fru* and *dsf* for *in vitro* experiments

Various coding sequences from the *fruitless* and *dissatisfaction* genes were cloned for use in *in vitro* biochemistry experiments (Figure 7). For *fru*, each of the 3' alternatively spliced, zinc-finger containing exons (A, B, C and D) and the male-specific, P1 promoter driven first exon (M) were cloned individually from genomic DNA. Due to initial difficulties in amplifying the full-length P1 transcripts, the 5 exon piece common to all *fru* variants (Comm) was cloned from cDNA with the intent of joining M, Comm and each of A, B, C or D. However, in later attempts, cloning of the full-length P1 to terminal exon versions was successfully accomplished.

For *dsf*, the full-length version and previously published (Pitman, 2002) truncation pieces were cloned. The truncation pieces chosen were amino acids 1-102 (DNA-binding domain, amino acids 322-692 (Ligand-binding domain and hinge) and amino acids 455-692 (Ligand-binding domain). These *fru* and *dsf* pieces were directionally cloned in frame into pT7-MAT-FLAG2® (Sigma-Aldrich).

Bacterial expression of *fru* and *dsf* for *in vitro* experiments

pT7-MAT-FLAG2® vectors containing *fru* and *dsf* constructs were overexpressed in BL21(DE3)pLysS *e.coli* by IPTG induction. Whole cell extracts were tested for expression by western blotting with an anti-FLAG antibody (Figure 8).

***in vitro* expression of *frus* and *dsf* for *in vitro* experiments**

Another method used to produce *fru* and *dsf* proteins for biochemistry experiments was *in vitro* transcription/translation from the pT7-MAT-FLAG2® (Sigma-Aldrich) using the TnT® Quick Coupled Transcription/Translation System (Promega). Expression was tested by western blotting with an anti-FLAG antibody (Figure 9).

Isolating *fru* and *dsf* proteins by nickel-column purification

Separating the FLAG-tagged *fru* and *dsf* pieces from the bacterial whole cell extract was deemed essential for downstream biochemical assays. This was accomplished by taking advantage of the metal affinity tag (MAT) in the pT7-MAT-FLAG2® (Sigma-Aldrich) vector and purifying the proteins using HisPur Ni-NTA Resin (Pierce/Thermo Scientific) (Figure 10A). Purification of proteins was found to be more efficient under denaturing conditions. As this presented a block to functionality in downstream *in vitro* experiments, different approaches were taken to cleaning up these samples including ethanol precipitation of proteins and desalting using Zeba Spin Desalting Columns (Thermo Scientific). Ultimately, this method of purification was abandoned for producing low yield and non-functional proteins.

Purifying *fru* and *dsf* by immunoprecipitation

Improved yield and biochemical activity was achieved by purifying bacterially-expressed and *in vitro* translated *fru* and *dsf* through immunoprecipitation using first an anti-FLAG M2 antibody (Sigma Aldrich) and Dynabeads® Protein A beads (Life

Technologies) (Figure10B) and later ANTI-FLAG® M2 Magnetic Beads (Sigma Aldrich) (Figure 11B).

The switch to conjugated beads was made following the observation that the anti-FLAG antibodies were co-eluting with the purified proteins. As the same antibody was used to detect these proteins on a western blot, the immunoreactivity of the secondary antibody used to the heavy and light chains of the anti-FLAG antibody made visualization of some of the proteins of interest difficult (Figure 10B, Figure 11A). Using the conjugated beads resolved this issue (Figure 11B). Initial denaturing elution with SDS and sodium bicarbonate also precluded use of these purified proteins in biochemical assays, but switching to competitive elution using a FLAG peptide (Sigma Aldrich) yielded biochemically active *fru* and *dsf* proteins.

Section 2.2: Assessing *in vitro* DNA Binding Activity

Systematic Evolution of Ligands by Exponential Enrichment (SELEX)

To test the unknown DNA binding activity of the various *fru* isoforms, *in vitro* translated, FLAG-tagged *fru* A, B, C and D and *dsf* FL and 1-102 were bound to SELEX probes containing a randomized 20-mer in three successive rounds of antibody purification and PCR amplification. Following the third round of SELEX, the recovered probes were sequenced and analyzed for the presence of any consensus binding sequence (Figure 12).

In this experiment, *dsf* served as both a technical positive and negative control, as an *in vitro* consensus binding sequence had been previously established (Pittman, 2002)

and *dsf* 322-692 and 455-692 lacked a DNA binding domain. Accordingly, in the first run 100% of probes sequenced from the *dsf* 1-102 binding reactions contained the canonical AAGTCA half site (Figure 13). Furthermore, this result was consistent across two technical replicates. Also consistent with published results and expectations, there was no DNA binding activity by *dsf* 322-692 and *dsf* 455-692 (data not shown). Strangely, SELEX binding reactions using the *dsf* FL protein also did not yield sequence specific DNA binding activity (data not shown). This could be attributable to the known difficulties in expressing large molecular weight proteins or molar difference between the amount of dsf-FL and *dsf*-102 used. Since the purpose of *dsf* in this experiment was to serve as a control, *dsf* FL, *dsf* 322-692 and *dsf* LBD were omitted in subsequent SELEX experiments.

Unfortunately, none of the *fru* isoforms provided as clear of a consensus binding sequence (Figure 14). The strongest *fru* SELEX result came from the C isoform, where in trial 2, 26 of 96 (27%) reads contained a sequence identified by MEME (Bailey, 1994) as being comprised of AGGnnAGG core. Combining the reads with a first trial of SELEX yielded 47 of 130 (36%) reads having the observed sequence. This confluence of sequences from two trials strengthened the confidence that AGGnnAGG could be a binding site for *fru*-C. The validity of this claim was to be determined by EMSA (see below). As for the other *fru* isoforms, the SELEX results were much less promising. *fru*-A SELEX recovered sequences contained in the second run showed 8 of 33 (24%) reads containing a weak AAGG motif, surrounded with highly variant sequences. Furthermore the same sequence was not seen when the experiment was repeated. The

lack of consistency between trials was also seen for *fru*-B and –D. The only thing that was consistent across all samples was the prevalence of a variant of an AGG motif.

The presence of the AGG trinucleotide in all of the *fru* samples, combined with its marked absence in the *dsf*-102 assay raises the possibility that this was an artifact of the SELEX protocol observed in absence of DNA binding. The divergence in amino acid sequence in the zinc fingers of the different *fru* isoforms (Figure 4) would predict that the different proteins would have disparate DNA binding specificities. Therefore, the prominence of a sequence seen across the isoforms calls the validity of the observed results in to question.

At best these SELEX results we inconclusive. Positively, *dsf*-102 did recapitulate the previously published DNA binding sequence (Pittman, 2002), which was determined by prediction by homology to *tlx*. In those experiments, only variants of the *tlx* binding sequence were examined, while this approach provided an unbiased assessment that validated the canonical *dsf* binding sequence as the highest affinity binding site. An additional benefit of the *dsf* result was confirmation of the experimental protocol. If *dsf* was able to find is cognate binding sequence, then in can be inferred that the probe, binding reaction, purification and iterative selection of target sequences were all suitable for uncovering a sequence specific protein-DNA interaction without experimental bias (Figure 13B). One variable that remains unanswered is proper folding of the protein domains involved in DNA binding. While it is true that *dsf*-102 was functionally active, in the absence of any other functional assay, the same cannot be said with certainty of the *fru* isoforms.

An alternative explanation for the absence of DNA binding seen with *fru* could be attributable to a difference in binding affinity between *fru* and *dsf*. The incomplete coverage of reads by potential consensus sequences in the *fru* samples was in contrast to the complete coverage seen with *dsf*-102. However, the inconsistency in sequences seen between experimental trials argues against this line of thinking. Seeing the same sequence consistently between trials, but only in a subset of reads would have provided more confidence in the sequences found.

In light of the recent findings that uncovered isoform-specific consensus sequences for *fru*-A, -B and –C using SELEX (Dalton, 2013) indicates that our approach was rendered inconclusive by an unresolved technical hurdle. Consistent with the argument that the various *fru* pieces have a lower binding affinity for DNA, the Dalton et al. study utilized 10 rounds of SELEX to our three. Our decision to limit SELEX to three rounds came from the observation that PCR introduced bias increases with additional rounds of selection (Schütze, 2011; Jolma, 2010; Djordjevic 2007). These findings indicate that the strongest binders are seen in the first three rounds of SELEX and low sequence complexity in the later rounds can lead to erroneous selection. Also mentioned was the incidence of *Taq* polymerase introduced mutations increasing as round number increased. Nevertheless, confirmation by EMSA of *fru* binding the sequences found (Dalton, 2013) argues that high repeat number SELEX may indeed be useful for uncovering binding sites for low affinity binding proteins. Only by measuring the dissociation constant for *fru* relative to a high affinity DNA binding protein could this question be answered. Alternatively, high-throughput sequencing, rather than Sanger sequencing, SELEX evolved ligands (Schütze, 2011) could potentially reveal low affinity

binders at low evolution rounds through sequencing coverage. We opted instead to use PB-Seq to address the *in vivo* DNA binding activity of *fru.*

Electrophoretic Mobility Shift Assay (EMSA)

The Electrophoretic Mobility Shift Assay (EMSA) was used to validate sequences gleaned from *in vitro* DNA-Protein interaction experiments (Figure 15). Combining bacterially-expressed, FLAG-immunopurified *dsf*-102 with a biotin-labeled probe containing the optimal direct repeat *dsf* binding site (aagtcagcaagtca) retarded probe migration in a native polyacrylamide gel (Figure 16). Supershifting was observed with the addition of an αFLAG antibody, indicating that *dsf*-102 binding the probe was responsible for the impeded migration. Again, the inclusion of *dsf* in this experiment served as a positive control for the binding reaction and the biochemical activity of the purified proteins.

Given the limited success of the *fru* isoform SELEX experiments, only two probes, representing the *fru*-C (gaaggtaaggtaaggtaagg) and *fru*-B (nngnnaggnnaggnn) sequences, were chosen. The *fru*-C probe specified the most represented variant nucleotides between the AGG repeats and the *fru*-B probe was ordered with random nucleotides in these positions to reflect the variability seen in SELEX. Furthermore, it was thought that an equal, random distribution of nucleotides might reproduce the degenerate nature of bases seen in this position and potentially provide for a positive gel-shift if the specified TA dinucleotide (*fru*-C) proved to be suboptimal for binding. Accordingly, all permutations of *fru*–B and –C probe and *fru*–B and –C protein were examined. Unfortunately none of them resulted in a gel-shift (Figure 17), indicating that

SELEX had failed to provide insight into the DNA-binding specificity of any of the *fru* isoforms. Inclusion of an EBNA positive control EMSA on the same gel accounted for any EMSA specific technical issues that may have led to this result.

Protein Binding-Sequencing (PB-Seq)

An alternate *in vitro* assay we used to test for *fru* and *dsf* DNA binding activity was protein/DNA binding followed by high throughput sequencing (PB-Seq) (Guertin, 2012) (Figure 18A). The advantage of this approach over SELEX was the biological relevance of sequences that comes from using the sonicated *Drosophila melanogaster* genome (Figure 18B), rather than a synthetic random probe. Also, positional information gleaned from this study could be useful for making predictions in subsequent experiments examining *in vivo* DNA binding by *fru* and *dsf*.

Briefly, fractionated genomic DNA was incubated with bacterially-expressed *fru*-A, -B, -C or -D and the DNA recovered after immunoprecipitation was sent for high-throughput sequencing. *dsf* 1-102 was also included as a positive control. Quantitation of the DNA recovered from the purification supported the notion that *dsf* bound DNA with a higher affinity than any of the *fru* isotypes (Figure 18C). However, despite passing quality control and receiving high-quality sequence reads, very few of the sequences aligned with the *Drosophila melanogaster* genome. Alternative sequence alignments indicated approximately 95% alignment coverage of the *e.coli* genome (data not shown). Unfortunately, the contamination from the source of the proteins used precluded finding any novel interaction of these proteins with *Drosophila* DNA. Peak calling and motif finding analysis on the *e. coli* aligned reads showed no consensus DNA

sequences for *fru* and even failed to recapitulate the known DNA binding activity of *dsf*. Rather than continue with *in vitro* approaches, we decided to search for isoform-specific *fru* DNA binding by ChIP-Seq using proteins expressed in S2 cells.

Section 2.3: Schneider 2 (S2) Cell ChIP-Seq

Expressing fru and dsf in Drosophila Schneider 2 (S2) Cells

Previously cloned *fru*-A, -B, -C and –D and *dsf* 1-102 were shuttled into an S2 cell expression vector backbone containing an Actin promoter and a C-terminal 3xFLAG tag (Figure 19A). Transiently transfected S2 cell cultures yielded FLAG immunoreactive proteins (Figure 19B).

S2 Cell ChIP

To test the DNA binding specificities of *fru* and *dsf* under semi-physiological conditions (Guertin, 2010), the *fru* and *dsf* proteins expressed in S2 cells were subjected to chromatin immunoprecipitation (ChIP). Although the proteins expressed well and produced high quality input chromatin (Figure 20), no DNA was recovered from immunoprecipitating *dsf*-102 or any of the *fru* isoforms (no data to show). It is surmised that this purification was unsuccessful due to the dilute nature of the chromatin prepared. In subsequent trials, a more concentrated chromatin preparation will be used to precipitate protein associated DNA molecules.

Section 2.4: Epitope Tagging *fru* and *dsf* by Homologous Recombination

Tagging and detecting *fru* and *dsf* with FLAG and HA tags *in vivo*

Initially, we set out to extensively tag the *fru* locus individually at the N-terminal P1 promoter and C-terminally at each of the alternative isoforms by homologous recombination (Figures 21, 22 and 23). The *dsf* protein was also targeted at the C-terminus and later at three distinct positions internally and the N-terminus (Fei Fei Ding and Edward Lee, unpublished results). For *fru*, successful targeting was accomplished for P1-FLAG, P1-HA and A-FLAG. The other *fru* targeting constructs were successfully cloned and assembled, but abandoned in various stages of completion due to limitations in time and money (Figures 23 and 24). For *dsf*, correctly integrated C-terminal and internal –FLAG and –HA tags were produced (Figure 23 and 25). Validation of targeting by sequencing of the genomic regions confirmed the inclusion of the aforementioned tags (Figures 24 and 25).

None of the tagged *fru* or *dsf* animals showed any overt courtship phenotype and all of the lines were fertile and fecund in the homozygous tagged genotype indicating that inclusion of these peptide tags did not interfere with normal *fru* or *dsf* function (data not shown). As downstream experiments were dependent on the expression of these epitope tags in the *fru* and *dsf* proteins, extracts form these animals were examined by western blotting using tag-specific antibodies. Unfortunately, none of the proteins from any of the *fru* or *dsf* tagged lines were immunoreactive to the corresponding FLAG or HA antibodies (Figures 26 and 27). These experiments were repeated in adult, pupal and larval life stages with consistently negative results. The protein extraction conditions

were controlled by extracting proteins from an ADAR-HA (Gift from the Reenan lab) tagged line (Figure 27C) and each blot was additionally controlled by inclusion of an *in vitro* expressed protein.

This result was not outside of the expectations for *dsf* given its restricted pattern of localization and known low expression. More surprising was the absence of signal from the *fru* tags. The expectation from its broader expression was that it would be more abundant and detectable by western blotting. Unfortunately this was not the case. Relative enrichment of tagged proteins by immunoprecipitation also failed to show tagged proteins on a western blot. Upon further review of the literature, the absence of western blots showing endogenous *fru* became readily apparent. Identification of the *fru* transcripts (Ito, 1996; Goodwin, 2000) and distribution of the isotypes (Billeter, 2006b) was accomplished by northern blotting. Even the temporal pattern of *fru* was described indirectly through changes in staining intensity in a cluster of *fru* positive neurons in the dorsal brain (Lee, 2000). The lack of this methodology in the literature highlighted the general difficulty of detecting *fru* by western blotting. Similarly, the western blots detecting endogenous *dsf* protein have never been published. Only mRNA levels were examined by temporally by non-quantitative RT-PCR requiring two rounds of amplification with stepped primers and spatially by *in situ* hybridization showing the aforementioned limited *dsf* localization (Finley, 1998). From the observed results, the extremely low level of *dsf* expression can be also extrapolated.

To confirm the incorporation of the respective FLAG and HA epitope tags and to test whether any unexpected transcriptional or splicing effects were to blame for the observed lack of signal on the protein level, mRNA from the tagged *fru* and *dsf* animals

was positionally sequenced around the tags (Figure 28). The results of this analysis validated the inclusion of the tag to its proper genomic position and the addition of the tag to the transcribed mRNA of these genes, but did not explain the lack of epitope tagged proteins by western blotting.

Although most of the localization of *fru* was described by *in situ* hybridization (Ito, 1996; Goodwin, 2000) or P1-Gal4 driven reporter genes (Stockinger, 2005; Manoli, 2005; Kimura, 2008 Yu, 2010), one paper did show anti-*fru* antibody staining detecting endogenous protein in the same domains observed with other methods (Lee, 2000). This led us to examine the functionality of the introduced epitope tags by immunohistochemistry.

Immunohistochemistry of tagged *fru* and *dsf* with FLAG and HA antibodies

In response to the aforementioned difficulties of detecting *fru* and *dsf* by western blotting, but also to examine the degree to which the tagged versions of these proteins recapitulated their known expression patterns, their localization was analyzed by immunohistochemistry approaches. *fru* P1-FLAG males stained with an anti-FLAG antibody (Figure 29) showed positive staining in the previously established P1 domain (Lee, 2000; Stockinger; 2005). Strangely, the same pattern was observed in *fru* P1-FLAG females (Figure 29). However, given the position of the tag and the known regulation of *fru* P1 transcripts in females, this is attributable to the persistence of the truncated P1 product in these animals. Further strengthening this argument is the observation that the circuit highlighted by P1-Gal4; UAS-GFP can be observed in females and is largely overlapping with the pattern of *fru* in males (Stockinger, 2005).

No immunofluorescence was observed in CS animals. This staining pattern validates the targeting by homologous recombination of the FLAG tag to P1 and confirms these animals as a viable reagent for downstream experiments using the FLAG tag. Examining the expression pattern of P1-FLAG in the larval and pupal brains would also be a worthwhile exercise. As a control flies in which *fru* P1-Gal4 was driving a membrane-bound GFP (UAS-mGFP) were also examined by immunohistochemistry. These flies reproduced the staining profile as published (Figure 30). To further show that the staining seen in *fru* P1-FLAG flies has fidelity to the known P1 domain, experiments are underway to doubly label w-; UAS-mGFP; P1-FLAG/P1-Gal4 flies to assess for the degree of co-localization. A nuclear GFP reporter would serve this purpose even better as coincident with its function as a transcriptional regulator, *fru* P1-FLAG staining is observed to be nuclear (data not shown).

Staining C-terminally tagged *dsf*-FLAG adult animals produced no discernable staining pattern (Figure 31). As previously described, the low level of expression of *dsf* is consistent with this finding. Immunohistochemistry with *dsf* has always been a difficult assay. A previous undergraduate in the lab, Kathleen Yan, did work analyzing the localization pattern of *dsf* using a line of flies where the endogenous coding sequence for *dsf* had been replaced with that of the LacZ reporter gene (Figure 5D). She was able to see a pattern similar to the published *in situ* localization of *dsf* only with homozygous LacZ/LacZ animals. While this approach did uncover a previously unknown staining pattern for *dsf* in the pupal VNC, the technique is limited by the *dsf*-null nature of these animals. If *dsf* is involved in patterning its circuit in the nervous system, analyzing animals absent of its gene product would likely give abnormal results. Unfortunately, the

heterozygous animals did not provide a strong enough signal for observation. In the future, it would be prudent to examine the pupal *dsf*-FLAG nervous system to see if the unpublished VNC pattern is consistent in this genotype.

An alternative approach to visualizing *dsf* would be to amplify the signal using one of the many molecular tools available in *Drosophila*. However, a *dsf*-Gal4 line does exist, but fails to show any signal driving a UAS-GFP reporter, even with an increase in Gal4 brought about by inclusion of a UAS-Gal4 in the same animal. A *dsf* enhancer trap, 1.8-Gal4 was published to rescue *dsf* function when driving a UAS-*dsf* in deficient background (Finley, 1998) and showed a pattern coincident with the *in situ* staining (Figure 5B). A drawback of this approach is the potential incomplete nature of the enhancer underestimating the distribution of *dsf*.

Given that the targeted *dsf* alleles contain FLAG or HA tags in their genome and these tags are transcribed, even in the absence of localization data these animals will be useful in downstream experiments probing *dsf* function *in vivo*.

***In Vivo* ChIP-Seq**

Currently, *in vivo* ChIP-Seq experiments are in the realm of future directions. They face considerable technical challenges, given the difficulties seen working with *fru* and *dsf* proteins, but in tandem with the S2 cell ChIP could produce novel findings about the *in vivo* transcriptional regulation of courtship behavior by *fru* and *dsf*.

Chapter 3: Discussion and Future Directions

Section 3.1: DNA binding by *fru* and *dsf*

There are few novel conclusions that I can draw, given the current state of the *fru* and *dsf* projects. The extraction of the *dsf* consensus binding sequence from a random pool from potentially 1.09 x 10^{12} unique sequences is a small step beyond the previously published results. Expanding on the SELEX experiments described in the Results section, the predicted binding sequences (http://zf.princeton.edu/; Persikov, 2013) for *fru*-A, B and C are shown in Figure 32. Comparing these predictions to our SELEX results (Figure 14) and those reported in Dalton, 2013 shows good correspondence with the AGG trinucleotide seen in our experiment for *fru*-A. This sequence differs from the result for *fru*-A from the Dalton paper. For *fru*-B, the predicted sequence (CCCT) matches well with the Dalton results and can also be observed in our analysis of this isoform, albeit at a low frequency (8 reads in 106 sequences; Figure 14). The predicted sequence (ATGG) for *fru*-C again fails to match with the Dalton results, but is also absent in our analysis. The Dalton paper ignores the D isoform and this aversion can be seen elsewhere in the literature examining the molecular functions of *fru* (Ito, 2012). Correspondingly, the prediction algorithm failed to find a putative binding sequence for this isoform. The single zinc finger present in this protein, compared to the tandem fingers found in the other *fru* proteins (Figure 4A) could partially explain why it is less well studied and fits in poorly with prediction algorithms, but does not completely exclude the possibility that this isoform binds DNA and directs and isoform-specific transcription program. Our SELEX experiments found extensively degenerate sequences lacking a discernable motif. One possibility of this result is that buried within these sequences is the *fru*-D binding motif. Alternatively, as described in the Results, the

presence of a repeating AGG tri-nucleotide in all four *fru* SELEX results could reflect this sequence as a technical background in the absence of specific DNA binding in this experiment. Nevertheless, confirmation of DNA binding by EMSA for the A and C isoforms in the Dalton paper show that, at least *in vitro*, these sequences, uncovered by SELEX are bona fide *fru* consensus motifs.

A second look at motif finding in our SELEX results showed similar findings (Figure 33). In this analysis, the window size for the consensus sequence was reduced to a minimum of 3 nucleotides and a maximum of 6 nucleotides in an attempt to better model DNA binding by zinc-finer proteins. With these changes, our results for *fru*-A more closely approached the predicted sequence with 47 of 96 reads containing the AGG motif. Strangely, *fru*-B became less coincident with no appearance of the predicted CCC sequence found in the Dalton paper and in our analysis with a wider sequence window. Although 34 of 152 reads from *fru*-C contained a GG, unlike the predicted sequence were not preceded by a T. Furthermore, the sequence for C in the Dalton paper did not contain a GG, but did shift in the EMSA, calling in to question the applicability of the prediction algorithm for *fru.* For *fru*-D, 39 of 144 reads had an AAGG motif again raising the question of background for the *fru* SELEX findings. Together, these additional analyses do not change the inconclusive nature of our SELEX results. Contrasting the success SELEX provided for *dsf*, the fact remains that for *fru*, an alternate line of experimentation needs to be performed to achieve a description of the isoform-specific DNA binding activity of this gene.

On this parallel track, the S2 cell ChIP experiments underway hold some promise in providing a profile of DNA binding by *dsf* in the context of chromatin and validating

(or challenging) the findings by Dalton, et al. for *fru*-A, -B and –C. Again, *fru*-D was notably excluded from their analysis, so any results obtained from S2 cell ChIP with this isoform would be novel. To this end, I am working to stably transfect the four *fru* zinc-finger exons and *dsf*-102 by co-transfecting the pCoBlast (Invitrogen) resistance vector and selecting with blasticidin to incorporate the constructs into the S2 cell genomes (Santos, 2007). Once this has been accomplished, the improved protein expression from these cells should provide for a more efficient ChIP analysis.

With the extensive work done to characterize the *fru* P1 neuronal circuit with both reporter genes (Stockinger; 2005; Yu, 2010) and antibodies against endogenous *fru* (Lee, 2000), the immunohistochemistry experiments using P1-FLAG do little beyond validating the homologous recombination in correctly targeting *fru.* Nevertheless, I am repeating both the P1-FLAG labeling to produce images that better represent the staining pattern I observed and the double-lableling with P1-Gal4; UAS-mGFP to place P1-FLAG in the known P1 domain. Unfortunately, the epitope tagged *dsf* flies did not produce a discernable signal under the conditions tested, as this would have been the most accurate representation of the endogenous protein pattern of *dsf* to date. However, as only adult brains were stained, examination of staining in the larval and pupal stages holds merit. These approaches remain the best method for validating the efficacy of homologous recombination in light of the difficulties in observing tagged *fru* and *dsf* proteins by western blotting.

