ABSTRACT

Title of Dissertation:	QUANTIFYING THE RELATIVE CONTRIBUTION AND FURTHERING QUALITATIVE UNDERSTANDING OF *FTZ CIS*-REGULATORY ELEMENTS IN *DROSOPHILA MELANOGASTER*
	Matthew Douglas Fischer, Doctor of Philosophy, 2022
Dissertation directed by:	Dr. Leslie Pick, Professor Department of Entomology University of Maryland, College Park

Embryonic development is coordinated by interactions within gene regulatory networks. This process is orchestrated at the level of transcription through the regulatory properties of enhancers, which direct spatiotemporal expression patterns when bound by specific *trans*-acting factors. Though enhancers can act upon promoters located at great distances irrespective of orientation, the contributions from these *cis*-regulatory elements (CREs) are limited by insulators and/or tethering elements that organize chromatin architecture. Much research has been conducted towards understanding the coordination of the segmentation genes that pattern the basic body plan of the fruit fly, *Drosophila melanogaster*, during embryogenesis. The pair-rule genes (PRGs) of this pathway, such as *fushi tarazu* (*ftz*), are expressed in seven alternating stripes across the embryo. These PRGs are required for the development of body segments, and the mis-regulation of a

single transcriptional domain can result in the loss of a segment. Here, I have investigated the *ftz* CREs to more precisely determine their sufficiency to direct expression within *ftz* stripe domains and their necessity for doing so in the native context of the gene. To investigate the sufficiency, I have generated 36 standardized reporter transgenes from 18 CREs, tested in both forward and reverse orientations. All CREs examined have been inserted into the same XbaI site of the reporter plasmid, and the transgenes have been inserted into the same genomic region. Through *in situ* hybridization experiments, I have determined that the qualitative patterns conferred by every CRE is orientation-dependent, and I have identified two putative insulators and/or tethering elements, proposed to explain this observation. To investigate their necessity, I targeted four genomic regulatory regions for precise deletion using the CRISPR/Cas9 system to generate seven deletion mutants. Though deletions were expected to cause lethality, most of the mutants are homozygous viable and fertile; only a mutant simultaneously removing two seven-stripe CREs was homozygous lethal. Quantitative gene expression analysis by fluorescent *in situ* hybridization chain reaction revealed that there is a critical threshold of *ftz* abundance required in each stripe for segmentation to proceed. In conclusion, I have determined that the *ftz* CREs are redundant and function together in a non-additive manner.

Quantifying the relative contribution and furthering qualitative understanding of *ftz cis*-regulatory elements in *Drosophila melanogaster*

by

Matthew Douglas Fischer

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Dr. Leslie Pick, Chair
Dr. Antony Jose
Dr. Sean Carroll
Dr. Sougata Roy
Dr. Stephen Mount

Dedication

To my parents; Wendy and Ken Marshall, Phil and Jackie Fischer. To my brothers; Michael, Kyle, Kevin, Patrick, and Dorian. To Gustavo. In loving memory to my grandparents, Frederick and Josephine Fischer.

Acknowledgements

First and foremost, I would like to thank my advisor, Leslie Pick, for her phenomenal mentorship throughout my graduate training. Leslie's constant support – both academically and personally – has been incalculable. I reflect back on our discussions fondly, as her active engagement and encouragement have fostered within me a stronger sense of self-confidence and admiration for scientific pursuit. She has cultivated a lab setting that is simultaneously rigorous and enjoyable. Thank you for advising me through this process.

I would also like to thank the many members of the Pick lab that I have met over the years: Patricia Graham, Katie Reding, Alys Jarvela, Ximena Gutierrez Ramos, Minh Le, Ashley Zavodnick, Abigail Heath, Catherine Trelstad, Omid Tabatabaee, Judy Wexler, Teresita Hernandez-Arriola, Daniel Zheng, and Abhigya Giri. All of you have contributed to a wonderful environment and have been so encouraging. In particular, I would like to thank Ebony Argaez and Jessica Hernandez for being so humorous and making the lab fun. I would also like to thank other members of the BISI program, and I extend a special shout-out to Zeke Gonzalez, Nicki Barbour, Molly Westbrook, Jessica Bowers, Custodio Nunes, Michael Lizzio, Dave Weber, Fuxi Wang, and Haleigh Eppler.

Finally, I would like to thank my parents: Wendy and Ken Marshall, Phil and Jackie Fischer – my brothers: Michael, Kyle, Kevin, Patrick, and Dorian – my dearest friends: Sean Aliff, Haley Wight, and Graham Perkinson – and my beloved fiancé, Gustavo Diaz. Their unconditional love, constant support, and endless encouragement have been so critical during this journey.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

1D	One-dimensional
3D	Three-dimensional
Bcd	Bicoid
CGI	CpG-island
ChIP	Chromatin immunoprecipitation
CNS	Central nervous system
CRE	*Cis*-regulatory element
CRISPR	Clustered regularly interspersed short palindromic repeats
CTCF	CCCTC-binding factor
DNA	Deoxyribonucleic acid
DPE	Downstream promoter element
en	*engrailed*
eve	*even-skipped*
FGF	Fibroblast growth factor
ftz	*fushi tarazu*
GFP	Green fluorescent protein
Gt	Giant
Hb	Hunchback
HCR	Hybridization chain reaction
HDR	Homology-directed repair
Hth	Homothorax
IDR	Intrinsically disordered region
Inr	Initiator element
Kni	Knirps
Kr	Kruppel
mRNA	Messenger ribonucleic acid
NHEJ	Non-homologous end joining
PCR	Polymerase chain reaction
PRG	Pair-rule gene
proExM	Protein-retention expansion microscopy
qPCR	Quantitative polymerase chain reaction
SEL-Seq	Synthetic enhancer library-sequencing
slp1	*sloppy paired 1*
svb	*shaven baby*
TAD	Topologically associating domains
TSS	Transcription start site

Ttk	Tramtrack
Ubx	Ultrabithorax
UPS	Upstream element
WRE	Wnt-repsonsive element
WT	wild type
Zld	Zelda

Chapter 1: Introduction

1.1 The transcriptional enhancer: a brief history

The discovery of the transcriptional enhancer was a mistake. Near the start of his professorial career, Dr. William Schaffner sought to express a rabbit β-globin gene from a genomic fragment in transfected HeLa cells (Schaffner, 2015). He had chosen to insert this fragment into two plasmids, one of which contained part of the SV40 viral genome. This choice was by sheer coincidence – his initial projects to express a sea urchin histone gene from an SV40 viral replicon were not successful. Consequently, he had an abundance of this plasmid readily available. As it turned out, β-globin expression was detectable by immunofluorescence only from HeLa cells that had been transfected with the plasmid containing both the viral genome and the globin gene (Banerji et al., 1981). One by one, the β-globin genomic fragment and the viral sequences were independently digested out the plasmid and re-inserted in the reverse orientation. The linear distances between the transcription initiation site of the β-globin gene and the viral sequences changed when the viral fragment was ligated into a downstream restriction site. Subsequent analyses confirmed that expression was still detected so long as the genomic fragment was on the same plasmid as the viral sequence, suggesting an interaction in *cis*. In determining that a small piece of the SV40 genome was responsible for the 200-fold increase of transcription from the β-globin promoter, the first enhancer had been identified.

In the intervening decades, research on *cis*-regulatory elements (CREs) has been conducted on a multitude of systems, ranging in complexity from single-celled organisms

to complex metazoans (Kyrchanova & Georgiev, 2021; Newlon & Theis, 1993). We know today that Schaffner's three properties of the enhancer – that it functions in both orientations, irrespective of distance, and from either up- or downstream of a promoter – are limited in the genomic context (Chen & Lei, 2019). Here, I will review the literature on eukaryotic mRNA transcriptional *cis*-regulation, presenting some of the historic, fundamental perspectives and more recent findings that have been spurred by advances in sequencing and imaging technology.

1.2 Clarifying convention: enhancers and promoters are distinct CREs

Enhancers affect the rate at which transcription occurs from a promoter. Whereas some publications have used the two terms interchangeably to describe upstream regulatory DNA (Porto et al., 2014), there are fundamental properties that differentiate these CREs into distinct classes. The primary function of the core promoter is to initiate transcription through association with the basal transcriptional machinery (Haberle & Stark, 2018). Most eukaryotic promoters are unidirectional in that stable transcripts are only produced from one strand (Mikhaylichenko et al., 2018). Though there are promoter-proximal elements that confer higher levels of expression in spatiotemporally restricted patterns, these elements are more like enhancers than they are like promoters (Small et al., 1996).

A recent study has revealed that in mammals, unidirectional promoters cannot enhance transcription rates from another promoter, but some enhancers can act as weak promoters (Mikhaylichenko et al., 2018). However, it has been generally observed that transcription directly from enhancers within the genome produces unstable transcripts, contrary to the stable transcripts produced from a promoter (Andersson et al., 2014;

Flynn et al., 2011). For the sake of convention, elements within 100 bp of a transcription start site (TSS) that directly associate with the basal transcriptional machinery to produce stable transcripts will be referred to here as core promoters. Any other elements, whether proximal or distal, that bind transcription factors primarily to alter transcription rates from a promoter will be referred to as enhancers.

1.3 Core promoters have evolved with a diversity of motifs with implications on enhancer specificity

There are distinct motifs within core promoters with which the transcriptional machinery can associate. The combination of these different motifs, generally referred to as promoter architecture, appear to have implications in the specific transcriptional properties of a promoter (Haberle et al., 2014; Haberle & Stark, 2018; Ohtsuki et al., 1998). The first eukaryotic promoter element to be discovered was the TATA box, which was identified as a consensus sequence in close proximity to the TSSs of *Drosophila melanogaster* histone genes (Lifton et al., 1978). The TATA box has since been identified in some actively transcribed core promoters across the genomes of many organisms, ranging from the plant *Arabidopsis thaliana* to complex animals such as humans (DGT, 2014; Bernard et al., 2010; Chen et al., 2013; Hoskins et al., 2011).

Other common core promoter elements among animals include the initiator element (Inr) and the downstream promoter element (DPE) (Burke & Kadonaga, 1997; Chen et al., 2013). The function of the DPE is analogous to the that of the TATA box, in that both elements associate directly with the basal transcription factor TFIID (Burke et al., 1998). In nature, the presence of a TATA box within a promoter is exclusive to the

presence of a DPE; thus, a general model emerged that a promoter can associate with the basal transcriptional machinery through different mechanisms (Haberle & Stark, 2018).

In flies, it has been reported that DPE-containing promoters are as prevalent as TATA box-containing promoters, and both promoter classes contain the Inr (Kutach & Kadonaga, 2000). Whether a promoter contains a TATA box or DPE has been implicated as one of the determinate factors for the specificity of enhancer-promoter interactions (Ohtsuki et al., 1998). This was demonstrated through the use of promoter competition assays in composite reporter transgenes (Calhoun et al., 2002; Ohtsuki et al., 1998). These P-elements contained two genes flanking a developmental enhancer that endogenously regulates a TATA-box containing gene. The promoter of one reporter gene contained a TATA box while the other had a DPE. Though both were proximal to the enhancer, only expression of the reporter gene with the TATA promoter was detected in the transcriptional domain conferred by the developmental enhancer. When the DPE promoter was mutagenized to contain a TATA box, expression of both genes was detected in the same domain.

Within the genome of an organism, the promoters of two different genes may be in close proximity to enhancers that direct different spatiotemporal expression patterns. Enhancer-promoter specificity is one mechanism by which enhancer interactions are limited and ectopic expression is prevented. This can cause experimental constraints for examining the sufficiency of an enhancer-containing genomic fragment to direct expression of a transgene. If the promoters for some of those genes have a TATA box and others contain a DPE, how can a common promoter in a reporter plasmid be used to test the enhancers of different genes? This has been circumvented in *Drosophila*, at least,

through the construction of the *Drosophila* synthetic core promoter (DSCP) (Pfeiffer et al., 2008). The DSCP contains the TATA box, Inr, and DPE, and it has been shown to be compatible with a wide variety of enhancers. Thus, it appears the mechanism of transcriptional initiation mediated by a TATA box is compatible with the transcriptional initiation mediated by a DPE despite their natural exclusivity.

Though early research on the DPE had identified it within the TATA-less promoter of the human *Interferon regulatory factor-1* gene (Burke & Kadonaga, 1997), the DPE is an uncommon TATA alternative in mammals. A more common motif within mammalian promoters is the CpG-island (CGI), which has been historically associated with housekeeping genes (Carninci et al., 2006). However, a recent study that analyzed genome-wide and transcriptome-wide TSS usage within human and mouse primary cells, cell lines, and tissues has suggested otherwise (DGT, 2014). Their analysis showed that 54% of CGI promoter-containing genes were transcribed in a highly cell-type-specific manner. As another study has shown that many of these specific genes are associated with promoter-proximal enhancers (Andersson et al., 2014), these authors conclude that cell-specific expression from CGI promoters is likely the result of interactions with a specific enhancer. Thus, it appears that previous generalizations about which types of mammalian genes (housekeeping vs cell-specific) are associated with which promoter architecture (TATA, CGI, non-CGI-non-TATA) will require re-evaluation and more vigilant study.

1.4 Enhancer architecture and function is determined by transcription factors

The capacity of an enhancer to direct a spatiotemporal expression pattern is bestowed by the *trans*-acting factors that bind to it. Often, these *trans*-acting factors are

transcription factors, which are proteins that have a DNA-binding domain and a regulatory domain that can associate with other cofactors (Spitz & Furlong, 2012). Transcription factors often have flexibility with which DNA sequences they can recognize (Stormo, 2000). The specific sequences to which a transcription factor can bind is often referred to as a motif or consensus binding site. Determining a consensus binding site requires comparing the profile of exact sequences to which a protein binds, as tested through biochemical analyses. Finding a predicted site within a CRE is insufficient to conclude a regulatory impact, for it must be experimentally confirmed *in vivo*. An example of determining direct protein binding to an enhancer was worked out for the proximal upstream element of the *Drosophila melanogaster* gene *fushi tarazu* (*ftz*) (Han et al., 1993). Through gel retardation assays, DNase I footprint assays, and methylation interference analyses, it was determined that the proteins Ftz-F1, Ftz, and Tramtrack (Ttk) directly bind to the enhancer. By following this with gene expression analysis of various reporter constructs that had specific binding sites mutagenized, alterations to the expression pattern were observed. Thus, it was determined that these binding sites were biologically relevant.

There are currently three models that rationalize enhancer functionality on the basis of the transcription factor motifs that compose them (Spitz & Furlong, 2012). This is referred to as enhancer architecture. Motif grammar describes the relative positioning of motifs within an enhancer, including the order, spacing, and orientation of each motif. The enhanceosome model states that motif grammar must be precise and is essential for proper downstream recruitment of the transcriptional machinery (Fig. 1-1A). This precision places transcription factors exactly where necessary along the helical turns of

DNA. The term enhanceosome was first used to describe the regulation of the *Xenopus* ribosomal promoter by a xUBF dimer (Bazett-Jones et al., 1994). More recently, an analysis of Sox2 and Oct4 transcription factor binding patterns in mouse embryonic stem cells followed an enhanceosome assembly process (Chen et al., 2014). An essential component of their methods was the use of a 3T3 engineered cell line that stably expressed halo-Sox2 along with an inducible system for Oct4. By tracking the binding of halo-Sox2 to DNA in the presence or absence of Oct4, it was determined that Sox2 and Oct4 bind in a specific order that alters the kinetics and thermodynamics of the next interaction. Thus, the ordered assembly of several transcription factors along the DNA in close proximity was necessary for functionality.

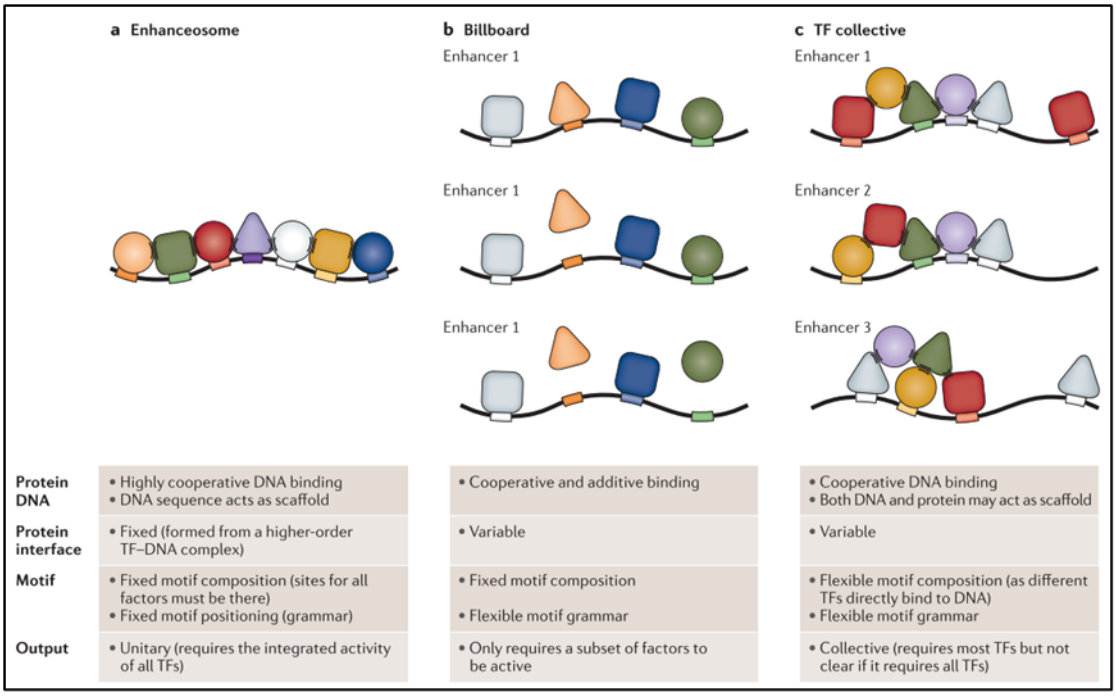


Figure 1-1. Three different models of enhancer function as determined by enhancer architecture. (A) The enhanceosome model calls for strict motif grammar that positions the transcription factors in specific locations along the helical turns of DNA. All transcription factors must be present for activity, and there is little flexibility. Transcription factors depend on association with DNA and likely do not form a complex

prior to binding. (B) The billboard model applies well to developmental enhancers that can direct expression in different spatiotemporal patterns when associated with different combinations of transcription factors. The motif grammar is flexible, but the presence or absence of a motif is consequential. (C) The transcription factor collective model states that protein complexes, with components that may or may not be contacting the DNA at a given genomic fragment, are primarily responsible for conferring regulatory properties. Enhancers that fit the collective model have flexible motif grammar and composition, but the transcription factors of the collective physically interact with each other. Figure taken from (Spitz & Furlong, 2012).

Some enhancers direct expression in multiple transcriptional domains by binding different sets of transcription factors. This is common for developmental enhancers, for which the enhanceosome model does not apply well. These enhancers more accurately follow the billboard model (Fig. 1-1B). In the billboard model, an enhancer acts as a substrate to which different transcription factors can bind in different combinations to confer spatiotemporal expression patterns. Enhancers that follow the billboard model thus have strict motif requirements but flexible motif grammar. The first implication of the billboard model was used to describe the transcriptional output of several transgenes in *Drosophila melanogaster* (Kulkarni & Arnosti, 2003). This study used a series of synthetic enhancers that contained different motif combinations at variable distances (variable motif grammar). While the exact arrangement of these motifs could be altered and still confer the same expression pattern, the presence or absence of a motif was consequential.

The billboard model was shown to be relevant in the genomic context from a study that compared the expression pattern conferred by *even-skipped* enhancers of six scavenger fly species (*Sepsidae*) to those of *Drosophila melanogaster* (Hare et al., 2008). Even though there is substantial sequence divergence between the enhancer counterparts

in these different species, there are small stretches of sequences with high identity that are enriched for pairs of adjacent binding sites. When these enhancers were fused to reporter transgenes and transformed into *Drosophila melanogaster* as a heterologous system, they directed expression in similar spatiotemporal patterns as their *Drosophila* counterparts. Thus, the main contribution to the expression pattern was the presence of motifs, not the motif grammar.

In some instances, a transcription factor complex consisting of multiple proteins is responsible for conferring the specific expression pattern of an enhancer. Certain components of these complexes do not necessarily need to directly bind to the DNA so long as they can associate with other DNA-binding factors. This mode of complex-dependent regulation is explained by the transcription factor collective model. It is currently unclear to the extent that all transcription factors of the collective need to be present, but theoretically the model has little flexibility for the absence of transcription factor collective components.

Recently, the transcription factor collective model has been implicated in the study of Wnt target enhancer regulation in HEK293T and HeLa human cell lines (Ramakrishnan et al., 2021). In this study, the factors that direct expression from a novel Wnt-responsive element (WRE) downstream of the *Axin2* gene were determined and analyzed. By surveying for predicted motifs in the enhancer, it was determined that this WRE contains predicted motifs for the factor CDX1 but not for TCF7. When Wnt signaling was activated with β-catenin, RNAi against either *CDX1* or *TCF7* resulted in reduced reporter activity, suggesting a role for both factors in the direct or indirect regulation of this enhancer. Direct association with the enhancer was confirmed through

independent ChIP-qPCR experiments with antibodies for the CDX1 and TCF7 proteins. Subsequent biochemical analyses suggested that these two proteins can interact. This physical interaction was determined to be required for enhancer function when amino acid mutations at the protein-protein interface disrupted the association and capacity to regulate transcription. In analyzing proximal motifs adjacent to the CDX1-TCF7 complex in this enhancer and other genomic locations, an enrichment of the sequence CAG was detected. The authors conclude that this TCF/CDX/CAG cassette represents a transcription factor collective that is implicated in Wnt response and thus cancer biology.

1.5 Motif sub-optimization is essential for some developmental enhancers

Interestingly, some developmental enhancers contain motifs that are sub-optimal for the transcription factors that bind them. This was found to be the case for the Otx-a enhancer of the sea squirt *Ciona intestinalis* (Farley et al., 2015). During embryonic development, *Otx* is expressed in the anterior neural plate and dorsal nerve cord. The binding sites for the transcription factors fibroblast growth factor (FGF) and GATA within this enhancer contain imperfect matches to the predicted consensus binding sites. Through the synthesis of custom oligos and Synthetic Enhancer Library-Sequencing (SEL-Seq), 2.5 million variants of the Otx-a enhancer were generated and fused to reporter transgenes. These variants included alterations to the two FGF and three GATA binding sites, and variable distances were added between the binding sites as well. Gene expression analysis of the reporter transgenes revealed that the enhancer mutants with motifs that more closely resembled the optimal consensus sequences expressed *Otx* in ectopic expression domains of the anterior endoderm and hindbrain. Further, enhancer mutants with altered spacing between the motifs reflected this ectopic expression,

suggesting the spacing between the native motifs is also suboptimal. Thus, the proper regulation of *Otx* in the developing neural plate and nerve cord of *Ciona intestinalis* relies on inefficient binding of the FGF and GATA transcription factors.

Sub-optimization was also discovered to be an important property of another enhancer within *Ciona* for the *Brachyury* gene (Farley et al., 2016). This enhancer (Ci-Bra) contains two low-affinity binding sites for FGF, but unlike the Otx-a enhancer, these binding sites are located in an optimal spacing and orientation. Despite the low affinity motifs, Ci-Bra confers high levels of expression in the notochord. In comparing these gene expression patterns to synthetic enhancers and motif mutants of the Ci-Bra enhancer, it was determined that the precise spacing and orientation of these two low affinity binding sites is necessary for strong expression specifically in the notochord.

That two development enhancers of *Ciona* are dependent upon sub-optimal motifs or spacing between the motifs highlights an importance for re-evaluating the parameters that are used in bioinformatic analyses to predict CREs throughout the genome. Many of these studies rely on searching for clusters of optimal predicted consensus binding sites (Royo et al., 2012; Schroeder et al., 2011; Zhang et al., 2005). That some naturally occurring developmental enhancers can function best with sub-optimal architecture likely means that these CREs remain largely understudied and underreported. A paradigm-shift within the field to focus on CREs with sub-optimal architecture may provide greater insights towards CRE function, redundancy, and evolution.

Though uncommonly reported, this phenomenon has been observed in other animals as well. In *Drosophila melanogaster*, three enhancers for the gene *shaven baby* (*svb*) contain low-affinity binding sites for the protein Ultrabithorax (Ubx). Through the

use of protein retention expansion microscopy (proExM) (Tillberg et al., 2016), high resolution confocal images revealed that transcriptionally active *svb* loci were enriched with high concentrations of Ubx and the co-factor Homothorax (Hth) (Tsai et al., 2017). The same enrichment of Ubx and Hth was observed around reporter transgenes on other chromosomes, suggesting that the low-affinity sites play a functional role in concentrating the transcription factors within a nuclear microenvironment.

In a separate study, the same scientists examined a deletion mutant that lacked one of the three low-affinity enhancers (Tsai et al., 2019). These mutants had a temperature sensitive denticle phenotype, and the nuclear microenvironment of the *svb* locus contained a relatively lower concentration of Ubx and Hth compared to wild type. Both the denticle and Ubx/Hth microenvironment concentration phenotypes were rescued in the presence of a *dsRed* reporter transgene on a separate chromosome that contained the three low-affinity enhancers. High resolution microscopy revealed that both the transgene and the native *svb* locus were localized within the same nuclear microenvironment. In conclusion, one mechanism by which sub-optimal motifs can be advantageous is through the recruitment of transcription factors and cofactors, thereby increasing their local concentration within a particular nuclear microenvironment. This concentration appears to confer flexibility towards environmental conditions that can fluctuate in nature, such as temperature.

1.6 There are multiple orders to chromatin organization

Mammalian enhancers have been discovered to function over a wide variety of distances. Some are located just a few kilobases from their target promoters, while others are located more than 800 kb away (Amano et al., 2009; Zhao et al., 2007). Schaffner's

proposed property of the enhancer to function over great linear distances is both limited and conferred by the three-dimensional (3D) architecture of the genome. Though many discoveries have been made towards understanding genomic organization, the precise mechanisms that dictate chromosomal architecture are not well understood.

Broad regions of chromosomes are generally classified according to their order of compaction during interphase (Fig. 1-2A) (van Steensel & Furlong, 2019). Regions that are considerably compacted are called heterochromatin and generally located near the nuclear lamina. The more open regions are called euchromatin, which are generally spread out in the nucleus and transcriptionally accessible. The difference in compaction is considerable, and the first discovery of this phenomenon was visualized in moss nuclei by microscopy (Heitz, 1928). Both the heterochromatic and euchromatic regions of the chromatin form higher-order domains, resulting from the bending and looping of the chromatin. Advances in sequencing technology and chromatin-capture experiments have facilitated the ability to demarcate these regions (Fig. 1-2B). These methods include Hi-C, which is capable of detecting intra-chromatin contacts over considerable distances, and the more recent Micro-C, which can resolve these contacts with single-nucleosome resolution (Hsieh et al., 2015; Lieberman-Aiden et al., 2009).

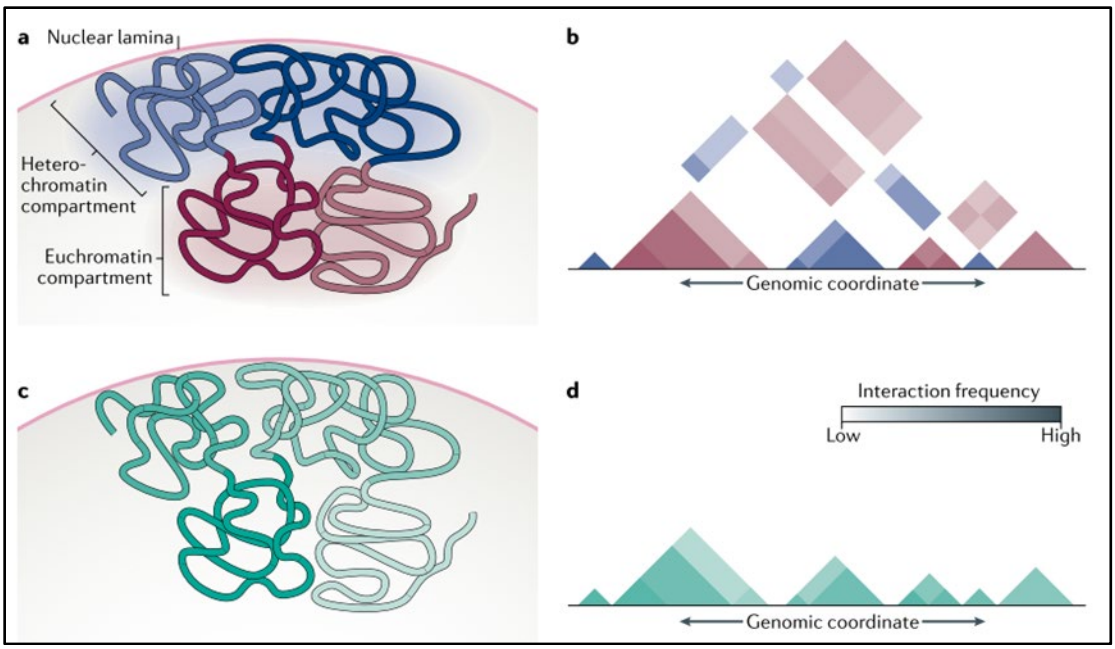


Figure 1-2. Chromosomes are organized into distinct domains. (A) Certain regions of the genome are highly compact, inaccessible to transcriptional machinery, and associated with the nuclear lamina (blue; heterochromatin). Other regions are open and more accessible to the transcriptional machinery (red, euchromatin). (B) Cartoon of Hi-C map; this experiment can reveal long-range intra-chromatin contacts and resolve higher-order domains. (C) Shades of green used to represent different topologically associating domains (TADs), characterized primarily by loops and intradomain contacts. (D) Cartoon of Hi-C map; intra-TAD contacts can be resolved through this analysis. Figure taken from (van Steensel & Furlong, 2019).

Though the euchromatin is generally spread out in the nucleus, there is a lower-order organization within specific regions of the chromatin. These are referred to as topologically associating domains (TADs), and they are characterized by the formation of loops and intra-chromatin contacts (Fig. 1-2C) (Jost et al., 2014). The boundaries that surround TADs can be resolved with current genome-wide chromatin capture experiments (Fig. 1-2D). The original analyses using Hi-C suggested that genomic architecture is conserved from *Drosophila* to mammals (Dixon et al., 2012; Sexton et al.,

2012). However, analyses conducted at higher resolution found that genomic organization is fundamentally different between flies and mammals.

In mammals, the boundaries of the loops that are fundamental to the TAD are generally enriched by insulator elements with the motif CCCTC and the protein factor that binds to it (CTCF) (Rhee & Pugh, 2011). The orientation of these motifs position CTCF to converge towards the anchor sites (Rao et al., 2014). Additionally, it was found that these anchor sites were also enriched for the cohesin complex. The site-specific and motif-driven enrichment of these proteins suggests that 3D architecture is encoded within the linear sequence of the mammalian genome. Another analysis that focused on the formation of loops has proposed that TADs are the direct result of loop extrusion events mediated by CTCF and cohesin complex proteins (Fudenberg et al., 2016). These authors reason that the formation of loops in this manner is a mechanism by which enhancer-promoter pairs are facilitated, consistent with earlier looping models (Su et al., 1991; Tolhuis et al., 2002).

Within *Drosophila*, there is limited evidence of CTCF-mediated TAD formation. Though there is a CTCF orthologue in *Drosophila*, homology is restricted to the Zn-finger domains (Moon et al., 2005). Association of CTCF with cohesin complex proteins requires the C-terminal domain of the mammalian CTCF protein, which is not conserved in *Drosophila*. Unlike in mammals, domain boundaries are not associated with convergent CTCF motifs in *Drosophila* (Rowley et al., 2017). Notably, the physical shape of the TADs is distinctly different. Whereas mammalian TADs are organized into clear boundaries, *Drosophila* TADs seem to contain a series of sub-TADs that are distinguishable in Hi-C analyses (Eagen et al., 2017; El-Sharnouby et al., 2017; Hug et

al., 2017; Rowley et al., 2017). More recently, high-resolution Micro-C analysis of the early *Drosophila* embryo has confirmed the presence of these sub-TAD associations between promoters and their distal regulatory regions (Batut et al., 2022). These authors propose that tethering elements, not insulators, are the functional CREs that organize these sub-TAD associations. Interestingly, some TAD boundaries in flies seem to be associated with Polycomb silencing complexes and less-conserved insulator binding proteins rather than CTCF (Eagen et al., 2017). Thus, it appears that genome organization and architecture is different across the animal kingdom in regards to the specific protein complexes involved. Regardless of these differences, in both mammals and flies, insulators play a general role in organizing the TADs of the genome.

Further support for a fundamental difference in genome organization between *Drosophila* and mammals comes from analyzing CTCF *Drosophila* mutants. Flies that lack both maternal and zygotic CTCF complete embryogenesis and larval development but die in the pupal case (Gambetta & Furlong, 2018). These dead adults display homeotic transformations resulting from ectopic expression of *Abdominal-B*. From this, the authors conclude that CTCF is required for proper regulation of *Hox* gene expression but not for other aspects of embryonic and larval development.

1.7 TADs and sub-TADs facilitate enhancer-promoter interactions

Though the distinct mechanisms that create TADs differ across the animal kingdom, the formation of loops is considered fundamental for contact and specificity between a promoter and an enhancer (Deng et al., 2012; Li et al., 2015; Tolhuis et al., 2002). Within these subdomains, one or more enhancers can be brought close to one or more promoters. Physical contact is made between an enhancer and a promoter through

the associations of the conserved mediator complex with the transcription factors bound to both CREs (Verger et al., 2019). The mediator is a large protein complex composed of multiple conserved subunits that was first visualized in yeast from an electron micrograph (Asturias et al., 1999). Recently, the association of many transcription factors and/or co-factors with intrinsically disordered regions (IDRs) within a specific nuclear microenvironment have been implicated in the formation of phase-separated condensates with the mediator complex (Fig. 1-3) (Boija et al., 2018; Shrinivas et al., 2019). The thermodynamics of these condensates seem to favor increased factor associations within a particular chromatin microenvironment. This results in a feedback cascade of increased associations with the basal transcriptional machinery, resulting in increased transcription bursts and output. Transcriptional bursts and their relation to phase-separated condensates remains an active field of study (Tunnacliffe & Chubb, 2020).