Experimentally, the usefulness of these targeted lines in precipitating specific DNA remains to be seen, but coupling this *in vivo* ChIP with the S2 cell ChIP leaves me optimistic that one or both of these approaches will reveal a previously unknown function

for either *fru* or *dsf*. Subsequently, any DNA binding profile from either *in vivo* ChIP or S2 cell ChIP would have to be functionally validated through quantitative PCR comparing expression of bound genes in wild-type vs. mutant (or acute RNAi mediated knock-down). Additionally interesting would be comparing the differential binding profiles and expression levels of bound target genes of *dsf* between males and females. Progress towards these ends is underway as you read this.

Section 3.2: The Spatial and Temporal Distribution of P1 Isoforms

One original intent in tagging *fru* at both P1 and each C-terminal zinc-finger containing exon was to address unanswered questions, not only about how the P1 proteins function, but also how each isoform is distributed in space and time and the implications these patterns have relative to their projected functions. Spatially, immunolocalization studies examining *fru* P1 transcripts (Lee, 2000; Stockinger; 2005; Yu, 2010) have focused solely on the P1 portion of the gene and largely ignored the differential representation of each isoform within this domain. Interestingly, systematic analysis of the distribution of the different isoforms within the P1 circuit has not been performed. Part of the difficulty arises in that the alternate zinc-finger exons (A, B, C and D) can be transcribed from any of the four *fru* promoters. As only P1 derived transcripts contribute to courtship behavior, the distribution of isoforms must first be restricted to those of this type. Previously (Billeter, 2006b) this was accomplished through double-labeling using antibodies directed against the endogenous P1 exon and each C-terminal alternate exon respectively. This approach was restricted to the neurons of the abdominal ganglion only, but revealed that transcripts of the C-type zinc-finger

completely overlapped the P1 domain in the adult. This result indicates that in this region, the function of *fru* P1-C is required in all of the adult neurons contributing to courtship behavior in the abdominal ganglion. *fru* transcripts of the A-type were not found in all P1-labeled neurons of the abdominal ganglion, suggesting a persistent function of A-type transcripts from other promoters in this region. Whether P1-derived A transcripts contribute to courtship in the abdominal ganglia and how the P1-A transcripts function relative to non-P1-A transcripts in this region remains to be seen. Furthermore, this approach does not preclude that even in the cells with P1 and A or C overlap, some contribution of A or C is derived from a non-P1 promoter. With regard to the other alternative exons, it was stated that they "could not currently survey expression of type-B" and that "Exon D could not be detected in adult CNS transcripts" (Billeter, 2006b). Together, this leaves open the spatial distribution of P1-B and any supposed P1-D transcripts and their functions in determining courtship behavior in the abdominal ganglia. As stated above, even in the overlap seen between P1 and A or C it cannot be said with certainty that both signals arise from the same type of molecule. Finally, this paper only looked at the representation of P1 transcripts by the A and C types in the abdominal ganglia. As, the P1 circuit is more expansive than just this anatomical region, a nervous-system wide sampling of differential *fru* isoform distribution would provide a map from which functional implications of isoform-specific DNA binding activity could be extrapolated. The same double-labeling experiments could be an adequate first pass to address this question.

Key information is also lacking about the proportion of P1 transcripts of each type that appears temporally throughout the life cycle. Much like the spatial analysis,

experiments probing the temporal behavior of *fru* did not differentiate between the distribution of isotypes, focusing instead on the male-type exon of P1 (Lee, 2000). Furthermore, as previously mentioned, the analysis was performed indirectly as the readout of P1 signal was extrapolated from the fluorescence staining intensity of a selected cluster of neurons in the anterior dorsal brain (Lee, 2000). This does not provide an accurate circuit-wide look at the temporal activity of the P1 promoter and as the individual isoforms were not examined, nothing is known about how each isoform is expressed throughout the life of the animal. Ideally, western blotting for each individual isoform relative to total protein and/or relative to P1 proteins at different time-points would provide the most quantitative information about when each isoform is likely to function. Unfortunately, western blotting for *fru* has proven difficult, both in the literature and in out hands. One approach that circumvented this somewhat was examining the profile of P1 and A, B and C transcripts from adult male and female heads by Northern blotting (Billeter, 2006b). Using a zinc-finger exon specific probe revealed P1-derived transcripts migrating at a distinct size relative to transcripts from the other promoters. However, this experiment was only performed at one time point. This same approach could have been expanded to include RNAs from different embryonic, larval and pupal stages to better understand, if not the absolute level, then at least the relative levels of P1 transcripts throughout development. We considered a version of this experiment by quantitative PCR, but the ever-present problem persists that primers targeting an individual zinc-finger exon do not differentiate the promoter from which that transcript arose. Purification of P1 transcripts by a biotin-labeled antisense oligonucleotide proved too little yield to amplify anything at the 3' end. In any case,

looking at the temporal dynamics of RNA is at best an incomplete picture, as it is not known if and how *fru* transcripts are post-transcriptionally regulated and what implications this has for *fru* function. A thorough understanding of the *fru* zinc-finger isoforms in space and time on the protein level remains a fundamental unanswered question in the field.

Epitope tagging of P1 and all of the zinc-finger exons would have provided reagents to create allelic combinations to track the temporal and spatial distributions of the P1-derived transcripts. One option for this would have been to target zinc-finger specific tagging in a P1-tagged background. Injection of C-terminal A, B, C or D targeting p{W-25} into a verified P1-tagged background would have produced lines of flies in which P1 and the appropriate C-terminal exon were differentially tagged. These single allelic lines would have been indispensible for double-labeling localization and purification experiments. Once the P1-tagged flies were properly targeted by homologous recombination, this line could have been used as a background into which subsequent C-terminal targeting constructs could have been injected. Taken to homozygosity, these doubly-targeted animals would have produced P1-transcripts that carried two different epitope tags on the same molecule. Alternatively, singly-tagged P1 and C-terminal isoforms could have been combined through breeding to produce animals with a P1 tag on one allele and a C-terminal tag on the other to localize P1 specific transcripts in an isoform specific fashion. In both cases, hypothetically, P1-FLAG and A, B, C or D-HA would have been used to examine the differential expression of P1-derived transcripts in the nervous system throughout development by double labeling with FLAG and HA antibodies. As in the case of immunohistochemistry using antibodies directed

against the endogenous proteins, the signal arising from the C-terminal exons cannon unequivocally be attributed to P1 alone, but cells immunoreactive to both P1 and a given C-terminal tag would describe how the different *fru* isoforms are distributed throughout the *fru* P1 circuit.

In the absence of antibodies against endogenous *fru*, we decided to undertake this extensive homologous recombination approach to build reagents not only for experiments aimed at uncovering the transcriptional function of these proteins, but to explore their spatial and temporal patterns as well. The use of monoclonal antibodies raised against epitope tags is an established method for ChIP experiments (Landt, 2012), although the use of polyclonal antibodies against endogenous targets seems to be more widespread and potentially more successful. Given our difficulties detecting P1-FLAG and *dsf*-FLAG by western blotting, using the cloned versions of these proteins intended for biochemical experiments to generate polyclonal antibodies might have been a prudent alternative. At the very least, tagging *fru* and *dsf* with a 3xFLAG, instead of the single tag, could have made these reagents more viable.

Section 3.3: *fru* Isoform-specific Control of Courtship Behavior

Another line of experimentation given the hypothesis that the different *fru* isoforms have distinct DNA binding specificities and therefore unique functions is the contribution each isoform has in establishing the P1 circuit and role it play in determining courtship behavior. Published results regarding this question showed that although no one *fru* isoform was sufficient to fully rescue courtship behavior, *fru*-C provided the greatest degree of normal courtship behavior in a *fru* mutant background (Billeter, 2006b).

Furthermore, only *fru*-C was necessary and sufficient in promoting mOL formation, indicating an unique role for this isoform in this function and strengthening the hypothesis that the different isoforms have non-overlapping functions in determining aspects of courtship behavior and the neuronal circuit that underlies it. A better understanding of how these isoforms localize to the *fru* circuit and what implications loss of an individual isoform has on courtship behavior is necessary.

One method for assessing how each alternative *fru* exon regulates courtship behavior could have been accomplished by introducing a premature stop-codon early into each exon by homologous recombination. Measuring the courtship in these animals deficient in only a single zinc-finger *fru* would have shown which aspect of courtship behavior that isoform is responsible for mediating. However, as the problem of which promoter gave rise to the transcripts persists, constitutive knockout of individual *fru* isoforms could have unintended effects by disrupting that isoform from non-P1 promoters. One way to circumvent this problem would be to introduce FRT sites flanking all or functional parts of each isoform individually. These lines could be combined with the *fru*-FLP line of flies to achieve disruption of an individual zinc-finger exon only in the P1 domain. This should leave the function of the other *fru* promoters undisturbed, maintaining non-courtship directed *fru* functions. Alternatively, isoform-specific disruption of P1 transcripts could be restricted to the P1 domain have through UAS-RNAi knockdown driven by the extant P1-Gal4 allele. This approach would target transcripts of a given isotype where P1 is expressed, but as described above, other *fru* promoters appear to have at least partial overlap with P1, likely leading to some off-target effects. Downregulating *fru* in this manner would have also required measuring the

degree of knock-down to correlate the observed phenotypes with the amount of *fru* targeted by RNAi. All together, removing individual *fru* isoforms genomically in the P1 circuit would have been the best approach to studying how each P1-derived isoforms contributes to what aspect of determining courtship behavior.

Ultimately, connecting these temporal and spatial locations and behavioral phenotypes of each *fru* P1 zinc-finger to its specific DNA binding specificity and the genes that that isoform regulates is the next fundamental challenge in fully understanding how *fru* regulates courtship behavior in *Drosophila melanogaster.*

Chapter 4: Materials and Methods

Section 4.1: General Lab Protocols

Biological Solvent

In all experimental reactions, TE 5 mM, pH 8.5 was used in place of ddH_2O unless otherwise specified. All other solutions were prepared using MilliQ ddH_2O.

Genomic DNA Extraction and Isolation

In the first years in the lab, an ethanol precipitation protocol written by Ken Livak, UCSD was used. Briefly, flies were homogenized in an appropriate volume (25 µL/1-4 flies) of Grind Buffer using a small pestle (Kontes) in a microcentrifuge tube. Samples were incubated at 65°C for 30 minutes. DNA was precipitated by addition of 8M potassium acetate to a final concentration of 1M, incubating on ice for 30 minutes, spinning out the insoluble fraction and precipitating in 3 volumes of ethanol at -20°C for one hour. Samples were centrifuged at 4°C at 13000 rpm for 10 minutes. The pellet was washed once with cold 70% ethanol, centrifuged again and allowed to air dry. Finally, the pellet was resuspended in a corresponding volume of TE (40 µL/1-4 flies).

Grind Buffer
0.1 M NaCl
0.2 M sucrose
0.01 M EDTA
0.03 M Tris-HCl, pH 8.0

Later, especially when genotyping the large numbers of flies from the HR crosses, better yield and purity were obtained using the NucleoSpin® Tissue Kit (Macherey-Nagel) according to manufacturer's recommendations.

mRNA

Both TRIzol® (Life Technologies) extraction followed by ethanol precipitation and the NucleoSpin® RNA kit (Macherey-Nagel) were used per manufacturer's protocol.

Complimentary DNA Generation

Using purified RNA, M-MLV Reverse Transcriptase (Invitrogen) was used to generate cDNA:

1 µg total RNA
2 µL oligo (dT) primer (250 µg/ml; Promega)
[Alternatively, 5 µM random primer (Promega)
or 40 ng gene-specific primer (IDT)]
1 µL 10 mM dNTP mix (Fisher)
to 12 µL TE

65°C for 5 minutes

4 µl 5X First-Strand Buffer
2 µl 0.1 M DTT
1 µl RNaseOUT

37°C for 2 minutes

1 µL reverse transcriptase

37°C for 60 minutes
70°C for 15 minutes
4°C hold

Nucleic Acid Quantitation

DNA and RNA quantitation was performed with a Nanodrop 2000 Spectrophotometer (Thermo Scientific). 2 µL of nucleic acid containing TE was placed on the pedestal and readings were taken at 260 nm. Concentration was determined using an OD conversion of 50 µg/mL for dsDNA and 40 µg/mL for RNA.

For chromatin immunoprecipitation sequencing experiments, DNA was quantified with a Qubit® 1.0 Fluorometer (Invitrogen) utilizing the Qubit dsDNA BR Assay or dsDNA HS

Assay (Life Technologies). 1 µL of DNA was added to the fluorescence reaction and the Qubit provided concentration as ng/mL.

Polymerase Chain Reaction

All PCR reactions performed were variants of the following general protocol:

28.5 µL TE (pH 8.0)
10 µL 5X Phusion HF Buffer
5 µL dNTP mix (2.5 mM; Fisher)
2 µL forward primer (20 ng/µL; IDT)
2 µL reverse primer (20 ng/µL; IDT)
2 µL DNA
0.5 µL DNA Polymerase
50 µL

Phusion® High-Fidelity DNA Polymerase (Thermo Scientific) and the associated $MgCl_2$ containing PCR buffer were used in the ABI GeneAmp 2700 or BioRad myCycler thermocycler running a version of the following program:

92°C for 5:00

92°C for 0:30
(Tm of primers - 2°C) for 0:30
72°C for (1 minute per kilobase of desired product)
Repeat x35-40

72°C for 5:00
4°C hold.

Nucleic Acid Visualization

Nucleic acids were run out on 1% agarose-TAE gels containing 10 µg/mL of ethidium bromide in water. Staining was visualized using a Kodak GelLogic100 Imaging System running Kodak 1D 3.6 software. Fragment size was compared to 1Kb+ ladder (Life Technologies).

Gel Extraction

Desired bands were excised from an agarose gels and purified through use of QIAquick Gel Extraction Kit (Qiagen) or NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) kits according to the provided protocol.

TOPO Cloning

Amplified DNA pieces were incorporated into pCR®2.1-TOPO® for propagation and verification. First, A-overhangs were added to the gel-purified pieces as Phusion® possesses 3'-5' exonuclease activity. In this Taq DNA Polymerase, Native (Life Technologies) reaction:

2 µL	10x PCR Reaction Buffer
2 µL	10 mM dATP (Fisher)
0.8 µL	50 mM $MgCl_2$ Buffer
0.2µL	Taq Polymerase
10 µL	Gel-purified DNA
20 µL	

was incubated at 72°C? for 20 minutes. Then, 4 µL of this A-addition reaction was added directly into the TOPO TA cloning reaction in the TOPO® TA Cloning® Kit with pCR®2.1 TOPO® (Life Technologies) and the remainder of the cloning protocol was followed. The entire 6 µL cloning reaction was then transformed into TOP10 chemically competent cells (Life Technologies). Transformants were tested for incorporation of the insert by digesting with EcoRI at the sites flanking the TA cloning sites.

Bacterial Transformation and Propagation

Chemically competent versions of the bacterial strains JM109 (Promega), XL1Blue (Gift from the Larschan Lab at Brown), BL21 (Promega) and TOP10 (Life Technologies) were used to grow plasmids. The cells were transformed in the following manner:

In 14 mL polypropylene culture tubes (BD), incubate 50 µL of bacteria with plasmid DNA on ice for 10 minutes
Heat shock in a water bath at 42°C for 45 seconds
Cool on ice for 2 minutes
Add 250 µL SOC Media (Invitrogen)
Recover at 37°C shaking horizontally at 250 rpm
Plate 50 and 200 µL on LB-Agar (+ 100 µg/mL ampicillin or +100µg/mL kanamycin) plates
Grow overnight at 37°C

Where appropriate, plasmid mediated blue-white selection was performed by adding 40 µg/plate X-Gal (USBiological) to screen for false antibiotic resistance. Individual colonies were picked with a sterile toothpick into 14 mL polypropylene culture tubes with 5 mL LB (+ 100 µg/mL ampicillin or +100µg/mL kanamycin) and grow overnight at 37°C shaking horizontally at 250 rpm.

gDNA/cDNA Sequencing for Cloning

Purified PCR product (50 ng) or plasmid DNA (500 ng) was combined with a gene-specific primer (5 µM) or primed at the sequencing facility with M13F or M13R universal primers and sequenced at Genewiz, Inc. (Cambridge, MA). Sequences were aligned to reference sequence in MacVector™ 7.2.3 (MacVector, Inc.), Serial Cloner 2.6 (Franck Perez) or Geneious 6.0.5 (Biomatters).

Plasmid Purification

Plasmid purification for all cloning experiments was done with either QIAprep Spin Miniprep (Qiagen) or NucleoSpin® Plasmid (Macherey-Nagel) kits as specified by the manufacturer.

Restriction Digestion

All restriction digestion was performed using New England BioLabs enzymes with their corresponding buffers at 37°C for 60 minutes and 65°C for 10 minutes (except for BsiWI at 55°C for 60 minutes and 80°C for 10 minutes

Restriction Digest Based Recombination – Ligation

Once the appropriate restriction sites for directional ligation were confirmed by sequencing to align the inserted coding sequence in-frame with the epitope tag in the destination vector, the clones were ready to by ligated into the vector. This was accomplished by cutting the insert from into pCR®2.1-TOPO® with the appropriate restriction enzyme (see below) and gel purifying the insert band. The destination vector was also cut with the same restriction enzymes and gel purified. The resultant linearized vector and insert were joined using T4 DNA Ligase (Thermo Scientific) in the following reaction:

Linear vector	50 ng
Insert DNA	3x molar excess relative to vector
10X Ligase Buffer	2 µL
T4 DNA Ligase	1 µL
TE	to 20 µL
	20 µl

Incubate at 16°C overnight and then at 22°C. Transform entire 20 µL reaction to chemically competent bacteria as described above.

Section 4.2: *in vitro* Sample Preparation

Cloning *fru* and *dsf* for Protein Expression

For *fru*, primers (see Appendix) were designed to amplify relevant pieces of the gene. For the individual alternative 3' exons the primers were positioned to amplify from the first nucleotide of divergence from *fru*-Comm to the last nucleotide 5' to the stop codon. For the M exon, the first nucleotide following the tandem ATG translation start site to the last nucleotide before *fru*-Comm. For *fru*-Comm, the first nucleotide following the ATG translation start site to the last nucleotide before divergence to the A, B or C isoforms were chosen. These pieces M to Comm to zinc-finger exon were to be joined, but full length *fru* was cloned from cDNA. A combination of the 5' fru-M primers and the 3' terminal exon specific primers were used to clone full-length *fru* P1-A, -B, -C and –D. For *dsf*, primers (see Appendix) were designed to amplify clones used in the Pittman, 2002 paper. Full-length *dsf* was cloned from a UAS-*dsf* line of flies from the the first nucleotide following the ATG translation start site to the last nucleotide before the stop codon. This product was used at a template for amplifying truncated proteins using primers specific to amino acids 1-102 (DNA-binding domain, amino acids 322-692 (Ligand-binding domain and hinge) and amino acids 455-692 (Ligand-binding domain) respectively.

The amplified pieces of *fru* and *dsf* were sequence verified and directionally cloned into T7-MAT-FLAG2 (Sigma) by introducing EcoRI and Acc65I (KpnI) restriction sites to the either end of the amplicon. Two additional nucleotides were included 3' to the EcoRI

site to maintain the reading frame throughout the inserted recombinant protein coding sequence to the N-terminal FLAG tag. The sizes of the *fru* and *dsf* pieces are as follows:

		DNA (nt)	Protein (AAs)	Protein (kDa)
fru zinc-finger exons	**A**	1020	340	37.4
	B	546	182	20.02
	C	522	174	19.14
	D	600	200	22
	M	297	99	10.89
dsf	**FL**	2073	691	76.01
	102	303	101	11.11
	322	1110	370	40.7
	LBD	711	237	26.07
fru P1 full-length	**M-A**	2868	955	105.05
	M-B	2391	797	87.67
	M-C	2370	789	86.79
	M-D	2301	767	84.37

fru P1 full-length also represents the predicted sizes for *fru-M* proteins *in vivo.*

in vivo dsf is, of course, full-length.

<u>*fru* and *dsf* Bacterial Protein Expression</u>

Confirmed *fru* and *dsf* pieces were ligated to pT7-MAT-FLAG2® (Sigma-Aldrich) using EcoRI and Acc65I (NEB) for protein expression. Successfully ligated constructs were transformed into BL21(DE3)pLysS Competent Cells (Promega) and plated as described. Proteins were expressed by IPTG induction as follows:

> 5 mL LB (+ 100 µg/mL ampicillin) cultures were innculated with fresh colonies and grown overnight at 37°C shaking horizontally at 250 rpm.
>
> Overnight cultures were diluted 1:100 new 5 mL LB (+ 100 µg/mL ampicillin) cultures and grown for 3 hours at 37°C shaking horizontally at 250 rpm.

2.5 mL media we removed for 'un-induced' sample and 2.5 mL LB (+ 100 µg/mL ampicillin) with 1 mM IPTG was added. The tubes were returned to 37°C shaking horizontally at 250 rpm and grown for 3 hours. (Schedl lab)

The 'induced' and 'un-induced' samples were pelleted at 5000 rpm for 5 minutes in a Sorvall RC6 Plus centrifuge. Pellets were snap frozen in liquid nitrogen and stored at -80°C.

Thawed pellets were re-suspended in Native Protein Extraction buffer (see below) supplemented with 1 mM PMSF and lysed in a SLM-Aminco French pressure cell press at 1260 psi in a single run. Lysates were centrifuged in an Beckman L8-80M ultracentrifuge at 100,000 x g for 35 minutes. The supernatant was then snap frozen in liquid nitrogen and stored at -80°C. Expression was tested by αFLAG immunoblotting (see below).

Native protein extraction buffer:
50mM Tris, 150mM NaCl in ddH_2O supplemented with 1mM PMSF

in vitro Protein Expression

fru and *dsf* constructs in pT7-MAT-FLAG2® (Sigma) were alternatively expressed in the TnT® Quick Coupled Transcription/Translation System (Promega). The expression reaction was per the protocol, using 1 µg of plasmid DNA and supplemented with 1 mM IPTG, incubated at 30°C for 120 minutes. Expression was tested by western blotting (see below) 3 µL of the TnT® reaction and detecting by αFLAG immunoblotting (see below). The red color (heme) in the reaction mix precluded quantitation of protein concentration by colorimetric methods. As an alternative detection approach, 2 µL of FluoroTect™GreenLys in vitro Translation Labeling System (Promega) was added to a TnT® reaction as described above. Fluorescence incorporation was examined by western

blotting (see below) 3 µL of the TnT® reaction and imaging the gel on a Molecular Dynamics Typhoon 9410 Variable Mode Imager.

Protein Concentration Quantitation

Protein concentration was measured with the BCA Protein Assay Reagent (Pierce). A standard curve (2000 µg/mL – 0 µg/mL) and 25 µL of the unknown sample were incubated with the assay reagent at 37°C for 30 minutes in a 96-well plate. Absorbance at 562 nm was measured on Beckman Coulter DTX-880 Multimode Detector. Concentration of the unknown samples was determined by comparison to the slope of the standard curve plotted on a graph of absorbance vs. concentration.

Western Blotting

Proteins were run on polyacrylamide gels with the following composition:

SDS Resolving Gel	8%	10%	12%
Acrylimide/bis (30% stock)	2 mL	2.5 mL	3 mL
Distilled Water	5.5 mL	5mL	4.5mL
1.5 M Tris-HCl pH 8.8	2.5 mL	2.5 mL	2.5 mL
0.5 M Tris-HCl pH 6.8	omit	omit	omit
10% SDS	100 uL	100 uL	100 uL
10% APS (100mg/mL stock)	50 uL	50 uL	50 uL
TEMED	5 uL	5 uL	5 uL

Most gels run were 10% considering the proteins to be resolved.

SDS Stacking Gel	4%
Acrylimide/bis (30% stock)	.65 mL
Distilled Water	3.17 mL
0.5 M Tris-HCl pH 6.8	1.25 mL
10% SDS	50 uL
10% APS (100mg/mL stock)	25 uL
TEMED	5 uL

The proteins were electrophoresed in the Mini-PROTEAN® Tetra Cell (BioRad) at a constant 200 V, limited at 150 mAmp at room temperature until the dye front reached the bottom of the gel, generally about 40 minutes.

The proteins were transferred to Whatman ProTran nitrocellulose membranes in a Mini Trans-Blot® Cell (BioRad) at a constant 100 V, limited to 400 mAmp at 4°C for 50 minutes.

The nitrocellulose was blocked in TBST with 1% Nestle Carnation powdered milk (Stop and Shop) for 1 hour at room temperature. The proteins were detected by staining with primary antibody at 4°C overnight:

αFLAG M2 (F3165; Sigma Aldrich) 1:1000 in TBST with 1% powdered milk
αHA (3F10; Roche) 1:1000 in TBST with 1% powdered milk

The blots were then washed in TBST 3 times for 5 minutes at room temperature.

Final labeling was done by incubating in secondary antibody at room temperature for 1 hour:

αmouse-HRP (Jackson Immunoresearch) 1:80000 in TBST with 1% powdered milk for FLAG
αrat-HRP (Jackson Immunoresearch) 1:10000 in TBST with 1% powdered milk for HA

The blots were washed again in TBST 3 times for 5 minutes each at room temperature.

HRP was developed by using ECL 2 Western Blotting Substrate (GE, Pierce, Thermo Scientific). Chemiluminescence was detected on CL-XPosure film (Thermo Scientific) developed in a SRX-101A film processor (Konica Minolta).

Gel running progression, transfer efficiency and protein size were monitored using EZ-Run™ Prestained Rec Protein Ladder (Fisher).

The following buffers were used in western blotting:

10x SDS Gel Buffer	**1 L**
Tris	30.3 g
Glycine	144 g
SDS	10 g

2x Sample Buffer	
10% SDS	4 mL
glycerol	1.5 mL
0.5 M Tris-HCl pH 6.8	2.5 mL
Bromophenol Blue	10 mg
Betamercapto ethanol	1µL/100mL before use

1x SDS Gel Running Buffer	**1 L**
10x SDS Gel Buffer	100 mL
10% SDS	10 mL
ddH2O	890 mL

1x SDS Gel Transfer Buffer	**1 L**
10x SDS Gel Buffer	100 mL
10% SDS	10 mL
ddH2O	890 mL
Methanol	200 mL

10x TBST, pH 7.4	**1 L**
NaCl	80 g
KCl	2 g
Tris Base	30 g
ddH2O	to 1L

1x TBST, pH 7.4	**1 L**
10x TBST, pH 7.4	100 mL
Tween 20	1 mL
ddH2O	to 1L

Coomassie Staining

Total protein was examined by running on polyacrylamide gels, as described above, and stained by Coomassie Brilliant Blue G-250 (Spectrum).

Gels were fixed in Coomassie fixing solution at room temperature for 1 hour:

Coomassie Gel Fixing Solution	1 L
Ethanol	500 mL
Glacial Acetic Acid	100 mL
ddH2O	400 mL

Stained in Coomassie staining solution at room temperature overnight:

Coomassie Staining Solution	1 L
Brilliant Blue R-250	1 g
Methanol	300 mL
Glacial Acetic Acid	100 mL
ddH2O	600 mL

Washed in Coomassie de-staining solution 3 times at room temperature for 5 minutes each:

Coomassie De-Staining Solution	1 L
Methanol	500 mL
Glacial Acetic Acid	100 mL
ddH2O	400 mL

Gels were rehydrated in Coomassie storage solution at room temperature for 1 hour:

Coomassie Storage Solution	500 mL
Glacial Acetic Acid	25 mL
ddH2O	475 mL

and imaged on a flat bed scanner and stored in Coomassie storage solution at room temperature.

Section 4.3: *in vitro* Sample Purification

Immunoprecipitation

Initially, αFLAG M2 antibody (Sigma Aldrich) and Dynabeads® Protein A (Life Technologies) were used to purify FLAG-tagged recombinant proteins from whole protein extracts. Briefly, up to 150 µg of total protein was incubated with 5 µg αFLAG

M2 antibody and 30 µL of Dynabeads® Protein A in 500 total µL Native Extraction Buffer nutating overnight at 4°C. The following day, the beads were washed 3 times with 500 µL Native Extraction Buffer for 5 minutes at 4°C and eluted in TE with 1% SDS and 100mM $NaHCO_3$ nutating at room temperature for 30 minutes.

To test immunoprecipitation efficiency, each step of the process was saved and 30 µL of 500 µL from each step was examined by western blotting and Coomassie staining. Unfortunately, elution of the antibodies from the beads occluded some of the bands of interest through immunoreactivity of the light and heavy chains of the αFLAG M2 antibody with the αmouse-HRP secondary antibody being used for detection. To circumvent this problem, conjugation of the αFLAG M2 antibody to the beads was attempted with little success and the use of ANTI-FLAG® M2 Magnetic Beads (Sigma Aldrich) was found to be a suitable solution.

The approach with the conjugated beads was similar to the approach described above with 30 µL of beads used, except without the additional antibody. No antibody background bands were observed using these beads.

Nickel Column Purification under Native Conditions

fru and *dsf* pieces in T7-MAT-FLAG2® expressed in BL21 cells or TnT® Quick Coupled Transcription/Translation System were purified using HisPur Ni-NTA Resin (Pierce/Thermo Scientific).

The Ni-NTA Resin was washed 2 times with Equilibration Buffer at room temperature for 5 minutes each. Up to 150 µg of total protein was combined with Equilibration Buffer to 300 µL and nutated overnight at 4°C. The resin was washed 6 times in 300 µL wash buffer for 5 minutes each at 4°C and eluted with 100 µL elution buffer at 4°C for 45

minutes. To test nickel-column purification efficiency, each step of the process was saved and 30 µL of 300 µL was examined by western blotting and Coomassie staining.

Nickel Column Buffers:
Equilibration Buffer: 20mM sodium phosphate, 300mM sodium chloride (PBS) with 10mM imidazole; pH 7.4
Wash Buffer: PBS with 25mM imidazole; pH 7.4
Elution Buffer: PBS with 250mM imidazole; pH 7.4

Nickel Column Purification under Denaturing Conditions

Protein purification using HisPur Ni-NTA Resin (Pierce/Thermo Scientific) was performed as described above, except with the inclusion of guanidine-HCl.

Buffers:
Equilibration Buffer: PBS with 6M guanidine-HCl and 10mM imidazole; pH 7.4
Wash Buffer: PBS with 6M guanidine-HCl and 25mM imidazole; pH 7.4
Elution Buffer: PBS with 6M guanidine-HCl and 250mM imidazole; pH 7.4

Desalting Ni-Column Purified Proteins for Biochemistry Experiments

Eluted proteins from immunoprecipitation or HisPur Ni-NTA Resin were cleaned up for downstream biochemistry experiments using Zeba Spin Desalting Columns (Thermo Scientific). The entire eluate from the purification reaction was loaded onto the desalting columns and centrifuged at 1500 × g for 2 minutes to elute.

Ethanol Precipitation of Proteins

Given the incompatibility of SDS-PAGE and high molar guanidine, the proteins purified with the HisPur Ni-NTA Resin (Pierce/Thermo Scientific) under denaturing conditions were recovered by ethanol precipitation. Briefly, samples were incubated with 9 volumes of cold 100% ethanol overnight at -20°C. To precipitate the proteins, samples were spun at 13000 rpm for 15 minutes at 4°C. Pellets were washed once with cold 90% ethanol, spun again for 5 minutes and allowed to air dry. Finally samples were resuspended in 30 µL of 2x SDS sample buffer and analyzed by western blotting.

Section 4.4: *in vitro* DNA Binding Assays

Systematic Evolution of Ligands by Exponential Enrichment (SELEX)

For the SELEX experiments, a single-stranded oligonucleotide probe was ordered with a constant flanking region on either side of a random 20-mer region:

3'- gggccaccaacgacattNNNNNNNNNNNNNNNNNNNNgttgatataaatagtgccca -5'

made double-stranded in a probe-generation PCR amplification using:

SELEX 5'2 cgcggatcctaatacgactcactatagggccaccaacgacatt
SELEX 3'2 cccgacacccgcggatccatgggcactatttatatcaac

and gel purified. 50 ng of purified probe was combined with 10 µL TnT expressed *fru* and *dsf* protein pieces and allowed to bind on ice for 20 minutes. The binding reaction was then immunoprecipitated overnight at 4°C with αFLAG M2 antibody (Sigma Aldrich) and Dynabeads® Protein A (Life Technologies) as described above. The subsequent day, the immunoprecipitation reactions were washed and eluted with 25µL TE with 1% SDS and 100mM $NaHCO_3$. 10 µL of eluate was loaded directly onto 10% polyacrylamide gel and western blotted as described. The selected probes in remaining eluted fraction were recovered with the NucleoSpin® Gel and PCR Clean-up kit and eluted in 25 µL EB. This entire elution was returned to the probe generation PCR reaction. The gel-purified, amplified product of that PCR reaction was again used in a binding reaction and the whole process was repeated three times in total.

The purified product of the final SELEX round was cloned into pCR®2.1-TOPO® and individual colonies were plasmid purified and EcoRI digested to test for probe incorporation. All examined clones had a probe ligated. These positive transformants were sent for Genewiz sequencing with M13F and M13 reverse primers.

The sequencing results all indicated the presence of SELEX probes and were aligned in Geneious 6.0.5 (Biomatters). Consensus sequences were aligned in MEME (Bailey, 1994) using a 6-14 nucleotide search window.

Electrophoretic Mobility Shift Assay (EMSA)

Based on the SELEX results, single-stranded EMSA probes (with binding sites in lowercase) were ordered for *dsf*-102, *fru*-B and *fru*-C:

dsfEMSAProbeF	ACGATGTTGTCCTCAaagtcagcaagtcaGCGCCGTAGTCCCAA
dsfEMSAProbeR	TTGGGACTACGGCGCtgacttgctgacttTGAGGACAACATCGT
B-EMSAProbeF	ACGATGTTGTCCTCAnngnnaggnnaggnnGCGCCGTAGTCCCAA
B-EMSAProbeR	TTGGGACTACGGCGCnncctnncctnncnnTGAGGACAACATCGT
C-EMSAProbeF	ACGATGTTGTCCTCAgaaggtaaggtaaggtaaggGCGCCGTAGTCCCAA
C-EMSAProbeR	TTGGGACTACGGCGCccttaccttaccttaccttcTGAGGACAACATCGT

5 pmol of single stranded probes were labeled with biotin using the Biotin 3' End DNA Labeling Kit (Thermo Scientific). Labeling efficiency was tested by dot-blot. The single-stranded, biotin labeled oligos were annealed by combing complimentary strands in a 1:1 molar ratio with 5 pmol each and incubated in a thermocycler at 95°C for 5 minutes followed by 70 cycles of one minute each decreasing 1°C each cycle to 25°C. 1 µg of each complimentary unlableled probe was also hybridized in the same fashion. Single-stranded probes were subsequently removed by treatment with Mung Bean Nuclease (Takara). Proteins used for EMSA were bacterially-expressed, French press-lysed pT7-MAT-FLAG clones. 125 µg of whole cell extract was bound to 30 µL of αFLAG M2 magnetic beads at 4°C overnight. The beads were washed 3 times with Native Extraction Buffer and to maintain biological activity, proteins were eluted with 100µL Native Extraction Buffer + 2µL FLAG Peptide (Sigma Aldrich; 100µg/mL) nutating at 4°C for 45 minutes.

For EMSA experiments, TBE-acrylamide gel (5%was found to be optimal) was poured:

Percent Gel	4%	5%	6%
ddH2O	6.7 mL	6.3 mL	6.05 mL
Acrylamide	1.3 mL	1.7 mL	2.0 mL
5x TBE	2.0 mL	2.0 mL	2.0 mL
10x APS	50 µL	50 µL	50 µL
TEMED	10 µL	10 µL	10 µL

and pre-run in 0.5X TBE running buffer in a Mini-PROTEAN® Tetra Cell (BioRad) at a constant 100 V, limited at 300 mAmp at room temperature for 40 minutes.

While the gel is pre-running, the EMSA binding reaction was prepared using LightShift Chemiluminescent EMSA Kit (Pierce) reagents:

Shift

Lane	1	2	3
Native Buffer	16	6	4
10x Rxn Buffer	2	2	2
dI:dC	1	1	1
Unlabeled Probe	-	-	2
Protein	-	10	10
Biotin Probe	1	1	1

20 fmol of biotin labeled probe and 4 pmol of unlableled probe were used.

Supershift

Same as above, except 1 µL (2-4 µg) of αFLAG M2 antibody was added to an additional Lane 2 binding reaction, bringing the total number of lanes for this type of EMSA to 4.

Positive Control

The control binding reaction included in the LightShift kit was performed to control for binding, gel/transfer and detection. This reaction used 4 pmol of unlabeled EBNA probe, 1 unit of EBNA protein extract and 20 fmol of biotin-labeled probe and incubated on ice for 30 minutes.

The wells of the pre-run gel were flushed and 20 µL of binding reaction was loaded per lane. Gels were run at a constant 100 V, limited at 300 mAmp at room temperature for 45 minutes. Transfer to a pre-soaked BioDyne nylon membrane (Pierce) in a Mini Trans-Blot® Cell (BioRad) at a constant 100 V, limited to 400 mAmp at 4°C for 30 minutes. Dry the nylon membrane and cross-link at 120mJ/cm^2 in a Stratalinker 2400 UV Crosslinker. Membranes were blocked in Blocking Buffer at room temperature for 15 minutes an labeled with Stabilized Streptavidin-Horseradish Peroxidase Conjugate in Blocking Buffer at room temperature for 15 minutes. The membranes were then washed 4 times for 5 minutes each at room temperature in Wash Buffer and once in Substrate Equilibration Buffer. Luminescence was induced with 1:1 Luminol/Enhancer Solution: Stable Peroxide Solution for 5 minutes at room temperature. Signal was detected by exposing membranes to CL-XPosure film (Thermo Scientific), which were then developed in a SRX-101A film processor (Konica Minolta).

Protein Binding – Sequencing

Protein Binding – Sequencing (PB-Seq) experiments were performed with bacterially expressed, FLAG- tagged *fru* A, B, C and D and *dsf* 1-102 incubated with fractionated *Drosophila melanogaster* genomic DNA. The protocol was adapted from the approach published by the Lis lab (Guertin, 2012).

gDNA was extracted using the NucleoSpin® Tissue Kit. 20 adult Canton S flies were collected per purification and 12 purifications were pooled. Concentration was determined on the Nanodrop. Genomic fractionation was performed by 5 rounds of sonication with a Model 505 Sonic Dismembrator (Fisher) of 30 seconds each with one

minute on ice between rounds. Sonication efficiency was monitored by agarose gel electrophoresis.

125 µg of French-press whole cell extracts were incubated with 5 µg of CS gDNA and 30µL αFLAG M2 magnetic beads in 500 µL Native Extraction Buffer nutating overnight at 4°C. The IPs were washed 3 times with 500 µL Native Extraction Buffer for 5 minutes each at 4°C and eluted with 100µL Native Extraction Buffer + 2µL FLAG Peptide (Sigma Aldrich; 100µg/mL) nutating at 4°C for 45 minutes

Protein-bound DNA was recovered with Gel and PCR Clean-up Kit (Macherey-Nagel) and re-suspended in 17µL elution buffer. The DNA was quantitated with the Qubit using the dsDNA HS Assay and libraries were prepared for sequencing.

To monitor immunoprecipitation efficiency, the eluted proteins were recovered by removing the column flowthrough from the clean up kit and adding 50µg BSA in Native Extratction Buffer as protein carrier and precipitating proteins using the NucleoSpin® RNA/Protein Kit (Macherey-Nagel). The entire recovered protein fraction was loaded onto a 10% acrylamide gel and western blotted as above.

Section 4.5: Illumina Sequencing

Illumina Library Preparation

Illumina libraries were prepared with NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (New England BioLabs) coupled with adaptor primers # 2, 4, 5, 6, 7 and 12 of NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1) using the manufacturer's protocol.

Illumina Quality Control and Sequencing

All quality control and sequencing was performed at the Brown Genomics Core Facility by Dr. Christoph Schorl and Hillary Hartalub. Very briefly, libraries were examined for quality and size distribution on the Fragment Analyzer (Advanced Analytical) using the High Sensitivity NGS Fragment Analysis Kit (Advanced Analytical) at 2ng/μL. Sample concentration was determined by adapter specific qPCR using theKAPA Biosystems' Illumina compatible library quantification Kit (KK4835) on an ABI HT7900 qPCR machine. Library clustering was done on the cBot (Illumina) with Illumina's TruSeq SR Cluster Kit v3 (GD401-3001) at 9.5 pM sample concentration followed by sequencing on the HiSeq 2000 in standard high-output mode using TruSeq SBS Kit v3 reagents for 50 cycles.

PB-Seq: Illumina Sequence Alignment and Peak Calling

The Illumina reads returned from the Sequencing Core were aligned to the *Drosophila melanogaster* genome (Berkeley Drosophila Genome Project, Release 5) in bwa (Li, 2009; version 0.7.4-r385) using default settings. With very little alignment of reads to the *Drosophila* genome, alignment was performed to the *Escherichia coli K12* substring *MG1655* genome (NCBI RefSeq Accession NC_000913). With improved coverage, both the *D. melanogaster* and *e. coli* alignments were analyzed for peak enrichment using MACS (Feng, 2011; version 1.4.2) using default parameters, an appropriate genome size for the organism used and a broad range of p-value thresholds. Potential enrichment sites were output as a FASTA file and analyzed in MEME (Bailey, 1994) using a 6-14 nucleotide search window.

Section 4.6: S2 Cell Experiments

Gateway Cloning

For expression in *Drosophila* Schneider (S2) cells (Gift from the Wharton lab), existing *fru* and *dsf* clones in T7-MAT-FLAG2® were PCR amplified using primers (see Appendix) containing att-B1 and att-B2 recombination sites, a *Drosophila* Kozak sequence (CACC) and an initiator methionine codon. PCR products were gel-purified and recombined into pDONR 221 (Gift from the Wharton lab) gateway entry vector with the BP Clonase II kit (Invitrogen). The cloning reaction was transformed to JM109 bacteria, kanamycin resistant colonies were miniprepped and the recovered plasmids were restriction digested with EcoRV and NheI (NEB). Insert containing vectors were sent for sequencing (Genewiz) with vector specific flanking primers.

Confirmed pDON221- *fru* and *dsf* constructs were transferred to pAWF (N-terminal 3x FLAG; Gift from the Wharton lab) using the LR Clonase II Kit (Invitrogen). This final cloning reaction was transformed to XL1Blue bacteria and colonies were again analyzed by diagnostic restriction digestion with EcoRV and NheI (NEB). Insert containing vectors were verified by sequencing (Genewiz) with vector specific flanking primers.

S2 Cell Culture

S2 cells (Gift from the Wharton lab) were maintained in 10 cm plates (Corning) or 75 cm^2 flasks (Company) in *Drosophila* S2 cell media (Gibco) supplemented with 10% (v/v) fetal bovine serum and 50 units of penicillin and 50 μg streptomycin sulfate per milliliter of medium at room temperature in Tupperware. Cell density was measured with a

scientific hemocytometer (Hausser Scientific) and cells were passed at 15-20x10^6 cells/mL at a ratio of 1:5 in a Forma 1400 Biological Safety Cabinet (Thermo Scientific).

S2 Cell Transfection

Once the *fru* and *dsf* pieces were confirmed in pAWF, 2 µg of plasmid was transfected into S2 cells at a density of 5x10^6 cells/mL using the Effectine Transfection Reagent (Qiagen). Transfected cells were kept at room temperature for 4 days and 100 µL of cells were spun down at 5000 rpm for 1 minute and re-suspended in 30 µL 2x loading buffer. Samples were assayed by western blotting with αFLAG antibody as above. One set of 10 cm plate transfections were harvested, re-suspended in 1 mL 2x loading buffer and stored at -80°C as a western blotting control.

S2 Cell Chromatin Immunoprecipitation

S2 cells were transfected with pAWF *fru* A, B, C and D and *dsf* 1-102 as described above. On day 4 after transfection, the cells were checked for protein expression and harvested for chromatin preparation.

The cells were fixed with 1% final concentration of formaldehyde at room temperature for 10 minutes. Glycine was added to a final concentration of 125 mM to quench cross-linking. Cells were then washed one time each with PBS-EDTA, ChIP wash buffer A and ChIP wash buffer B. A determination of cell number was made by hemocytometer to guide the subsequent volumes used. Cell pellets were snap frozen in liquid nitrogen and stored at -80°C.

To fractionate the chromatin, the thawed cell pellets were re-suspended in 1% SDS ChIP lysis buffer and nutated at 4°C for 10 minutes. The lysed cells were sonicated in a UCD-200 Bioruptor water bath sonicator (Diagenode) at 3 rounds of 5 minutes of 30 seconds

'on' and 30 seconds 'off', adding ice between each 5 minute round. The tubes were spun at 13, 000 rpm for 10 min at 4°C. 400 µL was removed for input clean up and the rest of the chromatin prep was snap frozen in liquid nitrogen and stored at -80°C.

The input sample was supplemented with 21.5 ul 20% SDS, 15ul 5M NaCl and 1ul RNaseA (5Prime) and incubated in a 65°C heat bloc overnight. The following day, 1 µL RNAse A was added and the prep was incubated at 37°C for 30 minutes. After RNAse treatment, 20 ul 1M Tris-HCl pH 6.8, 10 ul 0.5M EDTA pH 8.0and 3 ul 20mg/mL proteinase K (Invitrogen) was added and incubated at 42°C for 90 minutes. Once the protein was digested, the DNA was recovered by phenol/chloroform extraction and ethanol precipitation. The input sample was re-suspended in 30 µL of TE and quantitated with the Qubit using the dsDNA HS Assay. 30 ng of chromatin was run on a 1% agarose gel stained with ethidium bromide to check for fractionation efficiency.

To immunoprecipitate the chromatin-bound *fru* and *dsf* proteins, 30 µL of ANTI-FLAG® M2 Magnetic Beads were washed once with 10 ml RIPA 150 + 500mM NaCl and once with RIPA150 at 4°C and blocked with 5 mg/ml IgG-free BSA (Sigma) and 50 mg yeast tRNA (Invitrogen) in 15 mL RIPA150 buffer nutating overnight at 4°C. On the subsequent day, 200 µL of blocked beads were combined with 30 µg of each chromatin prep and bound nutating overnight at 4°C. On the third day, the beads were washed 2 times with RIPA150, 1 time with RIPA300, 1 time with LiCl buffer and 2 times with TE for 5 minutes each. Bound chromatin was eluted with 30 TE with 1% SDS and 100mM $NaHCO_3$ at room temperature for 15 minutes. The elution was repeated with another 30 µL and the samples were pooled.

To clean up the immunopurified chromatin, the elution was supplemented with 2.4 µL of 5M NaCl and incubated at 65°C overnight. Finally, the samples were treated with 1 µL of RNAse A at 37°C for 30 minutes. The chromatin bound protein was digested by addition of 2.4 µL 1M Tris-HCl pH 6.8, 1.2 µL of 0.5M EDTA pH 8.0 and 3ul 20mg/mL proteinase K. Once the protein was digested, the DNA was recovered by phenol/chloroform extraction and ethanol precipitation. The IP sample was re-suspended in [30] µL of TE and quantitated with the Qubit using the dsDNA HS Assay.

ChIP Solutions

ChIP Wash Buffer A
10 mM Hepes pH 7.6
10mM EDTA pH 8.0
0.5mM EGTA pH 8.0
0.25% Triton X-100
Filter Sterilize
1x PI
0.2mM PMSF

ChIP Wash Buffer B
10 mM Hepes pH7.6
100 mM NaCl
1 mM EDTA pH 8.0
0.5 mM EGTA pH 8.0
0.01 % Triton X-100
Filter Sterilize
1x PI
0.2mM PMSF

ChIP TE Wash Buffer
10mM Tris pH 8.0
1mM EDTA
0.01% SDS
Filter Sterilize
1x PI
0.2mM PMSF

RIPA 150 Buffer
50mM Tris-HCL pH8.0
1% NP-40

2mM EDTA
0.1% Sodium Deoxycholate
0.1% SDS
150mM NaCl
Filter Sterilize
1x PI
0.2mM PMSF
1mM DTT

RIPA 300 Buffer
50mM Tris-HCL pH8.0
1% NP-40
2mM EDTA
0.1% Sodium Deoxycholate
0.1% SDS
300mM NaCl
Filter Sterilize

LiCl/TE Wash Buffer
0.25M LiCl
1% NP-40
1% Sodium Deoxycholate
10 mM Tris-HCl pH 8.0
1mM EDTA pH 8.0

Sodium Bicarbonate Elution Buffer
1% SDS
0.1M Sodium Bicarbonate

Section 4.7: Homologous Recombination

***Drosophila melanogaster* husbandry**

Flies were kept on standard cornmeal food, made by the fly food facility, either at 18°C or 25°C with 50% relative humidity and a 12:12 hour day/night cycle. Virgin females of a given genotype were collected twice a day, 8 hours apart and kept 5 to a vial at 18°C for 3 days to check for the absence of larvae.

Directed Targeting of Epitope Tags by Homologous Recombination

FLAG or human influenza hemagglutinin (HA) epitope tags were added to the N- and C-termini of the *fru* and *dsf* proteins by ends-out homologous recombination (Gong, 2003; Maggert, 2008; Reenan lab, Homologous Recombination Episode I: The Clone Wars and Flygeddon protocols). Using the p{W25} (Gift from the Reenan lab) vector to target the *fru* and *dsf* loci by site-directed homologous recombination, proper in-frame integration of the epitope tags was accomplished. See Figure 23 for the strategy used to target the *fru* and *dsf* loci *in vivo*.

Approximately 6 kb of genomic sequence was the intended homology target for each construct. This distance was covered by two contiguous homology arms of about 3 kb each. One homology arm was centered on the initiator methionine (N-terminal) or the stop codon (C-terminal) for each gene, with targeting sequence extending in both directions from that position. These two pieces were then joined by introduction of the epitope tag. The end of this homology arm was directed to a low area of conservation across *Drosophila* species as indicated in the UCSC Genome Browser (http://genome.ucsc.edu/) with the intent of minimizing disruption from the introduction of the mini-*white* gene in this region. The second homology arm was then continued from the subsequent nucleotide through the next 3 kb. These stretches of genomic DNA amplified, and when appropriate joined, by PCR. Amplicons were subcloned, propagated and sequence verified in TOPO2.1 (Invitrogen). Correct clones were transferred to pW25 by restriction enzyme directed cloning. The orientation of the mini-*white* gene was antiparallel to the direction of transcription of the gene targeted, as ectopic expression of *white* has been shown to influence courtship behavior (Zhang, 1995). Once in pW25, the constructs were sent for injection into flies (Genetic Services Inc., Cambridge, MA). The

flies that were returned carried the targeting constructs that had been randomly integrated through P-element (foot) mediated recombination in their genomes.