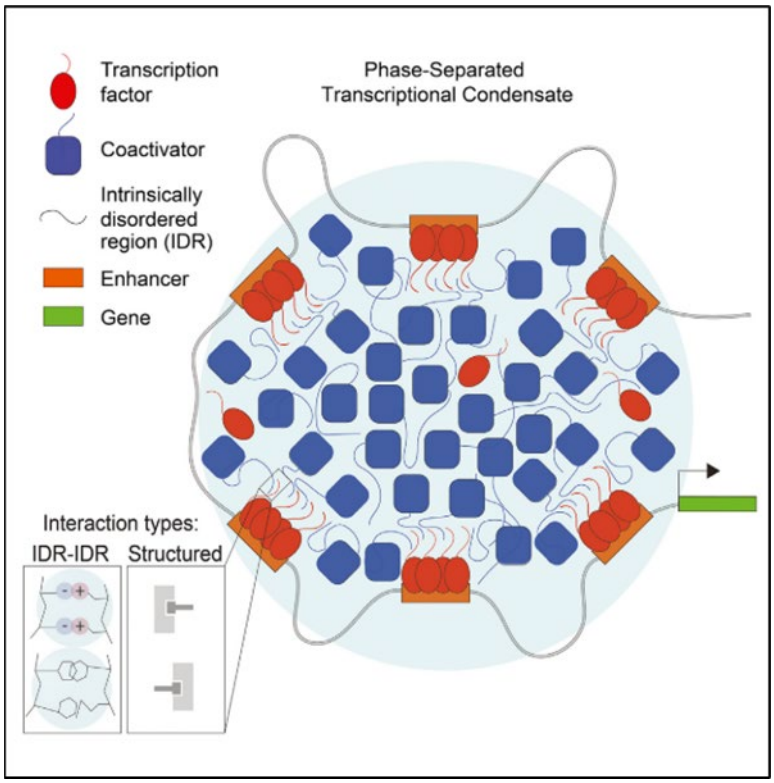


Figure 1-3. The associations between transcription factors, cofactors, and the mediator complex form a phase-separated condensate. Recent studies have found evidence that the intrinsically disordered regions of transcription factors and cofactors are capable of forming phase-separated condensates that favor the assembly of the transcriptional machinery. Higher concentration of factors within a given chromatin domain result in thermodynamically favorable conditions for increased associations, resulting in increased transcription bursts and output. Figure taken from (Boija et al., 2018).

1.8 CRE Conclusions

Over the last four decades, a great deal of knowledge has been discovered on *cis*-regulatory elements in eukaryotes. Advances in sequencing technologies have allowed for genome-wide analyses of promoter and enhancer motif compositions. Comparing this information to transcriptomics has facilitated the identification of multiple TSSs within broad promoters and the re-categorization of promoter architectures. High-resolution microscopy provides us with insights towards the regulatory events within specific loci and facilitates the study of phase-separated condensates. Chromatin capture experiments have revealed that the organization of the genome has regulatory consequences, with specific enhancer-promoter pairs relegated within the same TAD. Though these technologies have facilitated the acquisition of knowledge, they have simultaneously revealed that *cis*-regulation is exceptionally complex and that much remains to be determined.

Chapter 2: The *fushi tarazu* zebra element is not required for *Drosophila* viability or fertility

The work presented in this chapter is a minor adaptation of a collaborative publication between another member of the lab, Dr. Patricia Graham, a previous undergraduate student, Abhigya Giri, and myself. It was published in the journal *G3: Genes, Genomes, Genetics* (Graham et al., 2021). Dr. Graham and I are co-first authors of that article. The minor adaptation presented here is the inclusion of data collected from the *ftzΔZ-GAGA* mutant in Figures 2-4 and 2-7 and Table 2-5.

2.1 Abstract

Expression of genes in precisely controlled spatiotemporal patterns is essential for embryonic development. Much of our understanding of mechanisms regulating gene expression comes from the study of *cis*-regulatory elements (CREs) that direct expression of reporter genes in transgenic organisms. This reporter-transgene approach identifies genomic regions sufficient to drive expression but fails to provide information about quantitative and qualitative contributions to endogenous expression, although such conclusions are often inferred. Here we evaluated the endogenous function of a classic *Drosophila* CRE, the *fushi tarazu* (*ftz*) zebra element. *ftz* is a pair-rule segmentation gene expressed in seven stripes during embryogenesis, necessary for formation of alternate body segments. Reporter transgenes identified the promoter-proximal zebra element as a major driver of the seven *ftz* stripes. We generated a precise genomic deletion of the zebra element (*ftzΔZ*) to assess its role in the context of native chromatin and neighboring CREs, expecting large decreases in ftz seven-stripe expression. However, significant

reduction in expression was found for only one stripe, *ftz* stripe 4, expressed at ~25% of wildtype levels in *ftzΔZ* homozygotes. Defects in corresponding regions of *ftzΔZ* mutants suggest this level of expression borders the threshold required to promote morphological segmentation. Further, we established true-breeding lines of homozygous *ftzΔZ* flies, demonstrating that the body segments missing in the mutants are not required for viability or fertility. These results highlight the different types of conclusions drawn from different experimental designs and emphasize the importance of examining transcriptional regulatory mechanisms in the context of the native genomic environment.

2.2 Introduction

Precise control of gene expression by *cis*-regulatory elements (CREs) is critical to all aspects of embryonic development and organismal function (reviewed in Schaffner, 2015). CRE identification is often enabled by vectors derived from naturally occurring transposable-elements that allow for integration into the genome. Reporter genes, in which candidate CREs are placed upstream of a basal promoter and the coding region of an innocuous gene such as *lacZ* or *gfp,* allow for analysis of CREs *in vivo,* in the context of a developing organism, rather than *in vitro*, cell culture or heterologous systems. These approaches have been used extensively in so-called “promoter bashing” or “enhancer bashing” experiments in many model systems, including *Drosophila melanogaster*, where P-element mediated transformation enabled the identification of many cell-type-specific CREs (Rubin and Spradling, 1982; Rubin and Spradling, 1983).

One of the first genomic regions controlling early embryonic gene expression to be identified using the transgenic reporter gene approach was that of the pair-rule gene *fushi tarazu* (*ftz*) (Hiromi et al., 1985; Hiromi and Gehring, 1987). *ftz* is expressed in a

seven-stripe pattern in *Drosophila* embryos in the primordia of the alternate parasegments missing in *ftz* mutants (Hafen et al., 1984). The importance of precise control of *ftz* expression in stripes was highlighted by the finding that mis-expression of *ftz* throughout the embryo results in lethality (Struhl, 1985). In work that was groundbreaking at the time, Hiromi *et al.* (1985) identified three major CREs within a 10 kb genomic fragment that was sufficient to rescue *ftz* mutants (see Fig. 2-1A). The promoter proximal zebra element (Z) directed expression in seven *ftz*-like stripes, while the neurogenic element (N) directed expression in specific cells in the developing central nervous system (Hiromi et al., 1985; Doe et al., 1988). At the distal end of the 10 kb rescue fragment, Hiromi *et al.* identified an Upstream Element (UPS) that enhanced seven-stripe expression and which mediates autoregulation by Ftz and its partner Ftz-F1 (Hiromi and Gehring, 1987; Pick et al., 1990; Yu et al., 1997). Further analysis of the zebra element identified short regions that activate or repress striped expression (Dearolf et al., 1989b; Topol et al., 1991), as well as a portion of the zebra element that directs broad expression in the domain corresponding to *ftz* stripes 4-7 via direct binding of Caudal (Cad) (Dearolf et al., 1989a), which had previously been shown to be required for posterior *ftz* stripe expression (Macdonald and Struhl, 1986). These early studies led to the model that *ftz* expression is activated in all seven stripes via the zebra element, with interstripe repression playing a major role in stripe formation (Edgar et al., 1986; Carroll, 1990). Following this, striped expression is maintained by autoregulation via the UPS (Hiromi and Gehring, 1987; Pick et al., 1990).

Later studies questioned this simple model. For example, the demonstration that *hairy* and *ftz* stripes do not arise in complementary spatiotemporal patterns, along with a

lack of biochemical evidence that Hairy binds to the zebra element as a *trans*-acting factor, ruled out Hairy - the best candidate for interstripe repression - as playing a role in stripe establishment; rather, Hairy as well as Even-skipped (Eve) appear to maintain the *ftz* pattern once it is established (Yu and Pick, 1995). Identification of CREs that direct expression of single *ftz* stripes, similar to the classic *eve* stripe-specific elements (Small et al., 1991; Small et al., 1992; Arnosti et al., 1996; Small et al., 1996) raised further doubts about seven stripe elements establishing *ftz* stripes (Calhoun and Levine, 2003; Schroeder et al., 2004; Schroeder et al., 2011), as did detailed analysis of pair-rule stripes in wild type and mutant embryos (Clark and Akam, 2016). Rather, they supported the model that *ftz* expression is initiated in individual stripes by maternal or gap proteins interacting with stripe-specific CREs, followed by refinement and maintenance of stripes by pair-rule proteins interacting with 7-stripe CREs (Yu and Pick, 1995). Stripe-specific elements were identified for *ftz* stripes 1+5, 2+7 and 3 + 6/7 but not for stripe 4 (Schroeder et al., 2011). Schroeder et al. (2011) concluded that there is no stripe 4 CRE and proposed that the zebra element was the only CRE responsible for directing the initial expression of this *ftz* stripe. However, sequential deletions of the zebra element in reporter transgenes failed to identify any region of the zebra element directing expression in only stripe 4, as loss of expression in stripe 4 coincided with the loss of expression in the other stripes (Dearolf et al., 1989a).

Here, we evaluated conclusions drawn about *ftz* regulation from the use of reporter gene analysis in transgenic organisms, by precisely deleting the zebra element from its endogenous locus using CRISPR-Cas9 genome editing. Contrary to expectations that this CRE is either the major driver of expression of all 7 *ftz* stripes (Hiromi et al.,

1985) or that it is the sole driver of *ftz* stripe 4 initiation (Schroeder et al., 2011), we found that levels of expression of *ftz* stripe 4 were reduced but not fully lost in homozygous mutant embryos, while expression of other stripes was only marginally reduced. Animals homozygous for this genomic deletion were viable and fertile, with frequent defects observed in regions of larvae and adults corresponding to the stripe 4 expression domain. The *ftzΔZ*$^{-/-}$ homozygous line will be useful for further studies of Ftz function, as target gene expression was altered specifically in this portion of the embryo, providing an internally controlled environment – within a single embryo – to monitor gene expression. Overall, our results show that even for a well-studied gene like *ftz*, our understanding of the qualitative and quantitative contributions of CREs to endogenous gene expression is far from complete.

2.3 Materials and Methods

2.3.1 Transgenic plasmids

Genomic fragments for reporter constructs were amplified with NEB Q5 DNA polymerase (catalogue number M0491S) following product specifications; annealing temperatures were predicted with NEB Tm Calculator (https://tmcalculator.neb.com/#!/main). See Appendix A for the sequences of all primers and Appendix B for all probes used in this work. Zebra-73 was amplified with MF118 and MF119 and contains sequence from -670 to -74. Zebra-40 was amplified with MF118 and MF116 and contains sequences from -670 to -41. MF116, MF118, and MF119 all contain XbaI adapter sites on the 5' terminus. PCR products were inserted into the XbaI site of vector *placZ-attB* using Pyrite cloning (Fischer et al., 2018). Plasmids were

sequenced with SeqF and SeqR primers and sent to Rainbow Transgenic for embryonic microinjection into Bloomington Stock Center line 9740, which contains a 2nd chromosome PhiC31 docking site at Chromosome 2, 57F5, 2R:21645971.

2.3.2 Guide RNA and HDR template

Genomic targets for gRNAs matching the *ftz* zebra element were identified using CHOPCHOP (Labun et al., 2019) and the CRISPR Efficiency Predictor (Housden et al., 2015; Perrimon, 2020). DNA from *nos-Cas9* flies was sequenced to check for polymorphisms in the zebra element and surrounding regions to ensure a perfect match between gRNAs and the genome of flies in which targeting was to be carried out (Appendix C). Two gRNAs were selected and inserted into *pCFD4-U6:1_U6:3tandemgRNAs* (Addgene no. 49411) (Port *et al*., 2020), one matching *ftz* -634 to -614, the other matching -211 to -191. The Homology Directed Repair (HDR) template to generate *ftz*ΔZ was assembled by two PCR reactions using *nos-Cas9* genomic DNA and NEB Phusion HF DNA polymerase (catalogue number MD530S) to amplify the left and right homology arms. One reaction amplified sequence upstream of the zebra element (-2296 to -670), adding an AvrII site at the 3' end to produce the left homology arm. The second reaction amplified sequence from -73 to +1755, adding an AvrII site to the 5' end to produce the right homology arm. These fragments were annealed, amplified by PCR, inserted into pGEM and verified by sequencing. The plasmid was digested with AvrII and reannealed to generate a repair template containing 3455 bp of the *ftz* region 1627 bp upstream of the zebra element and 1828 downstream of the zebra element, with the zebra element itself (-671 to -74) replaced by a single AvrII site. The HDR template to generate *ftzΔZ-GAGA* was similarly assembled, but the second PCR reaction amplified

from -40 to +1755. The assembly was otherwise the same. Embryos were microinjected by Rainbow Transgenic Flies with an injection mix containing *pCFD4-U6:1_U6:3tandemgRNAs* (100 ng/µl), *pCFD3-ebony* (100 ng/µl), and *pGEM HDR* (250 ng/µl) vectors. The vector *pCFD3-ebony* (Addgene no. 83380) contains an *ebony* guide for Co-CRISPR (Kane et al. 2017). Selection of sgRNA, cloning of the *pCFD3* constructs, and cloning of the HDR templates for this chapter were conducted by Patricia Graham.

2.3.3 Genetics

Fly stocks used were: *y[1] M{w[+mC]=nos-Cas9.P}ZH-2A w[*]* (BDSC 54591), referred to as *nos-Cas9*; and *P{ry[+t7.2]=ftz-lacZ.ry[+]}TM3, e Sb[1] ry[*]/Dr[Mio]* (BDSC 3218), referred to as *Dr/TM3, Sb*. w^{1118} was used as wild type. Flies were reared at room temperature or at 25 °C. To generate deletion lines, each injected *nos-Cas9* adult was crossed to 3 *Dr/TM3, Sb* flies. If the G0 cross produced ebony progeny, ten F1 *Sb* individuals from that cross were used to setup F1 crosses (1 F1 *TM3, Sb* x 3 *Dr/TM3, Sb*). For G0 crosses producing no ebony progeny, six F1 crosses were set up. After individuals from the F1 crosses mated and laid eggs, each F1 *TM3, Sb* parent was screened for mutations. F2 *TM3, Sb* progeny from F1 individuals that had a deletion were crossed to each other to generate balanced lines. Two additional backcrosses were carried out to eliminate the *nos-Cas9* X chromosome.

To identify genomic deletions, gDNA was prepared from single flies (Gloor and Engels, 1992). 1 µl of the single fly preparation was used in a PCR reaction with one of the following primer sets, each of which each flank the zebra element: zebra1 (positions -754 to -733) and zebra2 (positions +710 to +731), or Dm zebra fullL (positions –1542 to

-1520) and Dm zebra fullR (+1121 to +1142). The wild type and the deletion PCR products were expected to be 1485 bp and 883 bp for the zebra1,2 primer pair (data not shown) and 2685 bp and 2089 bp with the zebraFullL,R primer pair. PCR bands indicating deletions were sequenced. The co-CRISPR strategy was effective in identifying HDR events: of 23 fertile G0 adults recovered, 9 (75%) that produced ebony offspring also produced *ftz* deletions as opposed to 5 (45%) for those not producing ebony progeny. For 10 of the 14 (71%) G0 that had an HDR event, at least half of the progeny carried the deletion. Screening of *ftzΔZ* and *ftzΔZ-GAGA* was conducted by Patricia Graham.

2.3.4 Gene expression analysis

For colorimetric *in situ* hybridization, digoxigenin-labeled probes were used following standard protocols (Kosman et al., 2004) and imaged using a Zeiss Microscope Axio Imager M1 microscope. For fluorescent *in situ* hybridization chain reaction (HCR), embryos were simultaneously stained for *ftz* and *eve* mRNA using a two-stage HCR protocol previously optimized for *Drosophila* (Choi et al., 2014; Surkova et al., 2019) except that the denaturing step with proteinase K was replaced by heating embryos barely submerged in PBST in a microcentrifuge tube to 90º C for 5 minutes. Regions of *ftz* and *eve* transcripts that do not match other transcripts in the *D. melanogaster* transcriptome (BLAST, FlyBase Dmel Release 6.26), were submitted to Molecular Instruments for probe design. Hairpin amplifier pairs were: *ftz*, Alexa Fluor 488; *eve*, Alexa Fluor 546. All HCR buffers were made in the lab using previously published specifications and storage conditions (Choi et al., 2014). Nuclei were stained with 5 ng/µl (1:1000 in PBST) Hoechst 34580 (ThermoFisher) for 8-10 minutes, rinsed 3x with PBST, and washed 3x in

PBST for 15 minutes. Embryos were then washed in PBS for 15 minutes before being mounted in Prolong gold antifade mountant (ThermoFisher).

All confocal images were taken on a Leica SP5X Laser Scanning Confocal microscope. The Leica Application Suite Advanced Fluorescence (LPA AF) software was used to control the hardware. Three 1 µm stacking optical sections, averaged 8 times, were scanned through the lateral surface of embryos to a total depth of 2 µm using a 20x objective. Embryo positions were marked using the "mark and find" feature within the LPA AF software. Following fluorescent image acquisition, sagittal scans with a 63x objective were taken with white light to visualize the degree of nuclear elongation and membrane deposition in all marked embryos. This allowed for determination of which samples had completed or were near complete cellularization (Surkova et al., 2013). These criteria were used to select six wild type embryos and five *ftzΔZ* embryos that were suitable for the quantitative dataset. The means of stripe intensity integration values from the wild type and *ftzΔZ* mutant datasets were tested for statistical significance with the Student t-test. P-values were adjusted with the Holm's method to correct for multiple comparisons. Normality of data was assessed statistically with the Shapiro-Wilk test.

Appendix D contains detailed methods on how the HCR images were visualized, background processed (Appendix E), and quantitated. These methods were based on various works (Surkova, Myasnikova, Janssens, et al., 2008; Surkova, Myasnikova, Kozlov, et al., 2008; Surkova et al., 2013; Surkova et al., 2019). In short, the fluorescent integration of each stripe was determined using the width of each stripe and the fluorescent intensity to calculate the area under each stripe in the central 10% of the

embryo. Thus, both stripe width and stripe intensity are an intrinsic part of the calculation.

2.4 Results

2.4.1 Deletion of the *ftz* zebra element

To test how the zebra element contributes to endogenous *ftz* gene expression, we used CRISPR-Cas9 to generate a precise deletion of this genomic region. A common strategy is to replace deleted regions with a reporter gene expressing a visible marker, such as *3XP3-GFP* (Berghammer et al., 1999), which greatly facilitates identification of genome editing events. However, we reasoned that insertion of a strong enhancer and promoter into the zebra-element region could impact expression of *ftz* itself and thereby confound effects of the zebra element deletion. Thus, although our strategy required screening individual flies by PCR, we designed an HDR template that removes the entire zebra element, replacing it only with a single AvrII site (Fig. 2-1B).

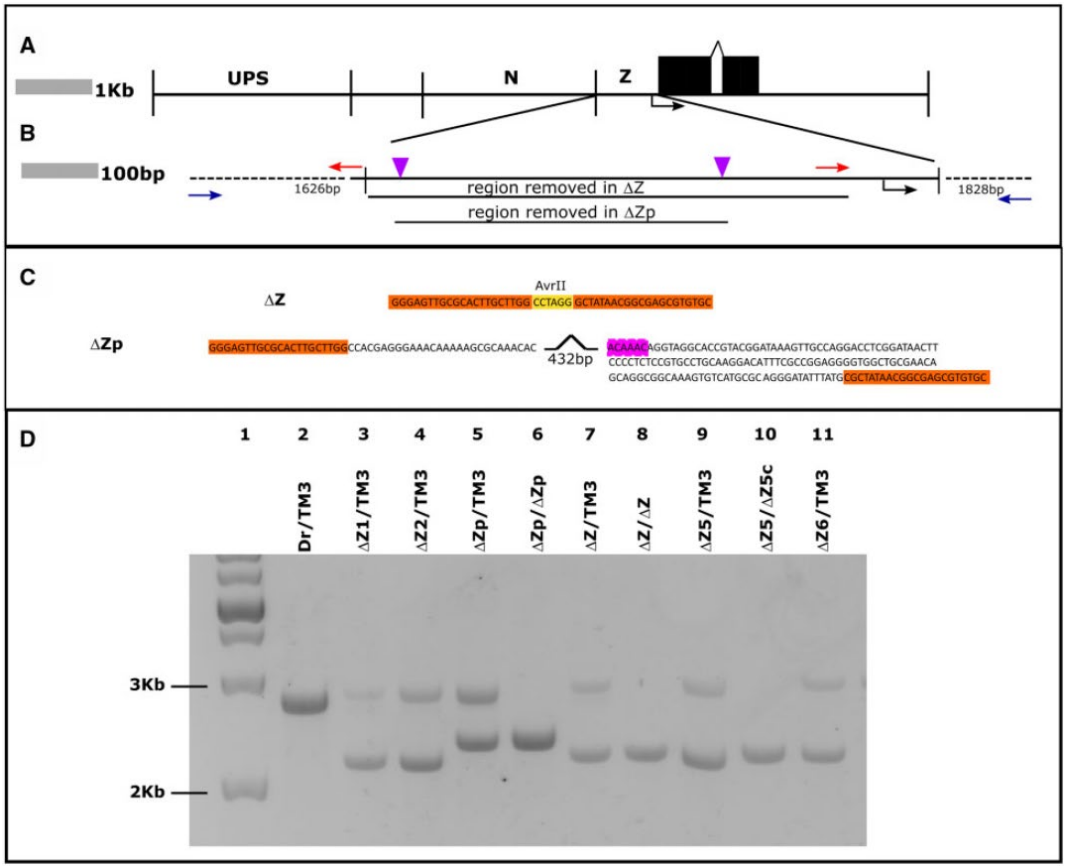


Figure 2-1. CRISPR deletion of the *ftz* zebra element. (A) Map of the *ftz* genomic region spanning the 10 kb KpnI fragment shown to rescue *ftz* mutants (Hiromi et al., 1985). UPS, upstream element; N, neurogenic element, Z, zebra element. Transcription start site indicated by the black arrow below the zebra element. Black boxes, *ftz* exons. (B) Expanded view of the zebra element. Positions of gRNAs used indicated with purple triangles. To produce the arms of the HDR templates, regions flanking the zebra element were amplified using the primers indicated with red and blue arrows. The regions removed in full *ftzΔ*Z and partial *ftzΔ*Zp deletions are indicated by black lines below the diagram of the intact region. (C) Sequences of the final full and partial deletions. In the full deletion, the zebra sequences are replaced by an AvrII site (gold box). The pink box contains bases that constitute an indel in *ftzΔ*Zp. (D) Representative gel showing the PCR products of the region surrounding the zebra element produced from single fly preparations. Lanes: 1, DNA ladder; 2, control fly (*Dr/TM3*) with no deletion; 3, heterozygous fly from deletion line 1; 4, heterozygous fly from deletion line 2; 5, heterozygous fly from the partial deletion line; 6, homozygous fly from the partial deletion line; 7, heterozygous fly from the viable full deletion line

(*ftzΔZ*); 8, homozygous fly from the viable full deletion line; 9, heterozygous fly from deletion line 5; 10, homozygous fly from deletion line 5; 11, heterozygous fly from deletion line 6. Design of sgRNA, HDR template, and PCR screen conducted by Patricia Graham.

The HDR template leaves intact the *ftz* basal promoter, including a GAGA site downstream of the deletion, in order to avoid impacts on basal promoter activity. The neurogenic element is left intact upstream of the deletion (Fig. 2-1B). The mutation generated in this way removes 597 bp (-670 to –74, Fig. 2-1B). G0 individuals were crossed to *Dr/TM3, Sb* flies, and F1 individuals were screened for the presence of a deletion, identifying 75 F1 flies carrying a deletion detectable by PCR. From these, we generated six balanced lines (Materials and Methods). Genotyping demonstrated that five lines carried the exact zebra element deletion described above (*ftzΔZ*, Fig. 2-1C). One line carried a partial deletion (*ftzΔZp*) of 432 bp removed from the 5' portion of the zebra element (-637 to -207). This deletion appears to have resulted from a non-homologous end joining event following cleavage by both of the gRNAs, as it removed most of the region between those cleavage sites as well as 3 bp 5' to the first gRNA (Fig. 2-1B, C). The 3' end of the deletion occurred within the second gRNA and removes 4 bp from the 5' end of the gRNA cleavage site.

Although we expected that deletion of the zebra element would result in homozygous lethality, we recovered homozygous viable adults for two lines with the complete deletion and one with the partial deletion. One of the full deletion lines was homozygous sterile, but the other was fertile. As these mutations were balanced over a *TM3, Sb* chromosome, the homozygotes were identified by the absence of *Sb* and verified with PCR (Fig. 2-1D; Materials and Methods); only the wild type fragment (2685 bp) was amplified from control flies (balancer line, Fig. 2-1D lane 2). Amplification of

genomic DNA from *Sb* adults expected to be heterozygous for *ftzΔZ* generated two fragments: the wild type and the expected smaller fragment (2089 bp; five independent *ftzΔZ* lines, Fig. 2-1D lanes 3, 4, 7, 9,11). Amplification of genomic DNA from heterozygotes for *ftzΔZp* generated the wild type fragment and an additional fragment that was slightly larger than that seen for *ftzΔZ* (2252bp, Fig. 2-1D lane 5). DNA isolated from flies identified as homozygotes by virtue of absence of the *Sb* marker produced only the smaller fragments (*ftzΔZ* lines, Fig. 2-1D lanes 8 and 10; *ftzΔZp,* Fig. 2-1D lane 6). Sanger sequencing was repeated on individual adult flies lacking the *Sb* marker, verifying homozygosity for the deletion (Appendices F and G). Complementation tests indicated that lethality and sterility seen in 4 of the 6 full deletion lines resulted from off-target effects (Table 2-1), and we proceeded to characterize the homozygous viable and fertile *ftzΔZ* and *ftzΔZp* lines. Taken together, these results demonstrate that the *ftz* zebra element is not necessary for *Drosophila* viability or fertility.

Table 2-1. Complementation tests suggest second site mutations in *ftzΔZ1c,2c,5c* and *6c*.

	ΔZ1c/TM3Sb % non-*Sb* (n)	*ΔZ2c/TM3Sb* % non-*Sb* (n)	*ftzΔZ* % non-*Sb* (n)	*ΔZ5c/Tm3Sb* % non-*Sb* (n)	*ΔZ6c/TM3Sb* % non-*Sb* (n)
ΔZ1c/TM3Sb	0 (175)				
ΔZ2c/TM3Sb	0 (74)	0 (354)			
ΔZ	40 (n=116)	47 (162)	100 (133)		
ΔZ5c/TM3Sb	12 (n=145)	30 (105)	51 (101)	13 (103)	
ΔZ6c/TM3Sb	30 (n=156)	22 (194)	40 (173)	30 (130)	0 (110)

2.4.2 The *ftz* zebra element is required for development of a single body segment

Since deletion of the *ftz* zebra element did not result in lethality, we were able to establish and maintain true-breeding lines of *ftzΔZ*$^{-/-}$ and *ftzΔZp*$^{-/-}$ for phenotypic analysis. Although many *ftzΔZ*$^{-/-}$ flies survive to adulthood, viability is reduced compared to wild type controls (Table 2-2). While ~60% of eggs laid by controls eclosed to adulthood, only

~30% of *ftzΔZ*$^{-/-}$eggs developed to adults, with losses at each stage – egg hatching, pupation, and eclosion – contributing to this decline. In contrast, *ftzΔZp*$^{-/-}$ homozygotes displayed hatching, pupation and eclosion rates similar to controls.

Table 2-2. Lethal periods of embryos homozygous for *ftz* deletions.

Strain	# eggs scored	% hatched[a]	% pupae[b]	% adult[c]	# adults (final %)[d]
w^{1118}	260	81	85	96	159 (61)
ftzΔZp	290	84	81	96	185 (64)
ftzΔZ	250	59	69	79	72 (29)

a. % hatched = (# empty eggshells/# eggs) x100

b. % pupae = (#pupae/# larvae) x100

c. % adult = (# adults/# pupae) x100

d. final % = (# adults/# eggs) x100

The *ftzΔZ*$^{-/-}$ and *ftzΔZp*$^{-/-}$ homozygous larvae displayed a range of phenotypes, most involving segment A3 (Fig. 2-2, Table 2-3). The most common phenotype for *ftzΔZ*$^{-/-}$ larvae was the absence of part (Fig. 2-2B) or all (Fig. 2-2C) of the denticle belt on segment A3, with other segments indistinguishable from wild type controls (n=117: 84% missing all of the A3 denticle belt, 1% missing part). Some larvae (15%) displayed defects in additional denticle belts (Fig. 2-2D, Table 2-3), most commonly partial or complete loss of A5 denticles (11%), and a few (3%) showing defects in multiple denticle belts or in just A6 or A8 (1%). No normal cuticles were observed for *ftzΔZ* homozygous larvae. For *ftzΔZp* homozygotes, an almost equal mix of partial and full loss of denticle belt A3 was observed (n=52: 44% each). None displayed defects in multiple segments and some cuticle preparations had wild-type-like appearance (12%).

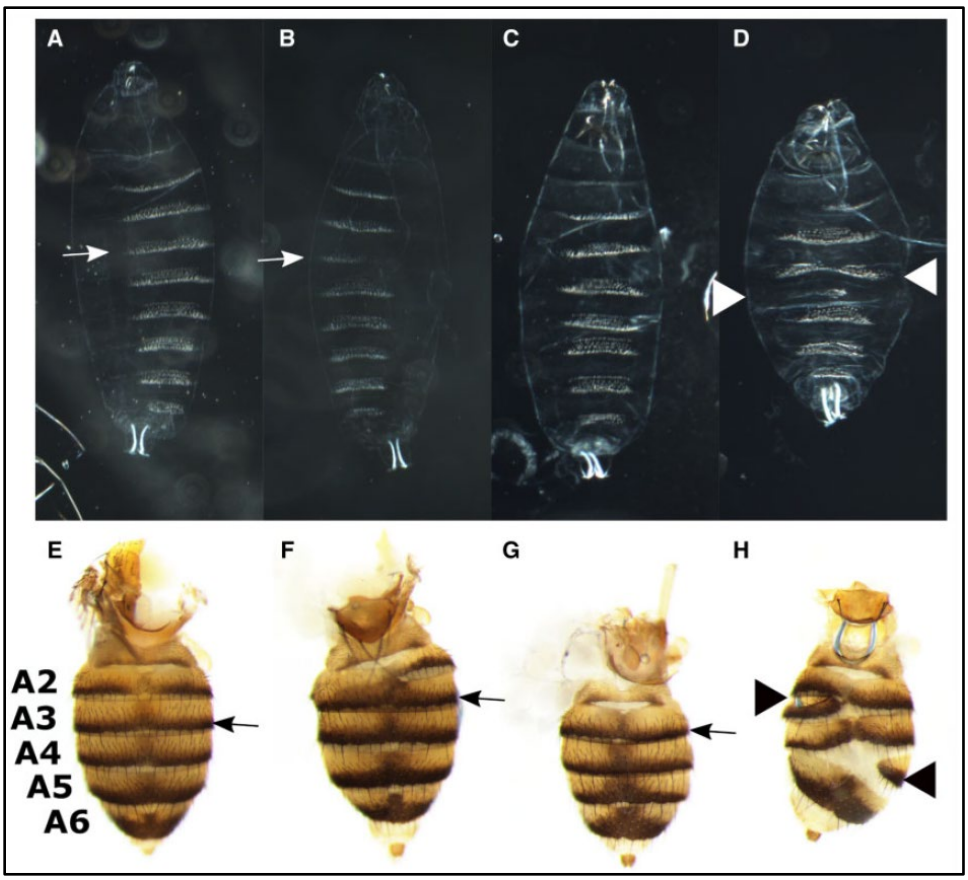


Figure 2-2. *ftzΔZ*$^{-/-}$ larvae and adults have abdominal defects and are missing a single segment. Cuticles (A-D) or abdomens (E-H), anterior is up. (A-D) Larval cuticles from *ftzΔZp*$^{-/-}$ (A-B), *ftzΔZ*$^{-/-}$ (C-D). White arrows indicate denticle belt A3, and white arrowheads indicate abnormal denticle belts. (A) *ftzΔZp*$^{-/-}$ flies produce some cuticles that are normal. (B-C) Both *ftzΔZp*$^{-/-}$ and *ftzΔZ*$^{-/-}$ produce cuticles with the A3 denticle belt partially (B) or completely (C) missing. (D) *ftzΔZ*$^{-/-}$ also produces cuticles with defects in multiple denticle bands. (E-H) Adult abdomens of w^{1118} controls **(**E**)** or *ftzΔZ*$^{-/-}$ homozygotes (F-H). Black arrows indicate the darkly pigmented region of segment A3, and black arrowheads indicate abnormal segments. Adult flies from *ftzΔZp*$^{-/-}$ and *ftzΔZ*$^{-/-}$ exhibit a similar range of phenotypes: wild-type-like (not shown), partially (F) or fully (G) missing portions of segments A2 and A3, or multiple segments affected (H). Larva cuticles and adults prepared and imaged by Patricia Graham.

Table 2-3. Cuticular phenotypes of larvae homozygous for *ftz* deletions.

Deletion strain		% Normal	% Missing part of A3	% Missing all of A3	% with defects in segments in addition to or other than A3
ftzΔZp (n=52)		12	44	44	0
ftzΔZ (n=117)		0	1	84	15

Similarly, adults homozygous for *ftzΔZ* or *ftzΔZp* eclosed with a range of phenotypes impacting segments A2 and A3, as well as more posterior segments (Fig. 2-2E-H, Fig. 2-3, and Table 2-4). *ftzΔZ*$^{-/-}$ adult flies were uniformly smaller in size than wild type adults, with smaller abdomens than wild type controls. Abdomens of *ftzΔZ*$^{-/-}$ flies (n=639) either partially (18%) or completely (60%) lacked either A2 or A3 (Fig. 2-2F-G; Fig. 2-3B-C; Table 2-4). The segment remaining most often had an A3 sternite bristle pattern, but sometimes had an A2 pattern or an ambiguous appearance. The ambiguity of identity seen here is reminiscent of the ambiguity seen in *hopscotch* (*hop*) mutants, which also survive to adulthood, missing a single segmental region, although *hop* mutants impact segments A4/A5 (Perrimon and Mahowald, 1986). Abdomens of other *ftzΔZ*$^{-/-}$ flies (20%) displayed defects in multiple segments, with the posterior segments most severely affected (Fig. 2-2H, arrowheads). A small number of adult abdomens had wild-type-like appearance (2%). *ftzΔZp*$^{-/-}$ adults (n=834) showed a similar range of phenotypes: 39% were missing part and 9% were missing all of A2 or A3, 50% were wild-type-like, and 2% had defects in segments in addition to or other than A2 and A3.