Moving the constructs from their random positions to the intended genomic positions was accomplished through a series of crosses. First, red-eyed females received from Genetic Services were crossed to male flies having the I-SceI endonuclease (Golic, 1989) and FLP recombinase under the control of the heat shock promoter. On days 3 and 4, the larvae of these crosses were incubated at 37°C for one hour, resulting in FLP-mediated excision and I-SceI-mediated linearization of the targeting construct. In some very small fraction of flies, this linear piece of DNA can homologously recombine into the intended genomic position in the germline for propagation into subsequent generations. The integration can be verified by a second cross in which white or mosaic eyed female progeny of the first cross are mated to *eyeless*-FLP males. Any red eye color arising from homologous recombination is resistant to FLP excision, while spurious red eye color from persistent random integration will be eliminated. In the third cross, the likely correct, red-eyed female progeny of the second cross are mated to TM3/TM6B balancer males to establish a stock. Once balanced, a heterozygote- heterozygote cross yielded homozygous progeny which were PCR sequence verified to establish that the epitope tags had been correctly targeted to their intended positions.

gDNA and RNA Sequencing of HR-tagged flies

Flies were genotyped for successful homologous recombination by extracting genomic DNA from a single homozygous (or Canton S for negative control) fly with the

NucleoSpin® Tissue Kit or total RNA with Trizol and reverse transcribed to cDNA with M-MLV Reverse Transcriptase as described previously. The genomic region of interest was amplified by flanking primers and the gel purified PCR product was sent for sequencing to Genewiz.

DNAse treatment

When examining P1-FLAG, *dsf*-FLAG and *dsf*-HA flies for incorporation of the epitope tag into mRNA transcripts, it was important to remove any contaminating genomic DNA as these would confound the results. For this purpose, TURBO™ DNase (Life Technologies) was added to mRNA prep and incubated at 37°C for 30 minutes. The DNAse was removed with NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel) and reverse transcription reaction proceeded as above.

Section 4.8: Work with HR Targeted Flies

***Drosophila melanogaster* Protein Extraction**

Drosophila (larvae, pupae and adults) were snap frozen in liquid nitrogen, ground to a powder with a pestle (Kontes) in a microcentrifuge tube and resuspended in Native Protein Extraction buffer. Protein concentration was quantitated with BCA Protein Assay Reagent and samples were diluted to 1 mg/mL and stored at -80°C.

Fly Western Blotting, Immunoprecipitation and Ni+ Column purification

All protein work involving fly derived protein was performed as described above, with the exception that flies were snap frozen in liquid nitrogen, homogenized in a

microcentrifuge tube with a pestle (Kontes) and resuspended in 100 µL/1-4 flies Native Protein Extraction buffer.

Immunohistochemistry of adult *Drosophila melanogaster* Brain and VNC

Adult *Drosophila* brains and ventral nerve cords were dissected and antibody stained as described in Wu et al. (2006). Briefly, brains were fixed in fresh 4% paraformaldehyde, washed in PBT, blocked with PBT+0.5% normal goat serum and stained with primary antibody nutating at 4°C for two days. The samples were washed with PBT and incubated with secondary antibody nutating at 4°C for two days. Finally, the samples were washed with PBT, allowed to settle in *SlowFade*® Gold Antifade Reagent with DAPI (Life Technologies) for 1 hour at room temperature and mounted on Colorfrost Plus slides (Fisher) with .15mm coverslips (Fisher). Brains were imaged on the Zeiss 510 confocal microscope.

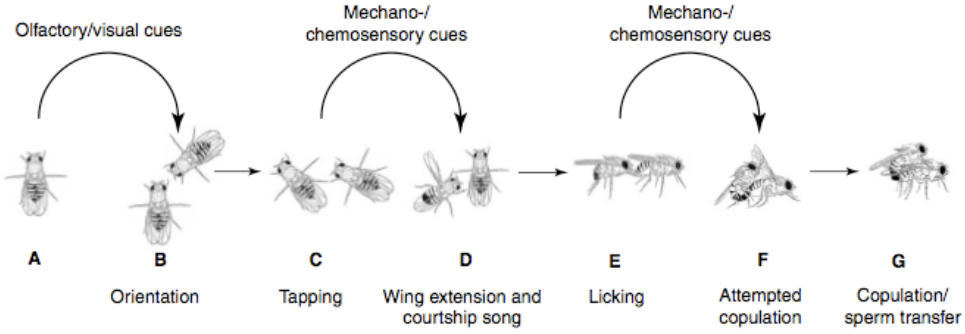
Antibody concentrations used:

αFLAG (Sigma) – 1:100
αMouse555 (Molecular Probes) at 1:200 for αFLAG

αGFP (Molecular Probes) – 1:1000
αRabbit488 (Molecular Probes) at 1:800 for αGFP

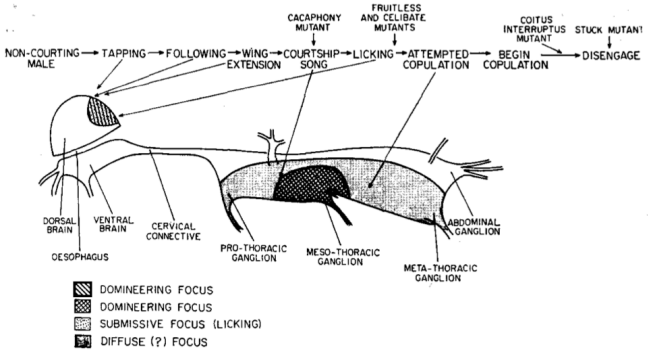
Figures

A.



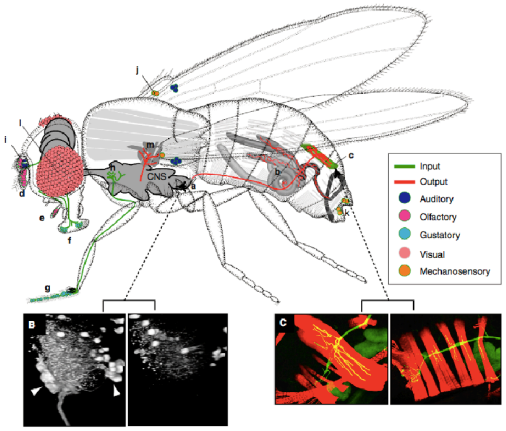
Billeter, 2006a

B.



Hall, 1979

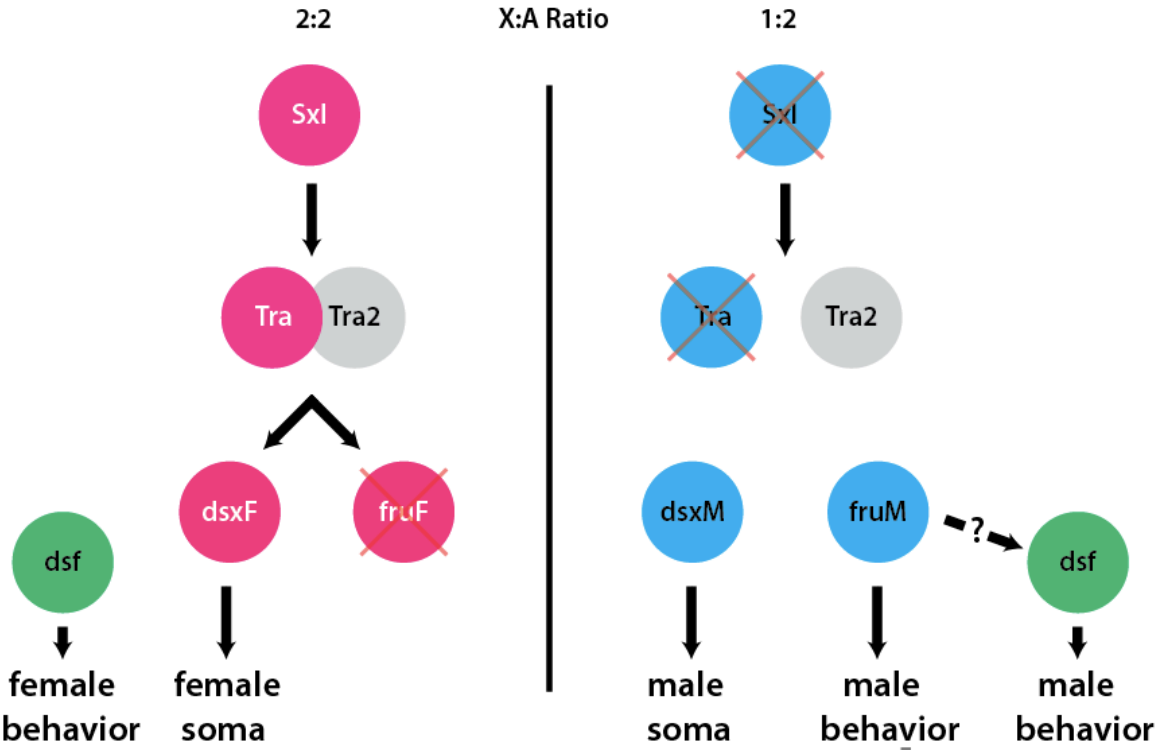
C.



Billeter, 2006a

Figure 1. – A. *Drosophila melanogaster* male courtship behavior sequence. **B.** The linear progression of courtship steps map onto the *Drosophila* nervous system from anterior to posterior. **C.** The *fru*-P1 neural circuit superimposes onto the known courtship centers of the brain and VNC. Inset: (left) fru-M serotonergic neurons in the abdominal ganigla innervate the internal reproductive organs; (right) mOL formation in males (left) and females (right) is dependent on a *fru*-M expressing neuron (**Figure 3C**).

A.

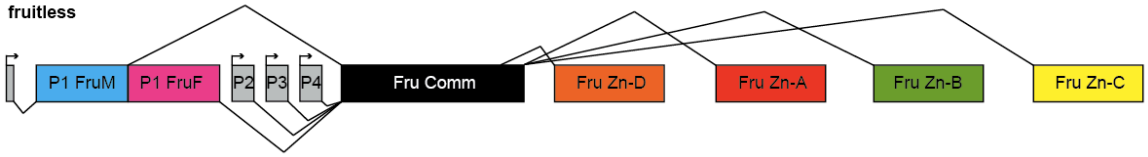


B.

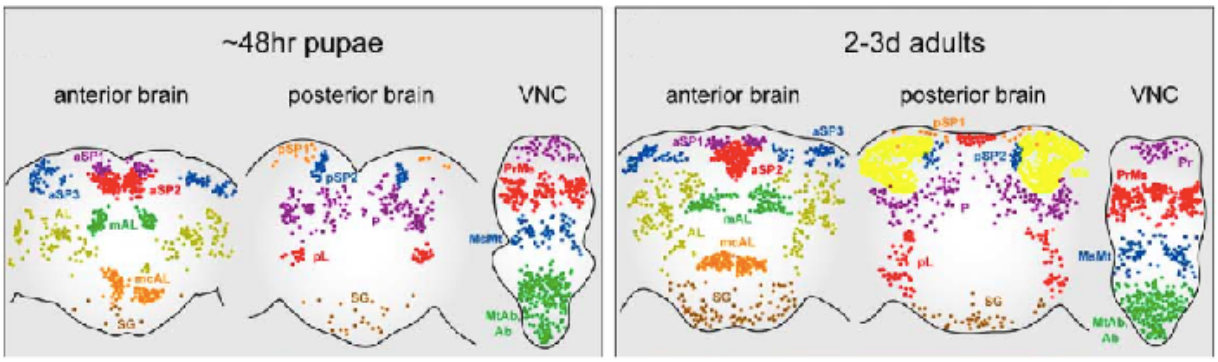
Yang, 2008

Figure 2. – A. Sex determination hierarchy in *Drosophila melanogaster*, determined by X chromosome to autosome ratio and drives an alternative splicing cascade controlled by *Sxl*, *tra* and the constitutively expressed *tra2* mediated alternative splicing of *dsx* and *fru*. *dsf* is not under direct *tra* splicing activity, but does appear to be epistatic to *fru* in a restricted subset of neurodevelopmental events (vA5). **B.** Alternative splicing of *dsxF* and *dsxM* at the C-terminus conferring different dimerization domains. *dsxF* and *dsxM* share the same DNA binding specificity.

A.

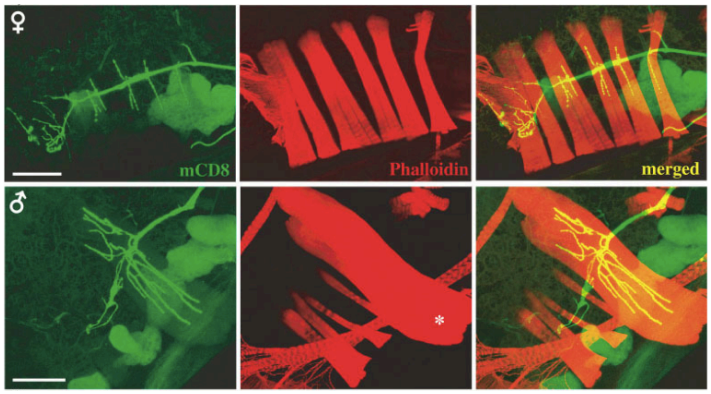


B.



Stockinger, 2005

C.

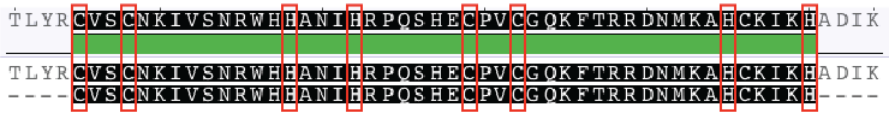


Billeter, 2004

Figure 3 – A. *fru* is a complex genetic locus, transcribed from one of four promoters (P1-P4). Only P1 is under the control of *tra* directed sex specific splicing, adding an additional 101 amino acids to the N-terminus of *fruM* and leading to translational repression of *fruF* through stop codons present in the included exon. At the 3' end, *fru* is alternatively spliced to one of four zinc finger containing exons (A, B, C or D).
B. Schematic of neuronal clusters in the brain and VNC expressing *fruM* (see text).
C. Differential muscle development dependent on its *fruM* expressing motor neuron. In the presence of *fruM*, longitudinal muscles in the 5th abdominal segment coalesce to form the muscle of Lawrence.

A.

Fru-A



Fru-B

Fru-C

Fru-D

B.

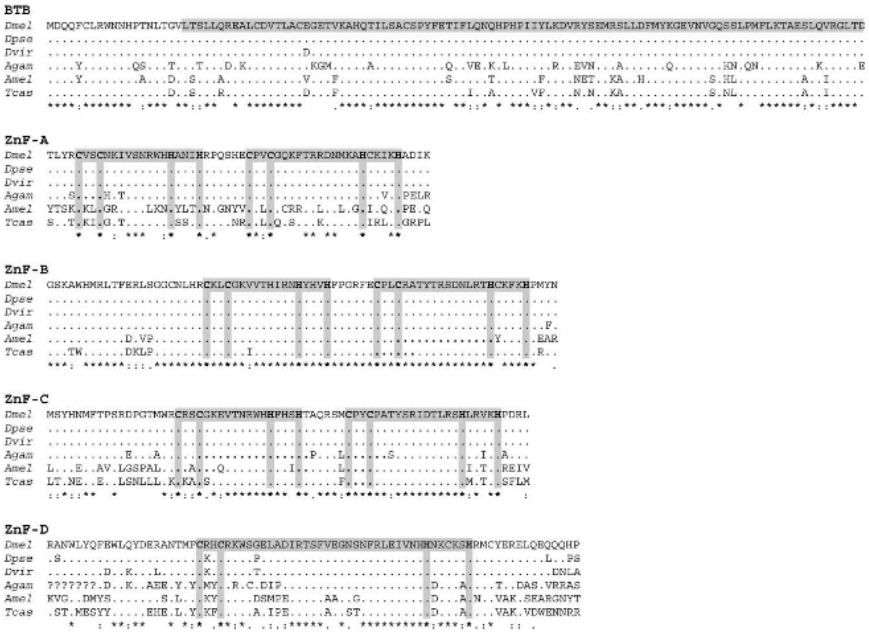
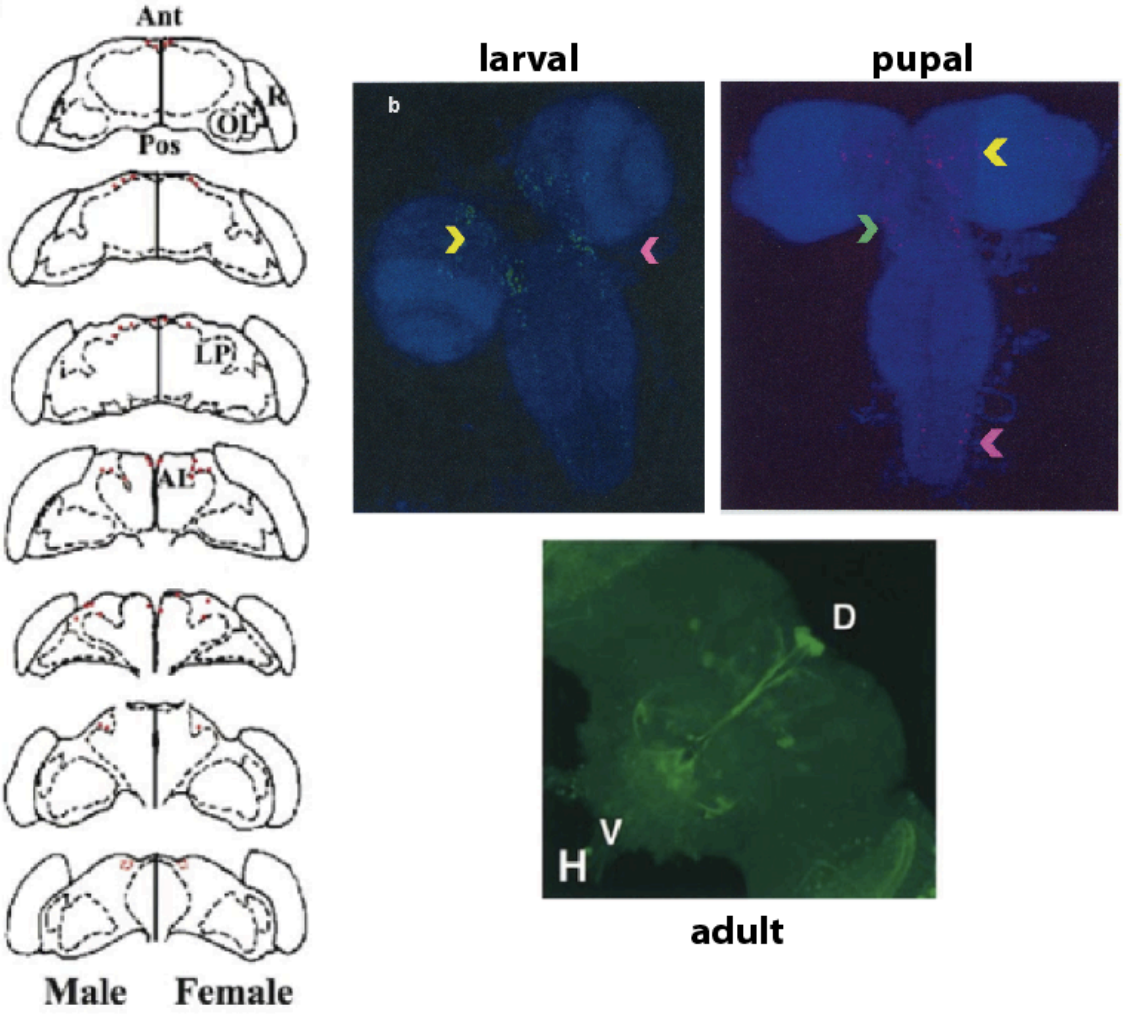


Figure 4 – A. Alternative C2H2 zinc fingers of *fru*. Key residues are boxed in red. *fru*-A, -B and –C contain a tandem set of fingers, while *fru*-D only possesses one. The lack of conservation between the fingers suggests different DNA binding specificities for each isoform. **B.** Evolutionary sequence conservation of the *fru* zinc fingers underlies a functional homology.

A.

B.



Finley, 1998; Kathleen Yan, unpublished results

Figure 5 – A. *dsf* is a sexually monomorphic NR2e orphan nuclear hormone receptor. Characteristic DNA-binding and ligand-binding domains are indicated. **B.** *dsf* is expressed in a limited number of neurons in both male and female. (left) Clusters of cells can be found in the antennal lobe (AL) and lateral horn of the protocerebrum (LP). (right) *dsf*-LacZ staining in male brain and VNC shows staining in the developing dorsal brain (larval - yellow and pink arrowheads; pupal - yellow and green arrowheads) and in the abdominal ganglia (larval – unlabeled; pupal - pink arrowhead). Enhancer trap *dsf*1.8-Gal4 driven UAS-mGFP also shows a similar expression pattern in the adult (anterior view). Additional staining can be observed in retinal neurons.

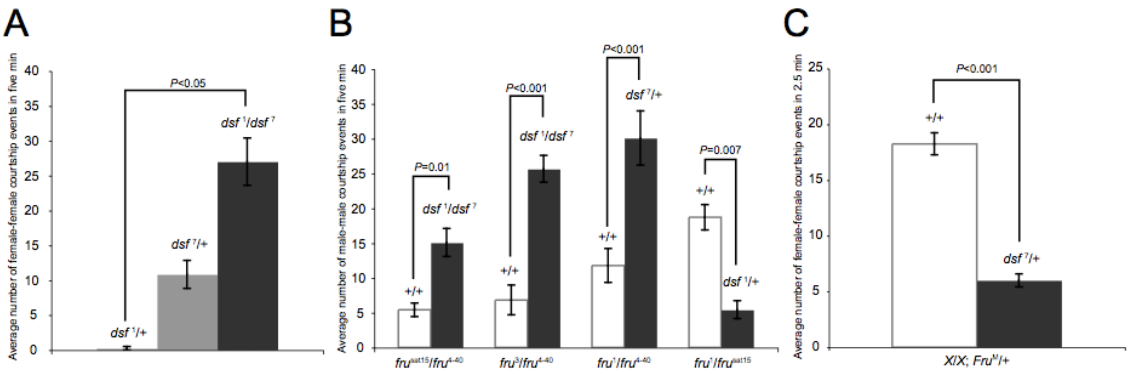


Figure 6 A, B, C – A *fru* dependent switch in *dsf* activity in determining courtship behavior A. Loss of *dsf* in females (*fru* OFF) increases courtship behavior. **B.** Progressively strengthening *fru* allelic combinations show a switch from anti-male to pro-male functions for *dsf* as *fru* increases. **C.** *fruM* expressing females exhibit decreases in courtship behavior in contrast to wild-type females.

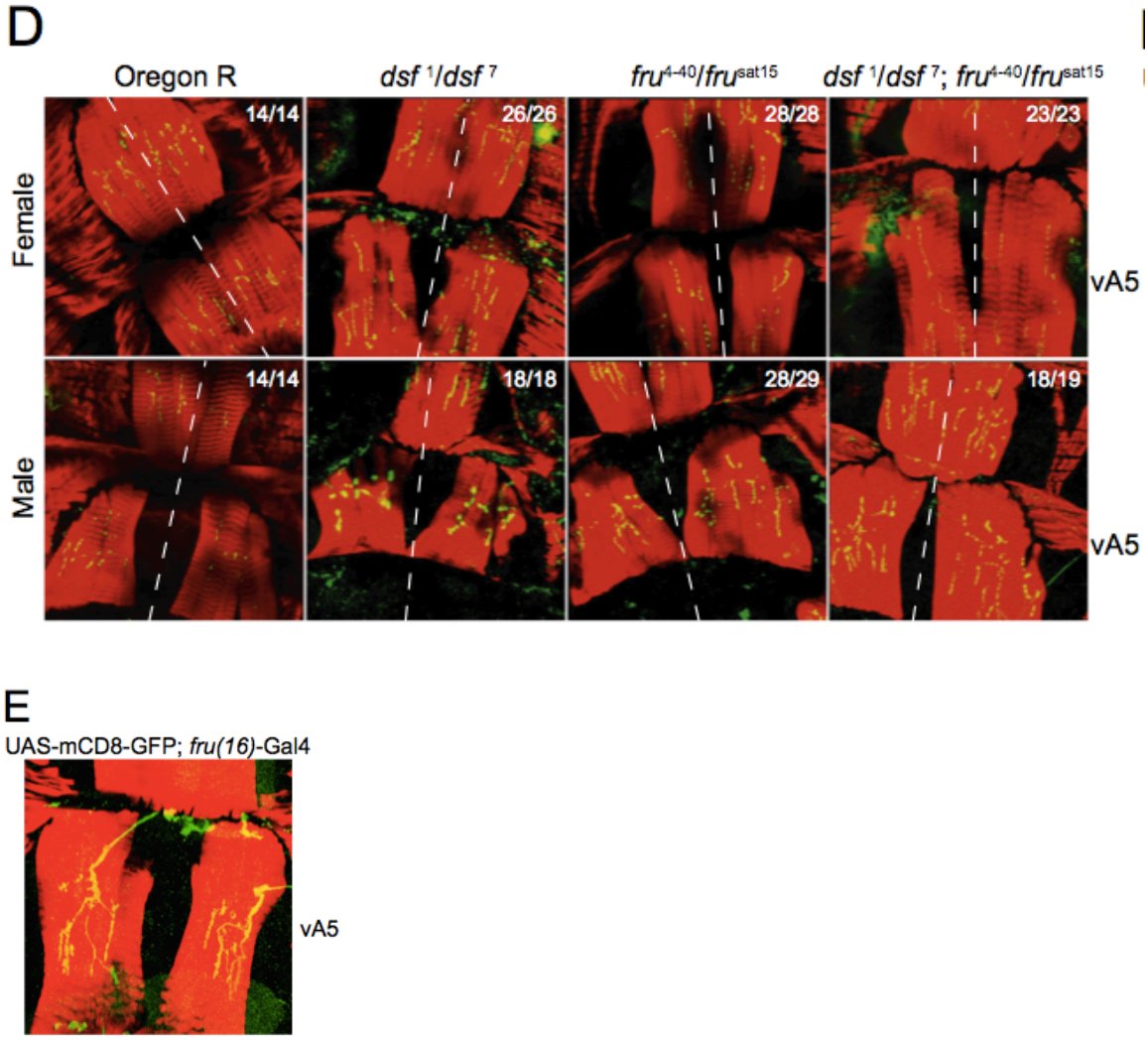


Figure 6 D, E – A *fru* dependent switch in *dsf* activity in determining neuromuscular innervation. D. The neuromuscular junction (synaptotagmin staining) at vA5 shows defects upon loss of *dsf* in the male (*fru* ON), but not in the female (*fru* OFF). Mutating *fru* alone does not produce this phenotype, but the double mutant relieves the dependence of *dsf* on *fru*. **E.** The muscles at vA5 are innervated by a *fru* P1 neuron.

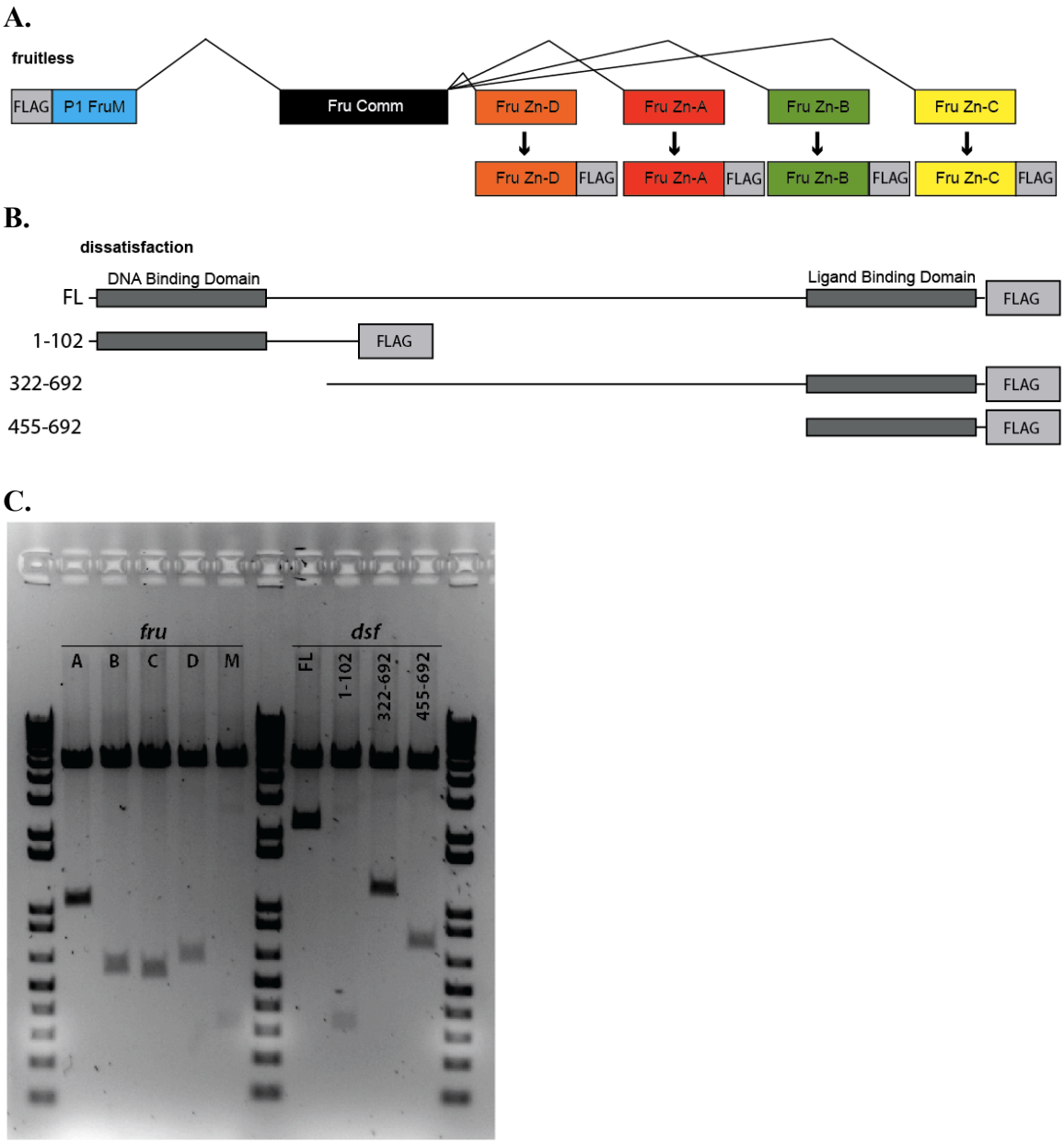
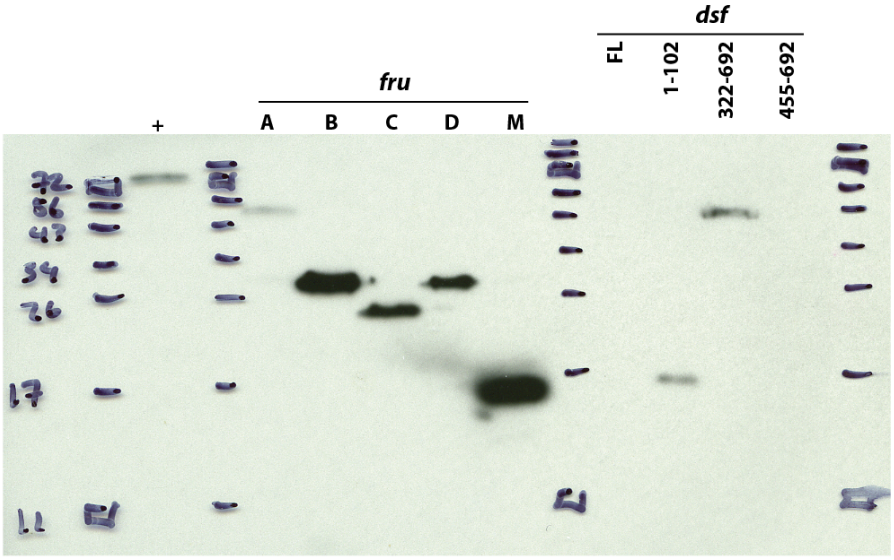


Figure 7 – Cloning *fru* and *dsf* for *in vitro* experiments. A. A schematic showing the cloning strategy for the individual *fru* zinc-finger containing exons (A, B, C and D) and the N-terminal M exon. **B.** A schematic showing the cloning strategy for *dsf* full length and truncations, aminio acids 1-102, 322-692 and 455-695. **C.** *fru* and *dsf* clones in pT7-MAT-FLAG2® excised by EcoRI/Acc65I digestion. For expected sizes, see Methods and Materials.

A.



B.

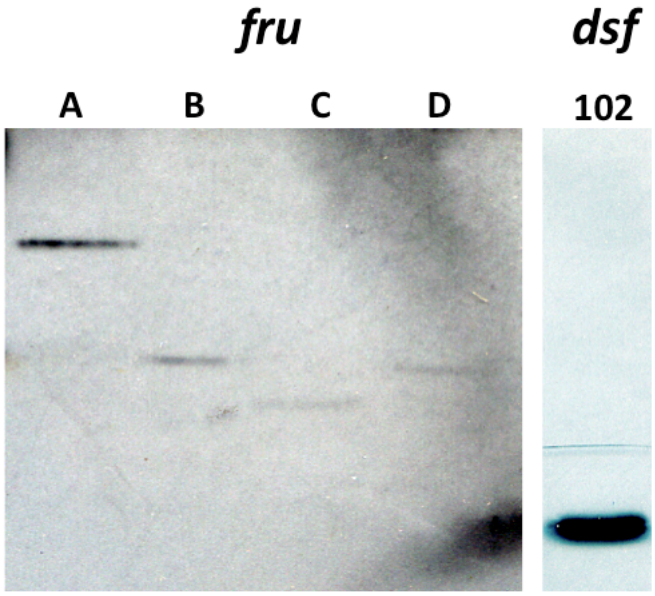
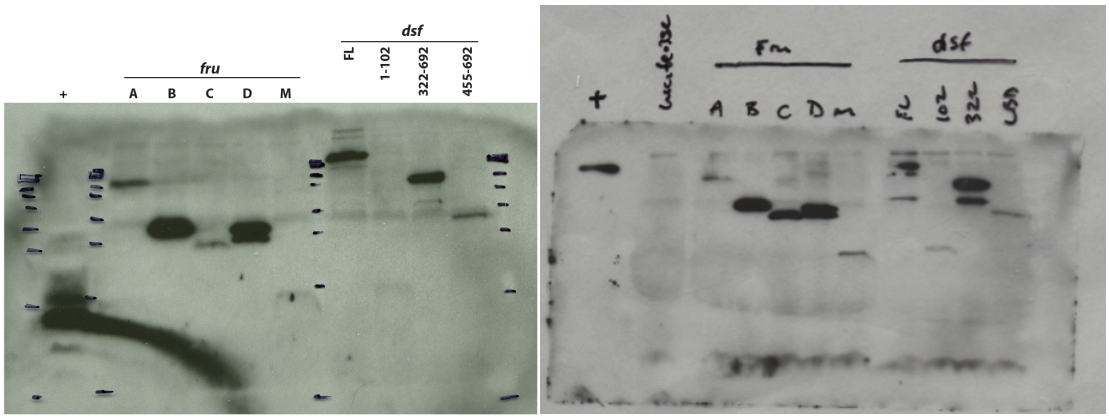


Figure 8 – Bacterial Expression of *fru* and *dsf*.
A. Anti-FLAG western blot showing IPTG induced bacterial protein extracts of *fru* and *dsf* clones. *dsf*-FL from a previous induction was used as a technical control. **B.** Anti-FLAG western blot showing French press lysed IPTG induced bacterial protein extracts of *fru* isoforms and *dsf*-102.

A.



B.

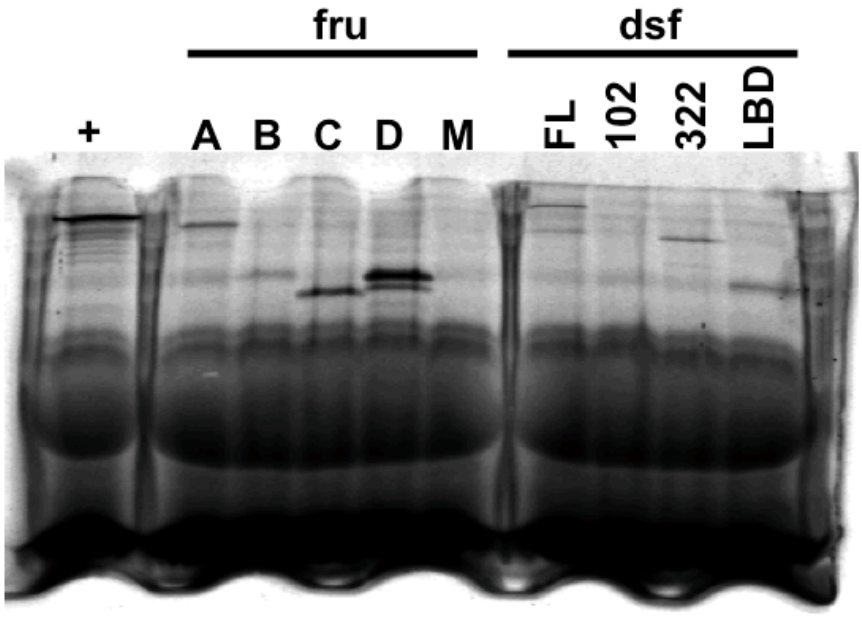
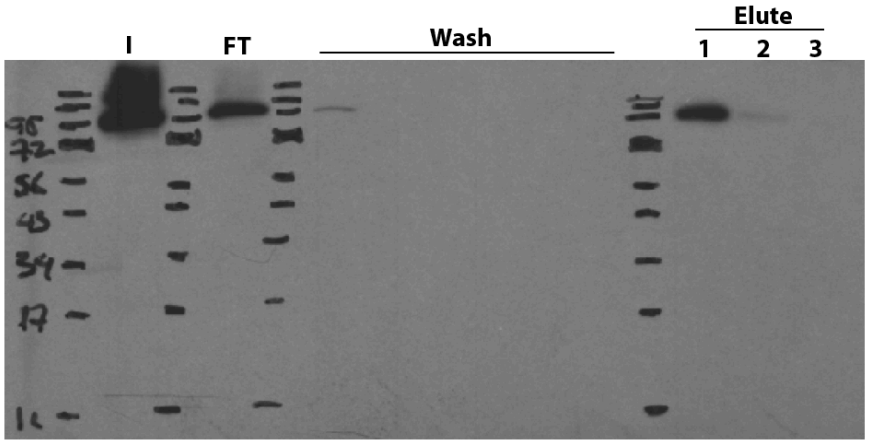


Figure 9 – *In vitro* transcription and translation of *fru* and *dsf*.
A. Anti-FLAG western blot showing *fru* and *dsf* clones expressed in TnT® Quick Coupled Transcription/Translation System. Bacterially expressed *dsf*-FL was used as a technical control (left gel). **B.** Fluorescence imaging of *in vitro* translated *fru* and *dsf* clones incorporating FluoroTect™GreenLys labeled lysine. Gel corresponds to left gel above.

A.



B.

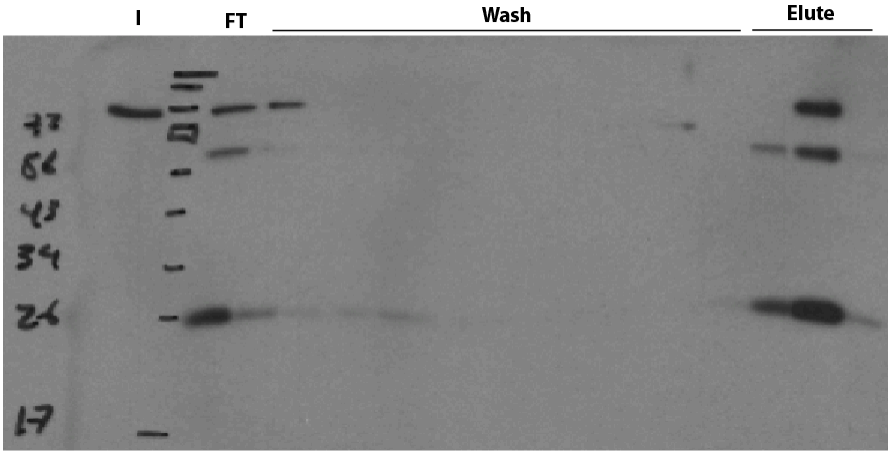
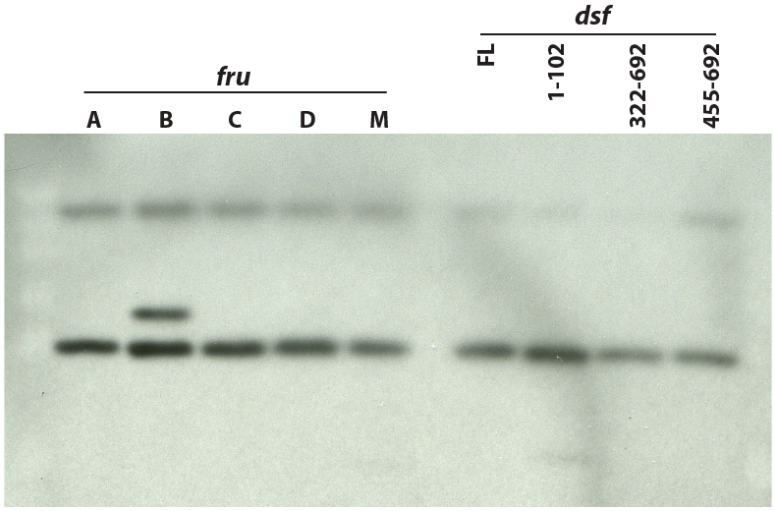


Figure 10 – Purifying *fru* and *dsf* for *in vitro* experiments.

- **A.** Representative anti-FLAG western blot showing nickel column purification of *dsf*-FL. (I – Input, FT – flowthrough)
- **B.** Representative anti-FLAG western blot showing anti-FLAG immunoprecipitation of *dsf*-FL. Note the antibody heavy and light chains co-eluting with the purified protein. (I – Input, FT – flowthrough)

A.



B.

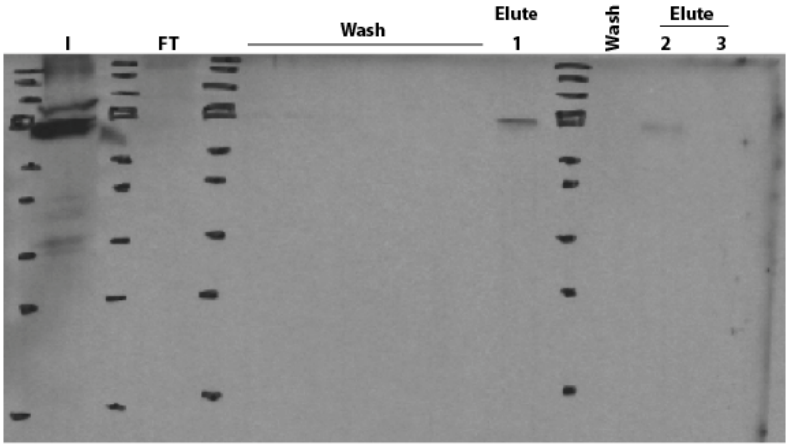


Figure 11 – Occlusion of IP by antibody bands

- **A.** Representative anti-FLAG western blot showing elution fractions from a SELEX round showing antibody band co-elution occluding proteins (Also seen in **Figure 10B.**). *fru*-B can be clearly seen in lane 2 and *dsf*-102 is faintly visible in lane 7.
- **B.** Immunoprecipitating *dsf*-FL using ANTI-FLAG® M2 Magnetic Beads with the antibody conjugated to the beads relieved co-elution of antibody bands. Note that the last wash and the first elution was accidentally switched on the gel.

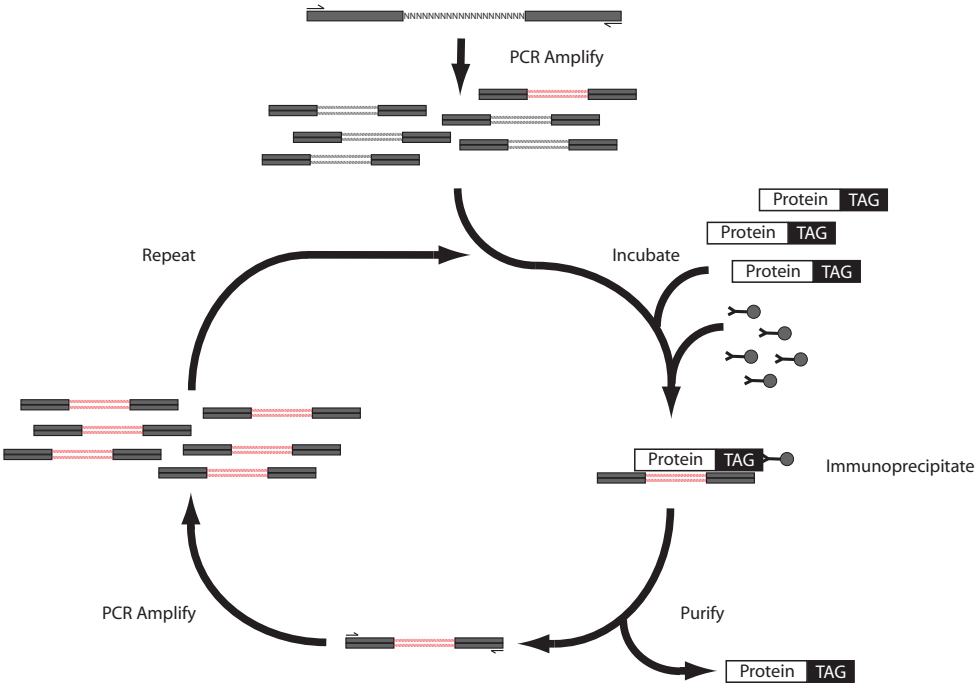
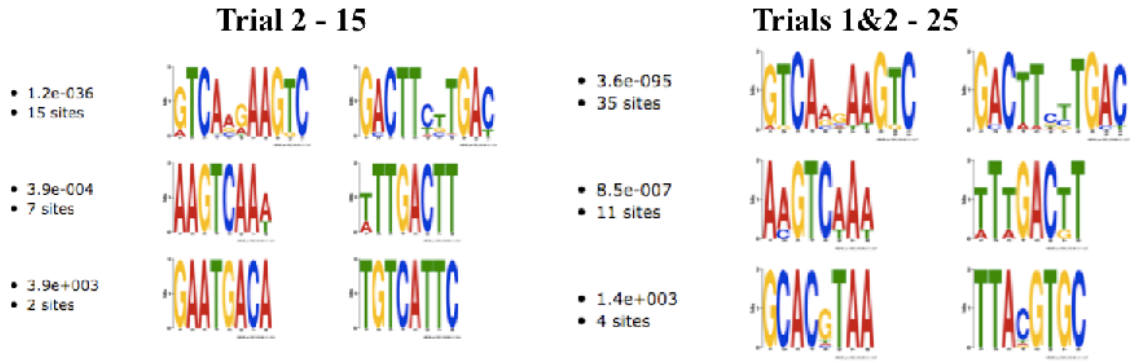


Figure 12 – Systematic Evolution of Ligands by Exponential Enrichment (SELEX) Double stranded 20N random oligo probes are generated by PCR amplification from the constant flanking sequences. This random pool of probes potentially contains a sequence of interest (red) that the target protein binds specifically. DNA-protein interactions are assayed by incubating the probes and an *in vitro* recombinant protein together and immunoprecipitating selectively bound probes. The recovered probes are amplified by PCR and the purification is repeated, iteratively enriching for specifically bound sequences in each round.

A.



B.

```
          GTCAAGAAGTC AACTCAGGA
AAGTCAAAC GTCAAGAAGTC
          GTCAAGAAGTC AACTCAGGA
AAGTCAAAC GTCAAGAAGTC
AAGTCAAAC GTCAAGAAGTC A
AAGTCAAAC GTCAAGAAGTC A
AAGTCAAAC GTCAAGAAGTC A
TACGGAACG GTCAAGAAGTC A
AAGTCAAAC GTCAAGAAGTC A
AAGTCAAAC GTCAAGAAGTC A
      GCG GTCAGGAAGTC ATGTTG
AGCACGTAA GTCACGAAGTC
GAGTGTCCC GTCAACAAGTC
AGCACGTAA GTCACGAAGTC
GAGTGTGCC GTCAACAAGTC
       TAA GTCACGAAGTC AATGTC
AGCACGTAA GTCACGAAGTC
CGCACNTAA GTCACGAAGTC
GATTGTGCA GTCAAAAAGTC
GATTGTGCA GTCAAAAAGTC
GGGTGGGAA GTCAGCAAGTC
GGGTGGGAA GTCAGCAAGTC
 GGACGGCC GTCAATAAGTC A
AGTGCCAAA GTCAAGAAGGC
GGTGGTAGA GTCAGTAAGTC
GGTGGTAGA GTCAGTAAGTC
ACGTCTAAA GTCAACATGTC
ACGTCTAAA GTCAACATGTC
 GGCAAGAG GGCAGGAAGTC A
 GGCAAGAG GGCAGGAAGTC A
GAATGACAA ATCAGAAAGTC A
GAATGACAA ATCAGAAAGTC A
GGGCAGNAA GTCNAAAAGTC
       AA GTCAAGTTGGC TGAAGTC
       AA GTCAAGTTGGC TGAAGTC
```

Figure 13 – *dsf* DNA binding specificity by SELEX

A. Consensus DNA binding sequence for *dsf*-102 determined by MEME is coincident with the previously published (Pittman, 2002) AAGTCA *dsf* half-site. Consistency between trials showed that every read sequenced contained at least a single *dsf* consensus sequence consistent with monomeric binding expected for *dsf*-102.

B. A sample alignment of *dsf*-102 SELEX results shows no positional or flanking sequence bias for *dsf* half-site placement.