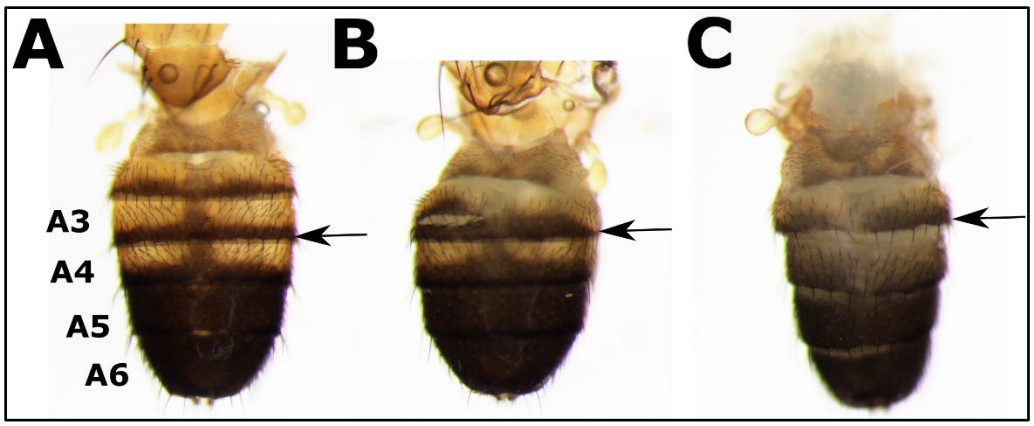


Figure 2-3. Adult phenotypes in males. Abdomens from w^{1118} (A) or *ftzΔZ* (B-C) male flies. Anterior is up. Arrows indicate the darkly pigmented region of segment A3. $ftz\Delta Z^{-/-}$ flies often partially (B) or fully (C) fail to develop portions of segments A2 and A3. Adults processed and imaged by Patricia Graham.

Table 2-4. Phenotypes of adults homozygous for *ftz* deletions.

Deletion strain	% Normal	% Missing part of A2/A3	% Missing all of A2/A3	% with defects in segments in addition to or other than A2/ A3
ftzΔZp (n=834)	50	39	9	2
ftzΔZ (n=639)	2	18	60	20

In sum, these experiments demonstrate that the *ftz* zebra element plays a major role in the development of segments A2 and A3 while also impacting regions posterior to them. However, it appears to be fully dispensable for the formation of the thoracic and A1 segments. $ftz\Delta Zp^{-/-}$ behaves as a weaker allele than $ftz\Delta Z^{-/-}$, but the types of defects were similar, consistent with distributed regulatory information within the zebra element (Dearolf et al., 1989a). Additional experiments indicate that a separate deletion of the zebra element to the -40 position ($ftz\Delta Z\text{-}GAGA^{-/-}$) results in the same range of phenotypes as observed for $ftz\Delta Z^{-/-}$ (Fig. 2-4 and Table 2-5). Finally, it appears that the presence of segment A2/A3 is not required for fly viability or for reproduction. The observed fertility is consistent with the gonads originating in segment A4 (Karch et al. 1985), which was not severely impacted in most $ftz\Delta Z^{-/-}$homozygotes.

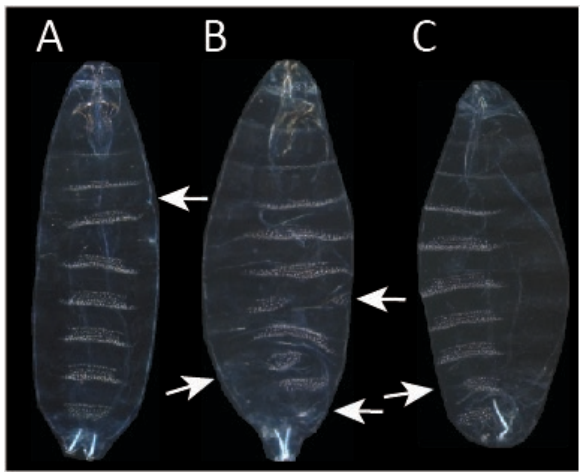


Figure 2-4. *ftzΔZ-GAGA*$^{-/-}$ larvae have segmental defects of A3 and other segments. (A) Most mutant larvae develop without segment A3. (B-C) White arrows indicate locations on larvae that develop with partially formed denticle bands in other segments in addition to missing A3.

Table 2-5. Cuticular phenotypes of larvae homozygous for *ftzΔZ-GAGA*$^{-/-}$.

Deletion strain	% Normal	% Missing part of A3	% Missing all of A3	% with defects in segments in addition to or other than A3
ftzΔZ-GAGA	2	0	91	7

2.4.3 *ftz* stripe 4 expression is reduced in *ftzΔZ*$^{-/-}$ mutants

Although the zebra element was previously shown to direct expression of all seven *ftz* stripes in reporter-transgenes, the *ftzΔZ*$^{-/-}$ phenotype suggested that stripes were differentially affected in this mutant. To assess this, we analyzed *ftz* expression by *in situ* hybridization using digoxigenin-labeled probes. At the cellular blastoderm stage when *ftz* stripes peak, the expression of *ftz* stripe 4 was greatly decreased in *ftzΔZ*$^{-/-}$ embryos (Fig. 2-5B, arrow). We next used fluorescent *in situ* hybridization chain reaction (HCR) to determine the extent to which stripe 4 expression in *ftzΔZ*$^{-/-}$ mutants differed from wild type. Expression of *eve* was used as an internal control in these reactions and appeared unaffected by the deletion, as expected (Fig. 2-5C, D, yellow). In contrast, the change in

ftz stripe 4 was evident by the gap in expression in the central region of the embryo (Fig. 2-5D, arrow).

Calculating the area under each stripe's fluorescent signal intensity (see Appendix D) revealed that stripe 4 was reduced to 27.5 ± 18.5% of wildtype levels in *ftzΔZ*$^{-/-}$ mutants (n=6 wild type, 5 *ftzΔZ;* p-value = 0.0176; Fig. 2-5E-G, Table 2-6). *ftz* stripe 4 is expressed in the primordia of segments A2 and A3, the segments most often affected in *ftzΔZ*$^{-/-}$ mutants (Fig. 2-2, Tables 2-3 and 2-4). These results suggest that for stripe 4, ~25% of wild type is very close to the lower threshold of *ftz* transcript that is both necessary and sufficient to direct segmental development. Comparing average stripe widths between wild type and *ftzΔZ*$^{-/-}$ mutants also revealed that only stripe 4 was significantly reduced to about 53% the width of wild type. Notably, there was no statistically significant difference detected for the average central position of each stripe when comparing wild type to *ftzΔZ*$^{-/-}$ mutants. Further, the positioning of *ftz* stripes relative to *eve* was the same in the *ftzΔZ*$^{-/-}$ mutants as compared to wild type (data not shown). However, it is possible that alterations in the timing of *ftz* expression may also affect the phenotypes produced.

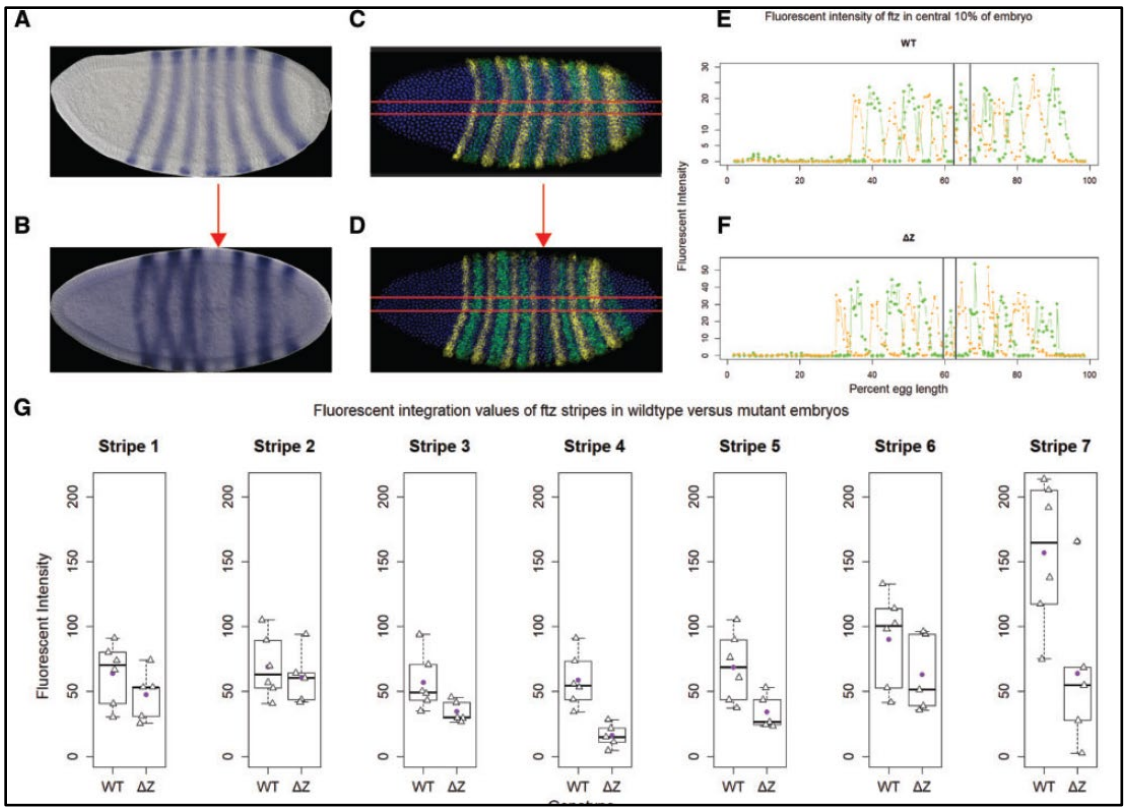


Figure 2-5. *ftz* stripe 4 is significantly reduced in *ftzΔZ*$^{-/-}$ mutant embryos. (A-B) Colorimetric *ftz* in situ hybridization. (A) In wild type embryos, *ftz* is expressed in seven alternating stripes approximately four cells wide. (B) Though all the stripes are present in *ftzΔZ*$^{-/-}$ embryos, stripe 4 is weak (arrow). (C-D) Fluorescent *in situ* hybridization chain reaction (HCR; *ftz*, green; *eve*, yellow; nuclei, blue Hoechst). Qualitative differences (A-B) were also observed in HCR-stained embryos (C, wild type; D, *ftzΔZ*). Red lines indicate the approximate central 10% of the embryo's dorsal-ventral axis from which centroid fluorescent signals are plotted in a one-dimensional analysis along the anterior-posterior axis (E**,** wild type; F, *ftzΔZ*$^{-/-}$). Points within line graph represent centroids. Vertical lines indicate the boundaries of *ftz* stripe 4. (G) Fluorescent integration for each stripe in each embryo. Analysis of each stripe is indicated in a separate box plot. Stripe 4 is significantly reduced to 27.5 ± 18.5% in *ftzΔZ*$^{-/-}$ embryos relative to wild type embryos (G; ± represents standard deviation; p-value = 0.0176; n=6 wild type, 5 *ftzΔZ*$^{-/-}$). Triangles represent mean stripe 4 integrations from each embryo. Purple circle represents mean of means.

Consistent with changes in *ftz* stripe 4, the expression of two genes known to be regulated by Ftz, *sloppy-paired1* (*slp1*) and *engrailed* (*en*), were perturbed in the *ftz* stripe

4 domain of *ftzΔZ*$^{-/-}$ mutants (Fig. 2-6). As predicted by the Gergen lab, Ftz is required to repress *slp1* expression (Prazak et al., 2010; Hang and Gergen, 2017). We found loss of *slp1* interstripe repression in *ftzΔ*$^{-/-}$ mutants, specifically in the domain impacted by *ftz* stripe 4, between *slp1* stripes 8 and 9 (Fig. 2-6B, C).

For *en*, a well characterized direct target of Ftz, a range of expression was observed: in some embryos, *ftz* stripe 4 expression was sufficient to generate a wild-type-like *en* pattern (Fig. 2-6E); in others, *en* stripe 8 was weak but detectable (Fig. 2-6F); and, finally, in some embryos, *en* stripe 8 was undetectable (Fig. 2-6G). Occasionally, embryos displayed additional abnormalities in *en* expression, such as missing *en* stripes 6 and 8 and weak expression of stripes 10 and 14 (Fig. 2-6H), or a slight shift in position of an *en* stripe (Fig. 2-6F, 6H). This is similar to shifts of *en* stripes after partial rescue of an *eve* mutant with stripe-specific CREs driving *eve* expression, which also correlated with loss of specific segments (Fujioka et al., 2002). Thus, the range of defects observed morphologically reflects the range of defects seen for target gene expression. The true-breeding *ftzΔZ*$^{-/-}$ mutant lines, in which patterning is disrupted differentially in different regions within single embryos, provide a tool to examine the role of *ftz* in regulating other target genes.

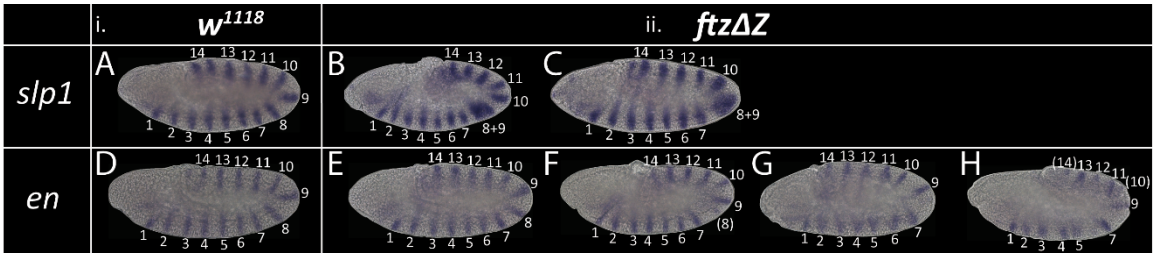


Figure 2-6. Perturbations of *slp1* and *en* gene expression in *ftzΔZ*$^{-/-}$ mutants. Colorimetric *in situ* hybridization to examine *ftz* target gene expression. (A-C) *slp1* expression in (A) wild type control and (B-C) *ftzΔZ*$^{-/-}$ mutants. (D-H) *en* expression in wild type control (D) and *ftzΔZ*$^{-/-}$ mutants (E-H). A range of

phenotypes was observed for *en* staining in *ftzΔZ*$^{-/-}$ embryos; E: wild-type-like, F: weak stripe 8, G: missing stripe 8, H: multiple abnormalities, including missing stripe 6 and 8 and a weak stripe 10 and 14. Weak stripes conveyed with numbers in parentheses.

2.4.4 The zebra element contributes marginally to expression of other stripes

While the decrease in *ftz* stripe 4 explained the phenotypic effects on A2/A3 in *ftzΔZ*$^{-/-}$ mutants, transgenic reporters containing the zebra element were expressed in seven stripes (Hiromi et al., 1985; Dearolf et al., 1989b). What role does the zebra element play for these other stripes? Although not evident in colorimetric *in situ* experiments, at peak expression, levels of expression in all seven stripes were lower in the mutant compared to wild type controls in HCR labeled embryos (Fig. 2-5G and Table 2-6). Interestingly, stripes 4, 5 and 7 showed the greatest reduction. Stripe 5 was reduced to 49.6 ± 27.2% and stripe 7 was reduced to 40.7 ± 42.2%. Though neither of these reductions were statistically significant after correcting for multiple comparisons, stripe expression level variance likely explains the range of defects seen in posterior abdominal segments of some *ftzΔZ*$^{-/-}$ mutant animals. For example, if *ftz* is reduced below the critical value for determination in either or both stripes 5 and 7 in addition to stripe 4 for some of the mutant embryos, it could result in the aberrant segmentation defects that are sometimes observed in segments other than A2/A3. Since the majority of mutant embryos develop with only A2/A3 parasegment defects, we suggest that the average *ftz* expression for stripes 5 and 7 are above the threshold level required for segmental development in most *ftzΔZ*$^{-/-}$ mutants, though the precise value of the threshold required for segmentation in these stripes cannot be determined with this dataset. Posterior defects

likely occur in the embryos expressing less than the critical value, although subtle alterations in timing of these posterior stripes may also impact morphology.

Table 2-6. Quantitative analysis of levels of individual *ftz* stripes.

Stripes	w^{1118} Integrations	*ftzΔZ* Integrations	Percentage (*ftzΔZ*/w^{1118})	P value	P adj
Stripe 1	63.9 ± 23.6	47.4 ± 19.6	74.2 ± 41.1	0.238	0.609
Stripe 2	69.1 ± 24.2	60.9 ± 21.1	88.2 ± 43.4	0.564	0.609
Stripe 3	56.9 ± 21.6	34.6 ± 8.37	60.8 ± 27.4	0.0546	0.218
Stripe 4	58.6 ± 20.6	16.1 ± 9.22	27.5 ± 18.5	0.00251	0.0176
Stripe 5	68.9 ± 26.5	34.1 ± 13.4	49.6 ± 27.2	0.024	0.144
Stripe 6	90.3 ± 35.7	63.2 ± 29.8	70 ± 43.1	0.203	0.609
Stripe 7	157 ± 55.5	63.9 ± 62.3	40.7 ± 42.2	0.0317	0.159

In summary, it appears that variation around the minimal level of *ftz* transcript required for segmentation causes the variation in phenotype. Given that the average level of expression in those stripes is higher than that of stripe 4, the number of embryos with expression low enough to cause an effect would be fewer, consistent with the observation that relatively few adults show defects in segments other than A2 or A3.

2.4.5 The order of *ftz* stripe activation is perturbed in *ftzΔZ*$^{-/-}$ mutants

The *ftz* seven striped pattern develops via a dynamic, sequential addition of stripes (Yu and Pick, 1995; Surkova, Myasnikova, Janssens, et al., 2008). We therefore considered the possibility that, in addition to impacting quantitative levels of stripe expression, the pattern of stripe development might be perturbed in *ftzΔZ*$^{-/-}$ mutants. Though wild type embryos show some variation in the order of *ftz* stripe activation (Surkova, Myasnikova, Janssens, et al., 2008), we found the generalized order to be as

previously observed (Yu and Pick, 1995): first, stripes 1 and 5 appear as the nuclei undergo elongation (Fig. 2-7i A). Stripes 2 and 3 appear next (Fig. 2-7i B), followed by a combined stripe 6-7 that appears initially on the ventral side of the embryo and progresses dorsally (Fig. 2-7i C). At this point, the nuclei are fully elongated and plasma membrane deposition begins at the apical pole, continuing towards the cortical cytoplasm. Stripes 1, 2, 3, and 5 become more established, while stripes 6 and 7 resolve into separate stripes (Fig. 2-7i E). Stripe 4 appears last, completing a pattern of all seven stripes (Fig. 2-7i F). Levels of expression in all seven stripes increase as the plasma membrane extends to the basal plane of the nuclei (Fig. 2-7i H). This pattern is maintained as membrane deposition completes cellularization (Fig. 2-7i I). When the ventral and cephalic furrows form, the embryo has entered the germband extension (GBE) phase, and *ftz* stripes begin to fade (Fig. 2-7i J). In summary, the *ftz* stripes form in the following order in a wild type background: 1+5, 2+3, a combined 6/7 which then resolves into separate stripes, and finally 4.

Initiation of *ftz* stripe formation in *ftzΔZ*$^{-/-}$ mutants differed in several ways from the wildtype pattern. First, the order in which stripes arose differed: stripes 1 and 3 appeared first (Fig. 2-7ii A), then stripe 2 (Fig. 2-7ii B) before posterior stripes 5 and 6 (Fig. 2-7ii C). Stripes 1, 2, 3, 5, and 6 continued to increase in intensity (Fig. 2-7ii D-E) and were joined by stripe 7 (Fig. 2-7ii F). Unlike wild type, in which stripes 6 and 7 arise as a fused stripe, stripe 7 appeared separately from stripe 6 in *ftzΔZ*$^{-/-}$ mutants. Stripe 4 appeared last, as in wild type embryos, completing the set of seven stripes that increase in intensity until GBE begins (Fig. 2-7ii G-J). In summary, *ftz* stripes arose in the following order in *ftzΔZ*$^{-/-}$ mutants: 1, 3, 2, 5+6, 7, with 4 being the last. A very similar pattern was

observed for *ftzΔZ-GAGA* embryos at the start of the activation (Fig. 2-7iii A-F). However, there was a divergence in that stripe 7 took slightly longer to activate (Fig. 2-7iii G-I). Notably, stripe 4 appeared faint even at peak signal (Fig. 2-7iii J). For *ftzΔZp*$^{-/-}$, the order was again slightly different: 1+5, 3+6, 2, 7, with 4 being the last one (data not shown). The major differences in the appearance of *ftz* stripes are in the posterior region of the embryo, most notably stripes 6 and 7, which arise in a coordinated fashion in wild type embryos but independently and in a different temporal order in *ftzΔZ*$^{-/-}$ mutants.

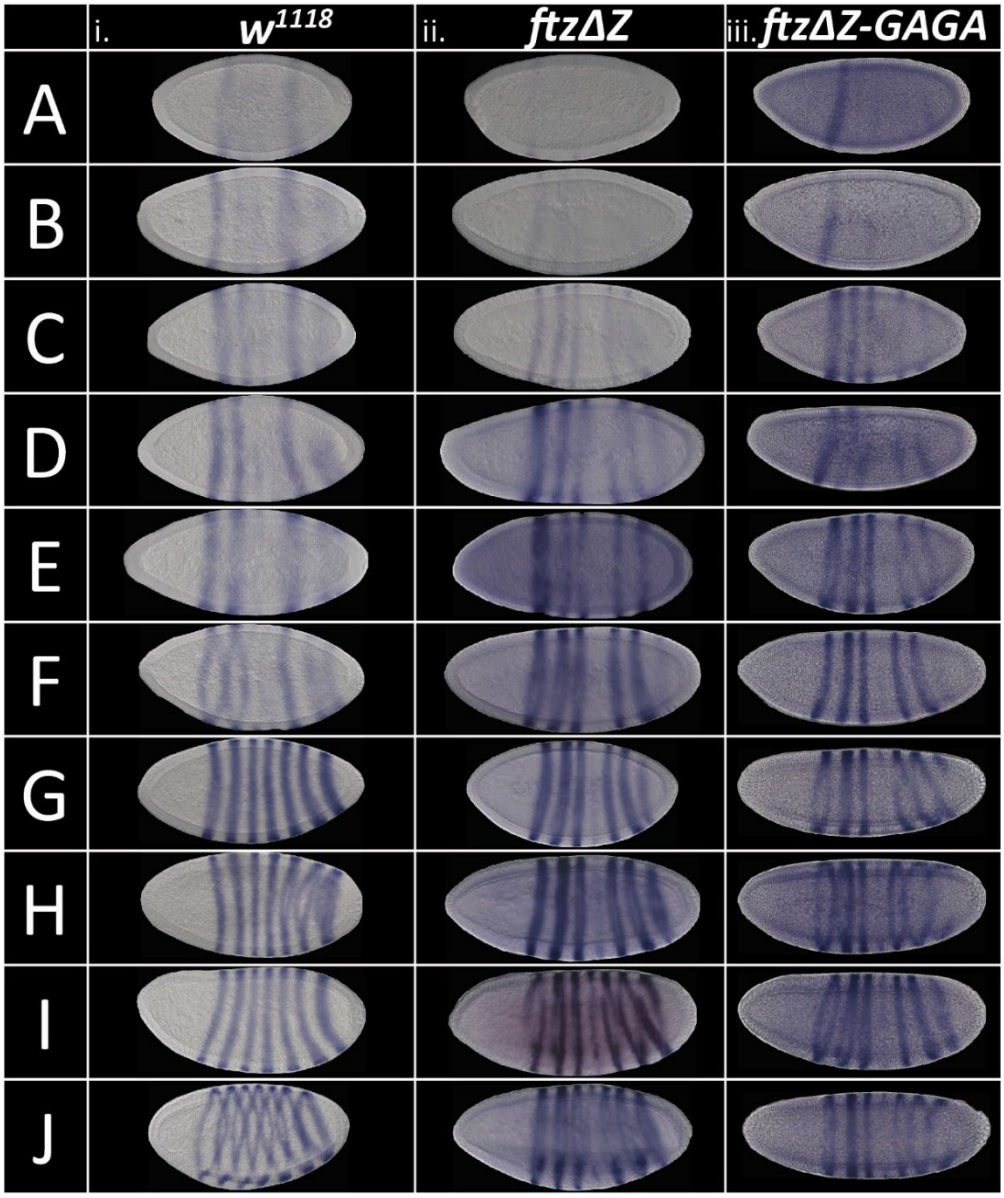


Figure 2-7. *ftz* stripes arise in different orders in *ftzΔZ*$^{-/-}$ mutants than in wild type embryos. Time course of *ftz* stripe activation with colorimetric *in situ* hybridization in (i) wild type, (ii) *ftzΔZ*$^{-/-}$ embryos, and (iii) *ftzΔZ-GAGA*$^{-/-}$ embryos. Embryos oriented anterior left, dorsal up with the following exceptions The relative *ftz* stripe intensity, degree of nuclear elongation, and degree of membrane deposition was examined to assess the developmental stage of each embryo so that embryos at similar stages could be compared. In wild type (i)**:** (A) stripes 1 and 5 start to appear. (B) stripes 2 and 3 start to appear. (C) A combined stripe 6/7 starts to appear without boundary separating 6 from 7. (D) stripe 2 is solid and 6/7 is stronger than C. (E) Stripes 6 and 7 start to resolve, and expression is stronger on the ventral side. (F) Stripes 6 and 7 are distinct and surround the embryo. Stripes 3 and 4 are very faint at this stage. (G) All stripes gain signal intensity. (H) Peak expression of all seven stripes. (I) *ftz* signal may start to degrade. (J) Invagination of the ventral and cephalic furrows. In *ftzΔZ*$^{-/-}$ (ii): (A) stripes 1 and 3 start to appear; unlike wild type (i), stripe 5 does not start to appear. (B) Stripe 2 starts to appear. (C) Stripes 5 and 6 appear very faintly. (D) Existing stripes become a little stronger. Unlike wild type, stripe 7 does not appear at this stage. (E) Stripes 1, 2, 3, 5, and 6 develop more. (F) Stripe 7 starts to appear. Unlike wild type, stripe 4 does not appear yet. (G) Stripe 4 starts to appear. The other six stripes become more defined. (H) Stripe 4 continues to develop. (I) Peak of *ftz* signal. (J) Invagination of the ventral and cephalic furrows. Unlike wild type (i), the zebra deletion mutant does not develop a strong stripe 4. In *ftzΔZ-GAGA*$^{-/-}$ (iii): (A-F) Beginning stages develop as described for *ftzΔZ*$^{-/-}$ (ii). (G) Stripe 7 develops more slowly than in *ftzΔZ*$^{-/-}$ (ii), and stripe 4 starts to activate. (H) Stripe 7 gains slightly more signal, and stripe 4 remains faint. (I) Stripe 7 gains much more signal, but stripe 4 remains faint. Peak observed signal is reached. (J) Start of gastrulation; signal is the same as in (I).

2.4.6 Refining the boundary between the zebra patterning element and the basal promoter

The *ftzΔZ* deletion described above removed the zebra element from position -670 at the 5' end, as defined by Hiromi et al. (1985) and Dearolf et al. (1989a), to position -74 at the 3' end (see Fig. 2-1). Although the *ftz* basal promoter had previously been defined as requiring sequences to -40 (Dearolf et al. 1989b), the genome edit that generated *ftzΔZ*

left intact additional sequence that includes a potential binding site for the chromatin remodeling protein GAGA Factor (Topol et al., 1991), which could influence basal promoter accessibility (Judd et al., 2021; reviewed in Chetverina et al., 2021). To ensure that the choice of the -74 position for the zebra element deletion did not inadvertently leave intact patterning information for *ftz* stripes, we generated two reporter genes: *zebra-40* includes sequences between -670 and -41 (*zebra-40*), while *zebra-73* includes sequences between -670 and -74. Each fragment was fused upstream of a basal *hsp70* promoter in vector *placZ-attB* and integrated into the *Drosophila* genome at a site on chromosome 2 previously shown to allow embryonic expression (Fig. 2-8 A-C).

These two reporter transgenes were expressed in seven stripe patterns indistinguishable from each other and from the pattern previously reported for *ftz* zebra element transgenes (Hiromi et al., 1985; Dearolf et al., 1989b). At the mid-blastoderm stage, expression was detected in stripes, with stripe 7 being the strongest, stripe 6 barely detectable, stripes 5 and 4 clearly formed, stripe 3 emerging, stripe 2 well resolved and stripe 1 not yet detectable (Fig. 2-8D, E). As previously noted, ectopic expression in the head region is observed for these zebra element reporters (Hiromi et al., 1985). At the cellular blastoderm stage, stripe 1 remained barely detectable, appearing more strongly on the dorsal side (Fig. 2-8F, G, asterisk) while all other stripes had clearly resolved, with stripe 7 remaining the strongest. By early- to mid-germband extension (Fig. 2-8H, I), stripes 2-7 were well developed, with expression stronger in the mesoderm than ectoderm, as previously reported (Hiromi et al., 1985). Stripe 1 remained faint throughout. These results corroborate the phenotypes observed in *ftzΔZ-GAGA*$^{-/-}$, which is the inverse of the *zebra-40* transgenes, in that they were very similar to those observed

from the *ftzΔZ*$^{-/-}$ mutant. Thus, the region between -73 and -40 does not contribute to the striped pattern directed by the zebra element.

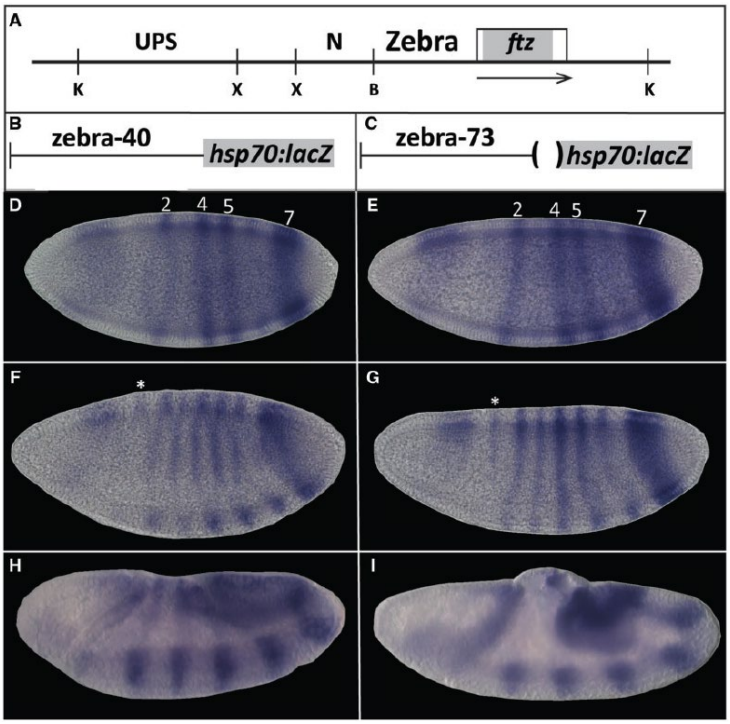


Figure 2-8. Zebra element transgenes striped expression pattern. (A) Schematic of *ftz* genomic locus with CREs identified by Hiromi et al. (1985) shown; upstream element (UPS), neurogenic element (N), and Zebra element. Restriction enzymes originally used to map these canonical elements are abbreviated below the line; KpnI (K), XbaI (X), and BalI (B). (B-C) Schematic of zebra element transgenes: Transgenes differ at the 3' end. *zebra-40* includes zebra element sequences up to 40 bp upstream of the *ftz* transcription start site; *zebra-73* includes zebra element sequences up to 73 bp upstream of the *ftz* transcription start site, omitting a GAGA site present in *zebra-40*. (D-I) Expression of *lacZ* for reporter genes indicated using colorimetric *in situ* hybridization. (D-E) mid blastoderm, *ftz* stripes labeled; (F-G) late blastoderm, asterisk above dorsal stripe 1; (H-I) germband extension. Embryos oriented anterior, left; dorsal, up. Embryos were staged according to degree of membrane deposition and morphology.

2.5 Discussion

Much has been learned about the regulation of transcription using reporter constructs expressed in transgenic animals. However, due to their very nature, these

experiments do not account for chromatin states that may impact utilization of identified CREs in their endogenous genomic locations. They also rarely address interactions between multiple CREs. These reporter transgenes identify genomic regions *sufficient* to direct spatiotemporal expression patterns but do not provide information about whether or to what extent these genomic regions are *necessary* for endogenous gene expression. Further, the fact that even exceedingly low levels of reporter gene expression can be detected above background in these assays can make the CRE appear quantitatively more important than it is in the native context.

This is exemplified by our studies here, using CRISPR-Cas9 to remove the *ftz* zebra element from its endogenous genomic location (Fig. 2-1). Based on previous experiments with transgenes, as well as those reported here (Fig. 2-8), we thought it likely that this CRE would be necessary for *ftz* function. However, we found that homozygotes developed into fertile adults, and we were able to establish stable *ftzΔZ*$^{-/-}$ stocks (Fig. 2-1D). Genetic manipulations of other segmentation genes also produced viable adult flies missing specific segments (Perrimon and Mahowald, 1986; Howard et al., 1988; Fujioka et al., 2002). Thus, flies can develop to and survive as adults missing whole body regions, although their fitness may be compromised. Despite this, because the zebra element deletion most strongly impacted segments A2/A3 and gonads develop from segment A4 (Karch et al., 1985), we were able to generate fertile, true-breeding lines for this particular mutant.

Since zebra element transgenes drive expression at some level in all seven stripes, deletion of this CRE was expected to reduce or eliminate expression to some extent in all seven stripes. The fact that zebra element transgenes are more strongly in the posterior

stripes suggested that a deletion might have a bigger effect on those stripes. In our analysis, we did find that expression was reduced to some extent in all stripes, with a slightly larger but non-significant effect on stripes 5 and 7 over stripes 1-3 and 6. Thus, the zebra element plays a role in modulating all of the stripes but the major impact of deleting it was a large decrease in levels of stripe 4 expression (Figs. 2-5, 2-7).

The original descriptions of the zebra element-directed transgenes revealed a mesodermal bias for these transgenes, with stripes being broader and less well-refined than upstream element-driven transgenes (Hiromi and Gehring, 1987; Dearolf et al., 1989a,b). The combination of the zebra and upstream elements generated a strong seven-stripe pattern with sharp stripes, more reminiscent of endogenous *ftz* expression, suggesting a synergistic interaction between these CREs. The level of expression of zebra plus upstream element transgene was also more than additive (see Yu and Pick, 1995). Other synergistic interactions between dispersed *ftz* CREs remain to be elucidated. In sum, while reporter studies were able to identify the zebra element as sufficient to drive expression in all seven stripes, they could not determine whether or to what extent this element was necessary for *ftz* expression; this serves as an example of the conflation of necessity and sufficiency. While we demonstrate this point here for the *ftz* gene, the same situation undoubtedly applies to "enhancer bashing" studies done for a wide cadre of developmentally-regulated genes in *Drosophila* and other organisms.