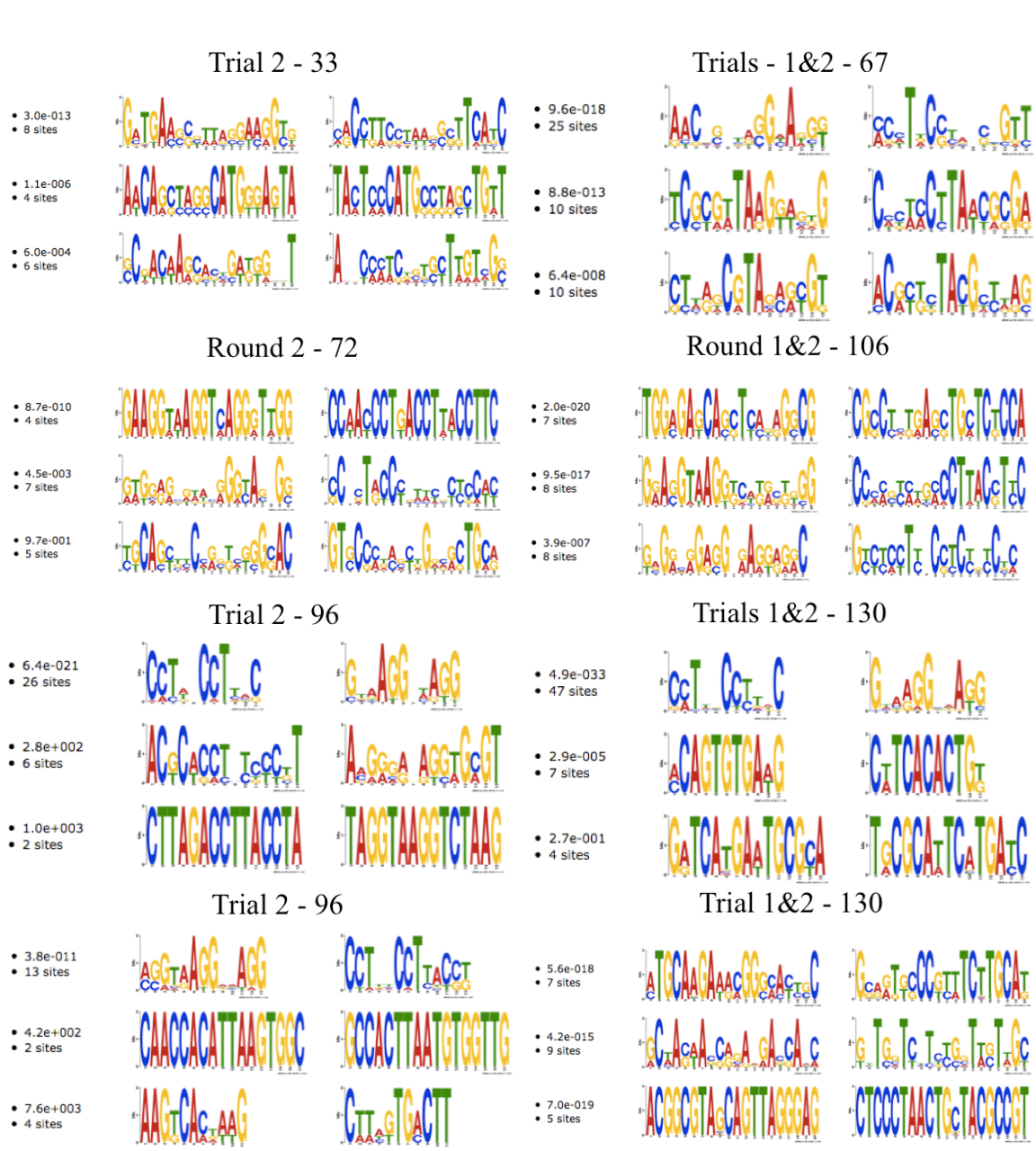


Figure 14 – Lack of *fru* DNA binding specificity by SELEX
MEME output for *fru*-A (**A.**), -B (**B.**), -C (**C.**) and -D (**D.**) SELEX experiments fails to show any sequence specific DNA binding activity for any of the *fru* isoforms. *fru*-C shows the strongest consensus sequence (AGGnnAGG) through confluence between trials. As the AGG motif appears in all of the samples, this could represent a background signal in the absence of specific DNA binding.

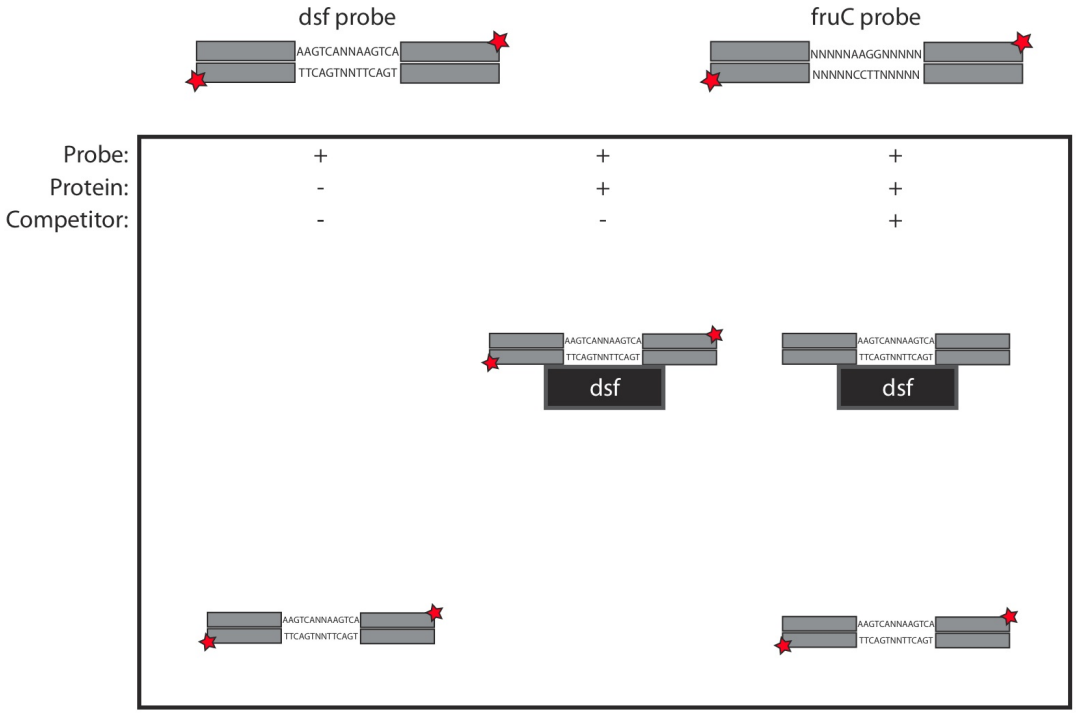


Figure 15 – Electrophoretic Mobility Shift Assay (EMSA)
EMSA provides a test for direct protein-DNA interaction *in vitro*. Migration of a labeled probe is limited in a native acrylamide gel when it is bound by protein. Specific probe-protein interaction is examined by inclusion of an unlabeled probe in molar excess. Additionally, antibodies specific to the protein can further restrict probe migration and directly identify the protein that the DNA is binding.

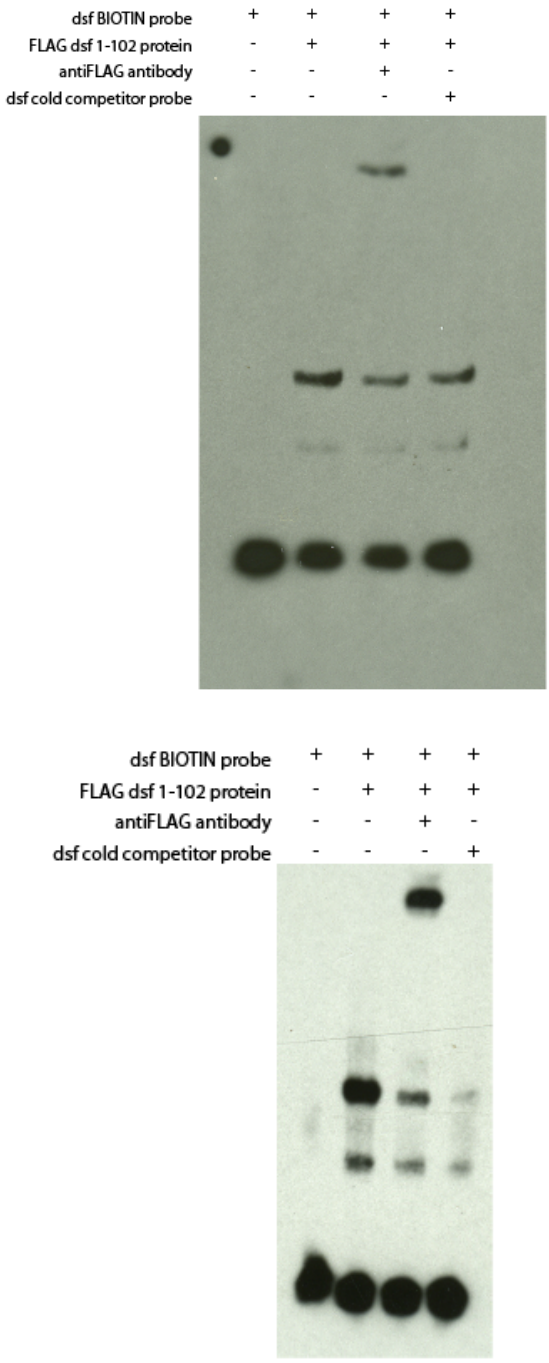
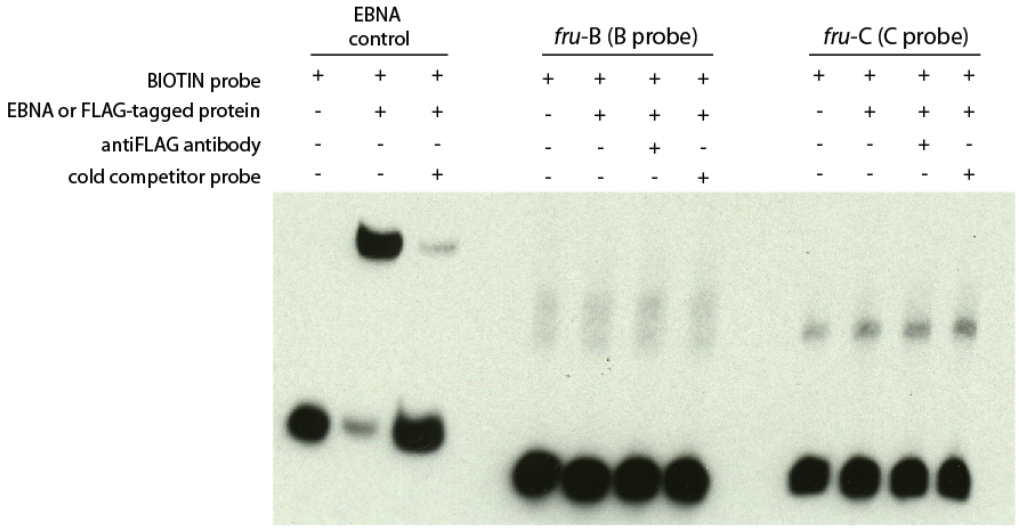


Figure 16 – *dsf*-102 EMSA
Consistent with published results (Pittman, 2002) and the findings of the *dsf*-102 SELEX experiment (**Figure 13**), *dsf*-102 specifically binds the AAGTCA half-site *in vitro*. Supershifting is observed with the anti-FLAG antibody.

A.



B.

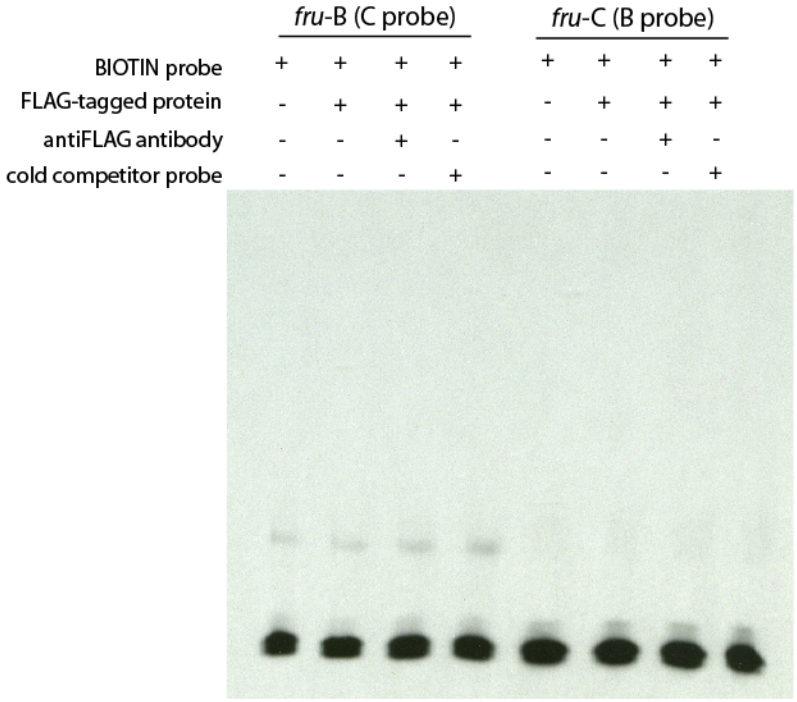
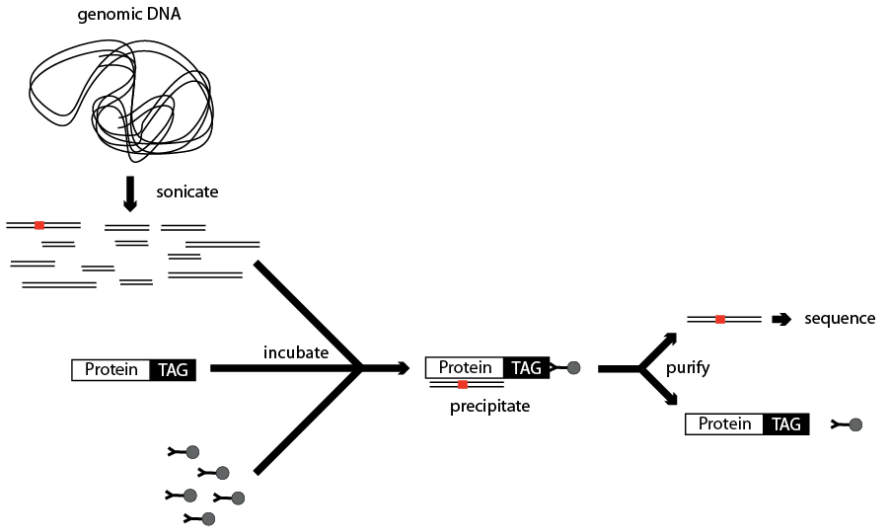
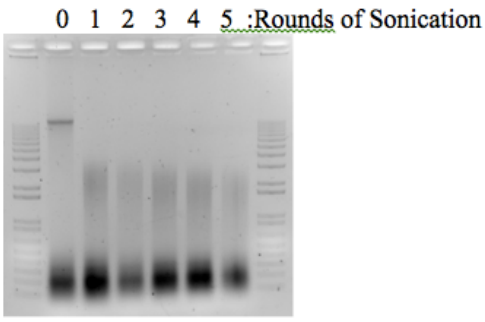


Figure 17 – *fru*-B and C EMSA
fru-B and –C fail to specifically bind putative consensus sequences from SELEX, regardless of if the intervening nucleotides are fixed (B probe: AGGtaAGGtaAGG) or variant (C probe: AGGnnAGGnnAGG). Positive control EBNA shifting was observed under the same conditions.

A.



B.



C.

Round 1

Sample	µg/mL	16µL (ng)
+	6.43	
A-	0.1	1.6
A+	0.307	4.912
B-	0.17	2.72
B+	0.309	4.944
C-	0.224	3.584
C+	0.371	5.936
D-	n/a	
D+	0.418	6.688
FL-	0.14	2.24
FL+	0.668	10.688
102-	0.218	3.488
102+	5.22	83.52

C- and C+ were accidentally mixed.

Round 1: C and 102

Sample	µg/mL	16µL (ng)
+	6.28	
C-	0.12	1.92
C+	1.09	17.44
102-	0.19	3.04
102+	7.48	119.68

Round 2

Sample	µg/mL	16µL
A-	0.12	1.92
A+	1.3	20.8
B-	0.14	2.24
B+	0.461	7.376
C-	0.434	6.944
C+	1.3	20.8
D-	0.408	6.528
D+	0.973	15.568

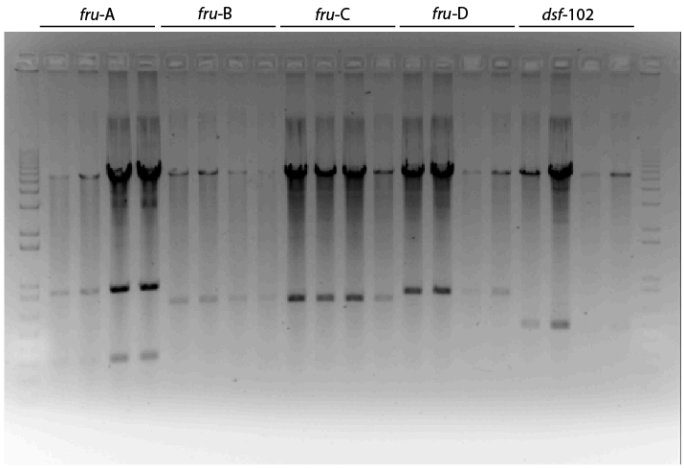
Figure 18 – Protein Binding-Sequencing (PB-Seq) Assay

A. Sonicated gDNA with a potential binding site (red) is incubated with bacterially expressed protein, purified by immunoprecipitation and sequenced.

B. Sonication efficiency increases with each round.

C. Sample Qubit quantitation data shows enrichment in the amount of DNA recovered with proteins included (+) over beads alone (-). Note the increase specifically seen with *dsf*-FL and especially *dsf*-102.

A.



B.

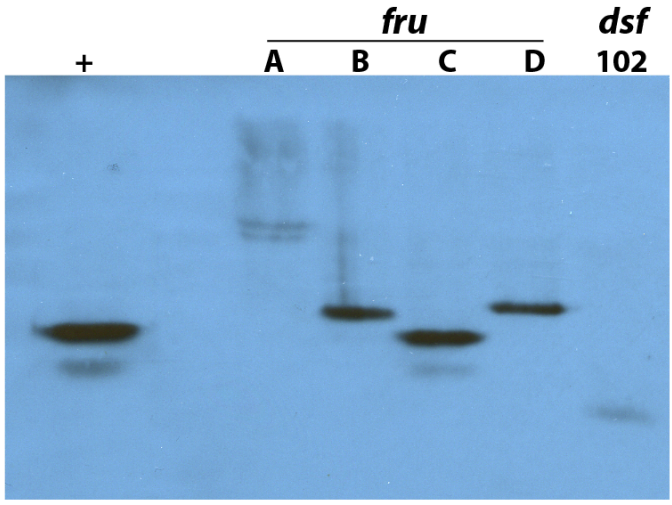
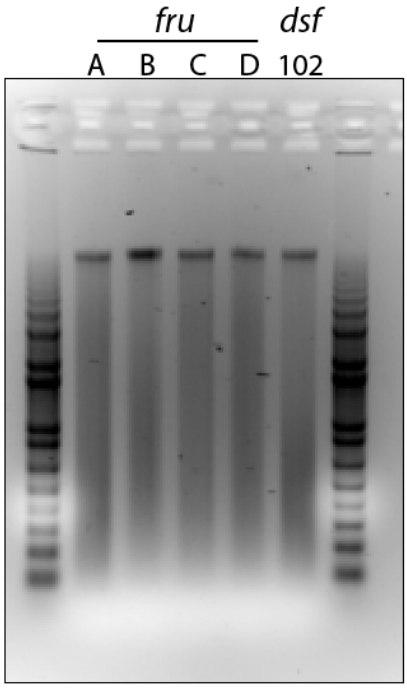


Figure 19 – Gateway Cloning and S2 Cell Expression of *fru* and *dsf*

A. *fru* and *dsf* clones excised from pAWF destination vector with EcoRV/NheI double digest. Note that EcoRV cuts in *fru*-A at position 259.

B. Representative anti-FLAG western blot showing expression of *fru* and *dsf* pieces transiently transfected in S2 cells. S2 cell derived *fru*-C from a previous transfection was used as a technical control.

A.



B.

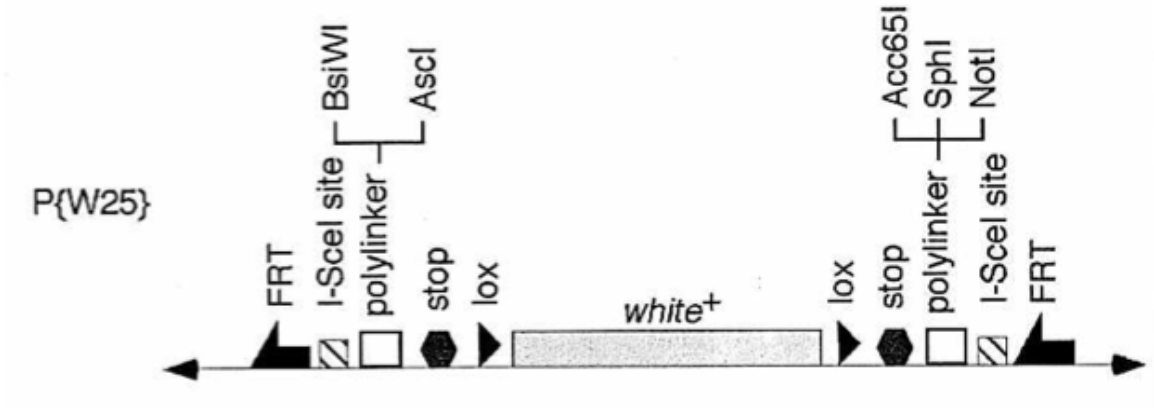
Sample	µg/mL	30µg
+	9.95	na
A	42.7	0.702576112
B	24.8	1.209677419
C	29	1.034482759
D	28	1.071428571
102	34.8	0.862068966

Figure 20 – Chromatin Immunoprecipitaion of S2 cell-expressed *fru* and *dsf*

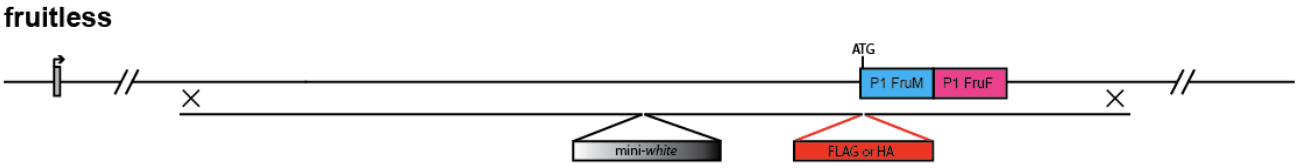
A. Sample input chromatin recovered from *fru* and *dsf* expressing S2 cell derived chromatin preparations

B. Qubit quantitation of concentrated ChIP input DNA

A.



B.



C.