Previous research using reporter transgenes identified stripe-specific CREs that direct expression of individual *ftz* stripes (Calhoun and Levine, 2003; Schroeder et al., 2011). Whether these stripe-specific elements are the primary drivers of *ftz* stripe establishment, play secondary roles in the maintenance of individual stripes, are

evolutionary remnants, and/or are redundant with the two 7-stripe elements (zebra and UPS) remains to be determined. Our findings here support the notion that the stripe-specific elements play the predominant role in the establishment of the *ftz* stripe pattern, as suggested by Schroeder and Gaul (2011). These researchers identified stripe-specific elements directing the expression of each of the *ftz* stripes, except for stripe 4. The only region tested that directed expression in stripe 4 was a portion of the zebra element (*ftz*-1). This fragment also directed expression in more posterior stripes, predominantly 5 and 7, with stripe 6 less apparent. This is similar to the results of Dearolf *et al*. (1989a) in that the stripes were modulated above a general broad expression in the posterior region. Thus, a strong impact on stripe 4 was expected upon deleting the zebra element, based on the work of Schroeder et al. (2011).

However, contrary to their prediction that zebra contained the only primary CRE for stripe 4 establishment, the deletion of the zebra element demonstrates that the zebra element contributes quantitatively to stripe 4 expression but is not the sole element directing its expression - the zebra element lacks full regulatory information for stripe 4. Additional CRE(s) directing the establishment of *ftz* stripe 4 must reside elsewhere, either within the autoregulatory UPS or, more likely, in a genomic region not yet studied. Clark and Akam (2016) have proposed novel trans-regulatory interactions among pair-rule proteins involved in refining and positioning *ftz*, as well as *odd*, stripe 4; perhaps the yet-to-be identified *ftz* stripe 4 CRE contains binding sites for these pair-rule proteins (Clark and Akam, 2016).

Interestingly, transgenes capable of rescuing *ftz* mutants do not accurately re-capitulate the order of *ftz* stripe activation (Yu and Pick, 1995), suggesting that although

they provide enough information to produce the final 7-stripe *ftz* pattern, additional information is missing from these transgenes. Similarly, for the three deletion mutants that we generated, *ftzΔZ*$^{-/-}$, *ftzΔZp*$^{-/-}$, and *ftzΔZ-GAGA*$^{-/-}$, the order in which stripes arose differed, and neither matched wild type (Fig. 2-7). Notably, though order was perturbed in both *ftzΔZ*$^{-/-}$ and *ftzΔZp*$^{-/-}$ lines, the phenotype was more severe in *ftzΔZ*$^{-/-}$ than *ftzΔZp*$^{-/-}$. Together, these results suggest that development is more flexible than might be expected – the wildtype order of stripe formation is not necessary for survival, since both *ftzΔZ*$^{-/-}$ mutants and *ftz* mutant flies rescued with the transgenes are still viable, despite neither recapitulating the wildtype order of stripe establishment.

Analysis of *ftz* null mutants, which behave as standard recessives, suggests that a 50% reduction of expression in heterozygous flies is sufficient for normal development. The finding that zebra element deletions have region-specific impacts on target gene expression levels (Fig. 2-6) allowed us to dissect the quantitative requirement for gene expression more finely. The loss of the zebra element reduces expression levels in *ftz* stripe 4 to a point very near the minimum needed to produce segment A2/A3. This teetering around the threshold is evidenced by variability in phenotype: most embryos fail to produce enough *ftz* in stripe 4 to make segment A2/A3, while others produce just enough to make all or part of A2/A3, and a few develop fully (Table 2-3 and 2-4). This would suggest that expression at ~25% of wild-type is the minimum level required for *ftz* to direct the development of segments.

The threshold for other stripes may not be the same, and indeed we did observe some variability in segments posterior to A3, suggesting that the level of expression in stripes 5-7 may also be coming close to a threshold. We hypothesize that the variation of

phenotypic variability from these mutants is a result of variation of *ftz* expression in stripes 5-7. It is possible that embryos with reduced *ftz* expression in one or more of these posterior stripes fails to specify segmental fate – much like what we believe is consistently occurring in the segment derived from *ftz* stripe 4. However, it is worth noting that the work shown here cannot rule out the possibility that subtle changes in the timing of *ftz* stripe establishment in *ftzΔZ* mutants may contribute to the phenotypes observed, though no statistical difference was detected for stripe positioning from this data set. Future studies, assessing the quantitative contributions of other *ftz* CREs, will allow us to more accurately refine the features of *ftz* gene expression that are critical for promoting morphological segmentation. This may reveal general rules about how the analog information of transcription factor gene expression connects to the digital information of developmental fate.

A great deal of redundancy is thought to be built into the regulation of expression for developmentally regulated genes (Hong et al., 2008). For *ftz*, in addition to the two seven-stripe elements, there are multiple stripe-specific elements that direct expression of each of the *ftz* stripes, with two stripe-specific elements identified to date directing expression in stripe 7, although the exact spatial and temporal dynamics of each CRE may differ. Future work is required to compare the necessity, sufficiency, and potential shared roles and synergistic roles of these CREs, as well as to understand how these mechanisms changed during evolution as *ftz* was incorporated into the pair-rule gene network, having arisen as a *Hox* gene with a very different ancestral expression pattern (Heffer et al., 2013). To this end, the *ftzΔZ* mutants, and perhaps future *ftz* enhancer deletion mutants, could serve as genetic systems with 'intra-embryo' internally controlled

environments for future studies on the role of *ftz* in regulating other target genes. Clearly, analysis of CREs using transgenes has provided much insight into this process. However, our studies, in keeping with those of others (Delker et al., 2019), highlight the importance of examining CRE function not only through transgenes, but also by functional analysis in the native context of the endogenous gene.

Chapter 3: Re-examining the role of the classic CREs of *ftz*

3.1 Abstract

The precise orchestration of gene regulatory networks during embryogenesis is essential for proper establishment of the body plan. At the level of transcription, where much regulation resides, this occurs through the binding of *trans* factors to *cis*-regulatory elements (CREs) in the genome. In the case of the *Drosophila melanogaster* gene *fushi turazu* (*ftz*), failure to regulate proper expression domains results in alternating segmental aberrations. Decades of research using transgenes has facilitated the study of the *ftz* CREs. However, many of the foundational conclusions were drawn from experiments that suffered from technical limitations. Here, I have re-examined the classical upstream element (UPS) using modern methodologies to re-evaluate both the sufficiency of this CRE to direct *ftz* expression patterns and its necessity for *ftz* function. Through examining reporter transgene expression, I have refined the boundary of this element and identified a possible insulator near a putative *Scr* CRE that is located between the *ftz* CREs. In analyzing a precise CRISPR/Cas9-induced deletion mutant of the UPS, I have determined that the UPS is not necessary for *ftz* function, but it is necessary for *ftz* expression after gastrulation. Finally, in analyzing a CRISPR/Cas9 mutant in which both the UPS and zebra element have been precisely deleted, I have determined that the stripe-specific CREs of *ftz* alone are insufficient for *ftz* function. This analysis has surprisingly revealed a probable role of the UPS in directing *ftz* stripe 4 expression in a Ftz-independent manner. These results highlight the importance of re-evaluating fundamental findings with modern methodologies.

3.2 Introduction

The dynamic process of embryogenesis requires the organization of a gene regulatory network (reviewed in Small & Arnosti, 2020). A vital aspect to this is the binding of transcription factors to specific sequences in *cis*-regulatory elements (CREs) of the genomic DNA. In turn, this determines which genes are transcribed or repressed within specific expression domains, conferring spatiotemporal control of gene expression. In the case of *Drosophila melanogaster*, the common fruit fly, this orchestration has been well studied. In particular, the development of segments is regulated by a hierarchy of genes (Nüsslein-Volhard & Wieschaus, 1980; Wieschaus & Nusslein-Volhard, 2016). From these studies, it was discovered that the pair-rule gene mutants have periodic segmental abnormalities. Additional analyses have concluded that the pair-rule genes function to determine the segmental primordia of the early embryo (Lawrence & Johnston, 1989). The pair-rule genes are expressed in alternating stripes throughout the cellular blastoderm stage of development (reviewed in (Lawrence, 1992)). To study the regulation of these initially unexpected expression patterns, reporter transgenes, which constitute a CRE-containing genomic fragment fused upstream of a basal promoter and reporter gene, have been used extensively. Many of these analyses relied on P-element based vectors, by which transgenes can be randomly inserted into the genome of the fly *in vivo* (Rubin & Spradling, 1982, 1983; Spradling & Rubin, 1982).

One such pair-rule gene, *fushi tarazu* (*ftz*), is expressed in the even parasegments of the embryo early in development (Carroll & Scott, 1985; Hafen et al., 1984; Lawrence et al., 1987). The loss of *ftz* results in animals that are missing the segments derived from

the segmental primordia in which *ftz* is normally expressed (Magrassi & Lawrence, 1988). Evidence for the necessity of transcriptional control of this locus was established when Hiromi created a P-element based rescue construct with a fragment digested from the genome with the restriction enzyme KpnI (Hiromi et al., 1985). This approximately 10kb genomic fragment contained the coding region of *ftz* in addition to the surrounding genomic DNA. When this was transformed into *ftz* mutants, partial rescue of the animal was reported. This suggested that within the 10kb fragment, the genomic regions surrounding the *ftz* transcription unit contained CREs that conferred the proper spatiotemporal control necessary to express *ftz* in the even parasegments of the early embryo. From this genomic fragment, subsequent enzyme digests produced additional rescue constructs to identify which regions could be removed from this KpnI genomic fragment and still result in rescue. In doing so, it was determined that there were three regions upstream of the *ftz* coding sequence that were necessary for *ftz* function. Notably, one of these regions was between the far KpnI and XbaI sites of the fragment, located approximately 6.5 to 4 kb upstream of the transcription start site (TSS). This CRE was named the upstream element (UPS). A fragment approximately 900 bp between the UPS boundary and the next XbaI site could be removed without affecting *ftz* rescue. This XbaI-XbaI fragment was thus determined to be dispensable for *ftz* rescue and function.

Through the subsequent generation and expression analyses of reporter transgenes by which these necessary fragments were fused to a basal promoter and *lacZ*, it was determined that the UPS and the zebra element reporter transgenes were detected in seven *ftz*-like stripes during the blastoderm and germband extension stages of development (Hiromi & Gehring, 1987; Hiromi et al., 1985). However, these transgenes

did not perfectly recapitulate the endogenous *ftz* expression pattern. These experiments also determined that adjacent to the zebra element was a neurogenic CRE that directed expression of *ftz* during a later stage of embryogenesis in cells of the developing central nervous system (CNS) (Doe et al., 1988; Hiromi & Gehring, 1987; Hiromi et al., 1985). In crossing these reporter transgenic lines to mutants harboring the amorphic ftz^{9H34} allele, it was found that expression of the UPS transgenes was drastically reduced. This was the case regardless of orientation of the UPS fragment relative to the promoter. Thus, the conclusion was drawn that the UPS is a Ftz-dependent enhancer.

Future analyses involved examining the expression of the transgene when the UPS had been digested to random lengths by 5' and 3' exonucleases (Pick et al., 1990). This dissection determined that the UPS contained a proximal enhancer on the 3' end and a distal enhancer near the middle of the UPS. While expression of the proximal enhancer reporter was detected in both mesodermal and ectodermal germ layers, expression of the distal enhancer reporter was detected only in the mesoderm germ layer. Regardless, these transgenes were otherwise detected in the seven *ftz*-like domains. Like the UPS, both of these enhancers were concluded to be Ftz-dependent. Later biochemical analyses provided evidence for the direct binding of the proteins Ftz, Ftz-F1, Adf-1, and Ttk to the UPS proximal enhancer (Han et al., 1993; Han et al., 1998).

Though these experiments were carefully designed, it is important to highlight that the detection methods from these studies were much less robust than the methodologies commonly employed in modern studies to detect gene expression. First, detection of RNA *in vivo* was restricted to the use of radioactive-labeled probes exposed to film to produce images from which quantitative differences in gene expression levels,

and in some cases qualitative patterning, were unable to be ascertained. Consequently, many of the reporter transgene studies described above used β-galactosidase assays instead of *in situ* hybridization. Therefore, they reported the activities of a protein rather than the abundance and/or pattern of a transcript. Second, the selection of genomic fragments to analyze was limited by the presence of restriction enzyme sites or the reaction rate of an exonuclease, which are not inherently biologically meaningful for determining CRE function. Finally, these experiments relied on randomly inserted P-elements that were subject to position effects.

Since the time that these conclusions were made, more robust means have been developed and have become common to detect transcripts *in vivo*. These include the use of digoxygenin-labeled probes and alkaline phosphatase-conjugated antibodies that catalyze a chromogenic reaction (Kosman et al., 2004; Tautz & Pfeifle, 1989). The signal developed from this reaction can be left to amplify or quenched to provide a modular degree of sensitivity. Additionally, fluorescent *in situ* hybridization chain reaction (HCR) combined with confocal microscopy provides images with fluorescent intensities that accurately and precisely reflect relative gene expression levels (Choi et al., 2014; Choi et al., 2010). In analyzing these images, quantitative differences in gene expression between samples or regions of interest within an embryo can be determined (Graham et al., 2021; Surkova et al., 2019).

Third, the analysis of gene expression at the time of the foundational studies described above was also limited to the use of transgenic constructs or alleles produced by random mutagenesis. While transgenic constructs are still used to conduct robust experiments today, the advent of CRISPR/Cas9 genome editing allows for the direct and

precise manipulation of genomic information (Gratz et al., 2013). Thus, the limitation to study the sufficiency of a transgene out of the native context of the gene can be circumvented to study the necessity of a genomic fragment on endogenous expression levels.

Here, I employ newer and more controlled methodologies to re-examine the past conclusions that were drawn about these "classic" *ftz* CREs to determine if and to what extent they can be validated. Through systematic comparisons of reporter transgene expression, I have re-evaluated the sufficiency of genomic fragments, including the UPS and "dispensable" region, to direct expression in a *ftz*-like pattern. I then re-examined the role of the UPS on endogenous expression patterns through the use of the CRISPR/Cas9 system to induce a precise deletion of the CRE. Finally, I examined the impacts on endogenous *ftz* expression in the absence of both the UPS and zebra element by precisely deleting both CREs in the same genetic background. In doing so, I have refined the boundaries of the UPS and found that while this CRE is not necessary for the qualitative expression of *ftz* during the blastoderm stage, the absence of both the UPS and zebra element is lethal and results in significant reductions of *ftz* expression in multiple stripe domains.

3.3 Materials and Methods

3.3.1 Transgenic plasmids

As in Chapter 2, genomic fragments for reporter transgenes were amplified with NEB Q5 DNA polymerase (catalog number: M0491S) following product specifications. Transformed colonies were screened with NEB Taq DNA polymerase (catalog number: M0273L). Annealing temperatures were predicted with NEB Tm Calculator

(https://tmcalculator.neb.com/#!/main) unless specified otherwise. Appendix A contains the sequences for all primers used in this work. All fragments were independently inserted in forward or reverse orientation into the XbaI site of vector *placZ-attB* using either Pyrite cloning (Fischer et al., 2018) or NEB HiFi Assembly (catalog number E2621S), to generate two reporter transgenes per fragment (forward and reverse). The consensus sequence for each fragment and all *in situ* probe sequences are given in Appendix B. Genomic DNA isolated from *nos-Cas9* flies (see Chapter 2 Materials and Methods) was used as the DNA template for all PCR reactions unless specified otherwise. The *Drosophila* genome version dm6 was used as the reference genome (Dos Santos et al., 2015). Sequencing was conducted using the service provided by Genewiz. Sequence fidelity of each construct was verified by comparing full sequence reads of each clone to an independently generated PCR product.

UPS^{7S} was amplified with the primers MF114 and MF115 using 0.625% DMSO. It contains sequence between -5434 to -3845 relative to the *ftz* TSS and is 1609 bp. The fragment was inserted into *placZ-attB* using Pyrite cloning with XbaI. Positive bacterial colonies were screened with colony PCR using the primers SeqF and SeqR, which were also used for fully sequencing the final transgene. MIDI prep DNA, used for *Drosophila* microinjections, was confirmed for orientation by sequencing with SeqF.

UPS^{0} lies between -6418 to -3845 and is 2565 bp. *UPS^{0}-R:lacZ* was generated by assembling three fragments together into *placZ-attB* linearized by XbaI with HiFi assembly. First, the *UPS^{7S}-R* construct was digested with EcoRI-HF and SacI to isolate a 595 bp fragment. Next, *UPS^{7S}-R* was used as template for PCR with the primers Dm_UES_3R and Dm_UES_3F to amplify a 767 bp fragment using 0.625% DMSO.

Finally, genomic DNA was used a template with the primers Dm_UES_2R and MF135 to amplify a 1317 bp fragment (without DMSO). For this reaction, a combined 67°C step was used for both the annealing and extension steps of each PCR cycle due to a AT-rich region within the amplicon. Colony PCR was conducted with the primers SeqF and NeuroSeq1. The isolated clone was fully sequenced with the primers SeqF, DmUES3R, MF24, and SeqR. The full length UPS^0 fragment was excised from this by XbaI/Kpn-HF digestion and fused back into the vector in the forward orientation with Pyrite cloning to generate *UPS^0-F:lacZ*. Colony PCR was conducted with the primers SeqR and NeuroSeq1. A diagnostic digest was conducted with KpnI-HF and SphI-HF to confirm the orientation of the MIDI preps used for *Drosophila* microinjections.

7SN-Z is between -5434 to -671 and is 4764 bp. *7SN-Z-F:lacZ* was generated by amplifying with the primers MF128 and MF132 with 0.625% DMSO. *7SN-Z-R:lacZ* was generated with the primers MF130 + MF132 with 0.625% DMSO. Both fragments were inserted into XbaI linearized *placZ-attB* with HiFi assembly. Colony PCR for both orientations was conducted with the primers DmFtz_UESSeq_Inner2F and DmFtz_UESSeq2R. Both constructs were fully sequenced with the primers SeqF, SeqR, MF22, DmNeuro1F, DmFtzNeurogenic201F, and DmFtzNeurogenic202F. MIDI prep DNA was verified for orientation by digesting with EcoRI.

7SN is between -5424 to -41 and is 5392 bp. *7SN-F:lacZ* was generated by assembling four fragments into *placZ-attB* linearized by XbaI with HiFi assembly. The first two fragments were formed by digesting *7SN-Z-F:lacZ* with BsaA1 and AgeI-HF. The first fragment was 2055 bp and the second was 1383 bp. The third fragment overlapped these two and was generated by amplifying a 401 bp fragment from the

genome with the primers SeqInner2F and DmFtz_UESSeq2R. The fourth fragment was 2181 bp and amplified from the genome with the primers DmFtzNeurogenic2.02F and MF131. Due to large overlaps between the fragments, the conditions of HiFi assembly were changed to 37°C for 10 minutes and 50°C for 50 minutes. Colony PCR was conducted with the primers DmFtz_UESSeq_Inner2F and DmFtz_UESSeq2R. The regions corresponding to PCR amplified fragments were sequenced with the primers DmUESSeq2F, DmZebraFullL2, DmFtzNeurogenic201R, ColonyPCRNandZ, and SeqR. *7SN-R:lacZ* was generated by assembling two fragments together. The first fragment was generated by digesting a part of the insert out of *7SN-Z-R:lacZ* with KpnI-HF and EcoNI and isolating the 14391 bp fragment. The second fragment was generated by PCR using the primers Dm_ftz_Neurogenic2.01F and MF131 to amplify a 3102 bp fragment. Due to the large overlap on the EcoNI side of the digested vector, HiFi assembly conditions were changed to 20 minutes at 37°C and 30 minutes at 50°C. Colony PCR was conducted with the primers DmFtz_UESSeq_Inner2F and DmFtz_UESSeq2R. Regions corresponding to those amplified by PCR were fully sequenced with the primers SeqF, DmZebraFullR, DmZebraFullL2, and DmFtzNeurogenic2.01F. This revealed a 34 bp deletion generated by the assembly, and the corresponding sequence is highlighted in the FASTA of *7SN-F:lacZ*, *7SN-Z-F:lacZ*, and *7SN-Z-R:lacZ* in Appendix B. The orientation of the MIDI prep DNA used for *Drosophila* microinjections was verified by digesting with PstI.

Scr^{ftz}-XbaI^{2} (*Scr*^{ftz}-X^{2}) is between -3850 to -2914 and is 932 bp. *Scr*^{ftz}*-XbaI*^{2}*-F:lacZ* and *Scr*^{ftz}*-XbaI*^{2}*-R:lacZ* were generated by digesting the fragment out of *7SN-Z-R:lacZ* with XbaI and inserting it into *placZ-attB* with Pyrite cloning using XbaI. Colony PCR and sequencing were conducted with the primers SeqF and SeqR. The orientation of

the MIDI prep DNA used for *Drosophila* microinjections was confirmed with diagnostic digestion by PstI.

Scr^{ftz}-XbaI/HindIII (*Scr^{ftz}-X/H*) is between -3850 to -3028 and is 834 bp. *Scr^{ftz}-XbaI/HindIII-R:lacZ* was generated by PCR using *7SN-Z-R:lacZ* as a template with the primers DZ1 and DZ3. The product was inserted into *placZ-attB* by Pyrite cloning with XbaI. Colony PCR was conducted with the primers SeqF and DmFtzNeurogenic2.01F, and the insert was fully sequenced with the primers SeqF and SeqR. To generate *Scr^{ftz}-XbaI/HindIII-F:lacZ*, the fragment was isolated by digestion with XbaI and re-assembled in the forward orientation using the oligos MF136 and MF137 with HiFi assembly. Several colonies were randomly selected and digested with PstI to confirm the forward orientation of the fragment. The orientation of the MIDI prep DNA used for *Drosophila* microinjections was confirmed with diagnostic digestion by XhoI.

As in Chapter 2, transgenes were integrated into the phiC31 docking site at Chromosome 2, 57F5, 2R:21645971 following embryonic microinjection into Bloomington Stock Center line 9740 by Rainbow Transgenic Fly.

3.3.2 Guide RNA and HDR template

As in Chapter 2, genomic targets for gRNA were selected using CHOPCHOP (Labun et al., 2019) and the CRISPR Efficiency Predictor (Housden et al., 2015). The work done in this chapter also utilized CRISPR Direct (Naito et al., 2015). Two gRNAs were selected to target the UPS. The left guide targeted -5428 to -5409 to induce a double-stranded break 23 bp from the 5' end of UPS^{7S}. The right guide targeted -3964 to -3945 to induce a double-stranded break 103 bp from the 3' end. The co-CRIPSR strategy described in Chapter 2 to target *ebony* was utilized here as well (Kane et al.,

2017). Guide RNAs were inserted into the *pCFD5* vector that processes all guides in a single transcript separated by an array of tRNA sequences (Port & Bullock, 2016). The vector was used as a template for PCR with primers MF120 and MF108 in one reaction and MF109 and MF110 in another. The 109+110 fragment was further amplified by MF109 and MF112. These two fragments were assembled into the *pCFD5* vector linearized with BbsI-HF using HiFi assembly. Colony PCR and sequencing of *pCFD5-UPS* was conducted with the primers CFDN-Seq and MF67.

The homology directed repair (HDR) template was assembled from two PCR-amplified fragments. The first reaction used primers MF40 and MF123 to amplify a 2399 bp fragment from -7791 to -5435. The second reaction used primers MF20 and MF124 to amplify a 3061 bp fragment from -3844 to -826. Both fragments were assembled into the vector *pUC19* linearized by KpnI-HF with HiFi assembly. Colony PCR was conducted with the primers MF25 and MF113. The assembled template, *pUC19-ΔUPS*, was sequenced with the primers M13F, DmFtzUESSeqInner1F, MF113, MF25, DmFtzNeurogenic2.01F, DmFtzNeurogenic2.02F, and M13R. Embryos were injected with both the sgRNA array construct and HDR template to generate *ftzΔUPS* by Rainbow Transgenic Flies. To generate *ftzΔUPSΔZ* flies, homozygous *nosCas9;ftzΔUPS* larva were sent to Rainbow Transgenic Flies for injection with *pCFD5-UPS* and the HDR template targeting the zebra element (see Chapter 2).

3.3.3 *Drosophila* genetics

Drosophila lines and methods for identifying CRISPR/Cas9 events were the same as described in Chapter 2 with exception to the primers used for screening. Deletion of the UPS was screened with the primers MF25 and MF113, which would amplify a 112 bp

fragment if a precise deletion occurred. Out of 40 fertile G0 adults, 13 produced offspring with precise deletions (32.5%). Of these lines, three had *ebony* progeny (7.5%), and all of these had siblings with precise deletions. DNA extracted from confirmed deletion lines was PCR amplified with primers MF25 and DmFtz_UESSeq2R and sent for sequencing with MF22 to confirm precise deletion by an HDR event. The conditions for detecting deletion of the zebra element in the *ftzΔUPS* background were as described for the original *ftzΔZ* deletion in Chapter 2. *Drosophila* harboring the balancer *TM3, Ser, hb8-lacZ* were used to maintain independent stocks for the amorphic allele ftz^{9H34} and the novel lethal allele $ftz\Delta UPS^{7S}\Delta Z$ (Schier & Gehring, 1993b) (Jurgens et al., 1984). The allele ftz^{9H34} (also known as ftz^{13}) had been previously determined to contain an amber mutation after 53 amino acids (Gln to stop) (Jurgens et al., 1984; Schroeder et al., 2011).

3.3.4 Analysis of adult abdomen segmentation

Approximately 200 flies were used in each experiment and were viewed with a dissection microscope and scored for segmentation defects. Flies were counted and categorized as having all body segments (wild-type-like), having partial fusions/deletions between segments, or missing one or more segments. Adults showing any segmentation defects were set aside and stored at -80°C for several hours. Abdomens were carefully cut away from frozen adults with a sharp razor blade and scalpel, placed in a drop of clear nail polish in a line on a microscope slide placed in petri dish, oriented with the dorsal side up. Once the polish cured, the specimens were covered with water and imaged under a dissection microscope.

3.3.5 Cuticle preparations

Fly embryos were collected in cups with yeast paste for 24 hours. Flies were flipped into fresh cups and the embryos aged for an additional 24 hours. Larva were poured into mesh baskets and rinsed several times to remove any residual yeast. The mesh baskets were treated with 50% bleach for 3 minutes to remove chorions, and the baskets were rinsed with water to remove bleach. The mesh was transferred to a 1.7 ml microfuge tube with 50% heptane and 50% methanol and shaken for one minute. The heptane was removed and replaced with methanol, and the tubes were inverted to ensure mixing. Methanol was decanted and samples were rinsed with methanol three more times. Specimen were transferred to a slide and allowed to dry completely. Lactic acid was pipetted onto the sample and covered. The slides were placed in a 60°C oven for 2-24 hours. Cuticles were imaged with darkfield microscopy using a Zeiss Microscope Axio Imager M1. Assessing the statistical significance of the proportion of aberrant cuticles to normal cuticles in the mutants compared to wild type was assessed with a Chi-squared test.

3.3.6 Gene expression analysis

Chromogenic *in situ* hybridization was conducted, imaged, and processed as described in Chapter 2. In addition to the antisense probes for *ftz*, *lacZ*, *en*, and *slp1* from Chapter 2, an antisense probe for *Scr* was synthesized using the primers MF152 and MF155 with embryonic cDNA as template. This 350 bp fragment was inserted into *pUC19* using Pyrite cloning with EcoRI and XbaI. The plasmid was linearized with EcoRI and used as a template for T7 RNA polymerase *in vitro* transcription to generate a digoxygenin-labeled probe. Embryos homozygous for the lethal alleles ftz^{9H34} and

ftzΔUPSΔZ were identified with a multiplex *in situ* using a blend of antisense probes targeting both *ftz* and *lacZ* to determine which embryos had the third chromosomal balancer *TM3, Ser, hb:lacZ*. Fluorescent *in situ* HCR was conducted and processed as described in Chapter 2 and Appendix D. In addition to the probes for *ftz* and *eve* for fluorescent *in situ* HCR described in Chapter 2, a *lacZ* probe was utilized along with a hairpin amplifier conjugated to Alex Fluor 647 (Appendix B). Both the *lacZ* probe and fluorescent hairpin amplifier were ordered from Molecular Instruments.

Embryos used for HCR *in situ* hybridization were imaged on a Zeiss LSM 900 Airyscan confocal microscope with a 40x immersion oil objective. Six stacking optical sections, averaged 8 times by line, were scanned through the lateral surface to a sufficient depth to detect all seven *ftz* stripes (between 9-12 μm). Images were taken at a resolution of 1024 px^2. Cellularization substage within stage 5 of embryogenesis was determined by the degree of cellular membrane deposition as viewed from the eyepiece with brightfield settings. Fluorescent intensity from each channel was quantified using the Prostack software and analyzed within R (described in detail in Appendix D). Eight cellularized blastoderm embryos were imaged for each genotype in the same confocal session using the same parameters for laser strength and master gain. Statistical significance was tested with the Kruskal-Wallis *H* test with a post hoc test using the criterium Fisher's least significant difference with the 'agricolae' package in R (de Mendiburu, 2021). The Holm-Bonferroni method was used to correct for multiple comparisons.

3.4 Results

3.4.1 *UPS^{0}:lacZ* and *UPS^{7S}:lacZ* transgenes are expressed during the blastoderm and germband stages

As mentioned in the introduction of this chapter, the original boundaries of the UPS were defined by the restriction sites KpnI on the 5' end and XbaI on the 3' end. Since restriction sites appear randomly throughout the genome, this designation was made by random chance rather than through systematic dissection of the CRE. As will be shown in Chapter 4, the ftz-6 stripe-specific CRE overlaps with the 5' boundary of the UPS. The *ftz-6:lacZ* reporter transgenes in either orientation are expressed in stripes 2 and 7. This suggests that the original UPS (UPS^{0}) contains a stripe-specific CRE that directs stripe 2+7 patterning separate from (or embedded within) the seven-stripe UPS (Fig. 3-1A). To test whether the UPS sequence without the stripe-specific region would still direct expression in all seven stripes, I generated a UPS transgene lacking the upstream ftz-6 fragment (UPS^{7S}), thereby moving the 5' boundary of the UPS downstream and compared it to a similar transgene carrying the original UPS (UPS^{0}). Fragments were fused upstream of a basal heat shock promoter (*hsp70*) and *lacZ* reporter in either orientation to generate two reporter transgenes for each version of the UPS. As in Chapter 2, these transgenes were integrated into the phiC31 docking site at Chromosome 2, 57F5, 2R:21645971 to control for position effect (Materials and Methods; Fig. 3-1B-E). Expression of the *UPS:lacZ* transgenes was examined in homozygous embryos using colorimetric *in situ* hybridization with a digoxigenin-labeled *lacZ* probe.

As shown in Fig. 3-1 Bi, the *UPS^0-F* transgene was expressed in seven clear, refined stripes during the blastoderm stage. This signal was maintained through germband extension (Fig. 3-1B ii). In the reverse orientation, the seven stripes were observed, and stripes 2 and 7 were particularly strong (Fig. 3-1C i). The strength of these stripes is likely due to the contribution from the portion of the 413/ftz-6 element that resides on the 5' end of UPS^0 (see Chapter 4 for ftz-6 reporter transgene expression). The stripe 4 signal seems slightly stronger as well. Interestingly, stripes 1 and 5 appeared qualitatively weaker in the reverse than the forward orientation (Fig. 3-1C ii). Stripes 3 and 6 appear about the same as in the forward orientation. In the germband extension stage, stripe 1 is very faint, and stripes 5 and 6 are relatively weaker than stripes 3, 4, and 7.

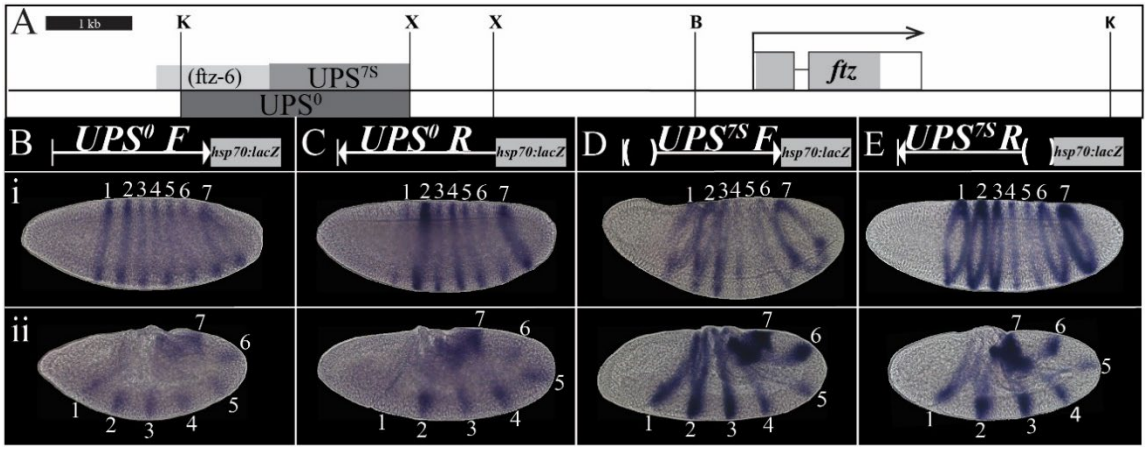


Figure 3-1. Reporter transgenes of UPS fragments in either orientation are detected in all seven stripes. (A) Schematic drawn to scale displaying the location of the UPS relative to the *ftz* transcription unit (*ftz*, black arrow represents transcription, boxes represent exons with CDS in gray, UTRs in white, black line represents the one intron; other grayscale boxes represent genomic fragments tested as transgenes). K: the KpnI sites surrounding *ftz* used in Hiromi et al. (1985); X: XbaI sites; B: BalI site. Scale bar represents distance of one kilobase. (B) Schematic of *UPS^0-F:lacZ* reporter transgene. (C) Schematic of *UPS^0-R:lacZ* reporter transgene. (D) Schematic of *UPS^{7S}-F:lacZ* reporter transgene. (E)

Schematic of *UPS^{7S}-R:lacZ* reporter transgene. Parentheses in (D) and (E) highlight difference of UPS^{7S} from UPS^{0}. (i-ii) Brightfield image of *lacZ* expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization in indicated transgenic embryos during the blastoderm (i) and germband extension (ii) stages. Numbers above embryos refer to the presumptive, corresponding *ftz* stripe. *Ftz*-6 stripe-specific CRE is labeled for reference.

When ftz-6 region was removed from the UPS, all seven stripes were still detected at the blastoderm stage for both forward and reverse orientations of this transgene. These results confirm that this region contains sufficient regulatory information to direct all seven stripes; accordingly, I named this region UPS 7 stripes (UPS^{7S}). Interestingly, stripes 4 and 5 were noticeably weaker in blastoderm stage embryos for both forward and reverse transgenes (Fig. 3-1D-E i). The signal appeared stronger in the germband extension phase, with all seven stripes having clear and strong signals for both orientations (Fig. 3-1D-E ii). Expression appeared stronger for the reverse orientation transgene in the blastoderm stage but equivalent during the germband extension stage. How would the endogenous *ftz* expression change if this region was removed from the genome?

3.4.2 Precise deletion of the UPS generated homozygous viable and fertile adult animals with a variety of segmental defects

Though the experiments above provide information on the sufficiency of the UPS^{7S} fragment to direct expression of transgenes in a *ftz*-dependent, seven-stripe pattern, they cannot address whether or not this CRE is necessary to establish the endogenous *ftz* seven-stripe pattern. To determine its endogenous role, I generated a precise deletion of UPS^{7S} using the CRISPR/Cas9 system (Fig. 3-2; Materials and Methods). CRISPR

sgRNAs were selected that target close to the termini of the UPS^{7S}; the left guide binds 23 bp from the 5' end and the right guide binds 103 bp from the 3' end (Fig. 3-2B, arrowheads). These guides were assembled into the *pCFD5* vector that expresses all guides as a single transcript. Because the guides are surrounded by tRNA sequences, this single transcript is processed into each individual guide, guaranteeing simultaneous expression of both guides. These conditions favor synchronous cutting on both sides of the CRE. The HDR template *pUC19-ΔUPS* was assembled by ligating homology arms surrounding the UPS^{7S} (2358 bp for the left arm, 3021 bp for the right arm; Fig. 3-2B) to the plasmid *pUC19* linearized with KpnI. Both the HDR template and sgRNA constructs were co-injected. Due to the structure of *pUC19-ΔUPS* and placement of the guide target sites, HDR of the CRISPR/Cas9-induced cut would result in precise deletion of the UPS^{7S} (Fig. 3-2C).

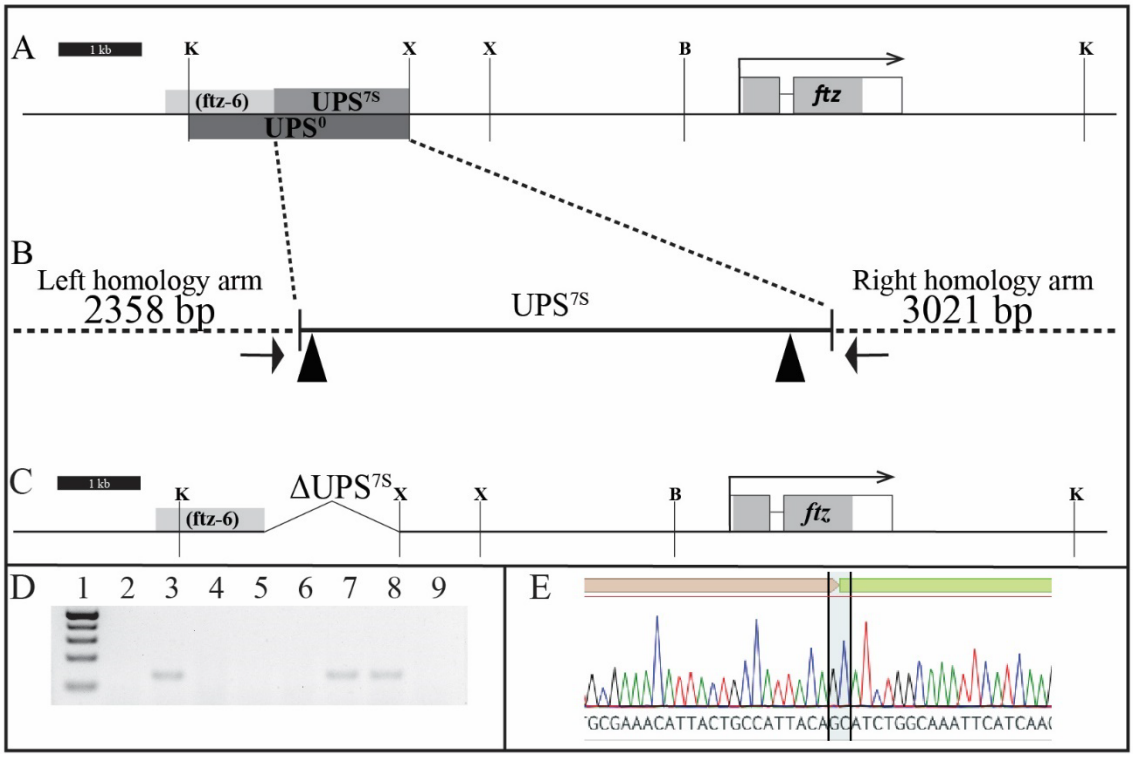


Figure 3-2. Deletion of the UPS^{7S} through HDR induced by CRISPR/Cas9. (A) Schematic of the genomic region surrounding *ftz* with the UPS^{7S}, UPS^{0}, and ftz-6 indicated in grayscale. Vertical lines indicate the following restriction enzyme sites: K, KpnI; X, XbaI; B, BalI. (B) The HDR template consisted of a left homology arm of 2358 bp and a right homology arm of 3021 bp, and sgRNAs were selected to cut close to both termini of the UPS^{7S} (B; black arrowheads represent gRNA target sites; arrows represent primers used for PCR screen). (C) Schematic of successful HDR event. (D) Agarose gel showing results of PCR screen of F1 progeny. A 112 bp product was expected for precise deletion events (D; lane 1: GeneRuler 1kb plus ladder; lanes 3, 7, and 8 show the deletion product; lanes 2, 4-6, and 9 are negative results). (E) Lines that tested positive were sent for sequencing to verify precise deletion of the element.

F1 progeny of injected G0 individuals were PCR screened for the deletion. Primers were MF25 and MF113 that flank the deletion site (Fig. 3-2B, arrows indicate primer locations), amplifying a 112 bp fragment from individuals with successful deletion events (Fig. 3-3D). DNA from individuals that tested positive was used as template in another PCR to be used for sequencing to confirm precise deletion of the element (Fig. 3-2E). Out of 40 fertile G0 adults, 13 produced offspring with precise deletions (32.5%). Though we hypothesized that *ftzΔUPS^{7S}* would be a homozygous lethal allele based on the rescue experiments conducted by Hiromi (1985), I was surprised to discover the contrary: homozygous *ftzΔUPS$^{7S-/-}$* mutants are viable, go through all stages of development, and are fertile.

The majority (91.1%, n = 236) of these *ftzΔUPS$^{7S-/-}$* adults developed with abdomens resembling wild type (Figure 3-3A, B). In the remaining 8.88% of adults (n = 23), a range of segmental defects were observed. Some mutants developed with naked cuticles in regions where pigmentation and bristles should be present (Fig. 3-3C; 6.9% of total adults, 78.2% of aberrant adults, n = 18). For others, partial fusions/deletions

between adjacent segments were observed (Fig. 3-3D – E; 1.9% of total adults, 21.7% of aberrant adults, n = 5). These fusions generally included cuticular separation within a segment, with one half disconnected and the other fused to an adjacent segment (indicated by the thick arrowhead in Figure 3-3). For some adults, these aberrations were observed between segments A4/A5 (Fig. 3-3 D). For others, the aberration was observed between segments A3/A4 (Fig. 3-3 E).

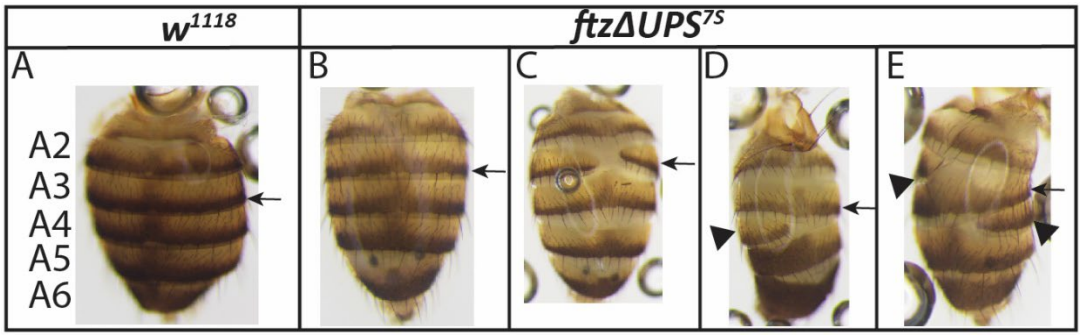


Figure 3-3. Segmentation was disrupted in a minority of *ftz*ΔUPS$^{7S-/-}$ adults. Brightfield images of adult abdomen segmentation phenotype of wild type (A) and *ftzΔUPS*^{7S} adults (B-E). While some *ftzΔUPS*^{7S} adults appear wild-type-like (B), others have partial defects in single segments (C; A3) or fusions/deletions between adjacent segments (D, A4/A5; E, A2/A3). Arrows indicates segment A3; arrowheads indicate segmentation defects.

3.4.3 *ftzΔUPS*$^{7S-/-}$ embryos can develop with disruptions to cuticle formation spanning multiple segments

We next examined larval cuticle preparations to determine if there were defective larvae that did not survive to adulthood. This possibility would result in an underestimate of the severity based on the adult phenotypes alone. Out of 243 wild type larvae, 98.7% developed with the expected denticle bands on the three thoracic segments and the eight abdominal segments (Fig. 3-4A; n = 240). The remaining 1.23% had minor segmental

defects, suggesting that a small fraction of larvae will fail to develop segments properly regardless of genotype (data not shown; n = 3).

Like the adults, the majority of *ftzΔUPS*$^{7S-/-}$ larvae had the expected segmentation (Fig. 3-4B; 96.4%, n=186/193). The remaining 3.6% developed with a range of segmental defects (Fig. 3-4C – F, arrows; n = 5). Notably, this proportion of defective cuticles is not statistically significant compared to wild type (p-value = 0.113) From four of these larvae, one or more segments was missing (Fig. 3-4C-E). One embryo was observed with pair-rule defects resembling weak alleles of other pair-rule genes (Fig. 3-4F, asterisks; Nusslein-Volhard and Wieschaus, 1980), but none showed the *ftz* pair-rule phenotype, lacking alternate segments (Fig. 3-4G). As the only segmental aberrations observed in the adults were partial fusions/deletions, it is assumed that larvae with these more severe segmental abnormalities did not develop to adulthood.

Regardless, I was surprised that a smaller percentage of larvae developed with cuticle defects compared to the percentage of adults with segmental abnormalities, for I hypothesized that the opposite would be observed. Thus, segmental defects in *ftzΔUPS*$^{7S-/-}$ animals appear to be infrequent.

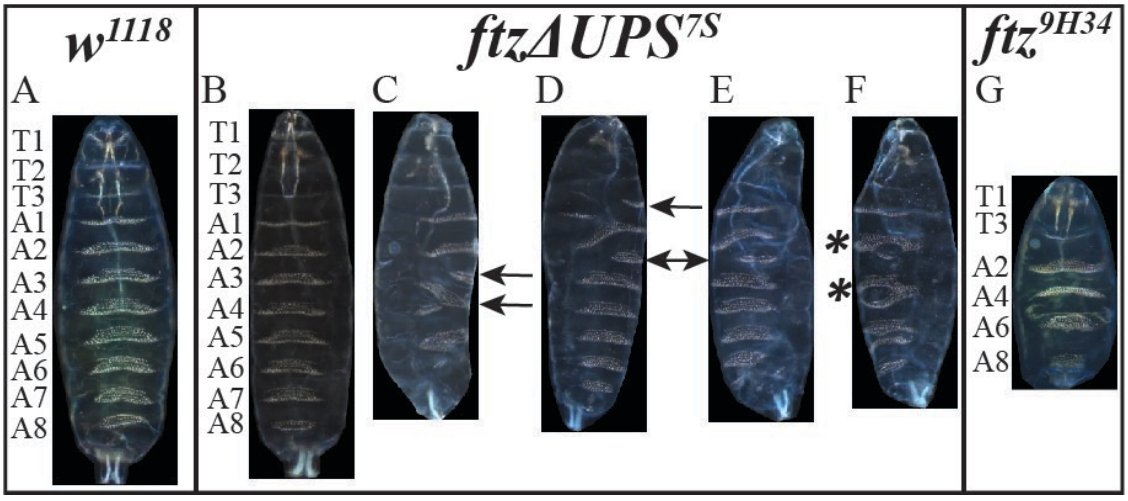


Figure 3-4. Segmentation was perturbed in a minority of *ftzΔUPS*$^{7S-/-}$ larva cuticle preparations. Darkfield microscopy of cuticle preparations. (A) In wild type, there are three thoracic segments (T1-T3) with relatively light denticle bands and eight abdominal segments (A1-A8) with thicker denticles. (B) The majority of *ftzΔUPS*$^{7S-/-}$ larval cuticles appear wild-type-like. (C-E) Some *ftzΔUPS*$^{7S-/-}$ larval cuticles show partial deletions and/or fusions between adjacent segments (arrow). (F) Total fusion between adjacent segments was observed from one *ftzΔUPS*$^{7S-/-}$ larval cuticle (asterisk). (G) Larvae homozygous for the amorphic *ftz*^{9H34} allele are missing alternating sets of segments.

3.4.4 Expression of *ftz* in *ftzΔUPS*$^{7S-/-}$ embryos rapidly deteriorates after gastrulation

To examine if the order by which *ftz* stripe domains were transcribed was different in the *ftzΔUPS*$^{7S-/-}$ background relative to wild type, *in situ* hybridization was conducted on staged embryos. In keeping with the largely wildtype morphology of *ftzΔUPS*$^{7S-/-}$ homozygotes, all seven *ftz* stripes were present in mutants. In addition, the order of stripe domain transcription during the blastoderm phase was nearly identical to that of wild type (Fig. 3-5). Briefly, *ftz* stripes 1, 2, and 5 appear first (Fig. 3-5A), joined by stripe 3 (Fig. 3-5B). This is followed by a combined 6/7 stripe (Fig. 3-5C-D). As stripes 6 and 7 resolve into individual expression domains, stripe 4 appears (Fig. 3-5E–F). The weak signal detected from the stripe 4 domain accumulates in intensity concurrently with all seven stripes (Fig. 3-5G). This change in intensity continues as the blastoderm approaches the end of cellularization (Fig. 3-5H). Finally, all seven stripes reach peak intensity (Fig. 3-5I). During this time, stripe 4 appears weaker than wild type in *ftzΔUPS*$^{7S-/-}$ embryos (Fig. 3-5G-I). The peak intensity holds as the ventral furrow forms and the embryo enters the next stage of development (Fig. 3-5J).

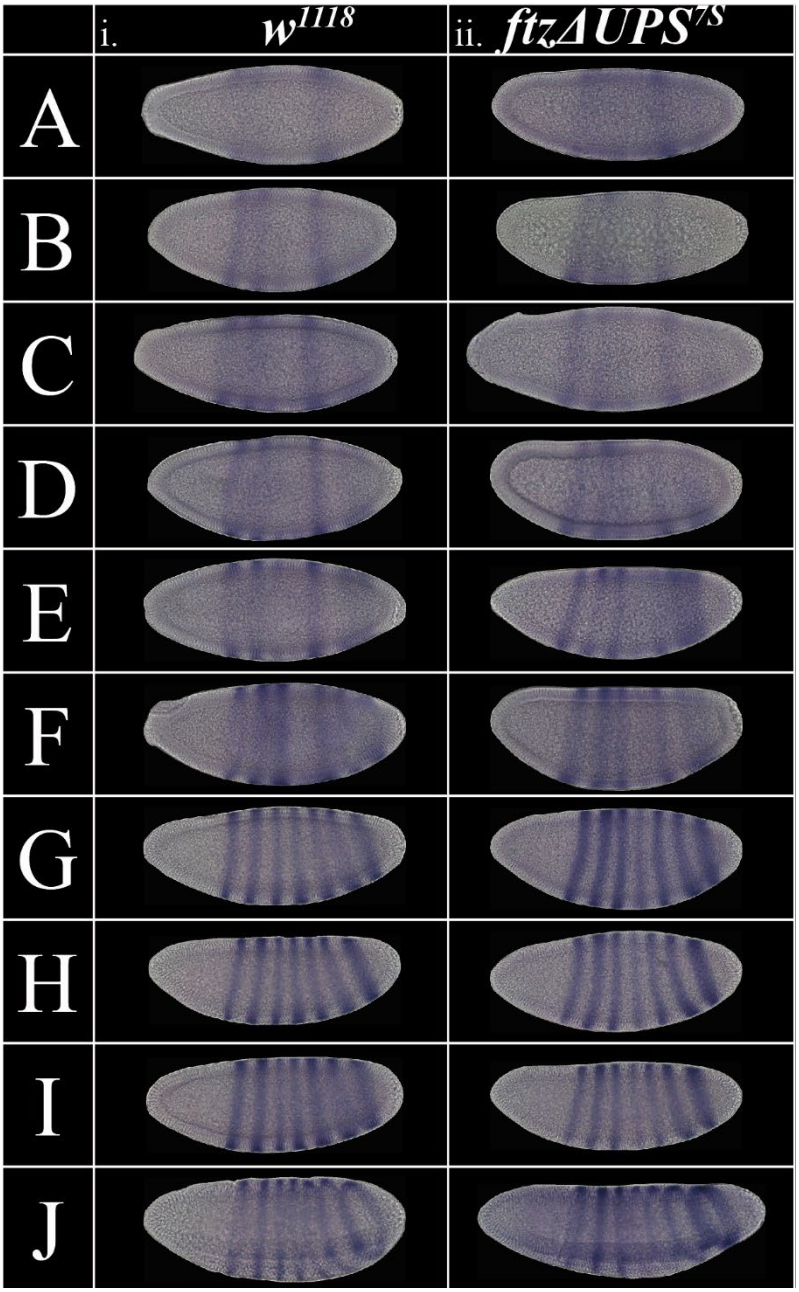


Figure 3-5. *ftz* stripes of *ftzΔUPS*$^{7S-/-}$ mutants develop in the same order as in wild type. Brightfield image of *ftz* expression throughout stage 5 (cellular blastoderm) for wild type (i) and *ftzΔUPS*$^{7S-/-}$ (ii) embryos using digoxygenin-labeled probes for chromogenic *in situ* hybridization. The order in which expression of *ftz* in specific stripe domains is detected is the same in *ftzΔUPS*^{7S} as in wild type. (A) Stripes 1 +2 + 5. (B) Stripes 1 +2 + 5 gain signal as stripes 3 and a combined 6/7 becomes detectable. (C) Stripes 3 and 6/7 gain signal. (D) Stripes 6 and 7 start to resolve from each other. (E) Stripes 1 + 2 + 3 + 5 + 6 + 7

gain intensity. (F) Stripe 4 expression becomes detectable. Stripes 3 + 6 + 7 are clear. (G) All seven stripes are clear. (H) Expression of *ftz* increases in all seven stripes as cellularization continues. (I) Peak *ftz* expression is reached as the blastoderm completes cellularization. Stripe 4 appears relatively weaker in the mutant background (ii). (J) Start of gastrulation.

In wild type, *ftz* expression is maintained through the formation of the germband (stage 6), maintaining seven stripes in the mesoderm and ectoderm layers as the germband extends (Fig. 3-6i). However, in *ftzΔUPS*$^{7S-/-}$ embryos, *ftz* expression decreased rapidly following gastrulation (Fig. 3-6ii). The reduction of signal was observed early during gastrulation, as the fourth and fifth *ftz* stripes diminished quickly (Fig. 3-6ii A). As germband elongation began, signal rapidly decreased in all stripes, but stripes 4 to 6 in particular nearly diminished (Fig. 3-6ii B). At the start of stage 7, only stripes 1, 2, 3, and 7 were detectable (Fig. 3-6ii C-D), while in wild type embryos at the same stage, all seven stripes were still strong (Fig. 3-6i C-D). As the embryo transitioned to stage 8, stripe 1 and 3 were no longer detectable, leaving a weak stripe 2 as well as stripe 7 apparent in mutant embryos (Fig. 3-6ii E). At this point in wild type embryos, stripe 6 becomes fainter, but the seven-stripe pattern is still visible.

To determine if *ftz* stripe domain transcription throughout gastrulation would similarly reduce in embryos lacking Ftz, *ftz* expression was analyzed in embryos homozygous for the amorphic translational null allele *ftz*^{9H34} in the same manner (Fig. 3-6iii). Lack of *lacZ* expression was simultaneously monitored by multiplexing with an antisense *lacZ* probe to unambiguously identify *ftz*^{9H34} homozygotes apart from embryos harboring the balancer *TM3, ser, hb8-lacZ*, which express *lacZ* in the anterior of the embryo. Homozygous *ftz*^{9H34} embryos express *ftz* in all seven stripe domains by the end of stage 5, suggesting that autoregulation is unnecessary for stripe establishment (Fig. 3-

6iii A). Interestingly, the *ftz* stripes appear wider with more intense signal in this mutant than is observed in wild type (see also below). Similar to what was observed for *ftzΔUPS*$^{7S-/-}$ embryos, *ftz* expression reduced greatly after the start of gastrulation (Fig. 3-6iii B). Though all seven stripes appear weaker than in the previous stage, stripes 3 and 6 were markedly weaker. Signal from all seven stripes continued to reduce, though stripes 1 and 6 seem to reduce more (Fig. 3-6iii C). By stage 8, there is no detectable signal of the *ftz* transcript (Fig. 3-6iii D-E). These results are consistent with a role for the UPS in *ftz* autoregulation (see Discussion).

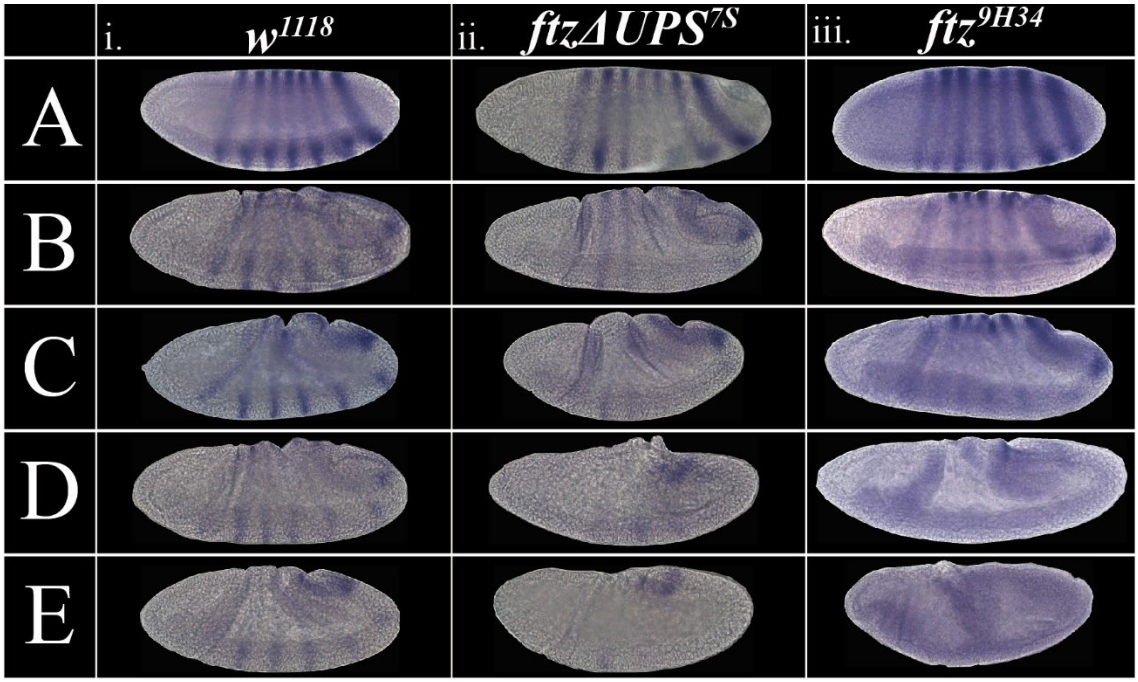


Figure 3-6. Expression of *ftz* after gastrulation was greatly reduced in *ftzΔUPS*$^{7S-/-}$ mutants. Brightfield images of *ftz* expression time series throughout stages 6-8 (gastrulation/germband extension) of wild type (i), *ftzΔUPS*^{7S} (ii), and *ftz*^{9H34} embryos using digoxygenin-labeled probes for chromogenic *in situ* hybridization. (A-D i) Though *ftz* expression reduces in wild type after gastrulation, it can be observed even after germband extension with the exception of stripe 6 (E i). (B ii) Expression of *ftz* in *ftzΔUPS*^{7S} embryos reduced immediately after gastrulation and was nearly undetectable before stage 8 (E ii). (C ii) Signal from

stripes 4-6 depleted first, followed by stripe 1 (D ii). Signal from stripe 3 then faded, leaving stripes 2 and 7 I. Stripe 7 was detectable but very faint at the end of gastrulation. (A iii) Expression of *ftz* in ftz^{9H34} embryos is strong in all seven stripes by the end of cellulairzation. (B iii) Signal from all stripes is lower than in the previous stage, though 3 and 6 are more so reduced. (C iii) Expression continues to reduce for all stripes, and stripes 1 and 6 signals are particularly faint. (D-E iii) *ftz* expression is undetectable by stage 8.

3.4.5 Downstream targets of *ftz* are perturbed in some $ftz\Delta UPS^{7S-/-}$ embryos

Is there a checkpoint at this developmental transition from stage 5 to 6 at which Ftz must be present at sufficient quantities to regulate downstream target genes in the segmentation pathway? Since *ftz* expression resembles wild type at early developmental stages but fades rapidly during germband extension, we analyzed expression of Ftz target genes in wild type and $ftz\Delta UPS^{7S-/-}$ mutants. Briefly, the even-numbered *engrailed* (*en*) stripes are activated by Ftz while *sloppy paired 1* (*slp1*) inter-stripe domains are repressed by Ftz (Graham et al., 2021; Hang & Gergen, 2017; Howard & Ingham, 1986; Prazak et al., 2010). The following *slp1* interstripe regions are regulated by the following *ftz* stripes: *slp1* expression in the region between *slp1*stripes 1 and 2 is repressed by Ftz stripe 1; between *slp1* stripes *3* and *4* repressed by Ftz stripe 2; and so on. These activating and repressing patterns allow us to assess the timing of each of these regulatory roles. In wild type embryos, only *slp1* stripe 1 and the even numbered stripe domains are expressed at the end of the blastoderm stage in a pair-rule like pattern of seven stripes (Fig. 3-7A). Just before gastrulation begins, even-striped expression of *slp1* begins faintly (Fig. 3-7B). These even stripes continue to gain intensity as the ventral furrow begins to form (Fig. 3-7C-D), and fully formed, faint stripes are visible once the ventral furrow begins invaginating (Fig. 3-7E). As this point, the odd numbered stripes are prominent

with a strong signal. As the germband begins to form, all 14 *slp1* stripes are visible (Fig. 3-7F).

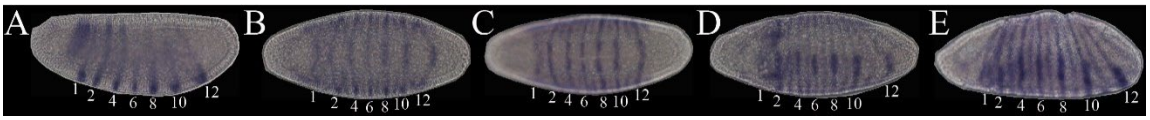


Figure 3-7. The 14-stripe patterning of *slp1* is established during gastrulation. Brightfield images of *slp1* expression pattern between stages 5-6 (gastrulation/germband extension) of wild type embryos using digoxygenin-labeled probes for chromogenic *in situ* hybridization. Stripe 1 and the even stripes 2-12 are labeled respective to the final 14 stripe pattern. (A) *slp1* is expressed in seven clear stripes at the end of stage 5. Cap expression is observed in the anterior. (B) Odd-numbered stripes begin to form in between the original even-numbered stirpes. (C) The original even-numbered stripes gain intensity while the odd-numbered stripes refine to two cell-wide domains as the ventral furrow forms. (D) The odd-numbered stripes become apparent, and the even-numbered stripes become more prominent. I A clear pattern of 13 *slp1* stripes coincides with gastrulation during the formation of the germband.

During the germband stage (stages 6-8), *slp1* and *en* are expressed in 14 stripes (Fig. 3-8A-D). While *slp1* was expressed in a wild-type fashion in 83% of *ftzΔUPS*$^{7S-/-}$ embryos (Fig. 3-8i E; n = 71), some stripes were slightly wider in 17% of embryos (asterisks; Fig. 3-8i F; n = 15), consistent with loss of repression by Ftz. This was most apparent for *slp1* stripes 7 and 8, and the interstripe region here is repressed by Ftz stripe 4 (Fig. 3-8i F). Interestingly, this coincides with the *ftz* stripe that dramatically loses signal at the start of gastrulation (Fig. 3-6ii A), suggesting that the decrease of *ftz* stripe 4 expression at the start of gastrulation of *ftzΔUPS*$^{7S-/-}$ embryos results in insufficient levels of Ftz protein to repress downstream genes such as *slp1* in 17% of embryos. The opposite trend was observed for *en*, which is activated by Ftz in the even-numbered parasegments. In wild type, *en* signal reaches strong expression in 14 stripes during germband extension

(Fig. 3-8ii C-D). This gene expression pattern is observed in 96% of *ftzΔUPS*^{7S} embryos as well (Fig. 3-8ii G, n=152). However, in about 3% of *ftzΔUPS*^{7S} embryos, reduced signal was observed in some even-numbered *en* stripes (Fig. 3-8ii H parenthetical numbers, n=4). In two embryos, no detectable *en* signal was observed for *en* stripe 2 (Fig. 3-8ii I), while the other stripes seem unaffected. In one case, several stripes had low expression and others were missing (Fig. 3-8ii J, number omitted).

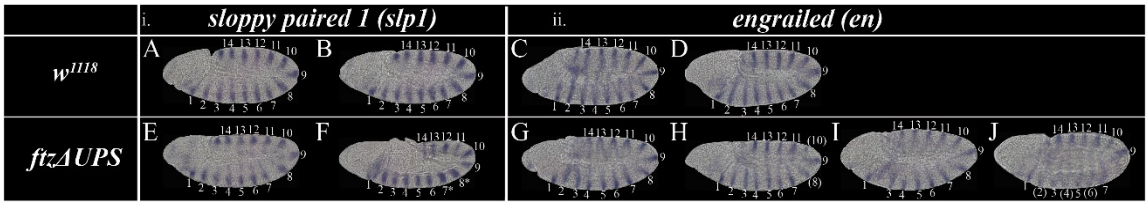


Figure 3-8. Expression patterns of genes downstream of *ftz* are perturbed only in a small number of *ftzΔUPS*$^{7S-/-}$ embryos. Brightfield images of *sloppy paired 1* (*slp1*; i) and *engrailed* (*en*; ii) expression in stages 7-8 (germband extension) of wild type (A-D) and *ftzΔUPS*$^{7S-/-}$ (E-J) embryos. While *slp1* expression is repressed by Ftz, *en* expression is activated by the protein. Both *slp1* and *en* are expressed in 14 stripes in wild type embryos (A-D) and 96% of *ftzΔUPS*^{7S} embryos (E and G). In some *ftzΔUPS*^{7S} embryos, *slp1* stripes were slightly expanded beyond the expected two-cell width (F; asterisks). Conversely, some embryos show weak signal in certain *en* stripes (H; parenthetical numbers) or are missing them entirely (I). From one embryo, a combination of stripes missing and others with weak expression was observed (J).

In summary, expression of downstream genes of *ftz* was wild-type-like in most *ftzΔUPS*^{7S} embryos, in keeping with the wildtype nature of most adults and larvae. However, in some *ftzΔUPS*^{7S} embryos, expression was perturbed within specific expression domains. The parasegments in which these perturbations were observed varied, much like the abdominal defects in adults (Fig. 3-3) and larval cuticle defects (Fig. 3-4).

3.4.6 *ftzΔUPS*^{7S}*ΔZ* double mutants are homozygous lethal

Contrary to hypotheses based on rescue experiments (Hiromi et al., 1985), homozygous *ftzΔUPS*^{7S} are both viable and fertile, and many adults appear wild-type-like. Similarly, animals homozygous for a deletion of the zebra element were homozygous viable and fertile (Chapter II and Graham et al., 2021). These two CREs, the UPS and zebra element, constitute the two known seven-stripe elements of *ftz*. The other known CREs direct expression in a stripe-specific manner (see Chapter 4).

While the individual deletion mutants of the seven-stripe elements were viable, we sought to determine if the stripe-specific CREs alone, in the absence of both seven-stripe CREs, would be sufficient for the transcriptional establishment of *ftz*. Would these embryos receive the necessary expression contributions for each *ftz* stripe to develop to adulthood? To answer these questions, a double deletion mutant of the UPS and zebra element was generated by precisely deleting the zebra element in the *ftzΔUPS*^{7S} background (Fig. 3-10).

Briefly, the *ftzΔUPS*^{7S} allele was crossed into a *nos-Cas9* background, and these homozygotes were used for CRISPR/Cas9 mutagenesis with the same guide RNA and HDR constructs that were used to generate the *ftzΔZ* allele (Materials and Methods). F1 progeny of injected G0 individuals were PCR screened for the deletion. Primers were zebra1 and zebra2 that flank the deletion site (Fig. 3-9B, arrows indicate primer locations), amplifying an 888 bp fragment from individuals with successful deletion events and a 1485 bp fragment in the presence of the zebra element (Fig. 3-9C). DNA from individuals that tested positive was used as template in another PCR to be used for sequencing to confirm precise deletion of the element (Fig. 3-9D).

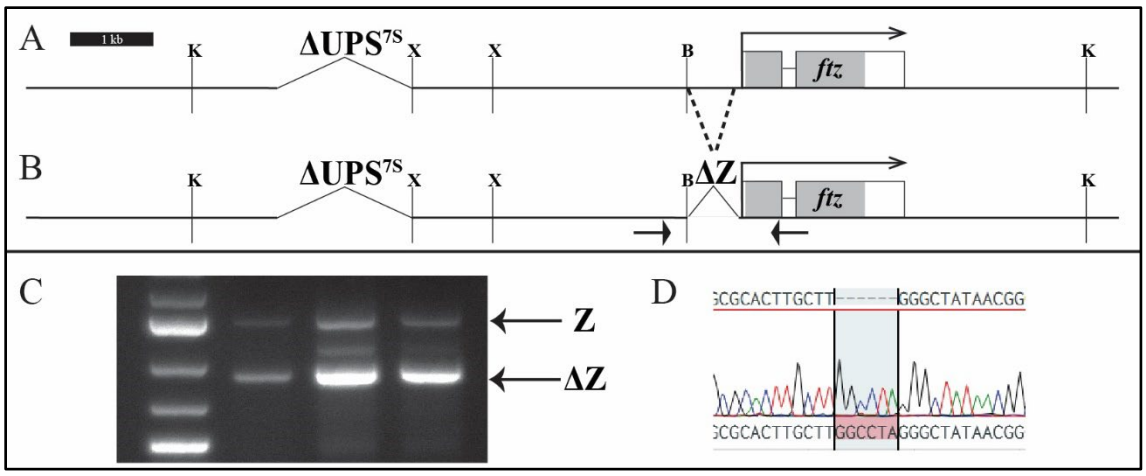


Figure 3-9. Precise deletion of the *ftz* zebra CRE from the *ftzΔUPS$^{7S-/-}$* background generated the homozygous lethal *ftzΔUPS^{7S}ΔZ* allele. (A-B) Schematic for deleting the *ftz* zebra CRE from the *ftzΔUPS^{7S}* background (A) to generate the double deletion mutant *ftzΔUPS^{7S}ΔZ* (B). Guide RNAs and repair template for HDR were the same as described in Chapter 2. Progeny were screened with PCR using the primers zebra1 and zebra 2 (B, arrows), which amplifies a 1485 bp fragment in the presence of the zebra element (C; Z arrow) or 888 bp fragment from precise deletion mutants (C; ΔZ arrow). Positive lines were sent for sequencing to confirm the replacement of the zebra element with an AvrII restriction site (D; red highlight indicates AvrII site).