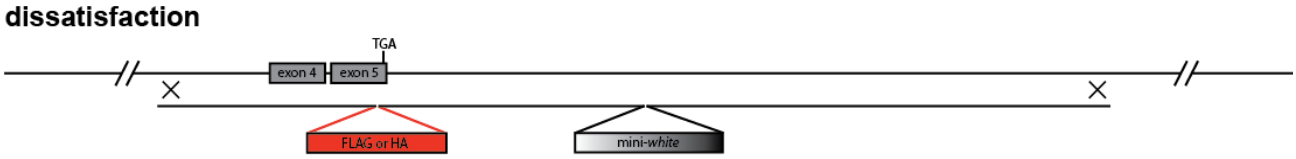


Figure 21 – Homologous Recombination (HR) Targeting Constructs

A. p{w25} vector map showing mini-*white* gene, targeting arm polylinker sites (BsiWI/AscI and Acc65I/SphI/NotI), FRT and I-SceI recognition sequences.
B. N-terminal targeting construct for the *fru* locus showing start site of translation and relative positions of the inserted FLAG or HA tag and mini-*white* gene
C. C-terminal targeting construct for the *dsf* locus showing stop codon and relative positions of the inserted FLAG or HA tag and mini-*white* gene

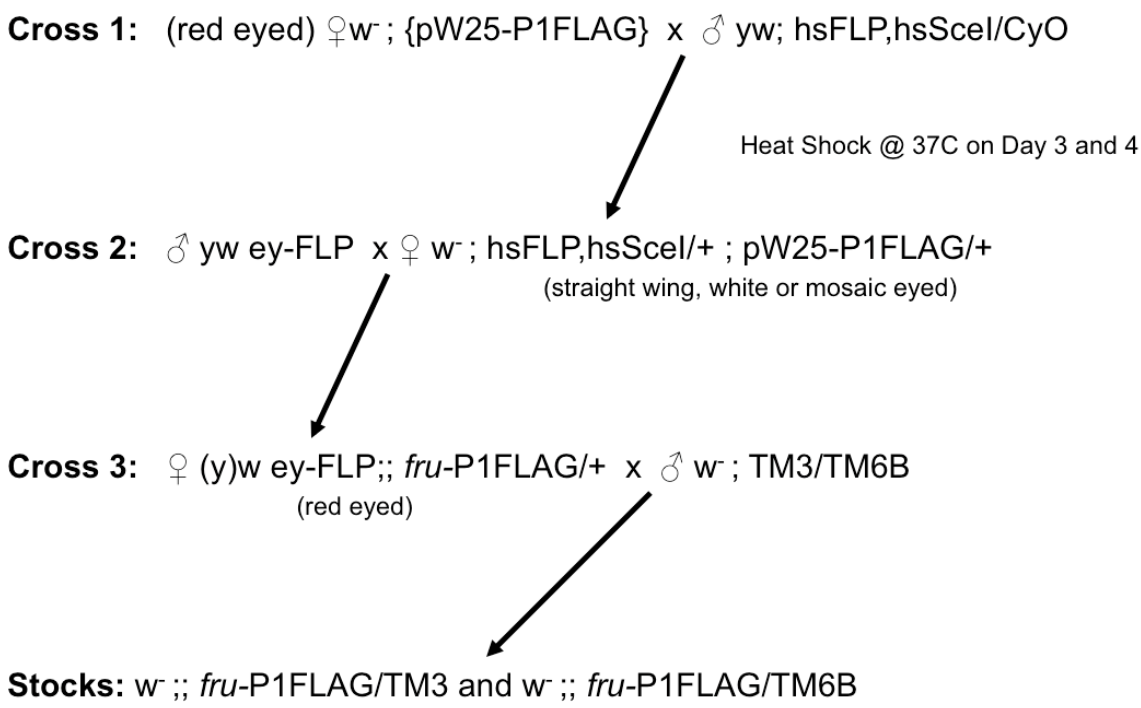
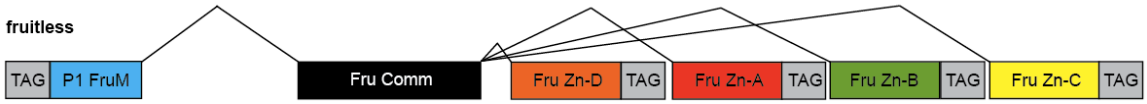


Figure 22 – Targeting Crosses for *fru* Homologous Recombination
The crossing scheme required to move the p{w25} *fru* targeting constructs from their randomly integrated injected positions to their corresponding genomic locations. **Cross 1** involves heat-shock promoter driven I-SceI and FLP mediated excision of the targeting arms from the genome. This extrachromosomal DNA is now competent to homologously recombine with its targeted locus. The loss of red eye color in the progeny of this cross is indicative of chromosomal excision of the mini-*white* gene located in between the homology arms. **Cross 2** uses *eyeless*-FLP to eliminate false positives from unsuccessful genomic excision of the targeting construct. Flies in which homologous recombination has taken place will have lost the FRT sites and their red eye color will be therefore resistant to FLP. The red-eyed progeny of this cross can be allelically balanced using third chromosome balancers in **Cross 3** to establish **Stocks** that can be genotyped to assess the accuracy of the targeting. The strategy is the same for *dsf*, except that owing to its location of chromosome 2, the h/s I-SceI, h/s FLP must be on the third chromosome to avoid targeting in that background and stocks are established using second chromosome balancers.

A.



B.

Tagging the *fru* locus by homologous recombination:

P1-FLAG: In flies and targeted by HR confirmed
P1-HA: In flies and targeted by HR confirmed

A-FLAG: In flies and targeted by HR confirmed
A-HA: Injected into flies and awaiting HR

B-FLAG: Cloned into pW25 and ready for injection
B-HA: Cloned into pW25 and ready for injection

C-FLAG: Cloned into pW25 and ready for injection
C-HA: Cloned into pW25 and ready for injection

D-FLAG: Injected into flies and awaiting HR
D-HA: Injected into flies and awaiting HR

C.

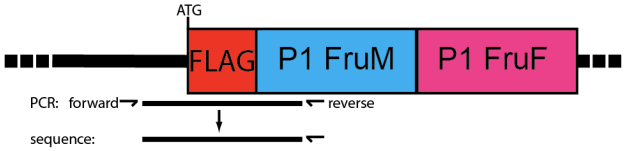
Figure 23 – *In Vivo* Targeting Strategy for *fru* and *dsf* HR

A. Schematic of the targeting strategy for tagging *fru*-P1 and each isoform individually with either a FLAG or HA epitope tag

B. Current state of *fru* HR

C. Schematic of the targeting strategy for tagging *dsf* N-terminally with either a FLAG or HA epitope tag (Fei Fei Ding).

A.

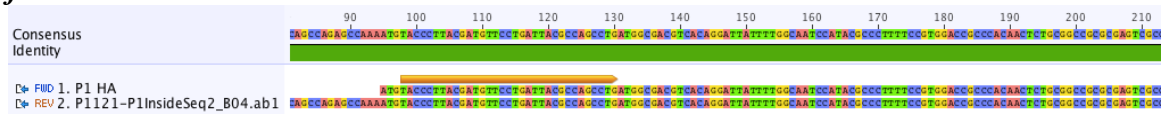


B.

Canton S

***fru* P1-FLAG**

***fru* P1-HA**



C.

***fru* A-FLAG**

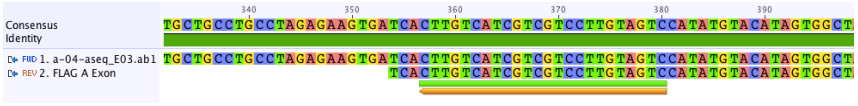
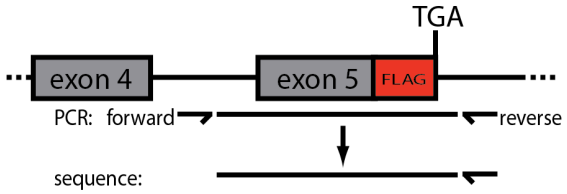


Figure 24: Genotyping *fru* Homologous Recombination

A. Sequencing primers for *fru* N-terminal integration
B. Sample sequence alignment for *fru* P1-FLAG and *fru* P1-HA PCR sequencing results. Note the absence of FLAG tag in the canton S sample.
C. Sample sequence alignment for the C-terminal *fru* A-FLAG PCR sequencing results.

A.



B.

***dsf*-FLAG**

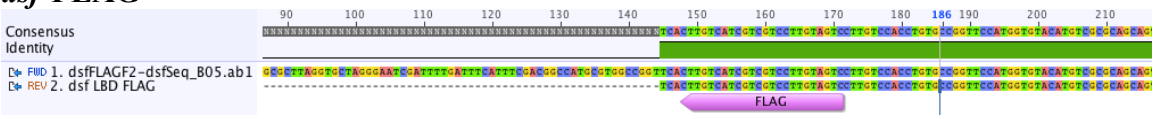
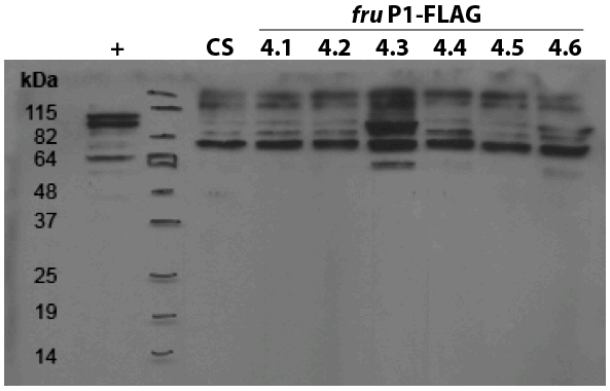


Figure 25: Genotyping *fru* Homologous Recombination

A. Sequencing primers for *dsf* C-terminal integration
B. Sample sequence alignment for *dsf*-FLAG sequencing results.

A.



B.

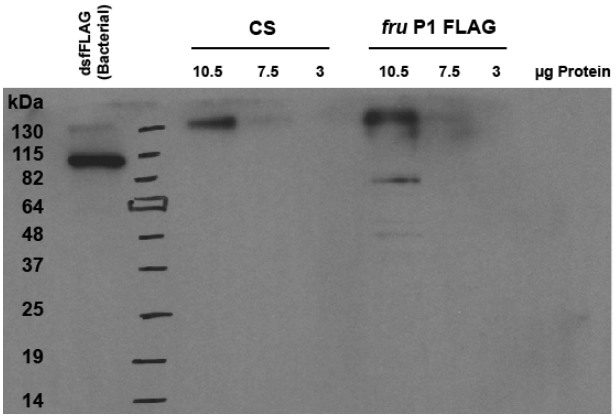
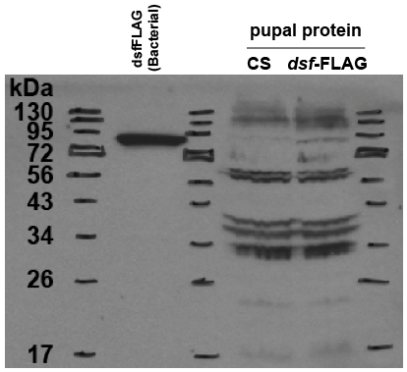


Figure 26 – Protein Expression in Epitope-tagged *fru* Flies

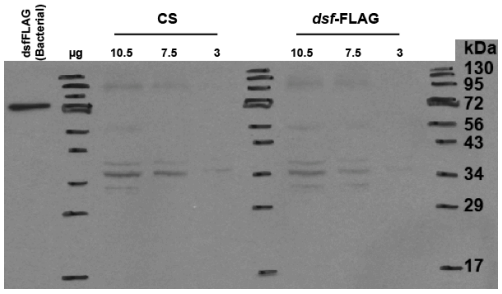
A. Representative anti-FLAG western blot showing lack of discernable FLAG immunoreactivity from *fru* P1-FLAG lines 4.1-4.6. Note background staining in CS sample (lane 2). While line 4.3 was promising in the sizes of the additional bands seen in this lane, this result could not be reproduced with decreasing amounts of protein loaded or subsequent immunoprecipitation experiments. Bacterial *dsf*-FL was used as a technical control.

B. Dilution anti-FLAG western blot to account for CS background seen in **A.** Unfortunately, as the background is reduced, so is the potential signal from line 4.3. Bacterial *dsf*-FL was used as a technical control.

A.



B.



C.

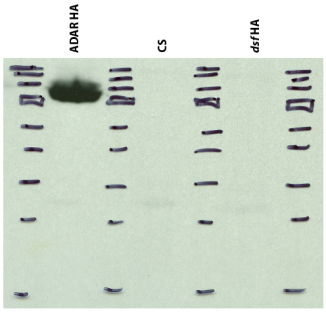
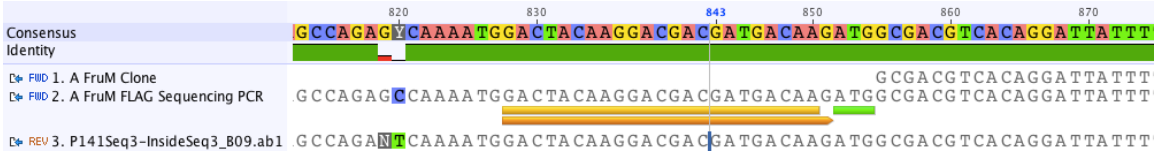


Figure 27 - Protein Expression in Epitope-tagged *dsf* Flies

- **A.** Representative anti-FLAG western blot comparing protein extracted from CS and *dsf*-FLAG pupae. A promising band migrating similarly to bacterial *dsf*-FL in the *dsf*-FLAG sample lane, but this result could not be reproduced with decreasing amounts of protein loaded or subsequent immunoprecipitation experiments. Bacterial *dsf*-FL was used as a technical control.
- **B.** Dilution anti-FLAG western blot to account for CS background seen in **A.** Unfortunately, as the background is reduced, so is the potential signal from *dsf*-FLAG. Bacterial *dsf*-FL was used as a technical control.
- **C.** Representative anti-HA western blot comparing protein extracted from ADAR-HA (Gift from the Reenan lab) adults and CS and *dsf*-FLAG pupae. ADAR-HA flies were used as a technical control for protein extraction, loading and antibody specificity.

A.

fru P1-FLAG mRNA



B.

dsf-FLAG mRNA

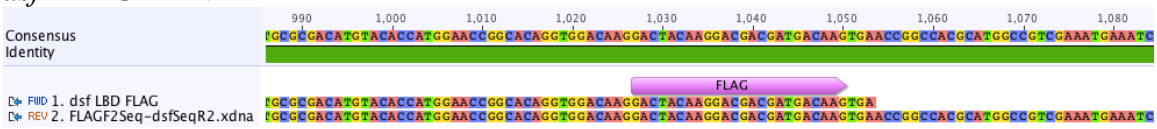
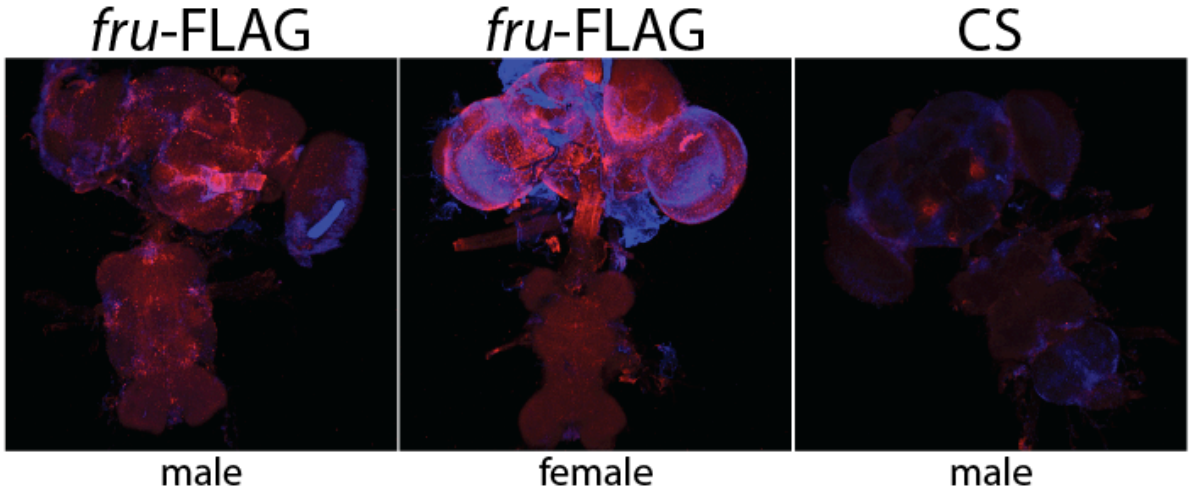


Figure 28 – mRNA Genotyping of *fru* P1-FLAG and *dsf*-FLAG Flies

- **A.** Representative sequence alignment for *fru* P1-FLAG fly derived mRNA sequenced as described in **Figure 24A.**
- **B.** Representative sequence alignment for *dsf*-FLAG fly derived mRNA sequenced as described in **Figure 25A.**

A.



B.

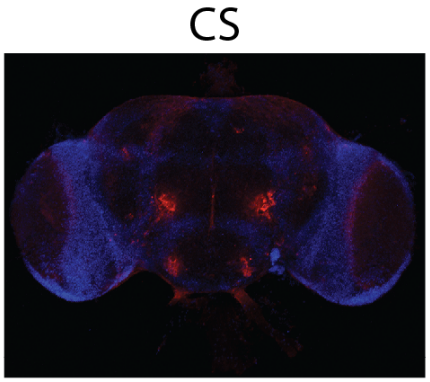


Figure 29 – *fru* P1-FLAG Immunohistochemistry

- **A.** Dorsal view of male and female (ventral) *fru* P1-FLAG adult brain and VNC stained with anti-FLAG antibody (red) shows specific staining in the published P1 domains, specifically [–paper]. No specific staining is observed in CS. Blue staining is DAPI nuclear stain.
- **B.** Anterior view of male CS adult brain stained with anti-FLAG antibody (red) shows no specific staining. Blue staining is DAPI nuclear stain.

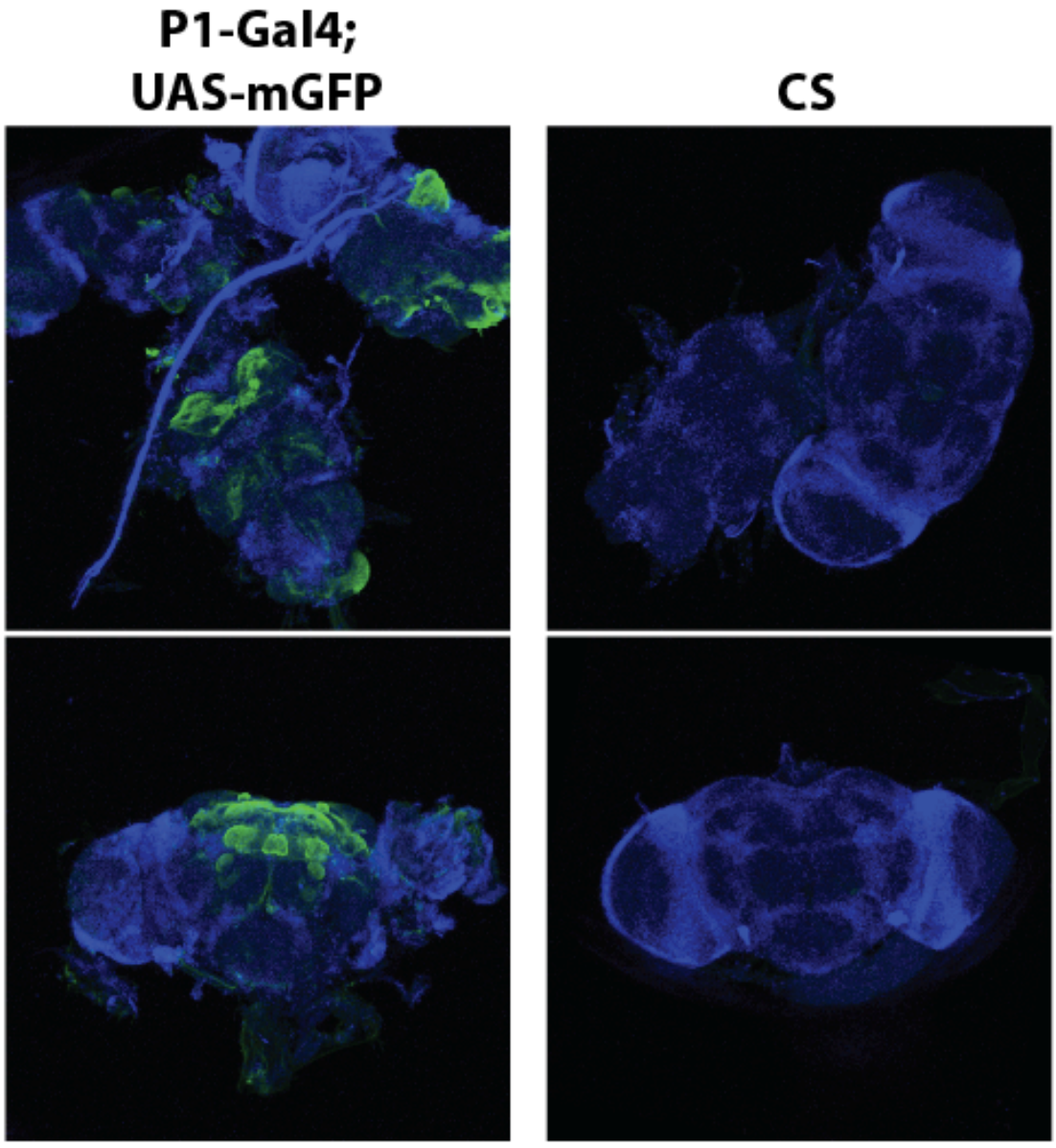


Figure 30 – *fru* P1-Gal4; UAS-mGFP Immunohistochemistry
Dorsal view of male *fru* P1-Gal4; UAS-mGFP and CS adult brain and VNC stained with anti-GFP antibody (green) labels neurites in the published P1 domains, specifically the DA1 glomerulus in the AL and the midline crossing observed in GRNs in the VNC. No specific staining is observed in CS. Blue staining is DAPI nuclear stain.

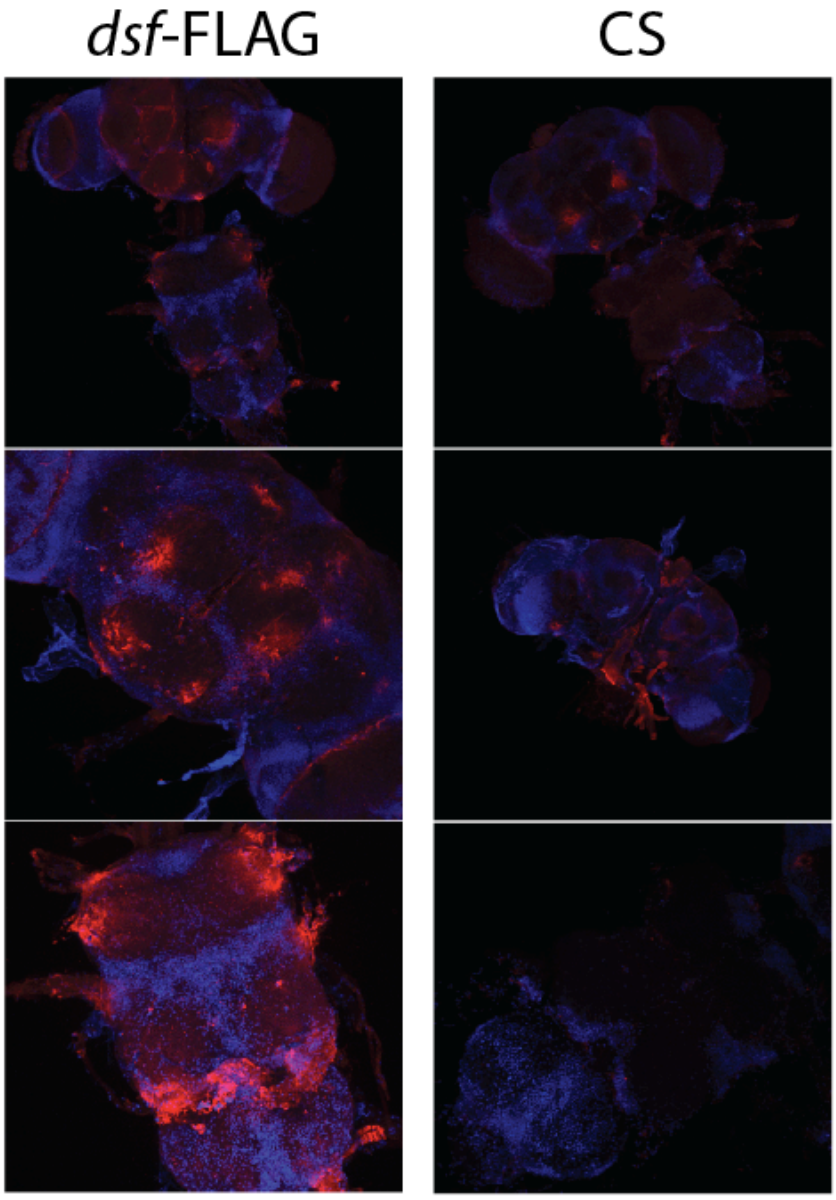


Figure 31 - *dsf*-FLAG Immunohistochemistry
Dorsal view of male *dsf*-FLAG and CS adult brain and VNC stained with anti-FLAG antibody (red) shows no specific staining in either *dsf*-FLAG or CS animals. Blue staining is DAPI nuclear stain.

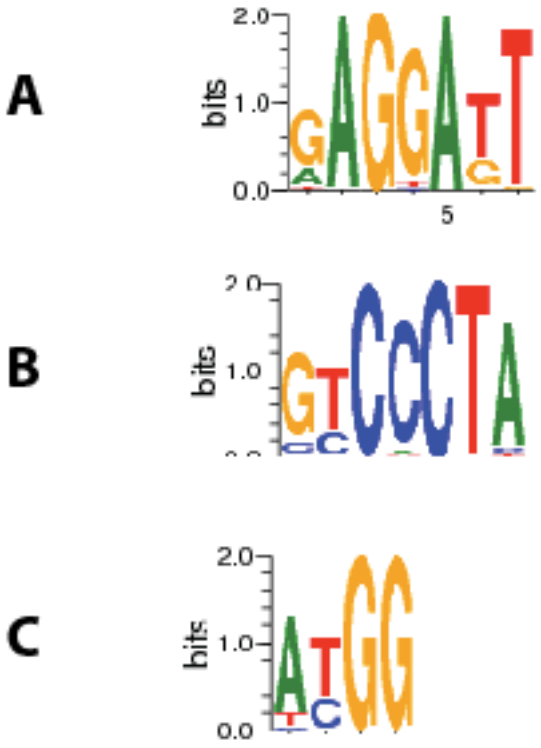


Figure 32 – Predicted DNA binding sequences for *fru* isoforms
The protein sequences for *fru*-A, -B and –C were input into a zinc-finger DNA binding prediction algorithm (http://zf.princeton.edu/).