Unlike the *ftzΔZ* and *ftzΔUPS^{7S}* single mutants, the *ftzΔUPS^{7S}ΔZ* allele is homozygous lethal. Crosses of *ftzΔUPSΔZ/TM3,Sb* did not yield any homozygous embryos developing to adulthood. Examination of cuticle preparations from these crosses revealed that 73% (n=140) had wild-type-like cuticles, 9% (n=17) developed with moderate segmental defects (defects to one segment), and 18% (n=35) developed with severe segmental defects (defects to two or more segments). Given that 75% of the progeny inherit the TM3 balancer with wild type *ftz* and 25% would be homozygous for the deletion, this data suggests that all homozygous *ftzΔUPS^{7S}ΔZ* embryos develop with segmental abnormalities (Fig. 3-10). Further, since ~75% of larvae had wild type cuticles

from both the double mutant and *ftz*^{9H34} experiments, it is evident that *Drosophila* homozygous for the balancer chromosome undergo all of embryonic development despite the homozygous lethal property of the balancer.

The exact nature of the segmental defects varied: in some cases, two abdominal segments were missing (Fig. 3-10A, arrow). In others, two segments were missing, with signs of fusion between two adjacent segments (Fig. 3-10B, arrow). Fusions between segments often appeared incomplete, creating a branch between two adjacent segments rather than two distinct bands (Fig. 3-10C, arrow). Other fusions had the appearance of crisscrossing cuticle between adjacent segments, suggesting an inability of the parasegments to remain properly compartmentalized and separated throughout development (Fig. 3-10D). In sum, *ftzΔUPS*^{7S}*ΔZ*$^{-/-}$ larva develop with defects that vary from moderate to severe. We hypothesized that this phenotypic range was a result of significant impacts on *ftz* transcriptions, which vary in their magnitude between homozygous embryos.

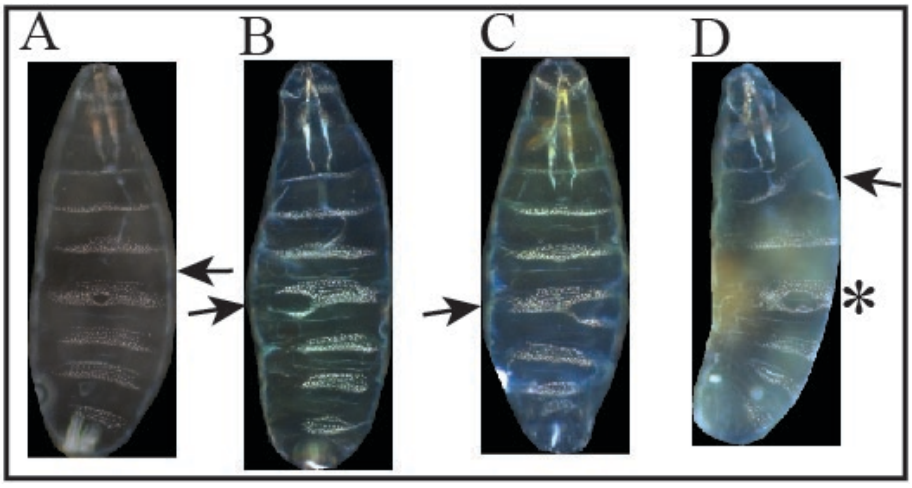


Figure 3-10. *ftzΔUPS*^{7S}*ΔZ*$^{-/-}$ larvae develop with moderate to severe segmental deformities. Darkfield images of *ftzΔUPS*^{7S}*ΔZ*$^{-/-}$ larval cuticle preparations. (A) Most larvae were missing two abdominal

segments (arrow). (B-C) Some other larvae were missing one segment and showed an additional partial fusion/deletion between adjacent segments (arrow). (D) In rare cases, multiple abnormalities were observed: crisscrossing between the T3/A1 segments (arrow), a missing abdominal segment, and pair-rule defect-like fusions between adjacent segments (D; asterisk).

3.4.7 Blastoderm expression of *ftz* is perturbed and rapidly decreases in *ftzΔUPS^{7S}ΔZ$^{-/-}$* embryos

In light of the homozygous lethal and perturbed larval cuticle phenotypes combined with the inability of the *ftzΔUPS$^{7S-/-}$* to transcribe *ftz* through gastrulation, I hypothesized that *ftz* expression is perturbed in at least two domains in these double mutant embryos. As was done for the *ftz^{9H34} in situ* hybridization experiment, a multiplexed chromogenic *in situ* hybridization was conducted with digoxygenin-labeled antisense probes for *ftz* and *lacZ*. This facilitated the identification of homozygous *ftzΔUPSΔZ$^{-/-}$* embryos from those harboring the *TM3,Ser,hb8-lacZ* balancer chromosome while simultaneously analyzing *ftz* expression. Wild type embryos were treated in the same experiment to verify that multiplexing would lead to detectable signal exclusively over *ftz* domains in the absence of *lacZ*.

This was indeed the case, and the detected stripe order was as follows: first, *ftz* is detected in stripes 1 and 5 (Fig. 3-11A i). Expression in stripes 2 is detected next along with very faint signal in stripes 3 and a combined 6/7 (Fig. 3-11B i). As stripes 1, 5, and the combined 6/7 gain intensity, stripes 2 and 3 are still faint (Fig. 3-11C i). Expression in stripe 2 increases substantially, stripe 6 resolves from stripe 7, and expression in stripe 4 becomes detectable (Fig. 3-11D i). Stripe 6 and 7 become more distinct while stripe 6 gains in intensity (Fig. 3-11E i). The full seven stripe pattern becomes distinct, though

expression levels in stripes 3 and 4 appear to remain low (Fig. 3-11F i). Transcription of *ftz* in all seven stripes becomes dramatically higher to form seven full stripes (Fig. 3-11G i). The seven-stripe pattern holds as membrane continues to deposit (Fig. 3-11H i), reaching peak expression at the end of cellularization (Fig. 3-11I i). Expression of *ftz* from all seven stripes is maintained during the transition from stage 5 and stage 6 (Fig. 3-11J i). In summary, striped *ftz* expression is observed in the following order during the blastoderm stage: 1 + 5, 2 + 3 + 6/7, 6 and 7 separate, and finally 4.

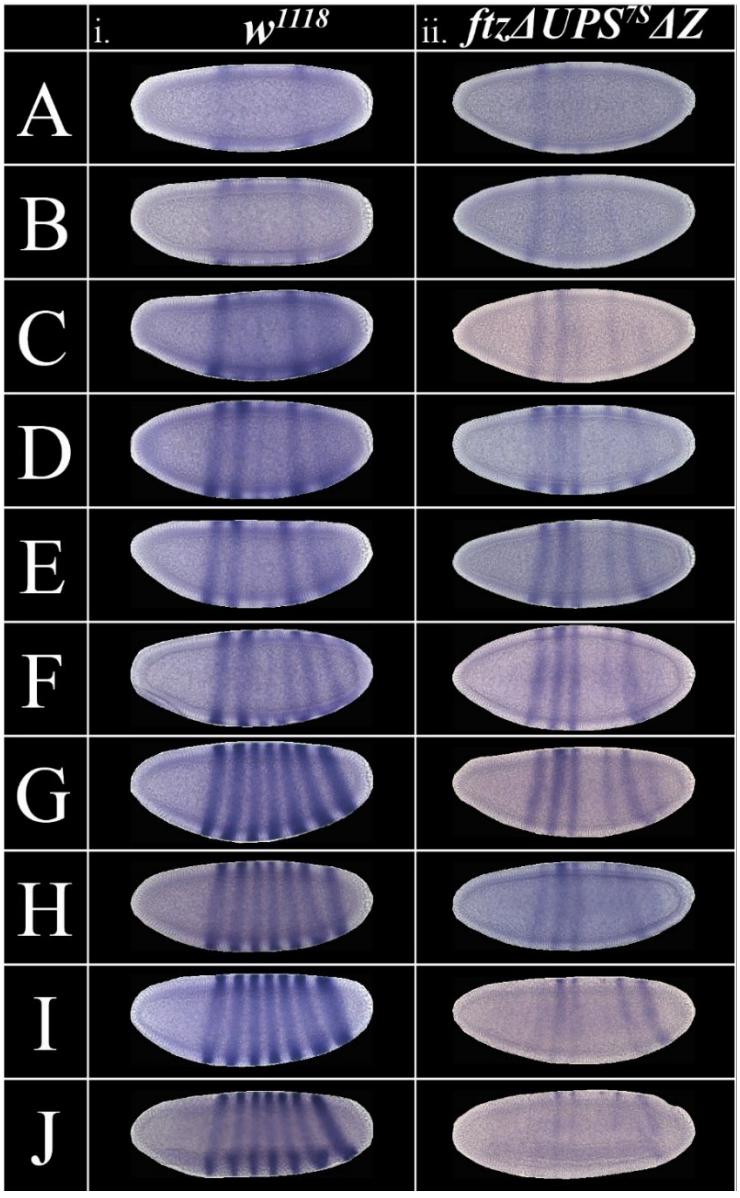


Figure 3-11. *ftz* expression is generally weak and undetected in the stripe 4 domain of *ftzΔUPS^{7S}ΔZ$^{-/-}$* blastoderm embryos. Brightfield images of *ftz* expression time series throughout stage 5 (cellular blastoderm) of wild type (i) and *ftzΔUPS^{7S}ΔZ$^{-/-}$* (ii) embryos using digoxygenin-labeled probes for chromogenic *in situ* hybridization. Expression of the *ftz* stripe domains occurs in the following order in wild type embryos: (A i) 1 +5 (both strong). (B i) 2 +3 + 6/7 (all faint). I 1 + 5 + 6/7 become stronger; 2 + 3 as well but still faint. (D i) 4 (faint) + 6/7 separate; stripe 2 is strong. (E i) 6 + 7 resolve more. (F i) All stripes are resolved, 3 + 4 are faint. (G i) All stripes are well established. (H i) Stripes are maintained as membrane continues to be deposited. (I i) Peak *ftz* expression as cellularization completes. (J i) Stripes are maintained as the embryo begins to transition from stage 5 to 6. Expression of the *ftz* stripe domains occurs in the following order in *ftzΔUPS^{7S}ΔZ$^{-/-}$* embryos: (A) Stripe 1 (strong) + 2 (faint) + 3 (faint). (B) Stripe 5 appears; 2 + 3 become more apparent. (C) Stripes 3 + 6 appear; 2 + 5 are stronger. (D) Stripes 1 + 2 + 3 + 5 + 6 gain signal. (E) Stripe 7 appears; 1-3 and 5-6 gain signal. (F) Stripe 6 broadens on ventral side. (G) Stripes 1-3 and 6 are strong, 7 remains faint. (H) All stripes become weaker (note stripe 4 is never present). (I) Decreases of *ftz* expression observed from each domain as cellularization completes. (J) Transition to stage 6 occurs with very low levels of *ftz* expression.

This pattern is further maintained through stage 6 (Fig. 3-12A-B i). At the start of stage 7, expression of *ftz* becomes noticeably weaker, particularly in the stripe 1 domain (Fig. 3-12C i). Stripe 1 continues to lose signal, but stripes 2-7 are still apparently (Fig. 3-12D i). As the embryo transitions from stage 7 to 8, *ftz* expression is still observed in stripes 2, 3, 4, and 7, but stripes 1, 5, and 6 are nearly undetectable (Fig. 3-12E i). By stage 8, only stripe 7 is clear, though stripes 2, 3, and 4 are faint (Fig. 3-12F i).

This contrasts to what was observed from the *ftzΔUPS^{7S}ΔZ$^{-/-}$* mutant embryos. Expression of *ftz* was first observed exclusively from stripe domains in the anterior, specifically a strong stripe 1 and weak stripes 2 and 3 (Fig. 3-11A ii). Stripe 5 is observed next while stripes 2 and 3 but remain faint (Fig. 3-11B ii). Expression of *ftz* in stripes 2, 3, and 5 increases, with stripe 2 gaining the most signal (Fig. 3-11C ii). At this same

point, stripe 6 expression begins – unlike wild type, it is a thin strip and is not concurrent with stripe 7. At the next point, *ftz* expression is strong in stripes 1, 2, and 3, stripe 6 becomes clearer, and stripe 5 remains faint (Fig. 3-11D ii). Faint expression of *ftz* stripe 7 becomes detectable while stripes 5 and 6 gain signal (Fig. 3-11E ii). While these signals remain consistent, *ftz* stripe 6 appears to broaden slightly on the ventral side (Fig. 3-11F ii). Stripe 5 and 7 expression increases slightly, though stripe 7 remains faint (Fig. 3-11G ii). From this point forward, *ftz* expression appears to decrease in the stripe 1, 6, and 7 domains, with a substantial decrease in stripe 5 (Fig. 3-11H ii). As the blastoderm becomes fully cellularized, stripes 1 and 5 are nearly undetectable, and stripes 2, 3, and 6 lose signal (Fig. 3-11I ii). Stripe 7 expression appears to be strongest at this point. As gastrulation begins with the ventral furrow, *ftz* expression is extremely low in all domains, with stripes 1 and 5 nearly undetectable. During gastrulation, only dorsal expression of *ftz* is observed for stripes 2, 3, 5, and 6 (Fig. 3-12A ii). Stripe 7 still extends across to the ventral side, though it is weak. Stripe 1 is undetectable at this point. For the rest of gastrulation, *ftz* expression is undetectable (Fig. 3-12 B-F ii).

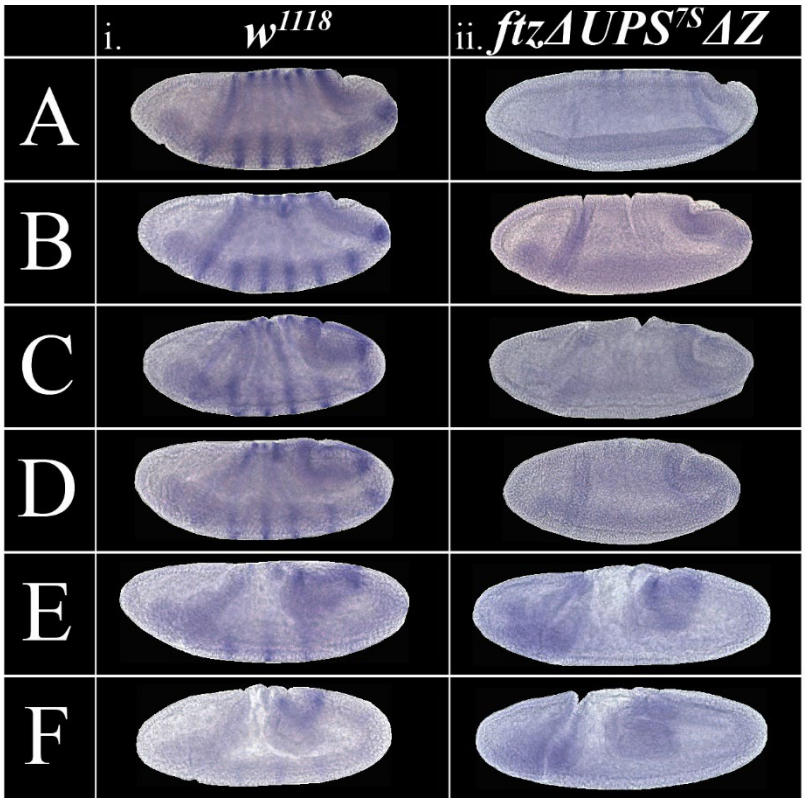


Figure 3-12. *ftz* expression is not maintained throughout gastrulation in $ftz\Delta UPS^{7S}\Delta Z^{-/-}$ embryos. Brightfield images of *ftz* expression time series throughout stage 5 (cellular blastoderm) of wild type (i) and $ftz\Delta UPS^{7S}\Delta Z^{-/-}$ (ii) embryos using digoxygenin-labeled probes for chromogenic *in situ* hybridization. (A) At the start of gastrulation in $ftz\Delta UPS^{7S}\Delta Z^{-/-}$ embryos, *ftz* expression is observed from stripes 2, 3, 5, and 6 exclusively on the dorsal side. Stripe 7 extends to the ventral side, but it is weak compared to wild type. Stripe 1 is undetectable. (B-F) *ftz* expression is not detected for the remainder of germband extension.

There are several differences that stand out when comparing *ftz* expression during the blastoderm between wild type and the double mutant embryos. First, expression of *ftz* begins in the anterior of the double mutant, while stripes 1 and 5 form first in wild type. Next, stripe 6 and 7 form together in wild type, but they are separate in the double mutant with stripe 7 expression occurring later and lower than in wild type. Furthermore, though *ftz* expression is at peak levels upon the completion of cellularization in wild type, *ftz* expression is decreasing in the double mutant by this point. By the time the ventral

furrow starts to form, *ftz* expression is nearly undetectable in multiple domains, whereas in wild type, all seven stripes are strong and maintained. Notably, expression of *ftz* is never detected in stripe 4 of the double mutant embryos. Thus, seven-striped expression is not observed from embryos lacking the UPS^{7S} and zebra elements of *ftz*. From this time series, I draw the following two conclusions: (1) the weak stripe 4 expression observed from the $ftz\Delta Z^{-/-}$ mutants (Chapter 2) was likely contributed from the UPS^{7S} CRE, suggesting that the contributions of this CRE to *ftz* expression are more than purely autoregulatory and (2) the UPS^{7S} and zebra elements constitute the sole seven-stripe CREs of *ftz*.

3.4.8 Expression of *slp1* and *en* is perturbed in $ftz\Delta UPS^{7S}\Delta Z^{-/-}$ embryos

As was done for the single mutants, we next examined expression of *ftz* target genes in $ftz\Delta UPS^{7S}\Delta Z$ embryos. Expression of *slp1* was expanded in 28% (n=23) of embryos (corresponding to the ~ 25% expected homozygotes). Fig. 3-13A shows an embryo in which *slp1* stripes 8/9 are combined and other stripes, such as 10 and 11, appear to be expanded as well. This pattern was observed in 21% of specimen (n=17). In 7% (n=6) of embryos, this loss of repression was observed for multiple *slp1* stripes, showing either complete fusion between stripes or a slight expansion of individual stripes (Fig. 3-13B). A similar trend is observed with *engrailed* expression, for which 19% (n=30) of total embryos showed defects in more than one expression domain. Though the exact expression pattern varied between embryos, there appeared to be two sets of phenotypes; one involved defects in stripes 8, 10, and 14, with a combination of one or two weak stripes, while others involved defects in all *ftz*-dependent *en* stripes (even-numbered stripes). Some showed weak expression of stripe 2 and were missing stripes 8,

10, and 14 (Fig. 3-13C). Other embryos were missing stripes 8, 10, and 14 completely but all other stripes appeared as in wild type (Fig. 3-13D). In the other class, some embryos showed defects in all *ftz*-dependent stripes, missing stripes 8, 10, and 14 while all other even *en* stripes were weak (Fig. 3-13E). In a small number of embryos, the *ftz*-dependent *en* stripes were undetectable (Fig. 3-13F).

In summary, downstream target genes of *ftz*, *slp1* and *en,* were altered in presumably all *ftzΔUPS^{7S}ΔZ* embryos. The prevalence of expression defects in this double mutant is a strong juxtaposition compared to the variability of expression observed from the *ftzΔZ$^{-/-}$* or *ftzΔUPS$^{7S-/-}$* embryos, which both had a range of phenotypes from wild-type-like to severe defect. However, the exact *slp1* and *en* domains that were impacted in the *ftzΔUPS^{7S}ΔZ* mutant embryos were variable. This suggests that the level to which *ftz* is reduced in this double mutant is likely below the threshold necessary for proper regulation of downstream segmentation pathway genes in some domains (for example *en* stripe 8, 10 14 or *ftz* stripes 4, 5, and 7) and near the cusp for others (for example *en* stripe 2 and 12 or *ftz* stripes 1 and 6).

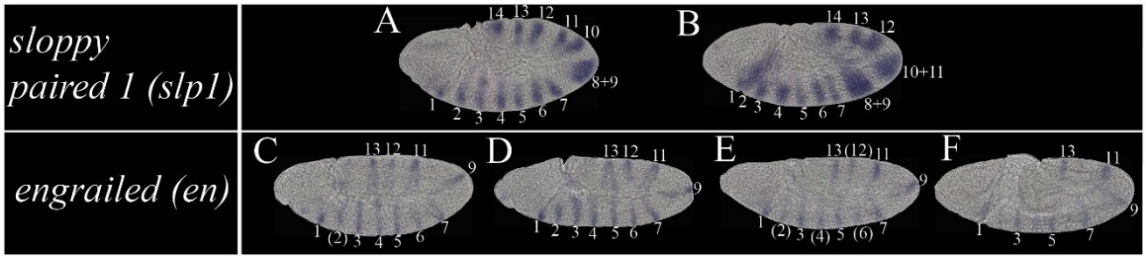


Figure 3-13. The expression domains of *slp1* and *en* are variably affected in *ftzΔUPS^{7S}ΔZ$^{-/-}$* embryos. Brightfield images of *sloppy paired 1* (*slp1*) and *engrailed* (*en*) expression in stages 7-8 (germband extension) of *ftzΔUPS^{7S}ΔZ$^{-/-}$* embryos using digoxygenin-labeled probes for chromogenic *in situ* hybridization. Stripe numbers for each gene are indicated. Expression of *slp1* expanded in *ftzΔUPS^{7S}ΔZ$^{-/-}$* embryos between some adjacent sets of stripes, represented with the plus sign. (A) 21% (n=17) of embryos

showed fusion between one set of domains (8+9). (B) 7% (n=6) showed expansion in multiple domains (8+9 ad 10+11). (C-E) Expression of *en* was consistently absent in multiple stripe domains (parentheses), frequently with weak expression in other stripes as well. (F) In one embryo, all even-numbered *en* stripes were undetectable.

3.4.9 *ftz* transcripts are significantly less abundant in multiple domains of *ftzΔUPS^{7S}ΔZ$^{-/-}$* embryos at the end of blastoderm cellularization

To test the hypothesis that *ftz* transcripts are less abundant in specific transcriptional domains of the *ftzΔUPS^{7S}ΔZ$^{-/-}$* mutant embryos relative to wild type, quantitative analysis was conducted on multiplexed fluorescent HCR *in situ* confocal images using a previously published pipeline (Fig. 3-14; channels are green, *ftz*; red, *eve*; blue, Hoechst) (Graham et al., 2021; Surkova et al., 2019). The *ftzΔUPS$^{7S-/-}$*, *ftzΔZ$^{-/-}$*, and *ftz^{9H34}* mutants were also analyzed in the same experiment to quantitatively compare the relative abundance of *ftz*. These comparisons also served to determine the extent to which Ftz quantitatively impacts *ftz* transcription through direct (autoregulation) or indirect (impact of Ftz on expression of other *ftz* regulators) means. The presence or absence of anterior *lacZ* expression was monitored to identify *ftz^{9H34}* and *ftzΔUPS^{7S}ΔZ$^{-/-}$* homozygotes from siblings with the *TM3, hb8-lacZ* balancer.

Qualitatively, the HCR images reflected the observations of *ftz* expression patterns from colorimetric *in situ* experiments (Fig. 3-11, 3-12), though greater detail could be ascertained. In wild type embryos, *ftz* was detected in seven strong stripes across the trunk (Fig. 3-14A-B). In *ftz^{9H34}* mutant embryos, *ftz* was detected in seven strong stripes across the trunk (Fig. 3-14C). The stripe 3-7 domains appeared to be slightly wider than in wild type (Fig. 3-14D), and stripe 7 width was significantly increased 132.7 ± 45.09% compared to wild type (Table 3-1, red text represents statistically significant

differences, Chi-squared p-value from Kruskal-Wallis test for multiple comparisons). Further, a significant increase of *ftz* abundance was detected in stripe 4 (137 ± 47.25%) and stripe 7 (163.5 ± 57.85%; Fig. 3-14K, Table 3-2, red text represents statistically significant differences, Chi-squared p-value from Kruskal-Wallis test for multiple comparisons). This suggests that the absence of Ftz resulted in an expansion of *ftz* transcriptional domains and greater accumulation of *ftz* transcript. These differences were counterintuitive, as I hypothesized that *ftz* would be less abundant in this mutation due to nonsense-mediated decay of the mutant transcript and lack of positive autoregulation from Ftz protein. However, a possible explanation becomes clearer when analyzing the CRE mutants (see Discussion).

In *ftzΔUPS*$^{7S-/-}$ embryos, expression of *ftz* in stripe 4 appeared qualitatively reduced. The stripe 5 domain appeared slightly thinner than in wild type, and stripes 1, 2, 3, 6, and 7 appeared like wild type (Fig. 3-14E-F). Unlike the *ftz*^{9H34} mutant, the stripe widths were not significantly different than wild type (Table 3-1). Significant reductions of *ftz* abundance relative to wild type levels were detected at 45.77 ± 13.46% in stripe 4 and 73.35 ± 23.13% in stripe 5 (Fig. 3-14K, Table 3-2).

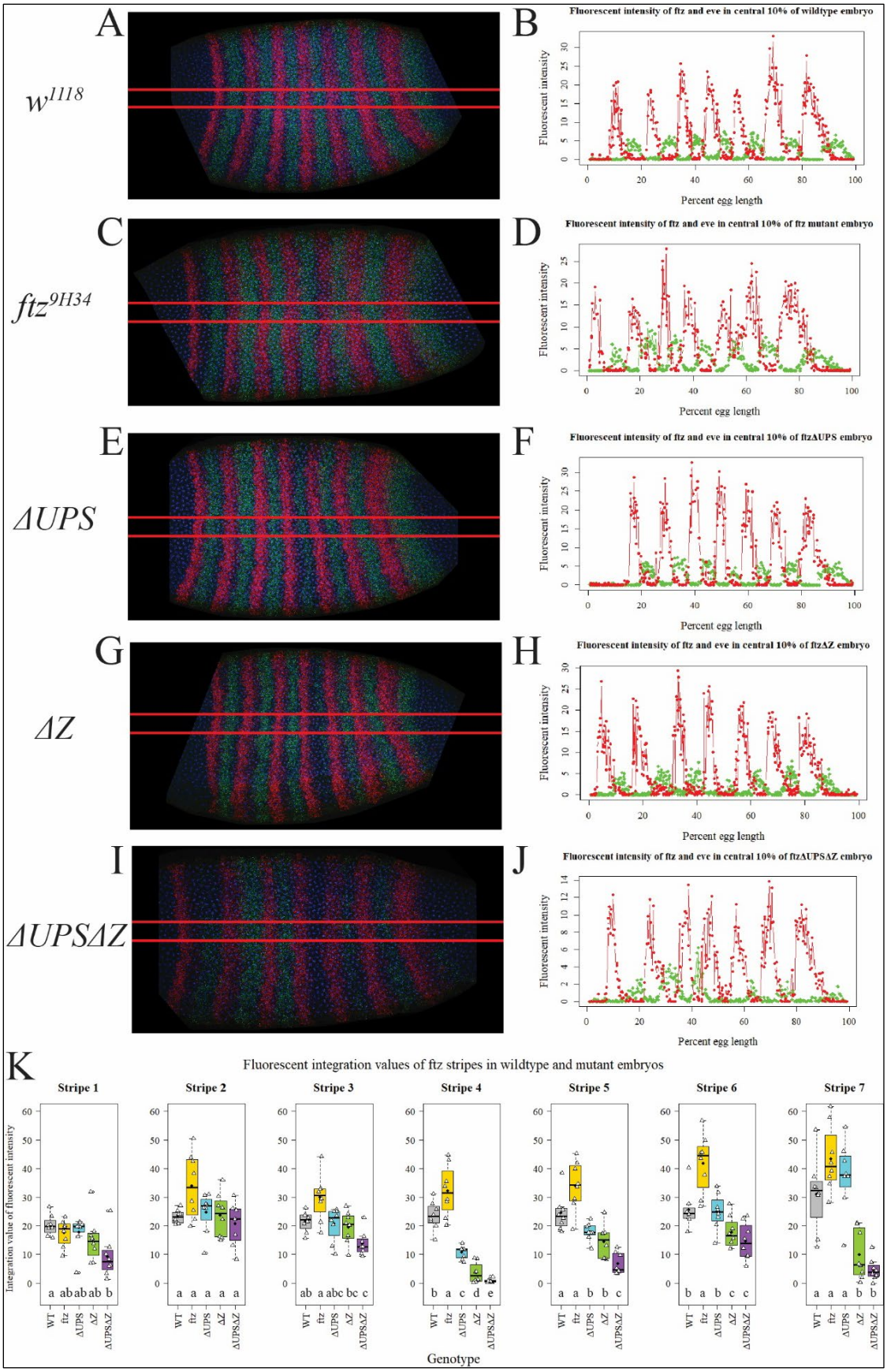

A
w^{1118}
B
Fluorescent intensity of ftz and eve in central 10% of wildtype embryo
Fluorescent intensity
Percent egg length
C
ftz^{9H34}
D
Fluorescent intensity of ftz and eve in central 10% of ftz mutant embryo
Fluorescent intensity
Percent egg length
E
ΔUPS
F
Fluorescent intensity of ftz and eve in central 10% of ftzΔUPS embryo
Fluorescent intensity
Percent egg length
G
ΔZ
H
Fluorescent intensity of ftz and eve in central 10% of ftzΔZ embryo
Fluorescent intensity
Percent egg length
I
ΔUPSΔZ
J
Fluorescent intensity of ftz and eve in central 10% of ftzΔUPSΔZ embryo
Fluorescent intensity
Percent egg length
K
Fluorescent integration values of ftz stripes in wildtype and mutant embryos
Stripe 1
Stripe 2
Stripe 3
Stripe 4
Stripe 5
Stripe 6
Stripe 7
Integration value of fluorescent intensity
a ab ab ab b
a a a a a
ab a abc bc c
b a c d e
a a b b c
b a b c c
a a a b b
WT
ftz
ΔUPS
ΔZ
ΔUPSΔZ
Genotype

Figure 3-14. The abundance of *ftz* and width of the stripe domains are different in deletion mutants relative to wild type and *ftz* null mutants. (A, C, E, G, I) Fluorescent HCR *in situ* hybridization images scanned on a confocal microscope. Channels are: red, *eve*; green, *ftz*; blue, Hoechst. Red line represents the central 10% of the embryonic lateral surface. Histogram for green was uniformly set to 0-75 for all confocal images in ImageJ. Histogram for red was uniformly set to 0-105 in ImageJ. (B, D, F, H, J) One-dimensional analysis line graphs of fluorescent intensity measured over each nuclear position in the central 10% of the lateral surface from the corresponding confocal image (red, *eve*; green, *ftz*). (K) Boxplot showing the integration of fluorescent intensity values under each *ftz* stripe domain from the line graphs. Triangles represent integrations from individual embryos (n=8). Line represents median, black dot represents mean. White circle signifies outlier determined by Boxplot function in R. Genotypes are labeled below the graph and plots are color coded as follows: gray, WT; yellow, ftz^{9H34}; blue, $ftz\Delta UPS^{7S-/-}$; green, $ftz\Delta Z^{-/-}$; purple, $ftz\Delta UPS\Delta Z^{-/-}$. Letters above genotypes represent statistical significance as determined by Kruskal-Wallis test for multiple comparison of treatments; post hoc test used the criterium Fisher's least significant difference. Multiple comparisons corrected with the Holm method.

In Chapter 2, *ftz* abundance in $ftz\Delta Z^{-/-}$ embryos was measured by this analysis in three Z-sections to a depth of approximately 2 μm from the ventral surface. Here, the analysis was done with six Z-sections to a depth of approximately 10 μm, capturing a different perspective that more completely encompassed the *ftz* expression domains. From these images, a general reduction of *ftz* abundance was detected qualitatively from all stripe domains. In comparing to wild type, it was found that significant reductions of stripe domain width were detected at 63.69 ± 26.33% for stripe 4 and 63.85 ± 30.29% for stripe 7 (Table 3-1). Further, significant reductions of *ftz* abundance relative to wild type were calculated at 15.09 ± 14.83% in stripe 4, 58.96 ± 28.59% in stripe 5, 68.91 ± 27.96% in stripe 6, and 31.63 ± 30.69% in stripe 7 (Table 3-2). These values are

consistent with the range of phenotypes on the posterior of the body plan observed from this mutant and corroborate the conclusions described in Chapter 2.

Table 3-1. Quantitative analysis of *ftz* stripe width in percent body length.

ftz stripe	Stripe 1	Stripe 2	Stripe 3	Stripe 4	Stripe 5	Stripe 6	Stripe 7
w1118 Width (percent)	7.75 ± 0.94	7.44 ± 0.91	6.68 ± 0.59	6.67 ± 1.17	7.38 ± 0.98	8.55 ± 0.8	10.62 ± 3.44
ftz9H34 Width (percent)	7.18 ± 1.48	7.75 ± 0.86	7.74 ± 0.96	8.02 ± 0.95	9.09 ± 1.28	10.06 ± 1.83	14.09 ± 1.45
Percentage (*ftz9H34*/WT)	92.67 ± 22.15	104.2 ± 17.2	115.8 ± 17.64	120.2 ± 25.45	123.2 ± 23.84	117.7 ± 24.07	132.7 ± 45.09
ftz ΔUPS Width (percent)	6.78 ± 1.5	7.82 ± 1.34	6.6 ± 1.184	5.78 ± 1.21	6.77 ± 0.89	8.71 ± 1.87	12.27 ± 2.46
Percentage (*ftz ΔUPS*/WT)	87.54 ± 22.07	105.1 ± 22.13	98.75 ± 19.76	86.59 ± 23.67	91.71 ± 17.14	101.9 ± 23.86	115.6 ± 44.01
ftz ΔZebra Width (percent)	7.29 ± 2.04	8.78 ± 1.04	6.87 ± 2.018	4.25 ± 1.59	6.56 ± 1.16	8.50 ± 2.14	6.78 ± 2.35
Percentage (*ftz ΔZebra*/WT)	94.06 ± 28.69	118.1 ± 20.09	102.8 ± 31.55	63.69 ± 26.33	88.89 ± 19.66	99.43 ± 26.7	63.85 ± 30.29
ftz ΔUPS ΔZ Width (percent)	7.56 ± 1.92	10.06 ± 1.84	7.397 ± 1.642	5.19 ± 0.75	5.85 ± 1.225	9.70 ± 2.17	7.69 ± 2.28
Percentage (*ftz ΔUPS ΔZ*/WT)	97.6 ± 27.45	135.3 ± 29.75	110.7 ± 26.46	77.74 ± 17.68	79.31 ± 19.71	113.5 ± 27.51	72.43 ± 31.8
p-value	0.5868	0.0058	0.2002	0.0002	0.0009	0.3352	<0.0001

In the double mutant *ftzΔUPS*^{7S}*ΔZ*$^{-/-}$ embryos, *ftz* expression appeared dramatically lower than in wild type for stripes 1, 3, 5, 6, and 7 (Fig. 3-14I-J). Expression in stripe 4 was undiscernible from background. Interestingly, stripe 2 was 135.3 ± 29.75% wider than observed in wild type, suggesting that the seven-stripe CREs contain sequences to which repressors of *ftz* expression in the neighboring parasegments 3 and 5 can bind (Table 3-1). Conversely, *ftz* stripe 4 width was significantly reduced to 77.74 ± 17.68% of wild type. Abundance of *ftz* was found to be significantly reduced in all stripes except stripe 2 in this mutant relative to wild type, and they are relative to wild type as follows: in stripe 1, *ftz* abundance was reduced to 45.48 ± 37.43%. In stripe 3, *ftz* abundance was reduced to 63.91 ± 22.6%. In stripe 4, *ftz* abundance was nearly undetectable at 3.87 ± 2.76%. In stripe 5, *ftz* abundance was reduced to 27.39 ± 16.89%.

In stripe 6, *ftz* abundance was reduced to 56.81 ± 29.02%. Finally, in stripe 7, *ftz* abundance was reduced to 14.98 ± 13.98%. In short, the most drastic reductions of *ftz* abundance were observed in the domains of stripes 4, 5, and 7. This is consistent with the majority of the defects observed from the phenotypic analyses of the double mutant.

Table 3-2. Quantitative analysis of *ftz* transcript abundance in each stripe domain.

ftz stripe	Stripe 1	Stripe 2	Stripe 3	Stripe 4	Stripe 5	Stripe 6	Stripe 7
w1118 Integrations	20.2 ± 3.52	23.18 ± 2.36	21.53 ± 3.53	23.53 ± 4.91	24.45 ± 6.46	25.63 ± 6.58	31 ± 12.8
ftz9H34 Integrations	17.47 ± 4.68	33.96 ± 11.38	29.88 ± 8.08	32.25 ± 8.85	34.02 ± 8.64	41.91 ± 10.2	43.34 ± 11.37
Percentage (*ftz9H34*/WT)	86.46 ± 27.64	146.5 ± 51.31	138.7 ± 43.89	137 ± 47.25	139.2 ± 50.99	163.5 ± 57.85	139.8 ± 68.39
ftz ΔUPS Integrations	17.76 ± 5.95	24.73 ± 7.02	20.64 ± 5.96	10.77 ± 2.23	17.93 ± 3.09	24.92 ± 6.26	37.39 ± 12.22
Percentage (*ftz ΔUPS*/WT)	87.92 ± 33.2	106.7 ± 32.17	95.83 ± 31.83	45.77 ± 13.46	73.35 ± 23.13	97.22 ± 34.92	120.6 ± 63.51
ftz ΔZebra Integrations	15.24 ± 7.87	23.68 ± 7.67	19.52 ± 5.57	3.55 ± 3.41	14.42 ± 5.86	17.66 ± 5.55	9.81 ± 8.61
Percentage (*ftz ΔZebra*/WT)	75.44 ± 32.26	102.2 ± 34.69	90.65 ± 29.84	15.09 ± 14.83	58.96 ± 28.59	68.91 ± 27.96	31.63 ± 30.69
ftz ΔUPS ΔZ Integrations	9.19 ± 7.39	20.68 ± 7.41	13.76 ± 4.31	0.91 ± 0.62	6.7 ± 3.73	14.56 ± 6.43	4.64 ± 3.84
Percentage (*ftz ΔUPS ΔZ*/WT)	45.48 ± 37.43	89.21 ± 33.23	63.91 ± 22.6	3.87 ± 2.76	27.39 ± 16.89	56.81 ± 29.02	14.98 ± 13.84
p-value	0.0216	0.1241	0.0017	<0.0001	<0.0001	<0.0001	<0.0001

3.4.10 Putative *Scr* CRE is intermixed with the *ftz* CREs

As discussed in the introduction, there is a genomic fragment nested between the *ftz* CREs that is dispensable for *ftz* rescue. Since then, it was proposed that this fragment may contain a CRE that regulates expression of the neighboring *Hox* gene, *Scr* (Gindhart et al., 1995). One mechanism by which the cross-talk between the CREs of these two genes may be regulated is by the presence of an inhibitory CRE such as an insulator or a tethering element. Though this putative *Scr* CRE (Scr^{ftz}-XbaI/HindIII; Scr^{ftz}-X/H) was first reported by Gindhart *et al.* (1995), it falls within a fragment (Scr^{ftz}-$XbaI^2$; Scr^{ftz}-X^2)

that Hiromi *et al.* had reported to be dispensable for *ftz* expression (1985) (Fig. 3-15A). These two fragments share the same 5' end but differ in the 3' position, with the Scr^{ftz}-X^{2} extending 114 bp downstream.

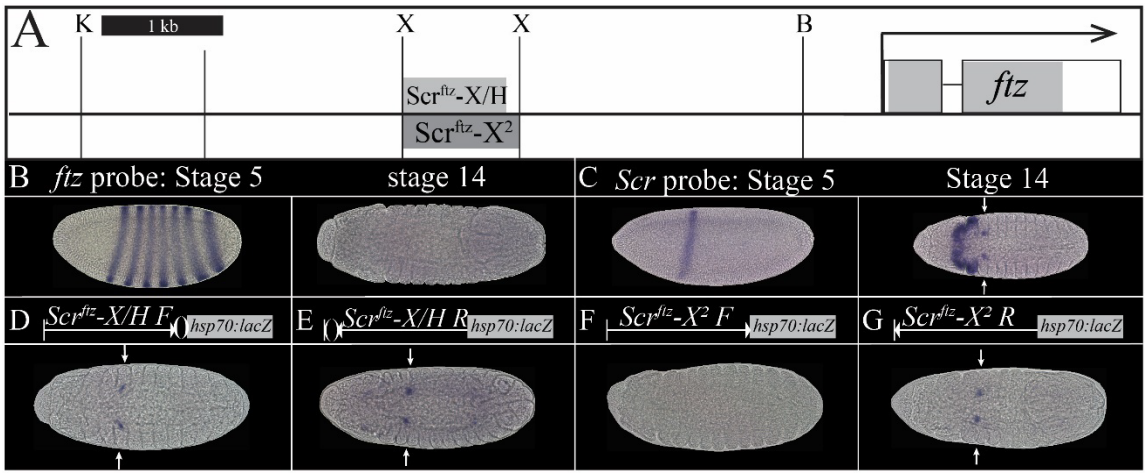


Figure 3-15. A putative *Scr* CRE, Scr^{ftz}-XbaI/HindIII, is inhibited from crosstalk regulation with *ftz*. (A) Schematic of genomic region upstream of *ftz*, highlighting the presence of a previously studied genomic fragment reported to direct *Scr*-like expression pattern in the anterior mid-gut that is intermixed with *ftz* CREs. (B) The putative *Scr* CRE does not appear to act on *ftz*. (C) The *ftz* CREs do not act upon *Scr*. (D-E) Schematic of the *Scr^{ftz}-XbaI/HindIII:lacZ* transgenes, which are expressed specifically in the anterior mid-gut of stage 14 embryos when tested in either orientation. Parentheses highlight the 113 bp difference with a slightly larger fragment, Scr^{ftz}-XbaI^{2}. (F-G) Scr^{ftz}-XbaI^{2} only directs expression of the *lacZ* reporter in the reverse orientation (G). (C, D, E, G) Arrows indicate position of anterior mid-gut expression pattern for *Scr* (C) and for the reporters (D, E, G).

To verify the specificity of gene regulation by this CRE, I examined endogenous *ftz* and *Scr* expression. As shown in Fig. 3-15B, *ftz* expression is not detected during stage 14. However, *Scr* is detected in stage 14 in the anterior midgut (Fig. 3-15C, arrows). *Scr* is also detected during the cellular blastoderm as a single stripe in the anterior of the trunk. Therefore, this CRE does not detectably impact transcription of *ftz* (Fig. 3-15B), nor do the *ftz* CREs direct expression of *Scr* (Fig. 3-15C).

To investigate if an inhibitory CRE is present in this genomic region to prevent this regulatory crosstalk, reporter transgenes were generated by fusing these fragments in both orientations upstream of a basal promoter and the *lacZ* reporter gene. I hypothesized that at least one of the transgenes would express in an orientation-dependent manner. As shown in Fig. 3-15D-E, the Scr^{ftz}-X/H fragment, in either orientation, was sufficient to direct expression of the reporter *lacZ* in the same anterior-midgut domain as *Scr* (Fig. 3-15D-E). However, this was not observed for the Scr^{ftz}-X^2 fragment: for the forward orientation, no expression was detectable (Fig. 3-15F), while the reverse orientation transgene was expressed in the anterior midgut domain (Fig. 3-15G). These results verify the earlier findings from both Gindhardt et al and Hiromi. Additionally, these results suggest that an insulator element is present in the 113 bp between the HindIII and XbaI restriction sites that separate these genomic fragments.

To test if this orientation-specificity depended on the context of the surrounding *ftz* CREs, two other, larger fragments were tested in reporter constructs in both orientations (Fig. 3-16A). One of these fragments, 7SN, contains both known seven-stripe enhancers of *ftz* (UPS^{7S} and zebra) and the neurogenic element. Since it spans the region from the UPS^{7S} through the zebra element, it includes the putative *Scr* CRE (Fig. 3-16B-I). The other fragment, UPS^{7S}-N, contains the same sequence but excludes the zebra element, making it a shorter fragment on the 3' end (Fig. 3-16J-Q). While it was hypothesized that all four fragments would direct *ftz*-like expression in both orientations (seven stripes in stages 5-8 and neurogenic in stage 10) and would not direct *Scr*-like expression (anterior mid-gut during stages 13-15), this was not the case.

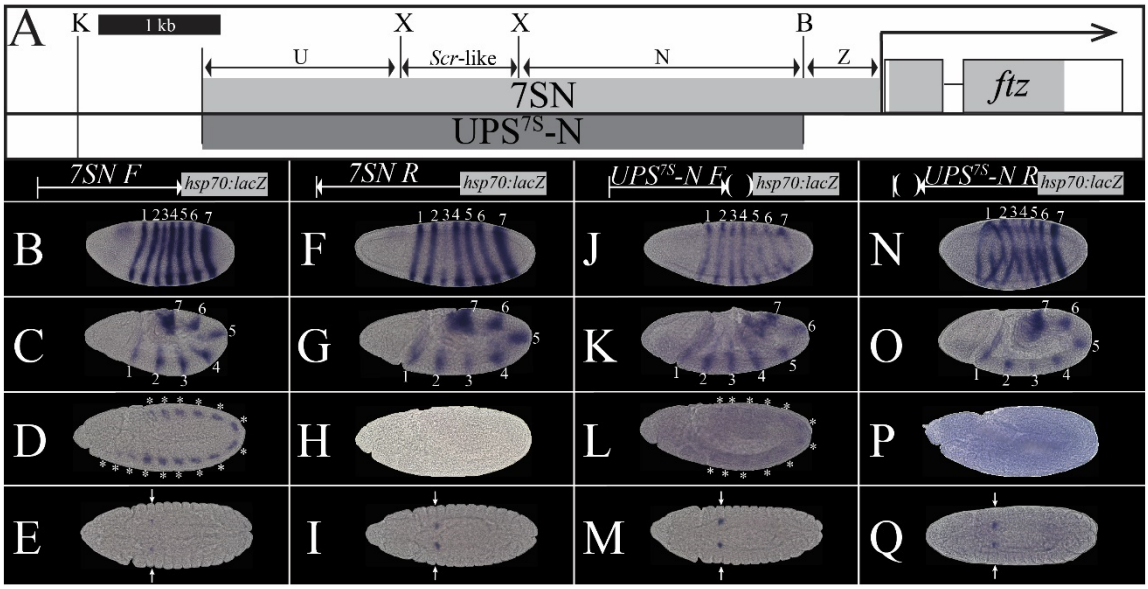


Figure 3-16. The composite 7SN and UPS^{7S}-N CREs direct transgene expression in both *ftz* and *Scr* expression domains. (A) Schematic of genomic region upstream of *ftz*, indicating the position of the 7SN and UPS^{7S}-N fragments and which CREs are contained within them (U, UPS; *Scr*-like, Scr^{ftz}XbaI^{2}; N, neurogenic; Z, zebra). Schematics of the *7SN:lacZ* and *UPS^{7S}-N:lacZ* transgenes; parentheses symbolize the size difference between the two. (D-Q) Brightfield images of *lacZ* expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization in indicated transgenic embryos during the stage 5 (B, F, J, N), stage 8 (C, G, K, O), stage 10-11 (D, H, L, P), and stage 14 (E, I, M, Q). (B-E, J-M) *7SN F:lacZ* and *UPS^{7S}-N F:lacZ* expression was detected in seven strong *ftz*-like stripes in the cellular blastoderm (B and J), during germband extension (C and K), in neuronal precursor cells (C and K), and in the *Scr*-like pattern in the anterior midgut during stage 14 (E and M; arrows). (F-I, N-Q) *7SN R:lacZ UPS^{7S}-N R:lacZ* were detected in seven *ftz*-like stripes in the cellular blastoderm (F and N) and through germband extension (G and O), though stripe 3 was relatively weaker. (H and P) Neither was detected in a *ftz*-like pattern during stages 10-11. (I and Q) Both were detected in the *Scr*-like anterior midgut during stage 14 (arrows).

The transgene *7SN F:lacZ* was expressed in all seven stripes during the blastoderm stage (Fig. 3-16B). The signal was strong, with refined stripes, and expression was maintained through the germband extension phase (Fig. 3-16C). Expression was also detected in the developing nervous system in a *ftz*-like pattern during stage 10 (Fig. 3-

16D). This pattern is consistent with the reported expression pattern of reporter transgenes fused to the *ftz* neurogenic element and of the *ftz* transcript itself (Doe *et al*., 1988). Faint expression was detected in the *Scr*-like pattern in the anterior mid-gut during stages 13-15 (Fig. 3-16E). In the reverse orientation, seven stripes were observed, though stripe 3 was weak (Fig. 3-16F). All seven stripes were observed through the germband extension phase (Fig. 3-16G). Unlike the forward orientation, no contribution to expression from the neurogenic element was detected (Fig. 3-16H). However, the *Scr*-like expression pattern was detected during stage 14 with a stronger signal than in the forward orientation (Fig. 3-16I).

The transgene *UPS^{7S}-N F:lacZ* directed expression in all seven stripes during the blastoderm stage (Fig. 3-16J). Expression continued throughout the germband stage (Fig. 3-16K). Though the signal was detected in the neuronal cells described above, the signal was very faint (Fig. 3-16L, asterisks). Strong signal was observed in the *Scr*-like anterior mid-gut domain during stages 13-15 (Fig. 3-16M). In the reverse orientation, *lacZ* was detected in seven stripes during the blastoderm stage, though stripe 3 was weak, similar to *7SN R:lacZ* (Fig. 3-16O). This expression pattern was maintained through the germband phase (Fig. 3-16P). As with *7SN R:lacZ*, there was no expression observed in the neuronal precursor cells during stage 10 in the *UPS^{7S}-N R:lacZ* line (Fig. 3-16P). Similarly, the *Scr*-like anterior mid-gut expression pattern during stages 13-15 was detected (Fig. 3-16Q).

In sum, orientation effects were observed for the *7SN-R:lacZ* and *UPS^{7S}-N R:lacZ* transgenes, as determined by the absence of signal in the neuronal precursor cells (Fig. 3-16H and P). Notably, the neurogenic element was previously tested only in the forward

orientation (Doe et al., 1988; Hiromi & Gehring, 1987; Hiromi et al., 1985). Though it was long presumed to function in both orientations like a classical enhancer, the results shown here suggest that may not be the case. Further, in the blastoderm and germband phases, the intensity of *lacZ* expression appeared partially dependent on orientation (Fig. 3-16F-G, N-O). Stripe 3 from the *7SN R:lacZ* and *UPS^{7S}-N R:lacZ* transgenic lines was weaker than what was observed from the *7SN F:lacZ* and *UPS^{7S}-N F:lacZ* transgenic lines. Generally, the stripes of *7SN F:lacZ* were more intense than those of *7SN R:lacZ*. Further, *UPS^{7S}-N R:lacZ* appeared stronger than *UPS^{7S}-N F:lacZ* during the blastoderm stage. The orientation dependence on neurogenic expression and blastoderm patterning may suggest that an insulator or other inhibitory CRE is located near the upstream boundary of the neurogenic element. This may help explain the relatively weak signal in the *Scr*-like anterior mid-gut domain of *7SN F:lacZ* animals relative to what was observed for the other three transgenes. Additionally, the presence of both the *Scr*-like anterior mid-gut pattern and *ftz*-like patterns from all four transgenic lines suggests that there are CREs within these constructs that override the transcriptional inhibition observed from the *Scr^{ftz}-X^{2}-F* transgene.

3.4.11 *Scr* expression is unaffected in *ftzΔZ$^{-/-}$*, *ftzΔUPS$^{7S-/-}$*, and *ftzΔUPS^{7S}ΔZ$^{-/-}$* embryos

Given that the *Scr^{ftz}-X/H:lacZ* transgenes were detected in the *Scr*-like anterior mid-gut domain but the *Scr^{ftz}-X^{2}-F:lacZ*transgene was not, combined with the detection of *Scr*-like signal from all four 7SN and UPS^{7S}-N constructs, I hypothesized that there are organizational CREs in the context of the *ftz* gene that prevent cross-talk of the *ftz* enhancers with *Scr* and vice-versa. To examine if any of these potential organizing elements, such as insulators or tethering elements, were disrupted by the CRISPR-

induced deletions of *ftz* CREs described above and in previous chapters, expression of *Scr* was assayed with *in situ* hybridization. As shown in Fig. 3-17, anterior mid-gut expression of *Scr* was detected in all three of the seven-stripe CRE deletion mutants. Accordingly, expression of *ftz* was undetectable in this expression domain in stage 14 embryos. This suggests that the interaction between the putative *Scr* CRE and the *Scr* promoter was not disrupted by these deletions.

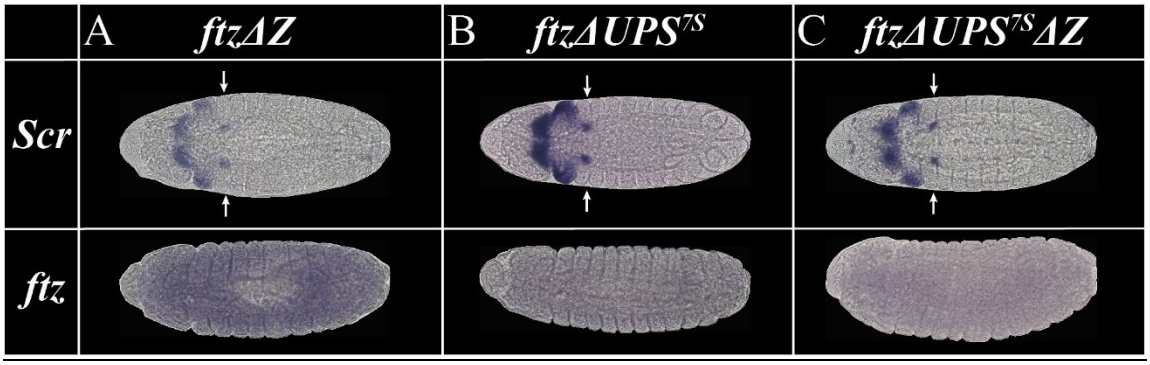


Figure 3-17. *Scr* is detected in the anterior mid-gut of stage 14 embryos in the *ftz* 7-stripe CRE deletion mutants. Brightfield images of *Scr* (top panel) and *ftz* (bottom panel) expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization in indicated deletion mutant embryos. (A) Expression of *Scr* and absence of *ftz* in stage 14 *ftzΔZ*$^{-/-}$ embryos in the anterior mid-gut expression domain suggests that deletion of the zebra element does not interfere with the specificity of interaction between the putative *Scr* CRE and the *Scr* promoter. (B) Expression of *Scr* and absence of *ftz* in stage 14 *ftzΔUPS*$^{7S-/-}$ embryos in the anterior mid-gut expression domain suggests that deletion of the UPS^{7S} similarly does not disrupt the *Scr* CRE-promoter interaction. (C) Expression of *Scr* and absence of *ftz* in stage 14 *ftzΔUPS*^{7S}*ΔZ*$^{-/-}$ embryos in the anterior mid-gut expression domain suggests that deletion of both seven-stripe CREs does not disrupt the *Scr* CRE-promoter interaction.

3.5 Discussion

The work presented in this chapter has been a re-examination of classic ftz CREs that contributes to the seven-striped expression of *ftz,* with a focus on the UPS. Though

much had been previously discovered about the UPS over the last four decades, prior studies were limited to the examination of transgenes (Han et al., 1993; Han et al., 1998; Hiromi & Gehring, 1987; Hiromi et al., 1985; Pick et al., 1990). These fusion genes, whether that be the rescue constructs tested by Hiromi or the numerous transgenic reporters that examined the expression domains affected by these genomic fragments, were inserted in a variety of constructs and randomly transformed into the genome through P-element-mediated mutagenesis. Here, I have equalized the molecular variables that may have been impacted by these approaches by fusing the examined genomic fragments to the same restriction site (XbaI) of the same plasmid (*placZ-attB*), which in turn were recombined into the same genomic location (PhiC31 docking site 57F5, 2R:21645971 on the second chromosome).

3.5.1 Transgene expression analysis has refined a classic *ftz* CRE

The expression of *lacZ* from these transgenic *Drosophila* embryos was then examined through chromogenic *in situ* hybridization with digoxygenin-labeled probes that were detected using alkaline phosphatase-conjugated α-digoxygenin antibodies. It is notable that this assay is more sensitive than the radioactive-labeled probes used in the original *in situ* hybridization experiments that monitored *ftz* expression (Hafen et al., 1984). This assay also better reflects the qualitative transcriptional output of these transgenes compared to the β-galactosidase assays used in previous publications to monitor transgene expression, as those previous experiments measured protein activity rather than presence of transcript (Hiromi & Gehring, 1987; Pick et al., 1990). Simply put, the chromogenic *in situ* hybridization experiments here are more sensitive and better reflect transcriptional patterning. The point made here is well illustrated by considering

that this methodology had previously determined that the striped expression of *ftz* occurs in a dynamic pattern rather than all seven stripes at once, as was previously believed from assays using radioactive-labeled probes (Yu & Pick, 1995).

In addition to these parameters, the experiments done here have taken into account that new CREs have been identified in the *ftz* locus since the original discovery of the UPS (Schroeder et al., 2011). Whereas the original boundaries of the UPS were determined through randomly placed KpnI/XbaI restriction sites that contained a genomic fragment sufficient to rescue *ftz* mutants when fused to the zebra element (Hiromi et al., 1985), the ftz-6 CRE was predicted through computational analysis and experimentally determined *in vivo* to be a stripe-specific enhancer that directs *ftz*-like expression in stripes 2 + 7 (more details to follow in Chapter 4). Accordingly, this information was utilized to refine the 5' boundary of the UPS to exclude the ftz-6 fragment (UPS^{7S}). Though this refinement excludes the ftz-6 region, it still contains the previously characterized distal enhancer region within the UPS^{7S} (Pick et al., 1990).

From monitoring *lacZ* expression of these UPS reporter transgenic *Drosophila*, I have determined that the refined UPS^{7S} is a more accurate reflection of the seven-stripe CRE. This claim is made on the basis that the *UPS^{7S}:lacZ* transgenes were detected in seven precise stripes during the blastoderm and germband extension phases, regardless of orientation (Fig. 3-1 D-E). It is worth noting that during the blastoderm, expression of *lacZ* in the stripe 4 and 5 domains appeared qualitatively weaker than the other stripes, though strong signal was detected from all seven stripes during the germband phase (Fig. 3-1 D-E). Future experiments examining these transgenes in a *ftz* mutant background would be able to ascertain whether or not this patterning is Ftz-dependent (purely

autoregulatory) as previously determined by less sensitive assays (Hiromi & Gehring, 1987; Pick et al., 1990).

3.5.2 Re-evaluation of UPS^{7S} necessity implicates redundancy

My work has also investigated the impacts on endogenous *ftz* expression by precisely deleting the seven-stripe CREs from the genome (Fig. 3-2). I was surprised to discover that animals homozygous for the precise deletion allele *ftzΔUPS^{7S}* are viable and fertile. The rescue experiments conducted by Hiromi provided strong evidence that both the UPS and zebra elements together are necessary and sufficient to rescue *ftz* mutants (Hiromi et al., 1985). In light of discovering that both *ftzΔUPS$^{7S-/-}$* and *ftzΔZ$^{-/-}$* animals (Chapter 2) successfully complete all stages of development, it becomes clearer that the conclusions drawn by Hiromi's experiments were limited to the context of the KpnI genomic fragment from which the rescue constructs were generated. Expanding on this, we now know that the KpnI fragment does not contain the majority of the stripe-specific CREs (Schroeder et al., 2011). Whereas loss of the UPS was not tolerated from Hiromi's rescue constructs, deleting the UPS^{7S} in the context of the *ftz* locus leaves the other stripe-specific CREs intact and still capable of directing *ftz* stripe expression.

A major conclusion from these results is that there is considerable redundancy in the *cis*-regulation of *ftz*: in the absence of the UPS, the zebra element and stripe-specific CREs are sufficient to enhance *ftz* transcription to the levels necessary for over 90% of homozygous deletion animals to develop with wild-type-like segmentation (Fig. 3-3 and 3-4). In monitoring the qualitative expression of *ftz* in this background, it was observed that the seven-stripe pattern of *ftz* is well established during the blastoderm stage (Fig. 3-5). It should be noted that *ftz* expression in stripe 4 appeared relatively lower than the

other stripes. Given that expression of both *UPS^{7S}:lacZ* reporter transgenes was strongly detected in stripes 1, 2, 3, 6, and 7 but weakly in stripes 4 and 5, this was unexpected. I had hypothesized that the loss of the UPS^{7S} CRE would result in the opposite effect on endogenous *ftz* expression. In other words, I expected a significant impact to the qualitative pattern of stripes 1, 2, 3, 6, and 7 but minor impacts to stripes 4 and 5. It may be that non-additive interactions between the *ftz* CREs in the context of this locus explain this contradictory phenomenon, but the data here alone is insufficient to draw this conclusion.

Regardless, *ftz* transcription fails to persist through gastrulation and germband extension in the absence of the UPS^{7S} region (Fig. 3-6). This is contrary to what is observed in wild type embryos, in which the seven-striped expression pattern of *ftz* is maintained until approximately stage 8. Notably, this process appears to be Ftz-dependent, for the same inability of *ftz* transcription through gastrulation was observed from embryos homozygous for the amorphic *ftz^{9H34}* allele, as previously reported for other amorphic *ftz* mutants (Lawrence & Johnston, 1989). Thus, the UPS^{7S} CRE plays a non-redundant role in the propagation of *ftz* transcription past stripe establishment. This is consistent with previous findings that the UPS is autoregulatory and Ftz-dependent (Hiromi & Gehring, 1987; Pick et al., 1990).

The expression patterns of downstream Ftz target genes (and the segment polarity genes in general) are established around this gastrulation transition point and through stages 6-8 (Fig. 3-7; DiNardo & O'Farrell, 1987; DiNardo et al., 1985; Fjose et al., 1985; Grossniklaus et al., 1992; Kornberg et al., 1985). It may be that the lack of *ftz* transcriptional propagation through this transition point exacerbates the relatively lower

levels of *ftz* transcript in the *ftzΔUPS*$^{7S-/-}$ mutant, as calculated from the fluorescent HCR *in situ* data (Fig. 3-14). This analysis revealed that *ftz* transcript levels were significantly reduced to around 46% in the stripe 4 domain and 73% in the stripe 5 domain.

While the adult abdominal analysis (Fig. 3-3), larval cuticle preparations (Fig. 3-4), and *en/slp1 in situ* hybridization experiments (Fig. 3-8) revealed that the majority of *ftzΔUPS*$^{7S-/-}$ animals were wild-type-like, a small fraction of embryos showed minor aberrations in the adult segments A3 and A5 and the embryonic parasegments 8 and 10. Stripes 4 and 5 of *ftz* are respectively expressed in parasegments 8 and 10, which include segments A3 and A5 (Lawrence & Johnston, 1989; Lawrence et al., 1988; Magrassi & Lawrence, 1988). Given the variance from the HCR quantitative analysis, it is possible that individual embryos in which *ftz* abundance was on the lower end of this range had insufficient quantities of Ftz protein to activate downstream genes of the segmentation pathway (Fig. 3-14K, small triangles represent stripe integrations from individual embryos). Thus, the majority of embryos would have wild-type-like levels of Ftz and would develop all Ftz-dependent segments, while only a fraction of embryos would be negatively impacted.

3.5.3 Stripe-specific CREs alone are insufficient for *ftz* function

In deleting the zebra element from the *ftzΔUPS*^{7S} background, it was found that the *ftzΔUPS*^{7S}*ΔZ* allele is homozygous lethal (Fig. 3-9). Monitoring *ftz* expression with *in situ* hybridization revealed that these homozygotes failed to express *ftz* in the stripe 4 domain, suggesting that the zebra element and the UPS^{7S} constitute the only seven-stripe CREs of *ftz* expression (Fig. 3-11). Analysis of *ftzΔZ*$^{-/-}$ mutants revealed that the zebra element is not the sole contributor to expression of *ftz* in the stripe 4 domain (Chapter 2;

Graham et al., 2021). This was contrary to predictions based on the absence of a known stripe 4-specific CRE (Schroeder et al., 2011), and it led to the conclusion that there is an as-of-yet undiscovered stripe 4-specific CRE elsewhere in the genome. However, taking into account the lack of detectable *ftz* in the stripe 4 domain of the double CRE deletion mutant, it appears that it was the UPS^{7S} fragment that enhanced transcription of *ftz* within this domain of the $ftz\Delta Z^{-/-}$ mutants and thus, the missing stripe 4 CRE resides within the UPS. It is likely that prior work suggesting that the UPS is purely Ftz-dependent and autoregulatory was limited by the methods of detection available at the time, highlighting the importance of re-investigating past conclusions with modern methodologies (Hiromi & Gehring, 1987; Pick et al., 1990).

The $ftz\Delta UPS^{7S}\Delta Z^{-/-}$ larval cuticles consistently had aberrations in two abdominal segments. The two most common segmental phenotypes were either the total loss of two segments or the loss of one segment with aberrations to another (Fig. 3-10). Presumably, the consistently missing segment was A3 due to the nearly total loss of *ftz* in stripe 4. The other affected segment is likely A5 due to reductions of *ftz* abundance in stripe 5 to approximately 27% (Fig. 3-14, Table 3-2). The presence or absence of A9 (derived from parasegment 14, which is the *ftz* stripe 7 domain) could not be ascertained from the cuticle preparations, since A9 cannot be observed from the ventral exterior of the larval cuticle. Future experiments examining the presence or absence of A9 specifically could determine if the reduction of *ftz* abundance to approximately 15% in the stripe 7 domain results in the inability to form this segment.

In examining *en* expression, it was observed that all the $ftz\Delta UPS^{7S}\Delta Z^{-/-}$ embryos were lacking *en* stripes 8, 10, and 14 (Fig. 3-13). Some embryos showed reductions or

loss of other Ftz-dependent *en* stripes as well (those expressed in the even parasegments). These observations, along with the calculation of lower *ftz* abundance in the stripe 5 and 7 domains, suggest that these reductions are below a threshold required for *ftz* function. The lower *ftz* abundance relative to wild type expression levels observed in stripe 1 (to approximately 45%), stripe 3 (to approximately 64%), and stripe 6 (to approximately 57%) may result in aberrant phenotypes. Specifically, though the average *ftzΔUPS^{7S}ΔZ$^{-/-}$* embryo has sufficient abundance of *ftz* in these stripe domains, the variance of the data suggests that some individual embryos express *ftz* on the lower end of this range. In those individuals with lower expression, the inability to propagate *ftz* transcription through gastrulation may make them more sensitive to downstream expression defects in these other even-numbered parasegments.

3.5.4 Severe phenotypes in the *ftzΔUPS^{7S}ΔZ$^{-/-}$* mutants implicate *ftz* threshold requirements

The threshold for *ftz* abundance necessary to accumulate sufficient Ftz protein to activate downstream segmentation genes in those domains is likely under (but close to) the percentages determined by these calculations. In analyzing the HCR data calculations from each seven-stripe CRE deletion mutant, it appears the threshold percentage for *ftz* abundance in each stripe relative to wild type levels is close to the following: under 45% for stripe 1, around 15-25% for stripe 4, around 27% for stripe 5, and between 15-32% for stripe 7. It is unlikely that the work presented here can provide a meaningful resolution for the *ftz* expression thresholds for stripes 2, 3, and 6. Deficiency mutants with large deletions can be maintained as heterozygotes, which presumably express *ftz* at around 50% of wild type. I note that for each of the CRE deletion mutants presented here,

expression of *ftz* was above 50% in stripes 2, 3 and 6. These values further suggest that in the absence of the seven-stripe CREs, the stripe-specific CREs alone are insufficient to express *ftz* at the sufficient quantities necessary for segmentation to proceed.