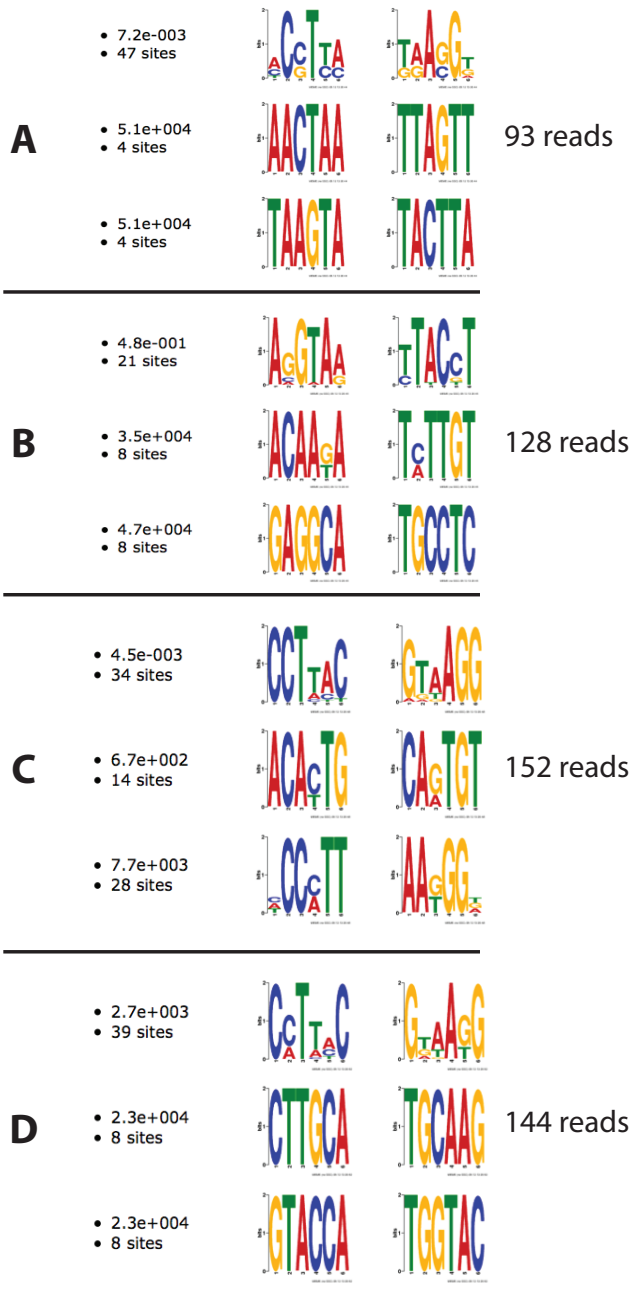


Figure 33 – Repeated motif finding for *fru* SELEX by MEME
The SELEX results for the four *fru* isoforms were reanalyzed in MEME with the parameters of 3 nucleotide minimum motif width and 6 nucleotide motif width.

Literature Cited

Albagli O, Dhordain P, Deweindt C, Lecocq G, Leprince D. (1995). The BTB/POZ domain: a new protein-protein interaction motif common to DNA- and actin-binding proteins. Cell Growth Differ. Sep;6(9):1193-8.

An W, Wensink PC. (1995). Integrating sex- and tissue-specific regulation within a single Drosophila enhancer. Genes Dev. Jan 15;9(2):256-66.

An W, Cho S, Ishii H, Wensink PC. (1996). Sex-specific and non-sex-specific oligomerization domains in both of the doublesex transcription factors from Drosophila melanogaster. Mol Cell Biol. Jun;16(6):3106-11.

Anand A, Villella A, Ryner LC, Carlo T, Goodwin SF, Song HJ, Gailey DA, Morales A, Hall JC, Baker BS, Taylor BJ. (2001). Molecular genetic dissection of the sex-specific and vital functions of the Drosophila melanogaster sex determination gene fruitless. Genetics. Aug;158(4):1569-95.

Bailey, Timothy L. and Elkan, Charles. "Fitting a mixture model by expectation maximization to discover motifs in biopolymers", *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pp. 28-36, AAAI Press, Menlo Park, California, 1994.

Baker BS, Ridge KA. (1980). Sex and the single cell. I. On the action of major loci affecting sex determination in Drosophila melanogaster. Genetics. Feb;94(2):383-423.

Bellaiche, Y., Mogila, V., Perrimon, N. (1999). I-SceI endonuclease, a new tool for studying DNA double-strand break repair mechanisms in Drosophila. Genetics Jul;152(3):1037-44.

Boerjan B, Tobback J, Vandersmissen HP, Huybrechts R, Schoofs L. (2012). Fruitless RNAi knockdown in the desert locust, Schistocerca gregaria, influences male fertility. J Insect Physiol. Feb;58(2):265-9.

Boggs RT, Gregor P, Idriss S, Belote JM, McKeown M. (1987). Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. Cell. 1987 Aug 28;50(5):739-47.

Billeter JC, Goodwin SF. (2004). Characterization of Drosophila fruitless-gal4 transgenes reveals expression in male-specific fruitless neurons and innervation of male reproductive structures. J Comp Neurol. Jul 19;475(2):270-87.

Billeter JC, Rideout EJ, Dornan AJ, Goodwin SF. (2006a). Control of male sexual behavior in Drosophila by the sex determination pathway. Curr Biol. Sep 5;16(17):R766-76.

Billeter JC, Villella A, Allendorfer JB, Dornan AJ, Richardson M, Gailey DA, Goodwin SF. (2006b). Isoform-specific control of male neuronal differentiation and behavior in Drosophila by the fruitless gene. Curr Biol. Jun 6;16(11):1063-76.

Bopp D, Bell LR, Cline TW, Schedl P. (1991). Developmental distribution of female-specific Sex-lethal proteins in Drosophila melanogaster. Genes Dev. Mar;5(3):403-15.

Burnet B, Connolly K, Kearney M, Cook R. (1973). Effects of male paragonial gland secretion on sexual receptivity and courtship behaviour of female Drosophila melanogaster. J Insect Physiol. Dec;19(12):2421-31.

Burtis KC, Baker BS. (1989). Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. Cell. Mar 24;56(6):997-1010.

Clyne JD, Miesenböck G. (2008). Sex-specific control and tuning of the pattern generator for courtship song in Drosophila. Cell. Apr 18;133(2):354-63.

Clynen E, Ciudad L, Bellés X, Piulachs MD. (2011). Conservation of fruitless' role as master regulator of male courtship behaviour from cockroaches to flies. Dev Genes Evol. May;221(1):43-8.

Couderc JL, Godt D, Zollman S, Chen J, Li M, Tiong S, Cramton SE, Sahut-Barnola I, Laski FA. (2002). The bric à brac locus consists of two paralogous genes encoding BTB/POZ domain proteins and acts as a homeotic and morphogenetic regulator of imaginal development in Drosophila. Development. May;129(10):2419-33.

Dalton JE, Fear JM, Knott S, Baker BS, McIntyre LM, Arbeitman MN. (2013). Male-specific fruitless isoforms have different regulatory roles conferred by distinct zinc finger DNA binding domains. BMC Genomics. Sep 27;14(1):659. [Epub ahead of print]

Datta SR, Vasconcelos ML, Ruta V, Luo S, Wong A, Demir E, Flores J, Balonze K, Dickson BJ, Axel R. (2008). The Drosophila pheromone cVA activates a sexually dimorphic neural circuit. Nature. Mar 27;452(7186):473-7.

Dauwalder B, Tsujimoto S, Moss J, Mattox W. (2002). The Drosophila takeout gene is regulated by the somatic sex-determination pathway and affects male courtship behavior. Genes Dev. Nov 15;16(22):2879-92.

Demir E, Dickson BJ. (2005). fruitless splicing specifies male courtship behavior in Drosophila. Cell. Jun 3;121(5):785-94.

Djordjevic M. (2007). SELEX experiments: new prospects, applications and data analysis in inferring regulatory pathways. Biomol Eng. 24(2):179-89.

Drapeau MD, Radovic A, Wittkopp PJ, Long AD. (2003). A gene necessary for normal

male courtship, yellow, acts downstream of fruitless in the Drosophila melanogaster larval brain. J Neurobiol. 2003 Apr;55(1):53-72.

Emery IF, Bedian V, Guild GM. (1994). Differential expression of Broad-Complex transcription factors may forecast tissue-specific developmental fates during Drosophila metamorphosis. Development. Nov;120(11):3275-87.

Erdman SE, Burtis KC. (1993). The Drosophila doublesex proteins share a novel zinc finger related DNA binding domain. EMBO J. Feb;12(2):527-35

Erdman SE, Chen HJ, Burtis KC. (1996). Functional and genetic characterization of the oligomerization and DNA binding properties of the Drosophila doublesex proteins. Genetics. Dec;144(4):1639-52.

Fabre CC, Hedwig B, Conduit G, Lawrence PA, Goodwin SF, Casal J. (2012). Substrate-borne vibratory communication during courtship in Drosophila melanogaster. Curr Biol. Nov 20;22(22):2180-5.

Feng J, Liu T, Zhang Y. (2011). Using MACS to identify peaks from ChIP-Seq data. Curr Protoc Bioinformatics. Jun;Chapter 2:Unit 2.14.

Ferri SL, Bohm RA, Lincicome HE, Hall JC, Villella A. (2008). fruitless Gene products truncated of their male-like qualities promote neural and behavioral maleness in Drosophila if these proteins are produced in the right places at the right times. J Neurogenet. Jan-Mar;22(1):17-55.

Ferveur JF, Störtkuhl KF, Stocker RF, Greenspan RJ. (1995). Genetic feminization of brain structures and changed sexual orientation in male Drosophila. Science. Feb 10;267(5199):902-5.

Finley KD, Taylor BJ, Milstein M, McKeown M. (1997). dissatisfaction, a gene involved in sex-specific behavior and neural development of Drosophila melanogaster. Proc Natl Acad Sci U S A. Feb 4;94(3):913-8.

Finley KD, Edeen PT, Foss M, Gross E, Ghbeish N, Palmer RH, Taylor BJ, McKeown M. (1998). Dissatisfaction encodes a tailless-like nuclear receptor expressed in a subset of CNS neurons controlling Drosophila sexual behavior. Neuron. Dec;21(6):1363-74.
Fujii S, Amrein H (2002) Genes expressed in the Drosophila head reveal a role for fat cells in sex-specific physiology. EMBO J 21: 5353–5363.

Gailey DA, Lacaillade RC, Hall JC. (1986). Chemosensory elements of courtship in normal and mutant, olfaction-deficient Drosophila melanogaster. Behav Genet. May;16(3):375-405.

Gailey DA, Taylor BJ, Hall JC. (1991). Elements of the fruitless locus regulate

development of the muscle of Lawrence, a male-specific structure in the abdomen of Drosophila melanogaster adults. Development. Nov;113(3):879-90.

Gailey DA, Billeter JC, Liu JH, Bauzon F, Allendorfer JB, Goodwin SF. (2006). Functional conservation of the fruitless male sex-determination gene across 250 Myr of insect evolution. Mol Biol Evol. Mar;23(3):633-43.

Garabedian MJ, Hung MC, Wensink PC. (1985). Independent control elements that determine yolk protein gene expression in alternative Drosophila tissues. Proc Natl Acad Sci U S A. Mar;82(5):1396-400.

Gill, KS. (1963). A mutation causing abnormal courtship and mating behavior in the males of *Drosophila melanogaster*. Am. Zool. 4: 507
Goldman TD, Arbeitman MN. (2007). Genomic and functional studies of Drosophila sex hierarchy regulated gene expression in adult head and nervous system tissues. PLoS Genet. Nov;3(11):e216.

Goldman TD, Arbeitman MN. (2007). Genomic and functional studies of Drosophila sex hierarchy regulated gene expression in adult head and nervous system tissues. PLoS Genet. 2007 Nov;3(11):e216.

Golic, K.G. and Lindquist, S. (1989). The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome. Cell Nov 3;59(3):499-509.

Goodwin SF, Taylor BJ, Villella A, Foss M, Ryner LC, Baker BS, Hall JC. (2000). Aberrant splicing and altered spatial expression patterns in fruitless mutants of Drosophila melanogaster. Genetics. Feb;154(2):725-45.

Greenspan RJ, Ferveur JF. (2000). Courtship in Drosophila. Annu Rev Genet. 34:205-232.

Guertin MJ, Lis JT. (2010). Chromatin landscape dictates HSF binding to target DNA elements. PLoS Genet. Sep 9;6(9):e1001114.

Guertin MJ, Martins AL, Siepel A, Lis JT. (2012) Accurate prediction of inducible transcription factor binding intensities in vivo. PLoS Genetics 8(3):e1002610.

Hall JC. (1978). Courtship among males due to a male-sterile mutation in Drosophila melanogaster. Behav Genet. Mar;8(2):125-41.

Hall JC. (1979). Control of male reproductive behavior by the central nervous system of Drosophila: dissection of a courtship pathway by genetic mosaics. Genetics. Jun;92(2):437-57

Hall JC. (1994). The mating of a fly. Science. 1994 Jun 17;264(5166):1702-14.

Hu S, Fambrough D, Atashi JR, Goodman CS, Crews ST. (1995). The Drosophila abrupt gene encodes a BTB-zinc finger regulatory protein that controls the specificity of neuromuscular connections. Genes Dev. Dec 1;9(23):2936-48.

Ito H, Fujitani K, Usui K, Shimizu-Nishikawa K, Tanaka S, Yamamoto D. (1996). Sexual orientation in Drosophila is altered by the satori mutation in the sex-determination gene fruitless that encodes a zinc finger protein with a BTB domain. Proc Natl Acad Sci U S A. Sep 3;93(18):9687-92.

Ito H, Sato K, Koganezawa M, Ote M, Matsumoto K, Hama C, Yamamoto D. (2012). Fruitless recruits two antagonistic chromatin factors to establish single-neuron sexual dimorphism. Cell. Jun 8;149(6):1327-38.

Johansson BG, Jones TM (2007) The role of chemical communication in mate choice. Biol Rev Camb Philos Soc 82: 265–289.

Jolma A, Kivioja T, Toivonen J, Cheng L, Wei G, Enge M, Taipale M, Vaquerizas JM, Yan J, Sillanpää MJ, Bonke M, Palin K, Talukder S, Hughes TR, Luscombe NM, Ukkonen E, Taipale J. (2010). Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. Genome Res. 20(6):861-73.

Karim FD, Guild GM, Thummel CS. (1993). The Drosophila Broad-Complex plays a key role in controlling ecdysone-regulated gene expression at the onset of metamorphosis. Development. Jul;118(3):977-88.

Keshishian H, Broadie K, Chiba A, Bate M. (1996). The drosophila neuromuscular junction: a model system for studying synaptic development and function. Annu Rev Neurosci. 19:545-75.

Kondoh Y, Kaneshiro KY, Kimura K, Yamamoto D. (2003). Evolution of sexual dimorphism in the olfactory brain of Hawaiian Drosophila. Proc Biol Sci. May 22;270(1519):1005-13.

Kimura K. (2011). Role of cell death in the formation of sexual dimorphism in the Drosophila central nervous system. Dev Growth Differ. Feb;53(2):236-44.

King-Jones K, Thummel CS. (2005). Nuclear receptors--a perspective from Drosophila. Nat Rev Genet. Apr;6(4):311-23.

Krstic D, Boll W, Noll M. (2009). Sensory integration regulating male courtship behavior in Drosophila. PLoS One. 4(2):e4457. doi: 10.1371/journal.pone.0004457.

Kurtovic A, Widmer A, Dickson BJ. (2007). A single class of olfactory neurons mediates behavioural responses to a Drosophila sex pheromone. Nature. 2007 Mar 29;446(7135):542-6.

Laudet V, Adelmant G. (1995). Nuclear receptors. Lonesome orphans. Curr Biol. Feb 1;5(2):124-7.

Lazareva AA, Roman G, Mattox W, Hardin PE, Dauwalder B. (2007). A role for the adult fat body in Drosophila male courtship behavior. PLoS Genet. Jan 26;3(1):e16.

Lee G, Foss M, Goodwin SF, Carlo T, Taylor BJ, Hall JC. (2000). Spatial, temporal, and sexually dimorphic expression patterns of the fruitless gene in the Drosophila central nervous system. J Neurobiol. Jun 15;43(4):404-26.

Lee G, Hall JC. (2001a). Abnormalities of male-specific FRU protein and serotonin expression in the CNS of fruitless mutants in Drosophila. J Neurosci. Jan 15;21(2):513-26.

Lee G, Villella A, Taylor BJ, Hall JC. (2001b). New reproductive anomalies in fruitless-mutant Drosophila males: extreme lengthening of mating durations and infertility correlated with defective serotonergic innervation of reproductive organs. J Neurobiol. May;47(2):121-49.

Li H, Durbin R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. Jul 15;25(14):1754-60.

Lours C, Bardot O, Godt D, Laski FA, Couderc JL. (2003). The Drosophila melanogaster BTB proteins bric à brac bind DNA through a composite DNA binding domain containing a pipsqueak and an AT-Hook motif. Nucleic Acids Res. Sep 15;31(18):5389-98.

Maggert, K.A., Gong, W.J., and Golic, K.G. (2008). Methods for homologous recombination in Drosophila. Methods in molecular biology (Clifton, NJ *420*, 155-174.)

Manning A. (1967). Antennae and sexual receptivity in Drosophila melanogaster females. Science. Oct 6;158(3797):136-7.

Manoli DS, Foss M, Villella A, Taylor BJ, Hall JC, Baker BS. (2005). Male-specific fruitless specifies the neural substrates of Drosophila courtship behaviour. Nature. Jul 21;436(7049):395-400.

Manoli DS, Meissner GW, Baker BS. (2006). Blueprints for behavior: genetic specification of neural circuitry for innate behaviors. Trends Neurosci. Aug;29(8):444-51.

Markow TA. (1987). Behavioral and sensory basis of courtship success in Drosophila melanogaster. Proc Natl Acad Sci. Sep;84(17):6200-4.

Mellert DJ, Knapp JM, Manoli DS, Meissner GW, Baker BS. (2010). Midline crossing by gustatory receptor neuron axons is regulated by fruitless, doublesex and the Roundabout receptors. Development. Jan;137(2):323-32.

Monaghan AP, Bock D, Gass P, Schwäger A, Wolfer DP, Lipp HP, Schütz G. (1997). Defective limbic system in mice lacking the tailless gene. Nature. Dec 4;390(6659):515-7.

O'Dell KM, Armstrong JD, Yang MY, Kaiser K. (1995). Functional dissection of the Drosophila mushroom bodies by selective feminization of genetically defined subcompartments. Neuron. Jul;15(1):55-61.

Pignoni F, Baldarelli RM, Steingrímsson E, Diaz RJ, Patapoutian A, Merriam JR, Lengyel JA. (1990). The Drosophila gene tailless is expressed at the embryonic termini and is a member of the steroid receptor superfamily. Cell. Jul 13;62(1):151-63.

Pitman JL, Tsai CC, Edeen PT, Finley KD, Evans RM, McKeown M. (2002). DSF nuclear receptor acts as a repressor in culture and in vivo. Dev Biol. 2002 May 15;245(2):315-28.

Pryor KD, Leiting B. (1997). High-level expression of soluble protein in Escherichia coli using a His6-tag and maltose-binding-protein double-affinity fusion system. Protein Expression and Purification. Aug;10(3):309-19.

Read D, Manley JL. (1992). Alternatively spliced transcripts of the Drosophila tramtrack gene encode zinc finger proteins with distinct DNA binding specificities. EMBO J. Mar;11(3):1035-44.

Rideout EJ, Billeter JC, Goodwin SF. (2007). The sex-determination genes fruitless and doublesex specify a neural substrate required for courtship song. Curr Biol. Sep 4;17(17):1473-8.

Rideout EJ, Dornan AJ, Neville MC, Eadie S, Goodwin SF. (2010). Control of sexual differentiation and behavior by the doublesex gene in Drosophila melanogaster. Nat Neurosci. Apr;13(4):458-66.

Roy K, Kuznicki K, Wu Q, Sun Z, Bock D, Schutz G, Vranich N, Monaghan AP. (2004). The Tlx gene regulates the timing of neurogenesis in the cortex. J Neurosci. 2004 Sep 22;24(38):8333-45.

Ryner LC, Baker BS. (1991). Regulation of doublesex pre-mRNA processing occurs by 3'-splice site activation. Genes Dev. Nov;5(11):2071-85.

Ryner LC, Goodwin SF, Castrillon DH, Anand A, Villella A, Baker BS, Hall JC, Taylor BJ, Wasserman SA. (1996). Control of male sexual behavior and sexual orientation in Drosophila by the fruitless gene. Cell. Dec 13;87(6):1079-89.

Sanders LE, Arbeitman MN. (2008). Doublesex establishes sexual dimorphism in the Drosophila central nervous system in an isoform-dependent manner by directing cell number. Dev Biol. Aug 15;320(2):378-90.

Schüpbach T, Wieschaus E. (1991). Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology. Genetics. Dec;129(4):1119-36.

Schütt C, Nöthiger R. (2000). Structure, function and evolution of sex-determining systems in Dipteran insects. Development. Feb;127(4):667-77.

Schütze T, Wilhelm B, Greiner N, Braun H, Peter F, Mörl M, Erdmann VA, Lehrach H, Konthur Z, Menger M, Arndt PF, Glökler J. (2011). Probing the SELEX process with next-generation sequencing. PLoS One. 6(12):e29604.

Scott K, Brady R Jr, Cravchik A, Morozov P, Rzhetsky A, Zuker C, Axel R. (2001). A chemosensory gene family encoding candidate gustatory and olfactory receptors in Drosophila. Cell. Mar 9;104(5):661-73.

Song HJ, Billeter JC, Reynaud E, Carlo T, Spana EP, Perrimon N, Goodwin SF, Baker BS, Taylor BJ. (2002). The fruitless gene is required for the proper formation of axonal tracts in the embryonic central nervous system of Drosophila. Genetics. Dec;162(4):1703-24.

Stockinger P, Kvitsiani D, Rotkopf S, Tirián L, Dickson BJ. (2005). Neural circuitry that governs Drosophila male courtship behavior. Cell. Jun 3;121(5):795-807.

Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Privé GG. (2005). Sequence and structural analysis of BTB domain proteins. Genome Biol. 6(10):R82.

Tauber, E, and Eberl, DF (2003). Acoustic communication in Drosophila. Behav. Processes 64, 197–210.

Taylor BJ, Villella A, Ryner LC, Baker BS, Hall JC. (1994). Behavioral and neurobiological implications of sex-determining factors in Drosophila. Dev Genet. 15(3):275-96.

Tompkins L, Siegel RW, Gailey DA, Hall JC. (1983). Conditioned courtship in Drosophila and its mediation by association of chemical cues. Behav Genet. Nov;13(6):565-78.

Usui-Aoki K, Ito H, Ui-Tei K, Takahashi K, Lukacsovich T, Awano W, Nakata H, Piao

ZF, Nilsson EE, Tomida J, Yamamoto D. (2000). Formation of the male-specific muscle in female Drosophila by ectopic fruitless expression. Nat Cell Biol. Aug;2(8):500-6.

Villella A, Hall JC. (1996). Courtship anomalies caused by doublesex mutations in Drosophila melanogaster. Genetics. May;143(1):331-44.

Villella A, Gailey DA, Berwald B, Ohshima S, Barnes PT, Hall JC. (1997). Extended reproductive roles of the fruitless gene in Drosophila melanogaster revealed by behavioral analysis of new fru mutants. Genetics. Nov;147(3):1107-30.

Villella A, Hall JC. (2008). Neurogenetics of courtship and mating in Drosophila. Adv Genet. 62:67-184.

von Kalm L, Crossgrove K, Von Seggern D, Guild GM, Beckendorf SK. (1994). The Broad-Complex directly controls a tissue-specific response to the steroid hormone ecdysone at the onset of Drosophila metamorphosis. EMBO J. Aug 1;13(15):3505-16.

von Schilcher F, Hall J. (1979). Neural Topography of Courtship Song in Sex Mosaics of Drosophila melanogaster. J. comp. Physiol. 129, 85-95

Vosshall LB, Stocker RF. (2007). Molecular architecture of smell and taste in Drosophila. Annu Rev Neurosci. 30:505-33.

Wu JS, Luo L. (2006). A protocol for dissecting Drosophila melanogaster brains for live imaging or immunostaining. Nature Protocols. 1(4):2110-5.

Yamamoto D, Ito H, Fujitani K. (1996). Genetic dissection of sexual orientation: behavioral, cellular, and molecular approaches in Drosophila melanogaster. Neurosci Res. 1996 Oct;26(2):95-107.

Yamamoto D. (2007). The neural and genetic substrates of sexual behavior in Drosophila. Adv Genet. 59:39-66.

Yamamoto D. (2008). Brain sex differences and function of the fruitless gene in Drosophila. J Neurogenet. 22(3):309-32.

Yang Y, Zhang W, Bayrer JR, Weiss MA. (2008). Doublesex and the regulation of sexual dimorphism in Drosophila melanogaster: structure, function, and mutagenesis of a female-specific domain. J Biol Chem. 2008 Mar 14;283(11):7280-92.

Yu JY, Kanai MI, Demir E, Jefferis GS, Dickson BJ. (2010). Cellular organization of the neural circuit that drives Drosophila courtship behavior. Curr Biol. Sep 28;20(18):1602-14.

Zarkower D. (2002). Invertebrates may not be so different after all. Novartis Found Symp. 244:115-26

Zhang , S.D. and Odenwald, W.F. (1995). Misexpression of the white (w) gene triggers male-male courtship in Drosophila. Proceedings of the National Academy of Sciences of the United States of America June 6; 92(12): 5525–5529.

Zollman S, Godt D, Privé GG, Couderc JL, Laski FA. (1994). The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila. Proc Natl Acad Sci U S A. Oct 25;91(22):10717-21.