3.5.5 Ftz protein may indirectly refine *ftz* transcriptional domains

In comparing the ftz^{9H34} mutant HCR analysis to those of the CRE deletion mutants, it is interesting to note that opposite trends were observed (Fig. 3-14, Table 3-1, Table 3-2). While *ftz* expression domains are expanded and have a higher abundance of *ftz* transcript in the ftz^{9H34} mutant, they are generally reduced in the CRE mutants (with the exception of stripe 2 expansion in the double CRE deletion mutant). I hypothesize three possible explanations for the phenomenon observed from the ftz^{9H34} mutant: one, the loss of Ftz indirectly results in the increased abundance of transcriptional activators that act in *trans* to increase *ftz* expression. Two, the loss of Ftz indirectly decreases the abundance of transcriptional repressors of *ftz* expression around these domains, resulting in increased expression of *ftz* in the even parasegments and/or ectopic expression of *ftz* in the cells immediately surrounding them. Three, the loss of Ftz results in a combination of increased expression of *ftz* transcriptional activators and decreased expression of *ftz* transcriptional repressors. If any of these are true, a paradox is presented from the perspective of the seven-stripe CRE deletion mutants: in all three mutants, *ftz* abundance is significantly reduced in at least two stripe domains, and the occurrence of both segmental and parasegmental phenotypes suggests that this resulted in a decrease of Ftz protein abundance.

Stripe 4 expression stands out as a consistent example between the three deletion mutants. In the double CRE deletion mutant $ftz\Delta UPS^{7S}\Delta Z$, expression of *ftz* within this

domain was essentially undetectable. If the above hypotheses were true, it would be expected that loss of *ftz* would paradoxically result in increased or stable expression compared to wild type rather than total loss of *ftz* expression and/or transcript abundance.

However, this paradox is resolved in considering the fundamental fact that genomic fragments known to bind transcriptional regulators of *ftz* have been removed in these deletion mutants (Dearolf et al., 1989a; Han et al. 1993; Han et al., 1998; Pick et al., 1990). Thus, even if the presence or absence of these transcriptional regulators of *ftz* are impacted by reduced Ftz protein levels, there would not be a way for these *trans* factors to interact with the *ftz* locus. In the absence of these genomic fragments, the CRE deletion mutants appear to be refractory to these proposed mechanisms that would result in the buffered/increased *ftz* abundance in the ftz^{9H34} mutant. In other words, I propose that the factors altered in the ftz^{9H34} mutants bind to the UPS^{7S} and/or the zebra CREs. Given that the expression of *ftz* in stripe 7 is much higher and covers a significantly wider domain in the $ftz\Delta UPS^{7S-/-}$ and ftz^{9H34} mutants compared to the $ftz\Delta Z^{-/-}$ and double CRE deletion mutants, it is possible that *trans* factors that are present in this specific domain may respond in such a way to altered Ftz levels and strongly activate *ftz* expression when bound to the zebra element.

The impact on *ftz* expression in the absence of Ftz has been monitored over decades, including with the use of digoxygenin-labeled probes and the alkaline phosphatase reaction (Hiromi & Gehring, 1987; Lawrence & Johnston, 1989; Pick et al., 1990; Schroeder et al., 2011). This determined that the seven *ftz* stripes can be established in the mutant background during stage 5, as has been described here (Fig. 3-6). However, presumably due to the amplification nature of the alkaline phosphatase reaction, the slight

alterations to *ftz* stripe width and the changes to *ftz* abundance that were calculated from the HCR analysis have not been reported until now (Fig. 3-14, Table 3-1, Table 3-2). The intensity of the fluorescent output from HCR *in situ* hybridization has been shown to correlate closely with the number of transcripts present (Choi et al., 2014). It seems that this precision has allowed for the careful evaluation of increased *ftz* transcript abundance in the ftz^{9H34} mutant embryos, which could not be discerned from *in situ* hybridization experiments using other methodologies. This finding further highlights the benefits of using new methodologies to re-evaluate priorly drawn conclusions.

3.5.6 *ftz* CRE synergy is non-additive

While reporter transgenes can provide valuable information on the sufficiency of a genomic fragment to direct expression of a reporter, these experiments sometimes fail to recapitulate endogenous expression patterns. This phenomenon has been reported for various *ftz* CRE reporter transgenes, specifically those involving the UPS, zebra element, or a combination of the two (Chapter 2, Graham et al., 2021; Hiromi & Gehring, 1987; Pick et al., 1990; Yu & Pick, 1995). Though *ftz*-like expression patterns were detected from *7SN F:lacZ* and UPS^{7S}*-N F:lacZ* transgenic embryos, disparities in regards to the endogenous *ftz* expression pattern were also observed (Fig. 3-16). From the *7SN:lacZ* transgenic embryos, an anterior expression domain was detected. This has been previously reported for reporter constructs that contain the zebra element, but it does not accurately reflect endogenous *ftz* expression. Further, expression of *lacZ* in *ftz*-like stripes was very weak in UPS^{7S}*-N F:lacZ* embryos. Additionally, the contribution of the neurogenic CRE to neuronal *lacZ* expression during stage 10-11 in these transgenic embryos was also weak.

Most puzzling was the detection of *lacZ* expression in the *Scr*-like anterior mid-gut pattern of both *7SN:lacZ* and *UPS^{7S}-N:lacZ* embryos, especially if we consider that this pattern was not detected in *Scr^{ftz}-X^{2} F:lacZ* transgenic embryos (Fig. 3-15). If the transcriptional enhancement from this putative *Scr* CRE is inhibited in *Scr^{ftz}-X^{2} F:lacZ* animals, how does it produce the pattern in the surrounding context of the *ftz* genomic fragments of 7SN and UPS^{7S}-N when there is no evidence of endogenous *ftz* expression in the anterior mid-gut? A simple answer to this question would suggest the presence of yet another insulator between the neurogenic element and the zebra element that cancels out the inhibition of expression observed in *Scr^{ftz}-X^{2} F:lacZ* embryos, but this does not make sense in the endogenous context of the gene. This further fails to explain how reporter transgenes of the neurogenic element direct detectable expression. Regardless of the cause, it appears that the expression patterns of reporter transgenes fused to large genomic fragments that contain multiple *ftz* CREs and the intervening genomic sequences between them have historically been difficult to rationalize (Yu & Pick, 1995). Future studies aimed to dissect the complexity of enhancer-promoter crosstalk within this locus may be facilitated with the use of CRISPR/Cas9-induced genome edits to study endogenous expression patterns, as was done with the deletions investigated here, rather than reporter constructs.

3.5.7 Conclusions

Overall, I have shown that the use of newer methodologies to re-evaluate findings made from experiments over four decades has provided several unexpected results. The UPS^{7S} probably directs expression of *ftz* in stripe 4 in a Ftz-independent manner, the UPS^{7S} is unnecessary for seven-striped *ftz* expression during the blastoderm, the loss of

ftz expression during gastrulation from loss of autoregulation will still produce wild-type-like animals over 90% of the time, and *ftz* expression during the blastoderm is significantly increased in the absence of Ftz. The stripe-specific CREs of *ftz* are insufficient for *ftz* function. Finally, an insulator appears to be present within the dispensable region that is somehow circumvented by the upstream CREs. Taken together, these findings further our understanding of how transcription is regulated in a complex genomic region and emphasize that re-examining prior conclusions, even those that have been carefully and extensively studied, can still be elucidating.

Chapter 4: The role of stripe-specific elements in regulation of *ftz* gene expression

4.1 Abstract

Pair-rule genes (PRGs) are expressed in alternating sets of stripes across the embryo of *Drosophila melanogaster*. Our understanding of the regulatory mechanisms that result in refined striped expression is convoluted and inconsistent. Misregulation of a single stripe can have developmental consequences and result in aberrant segmentation. Here, I conducted a comprehensive analysis of stripe-specific CREs of the PRG *fushi tarazu* (*ftz*). Two approaches were used to determine their contributions toward the *ftz* stripe expression pattern. First, I systematically monitored the sufficiency of these genomic regions, in both forward and reverse orientations, to direct expression in *ftz* stripe domains. I have found that all of the examined stripe-specific CREs are impacted by orientation, and I propose the location of a tethering element directly downstream of the *ftz* transcription unit is in part responsible for these orientation effects. I also confirmed that these CREs direct expression specifically in the *ftz* domains. For the second approach, I targeted a nested pair of *ftz* stripe 2 CREs for precise deletion with the CRISPR/Cas9 system to create the *ftzΔ413* and *ftz413Δftz-6* alleles. The peak *ftz* stripe pattern of these mutants resembles the wild type background despite perturbations of *ftz* stripe establishment. Both deletion mutants are viable, fertile, and are maintained as homozygous lines, though the *ftz413Δftz-6*$^{-/-}$ females appear to lay fewer embryos. This proof of concept on the necessity of a stripe-specific CRE towards *ftz* function suggests

that stripe establishment is less consequential than stripe maintenance due to *cis*-regulatory redundancy.

4.2 Introduction

Segmentation of the insect body plan is established during embryogenesis by a hierarchy of segmentation genes (Nüsslein-Volhard & Wieschaus, 1980; Wieschaus & Nusslein-Volhard, 2016). Proteins translated from four maternally-contributed transcripts are critical for establishing the anterior-posterior axis of the early embryo (reviewed in (Johnston & Nüsslein-Volhard, 1992)). Briefly, the gradients of these maternal morphogens determine the broad domains in which the gap genes are transcribed. Within the trunk of the embryo, subsequent inter-gap gene repression results in the formation of gap gene product gradients. They are as follows, from anterior to posterior: Hunchback (Hb), Giant (Gt), Kruppel (Kr), Knirps (Kni), and then relatively lower levels of Gt and Hb again (Knipple et al., 1985; Mohler et al., 1989; Pankratz et al., 1990; Tautz, 1988).

Through a combination of genetic analysis and gene expression studies, it was determined that the gap gene products regulate the periodic striped expression of the pair-rule genes (PRGs) during the cellular blastoderm stage of development (reviewed in (Carroll, 1990; Jaeger, 2011; Pick, 1998)). For the PRGs *ftz* and *eve*, these stripes form registers of approximately four cells wide in alternating parasegments across the trunk of the embryo (Carroll & Scott, 1985; Hafen et al., 1984; Lawrence et al., 1988; Macdonald, Ingham, & Struhl, 1986). Despite decades of diligent study, the exact mechanisms explaining how the non-periodic gap gene expression domains induce the precise periodic expression of each pair-rule register remain a mystery today.

Early studies of PRG stripe-specific CREs determined that they can directly bind maternal and gap gene products *in vitro* (Pankratz et al., 1990; Small et al., 1991; Štanojević et al., 1989; Stanojevic et al., 1991). This was initially done through biochemical analysis. The clustering and quantity of these binding sites, and their affinities for specific *trans* regulatory factors, were proposed to be the fundamental mechanisms that direct stripe-specific *cis*-regulation. Direct evidence for this was first experimentally determined through the careful dissection of the *eve* stripe 2 CRE (Small et al., 1992; Small et al., 1991; Stanojevic et al., 1991). To summarize, mutant analysis revealed that *eve* stripe 2 expression is activated by Hb and the maternal effect protein Bicoid (Bcd) (Frasch & Levine, 1987; Goto et al., 1989). Repression by Gt and Kr inhibit expression around the stripe 2 domain on the anterior and posterior borders, respectively. By systematically mutagenizing specific binding sites for these factors in reporter transgenes, stripe 2 expression was reduced and/or lost when the Hb and/or Bcd binding sites were mutated (Small et al., 1992). Likewise, expression was expanded beyond the stripe 2 domain if the Gt and/or Kr binding sites were mutated (Small et al., 1992). This provided evidence that the biochemical studies conducted *in vitro* reflected the binding of these factors to the CREs *in vivo* to result in regulation of gene expression. Expanding on this, a model was developed that PRG stripe establishment was a function of binding site affinity and *trans* factor concentration within certain regions of the embryo.

Similar mechanisms were worked out for the *eve* stripe 3/7 CRE, though they were less deterministic. Unlike the *eve* stripe 2 CRE that is activated by gradients of Bcd and Hb, the *eve* stripe 3/7 CRE is repressed by Hb and Kni (Small et al., 1996). Activation of this CRE is believed to be mediated by factors that are ubiquitous

throughout the embryo. The pioneer factor Zelda (Zld) has been implicated as an activator, but definitive evidence for this has not been found (Vincent et al., 2018). The model for *eve* stripe 3/7 is thus ubiquitous activation that is refined to stripes 3 and 7 by gap gene-mediated repression elsewhere.

Interestingly, recent studies have re-evaluated the Hb-mediated regulation of the *eve* stripe 2 CRE (Vincent et al., 2018). In doing so, it was determined that the original proposed activation by Hb is actually mediated by a complicated counter-repression of Hb. Between the contradictory roles of Hb in *eve* stripe-specific regulation and the inability to identify a specific transcriptional activator for the *eve* stripe 3/7 CRE, the generalized logic presented by the *eve* stripe 2 model starts to break down. In other words, specific concentrations of non-periodic regulatory gradients of the maternal effect and gap gene products fail to explain the periodic expression of all pair-rule registers. This point has been previously discussed in some of the *eve* CRE publications summarized above (Small et al., 1996; Vincent et al., 2018).

Other studies around the same time of the *eve* stripe 2 work identified stripe-specific CREs for the PRGs *hairy* and *runt* (Howard et al., 1988; Howard & Struhl, 1990; Klingler et al., 1996; Pankratz et al., 1990; Riddihough & Ish-Horowicz, 1991). Through analysis of PRG expression in PRG mutants and gap gene mutants, genetic interactions were identified (Carroll & Vavra, 1989; Carroll & Scott, 1986; Frasch & Levine, 1987; Howard & Ingham, 1986; Ingham & Gergen, 1988). This determined that the expression of *ftz*, *odd*, *paired*, and *sloppy-paired* was perturbed in *hairy*, *runt*, and *eve* mutants. The inverse was not observed; *hairy*, *runt*, and *eve* were epistatic to the other PRGs. This led to the characterization of PRGs into two groups on the basis of genetic interaction and

presence of stripe-specific CREs. Primary PRGs received direct input from the maternal effect and gap gene products, whereas secondary PRGs received input from the primary PRGs. It was believed that once Hairy, Runt, and Eve stripes were established first, the seven stripes of the secondary PRGs would be established simultaneously. The stripes would be positioned by repression in the interstripe registers by Hairy, Runt, and Eve. However, advances in detection methods revealed that this model does not justify the subdivision of PRGs into primary and secondary classes.

When the first chromogenic *in situ* hybridization experiments were conducted to detect *ftz* with digoxygenin-labeled probes, it was found that *ftz* striped expression occurred in a dynamic, sequential order (Yu & Pick, 1995). Due to the timing of these stripes compared to that of the primary PRGs, repression in the interstripes could not rationalize the sequential order of *ftz* stripe evolution. Further, it was found that the full set of *ftz* stripes were properly expressed in *hairy*, *eve*, and *runt* mutants during the blastoderm stage. From this study, it was concluded that *ftz* stripe establishment was independent of the primary PRGs. Later evidence suggested that there were genomic fragments downstream of the *ftz* transcription start site that could direct transgene expression in the same domains as *ftz* stripes 1 and 5 (Alonso and Pick, unpublished; Calhoun & Levine, 2003). It is noteworthy that the first two endogenous *ftz* stripes that become detectable using the chromogenic *in situ* hybridization method are also stripes 1 and 5.

This finding suggested that there could be other genomic fragments yet to be discovered that were sufficient to direct stripe-specific transgene expression. On the assumption that the stripe-specific CREs of the primary PRGs receive input from

maternal effect and gap gene proteins, a genome-wide computational analysis was conducted to search for genomic fragments that were enriched for clusters of binding sites for these factors (Schroeder et al., 2011). These Stubb runs were conducted using position weight matrices of consensus binding sites. The genomic regions identified from the analysis were isolated by PCR and fused into transgenic reporter constructs. Expression analysis revealed that these genomic fragments that were sufficient for directing expression in PRG-like domains. Taken together, individual, novel CREs could direct transgene expression in the same domains as *ftz* stripes 2+7, *ftz* stripes 3+6, *runt* stripes 2+7, *runt* stripe 3, *runt* stripe 4, *runt* stripe 7, *slp* stripe 1, *odd* stripes 3+6, and *odd* stripes 1+5. Focusing on *ftz* in particular, this meant there were stripe-specific elements identified for every stripe except stripe 4. Taking together the dynamic order of *ftz* stripe activation, the existence of these stripe-specific CREs, and the independence of *ftz* stripe establishment from the primary PRGs, it became clear that the PRG hierarchy is more complex than a dichotomy.

Here, I present a careful analysis of the *ftz* stripe-specific CREs to better discern how they contribute to the spatiotemporal expression of this developmental gene. To this end, I have conducted a systematic gene expression analysis in a uniform context to observe the sufficiency of these genomic fragments to direct transgene expression. This was achieved by controlling the specific molecular variables of the reporter transgenes. All examined CREs were inserted in both orientations into the same restriction site directly adjacent to the promoter. Additionally, position effects were normalized by inserting all transgenes into the same genomic location. From these analyses, I have determined that all of the tested stripe-specific CREs of *ftz* reflect a degree of orientation

dependence, and there is likely a tethering element directly downstream of *ftz*. Further, I targeted a pair of these stripe 2 CREs for precise deletion with the CRISPR/Cas9 system to generate the *ftzΔ413* and *ftz413Δftz-6* alleles. Both deletion mutants are viable, fertile, and resemble wild type throughout development. Overall, my results suggest a high degree of redundancy between stripe-specific and seven-stripe CREs for *ftz* gene.

4.3 Materials and Methods

4.3.1 Transgenic plasmids

Generation of genomic fragments, screening of positive bacterial transformants, and sequencing of constructs was done as described in Chapters 2 and 3. Appendix A contains the sequences for all primers used in this work. The consensus sequences for each fragment and each probe sequences are specified in Appendix B.

The ftz-7 CRE is between -8222 to -6669 relative to the *ftz* TSS and is 1554 bp. *ftz-7-F:lacZ* was generated by amplifying with the primers MF41 and MF42. *ftz-7-R:lacZ* was generated by amplifying with the primers MF43 and MF44. Both purified fragments were assembled into XbaI linearized *placZ-attB* with HiFi assembly. Colony PCR for both orientations was conducted with the primers SeqF and SeqR. Both constructs were fully sequenced with the primers SeqF, MF40, and SeqR. MIDI prep DNA for microinjection was verified by two separate digestion reactions; EcoRI-HF and XbaI.

The ftz-6 CRE is between -6674 to -5435 and is 1240 bp. *ftz-6-F:lacZ* and *ftz-6-R:lacZ* were generated by amplifying with the primers MF3 and MF4. The fragment was inserted into *placZ-attB* with Pyrite cloning using XbaI. Colony PCR for *ftz-6-F:lacZ* was conducted with the primers SeqF and SeqR. Colony PCR for *ftz-6-R:lacZ* was conducted

with the primers SeqF and MF25. Both constructs were fully sequenced with the primers SeqF and SeqR. MIDI prep DNA for microinjection was verified by digesting with KpnI-HF.

The ftz413 CRE is between -6418 to -6006 and is 413 bp. *ftz413-F:lacZ* and *ftz413-R:lacZ* were generated by amplifying with primers MF1 and MF2. The fragment was inserted into *placZ-attB* with Pyrite cloning using XbaI. Colony PCR for *ftz413-F:lacZ* was conducted with the primers SeqF and SeqR. Colony PCR for *ftz413-R:lacZ* was conducted with the primers SeqR and MF2. Both constructs were fully sequenced with the primers SeqF and SeqR. MIDI prep DNA for microinjection was verified by digesting with EcoRI.

The ftz3' element (ftz3'E) is between +1876 to +4008 and is 2133 bp. *ftz3'E-F:lacZ* was generated by amplifying with the primers MF45 and MF46. *ftz3'E-R:lacZ* was generated by amplifying with the primers MF47 + MF48. Both fragments were inserted into XbaI linearized *placZ-attB* with HiFi assembly. Colony PCR for both was conducted with the primers SeqF and SeqR. Both constructs were fully sequenced with the primers SeqF, MF125, MF127, and SeqR. MIDI prep DNA for microinjection was verified for orientation with two different enzyme digestions; EcoRI-HF with SphI-HF and XbaI.

The ftz+3 CRE is between +2571 to +4333 and is 1763 bp. *ftz+3-F:lacZ* was generated by amplifying with the primers MF63 and MF64. *ftz+3-R:lacZ* was generated by amplifying with the primers MF65 and MF66. Both fragments were inserted into XbaI linearized *placZ-attB* with HiFi assembly. Colony PCR for both was conducted with the

primers SeqF and SeqR. MIDI prep DNA for microinjection was verified for orientation with KpnI-HF.

As in previous chapters, transgenic constructs were integrated into the phiC31 docking site at Chromosome 2, 57F5, 2R:21645971 following embryonic microinjection into Bloomington Stock Center line 9740 by Rainbow Transgenic Fly.

4.3.2 Guide RNA and HDR templates

Genomic targets for gRNAs were identified as described in Chapters 2 and 3. Two gRNAs were selected to target the ftz413 region. The left guide targeted the antisense strand from -6393 to -6412 to induce a double-stranded break 10 bp from the 5' end. The right guide targeted from -6059 to -6040 to induce a double-stranded break 36 bp from the 3' end. The *pCFD5* vector was used as a DNA template for PCR with the primers MF138 and MF139 in one reaction and MF140 and MF110 in another. The MF140+110 fragment was further amplified by MF140 and MF112. These two fragments were assembled into the *pCFD5* vector linearized by BbsI-HF using HiFi assembly to create *pCFD5-ftzΔ413*. Colony PCR was conducted with the primers CFDN-Seq and MF67. Sequencing of the isolated *pCFD5-413* plasmid was conducted with the primer CFDN-Seq. The MIDI prep DNA used for microinjection was verified by digestion with XbaI and EcoRI-HF.

The HDR template for the precise deletion of ftz413 (*pUC19-ftzΔ413*) was assembled from three PCR-amplified fragments. The left homology arm spans from -7792 to -6420 and was amplified by MF141 and MF40 using 0.625% DMSO and *pUC19-ΔUPS* as the DNA template to form a 1432 bp fragment. The right homology arm spans from -6006 to -3846, is 2153 bp, and was fused from two separate PCR reactions

using *UPS^{0}-F:lacZ* as the DNA template. The first reaction used the primers MF142 and MF143 with an adjusted extension temperature of 68°C and amplified a 572 bp fragment. The second reaction used the primers MF25 and MF144 with 0.625% DMSO and amplified a 1659 bp fragment. All three PCR fragments were assembled into *pUC19* linearized with KpnI with HiFi assembly to generate *pUC19-ftzΔ413*. Colony PCR was conducted with the primers Dm_UES_2F and MF143. *pUC19-ftzΔ413* was fully sequenced with the primers M13F, DmFtzUESSeq1F, MF142, MF25, DmUES3F, and M13R. The MIDI prep DNA for microinjection was verified by digestion with EcoRI-HF.

Two gRNAs were selected to target the ftz-6 region. The left guide targeted the antisense strand from -6651 to -6670 to induce a double-stranded break 8 bp from the 5' end. The right guide targeted the antisense strand from -5442 to -5461 to induce a double-stranded break 23 bp from the 3' end. The *pCFD5* vector was used as a DNA template for PCR with the primers MF145 and MF146 in one reaction and MF147 and MF110 in another. The MF147+110 product was further amplified by MF147 and MF112. These two fragments were assembled into the *pCFD5* vector linearized with BbsI-HF using HiFi assembly to create *pCFD5-Δftz-6*. Colony PCR was conducted with the primers CFDN-Seq and MF67. Sequencing of *pCFD5-Δftz-6* was conducted with the primer CFDN-Seq. The MIDI prep DNA for microinjection was verified by digestion with XbaI and EcoRI-HF.

The HDR template for the precise replacement of the full ftz-6 fragment with only the ftz413 sequence (*pUC19-ftz413Δftz-6*) was assembled from three PCR-amplified fragments; a left homology arm, the ftz413 fragment, and a right homology arm. The left

homology arm spans from -7792 to -6676, is 1117 bp, and was amplified by the primers MF40 and MF94 with 0.625% DMSO using the *pUC19-ΔUPS* construct as a DNA template. The ftz413 fragment was amplified by the primers MF95 and MF96 with 0.625% DMSO using *UPS^{0}-F:lacZ* as a DNA template. The right homology arm spans from -5435 to -3850, is 1586 bp, and was amplified by the primers MF97 and MF144 with 0.625% DMSO using *UPS^{0}-F:lacZ* as a DNA template. All three fragments were assembled into the *pUC19* vector linearized by KpnI-HF using HiFi assembly. Colony PCR was conducted with the primers MF21 and MF96. The construct was fully sequenced with the primers M13F, DmFtzUESSeq1F, MF24, MF97, DmUES3R, and M13R. The MIDI prep DNA for microinjection was verified by digestion with EcoRI-HF.

4.3.3 *Drosophila* genetics and phenotypic analysis

Drosophila lines and methods for identifying CRISPR/Cas9 events were the same as described in Chapter 2. Deletion of ftz413 to generate the *ftzΔ413* allele was screened by PCR with the primers MF143 and Dm_UES_2F, which would amplify a 786 bp fragment from precise deletion events and 1198 bp from wild type. Out of 32 fertile G0 lines, 12 produced progeny that tested positive for the precise deletion (37.5%). Precise replacement of the ftz-6 region by ftz413 to generate *ftz413Δftz-6* was screened by PCR with the primers MF21 and MF149, which would amplify 149 bp from HDR events and 405 bp from wild type. Out of 24 fertile G0 lines, 5 produced positive progeny (20.8%).

Adult segmentation was counted and recorded as described in Chapters 2 and 3. Larval cuticles were processed, scored, and imaged as in Chapters 2 and 3. Chromogenic *in situ* reactions for *ftz*, *lacZ*, *en*, *Slp1*, and *Scr* were conducted as described in Chapters 2

and 3. Fluorescent *in situ* HCR was conducted as described in Chapter 3 with the *ftz*, *eve*, and *lacZ* probes.

4.3.4 Analysis of transcription factor binding sites in ftz413 and ftz-6 CREs

The location of transcription factor binding sites shown to regulate the expression of the *eve* stripe 2 and *eve* stripe 3/7 domains were searched for in the ftz413 and ftz-6 CREs DNA sequences using RSAT (Medina-Rivera et al., 2015; Nguyen et al., 2018; Small et al., 1992). Consensus binding site matrices for Bicoid, Hunchback, Giant, and Kruppel were input from the Database of Drosophila DNA-binding Specificities. *Drosophila melanogaster* BDGP 6.34 was used to calculate the background model estimation with sequence type set to upstream-noorf and a Markov order of 1. The threshold for weight was set to values greater than 4.65 in order to best match the output of the *eve* stripe 2 search to that of the original publication (Small et al., 1992) to facilitate comparisons to ftz413 and ftz-6.

4.4 Results

4.4.1 A systematic approach to transgenesis facilitated comparison between *ftz* CRE reporter transgenes

Prior research has identified six stripe-specific CREs of *ftz* (Fig. 4-1). While two of these are based on fragments within the genomic region of the original *ftz* rescue construct (Alonso & Pick, unpublished; Pick et al., 1990), most were identified through a computational analysis of genomic regions in which clusters of maternal and gap gene factor binding sites were identified (Schroeder et al., 2011). The five stripe-specific CREs of *ftz* reported from these studies are: 1) ftz-7, located approximately 7 kb upstream of the

ftz transcription start site (TSS) (Schroeder et al., 2011). 2) ftz-6, located approximately 6 kb upstream (Schroeder et al., 2011). 3) ftz413, consisting of 413 bp within the ftz-6 sequence and bound by the upstream KpnI restriction enzyme site (Pick et al., 1990). 4) the 3' element (ftz3'E) that starts near the end of the *ftz* 3' UTR and extends approximately 2.1kb downstream and is similarly bound by a KpnI restriction site (Alonso & Pick, unpublished). 5) ftz+3, located approximately 3 kb downstream of the TSS. The ftz3'E and ftz+3 genomic regions overlap in the middle but have distinct termini. One additional downstream CRE, ftzDE, has been previously studied (Calhoun & Levine, 2003). This CRE is located approximately 3.5 kb downstream of the *ftz* TSS and is 1.25 kb. Taken together, the identification of these elements has furthered the understanding of *ftz* transcriptional regulation.

However, these studies were conducted in different contexts that confound comparisons between the results. Specifically, the CREs were inserted into different reporter transgene expression plasmids and randomly inserted into the genome by P-element mutagenesis. It has been reported that bacterial plasmid fragments between a CRE and a promoter can confound the expression patterns of reporter transgenes (reviewed in Schaffner, 2015). Since these previous studies used different restriction enzymes to insert the studied CREs into reporter plasmids, there were spacers of variable length and sequence between the CRE and the promoter-reporter gene fusion. Additionally, transgenes randomly inserted by P-element mutagenesis are variably subject to position effects from endogenous CREs around the insertion site. Finally, it is unknown if all of these prior studies examined the stripe-specific CREs in both orientations. Though transcriptional enhancers are classically defined by the ability to

regulate gene expression regardless of distance and orientation, there are inhibitory CREs such as insulators or tethering elements that override this property (Banerji et al., 1981).

The *ftz* locus is a complicated genomic region that contains several *ftz*-specific CREs and at least two *Scr*-specific CREs (Calhoun & Levine, 2003; Calhoun et al., 2002; Gindhart et al., 1995). Through analyzing the qualitative differences of orientation-dependent transgene expression patterns, putative insulators or tethering elements can be identified. Out of the *ftz* stripe-specific CREs, explicit orientation-dependent analysis has only been conducted on ftz413 and ftz3'E, from which it was observed that only one orientation of the CRE could transmit regulatory cues (Pick et al., 1990; Pick & Alonso, unpublished). These results led to the conclusion that inhibitory CREs are located around the KpnI restriction sites that surround *ftz*. In sum, inconsistencies among these prior studies in regard to these molecular variables confound the comparison between results.

Here, I have normalized these molecular variables in order to conduct a systematic analysis by which the stripe-specific regulation of *ftz* could be better discerned and compared (Materials and Methods). To verify prior conclusions on the expression patterns of these CREs, genomic regions corresponding to each of the elements were amplified by PCR and inserted into the XbaI site of the reporter plasmid *placZ-attB*. This restriction site is directly upstream of a basal *hsp70* promoter and *lacZ* reporter. In doing so, the only plasmid sequence between the CRE and *hsp70* promoter was the XbaI site. To determine orientation-dependence, each fragment was separately inserted into the same site in forward and reverse orientations. Each reporter transgene was then separately integrated into the same genomic phiC31 docking site at Chromosome 2, 57F5, 2R:21645971. Since all transgenes were inserted into the same location, the

position effect from this genomic site was equalized for all transgenic lines. Thus, this approach was molecularly controlled at the level of the reporter plasmid and genomic location. These parameters facilitated the systematic comparison between transgenic lines.

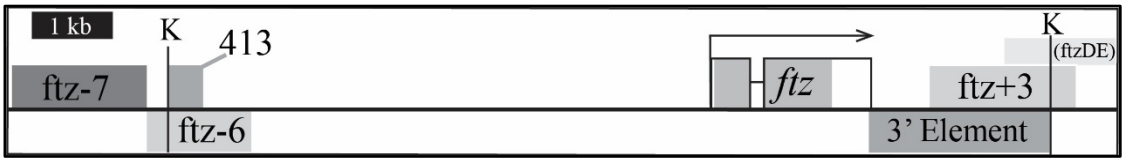


Figure 4-1. Organization of the stripe-specific *ftz* CREs analyzed in this work. Schematic drawn to scale displaying the relative location of each stripe-specific element within a 12.6 kb genomic region. Black arrow indicates the *ftz* transcription unit, boxes represent exons with coding sequence in gray and UTRs in white; black line represents the one intron; other grayscale boxes outside *ftz* represent the five genomic fragments tested. Vertical lines labeled with K represent the KpnI sites surrounding *ftz*. Scale bar represents distance of one kilobase. Five stripe-specific CREs have been identified in the genomic region of *ftz*: ftz-7 is located approximately 7 kb upstream and is 1.6 kb. ftz-6 is located approximately 6 kb upstream and is 1.2 kb. ftz413 consists of 413 bp, has the upstream KpnI site on the 5' end, and it starts 256 bp into the ftz-6 fragment (shown as 413). The 3' element (ftz3'E) is approximately 2 kb downstream and is 2 kb. ftz+3 is approximately 3 kb downstream and is 2 kb. Ftz3'E and ftz+3 overlap for 1.4 kb, and the overlap starts 700 bp into the ftz3'E. ftz+3 continues for 325 bp after the ftz3'E. The furthest downstream CRE, ftzDE, is not examined in this study. The ftzDE CRE is located approximately 3.5 kb downstream and is 1.25 kb.

4.4.2 Stripe-specific CREs upstream of the *ftz* transcription unit display orientation-dependence

The expression of *lacZ* was detected from transgenic embryos using a digoxygenin-labeled probe for chromogenic *in situ* hybridization reactions. The results for the upstream stripe-specific CREs in both orientations are shown in Fig. 4-2. Expression of the *ftz-7F:lacZ* transgene was detected in one strong stripe and another

broad band (Fig. 4-2A), presumably corresponding to *ftz* stripe 3 and a band over stripes 6 and 7. For the reverse orientation transgene *ftz-7R:lacZ,* expression was detected at relatively lower levels distinctly in the stripe 3, 6, and 7 domains (Fig. 4-2B). Expression over the stripe 7 domain was barely detectable from the reverse orientation. In comparing the orientations, stripes 6 and 7 were distinct in the reverse orientation but appear fused in the forward orientation. Further, the expression of *ftz-7F:lacZ* more closely resembles the reported expression pattern of this CRE compared to that of *ftz-7R:lacZ* (Schroeder et al., 2011).

Expression of the *ftz-6F:lacZ* transgene was detected faintly in stripe 2 (Fig. 4-2C). In the reverse orientation, stronger expression was observed in stripes 2 and 7 (Fig. 4-2D). The expression in stripe 2 appeared slightly stronger than in the forward orientation, and stripe 7 was very faint. Weak ectopic expression was detectable in the head region, which is not a native *ftz* expression domain (asterisk; Fig. 4-2D). The expression of *ftz6R:lacZ* more closely resembled the published expression pattern for this CRE compared to that of *ftz-6F:lacZ* (Schroeder et al., 2011). Strong expression of the *ftz413F:lacZ* transgene was detected in stripe 2 (Fig. 4-2E). In the reverse orientation, expression in stripe 2 was substantially stronger and extended over the surrounding parasegments, and a faint band was detected in the posterior of the trunk around the stripe 7 domain (Fig. 4-2F).

Previous studies on ftz413 reporter transgenes suggested that this fragment can only function as an enhancer when fused in the reverse orientation (Pick et al., 1990). The method to monitor *lacZ* expression with β-galactosidase activity used in that study is less sensitive than the chromogenic *in situ* reaction used here. This lower sensitivity

combined with the stark difference of expression levels comparing *ftz413F:lacZ* to *ftz413R:lacZ* likely explains the prior conclusion that ftz413 only functions as a stripe-specific enhancer in the reverse orientation (Pick et al., 1990) (see Discussion). The ftz413 CRE is located within ftz-6 in the genome, with ftz-6 beginning 256 base pairs 5' of ftz413 and extending 571 base pairs 3' of ftz413 (Fig. 4-1). Simply put, ftz-6 is a relatively larger genomic region that extends on both sides of the ftz413 fragment.

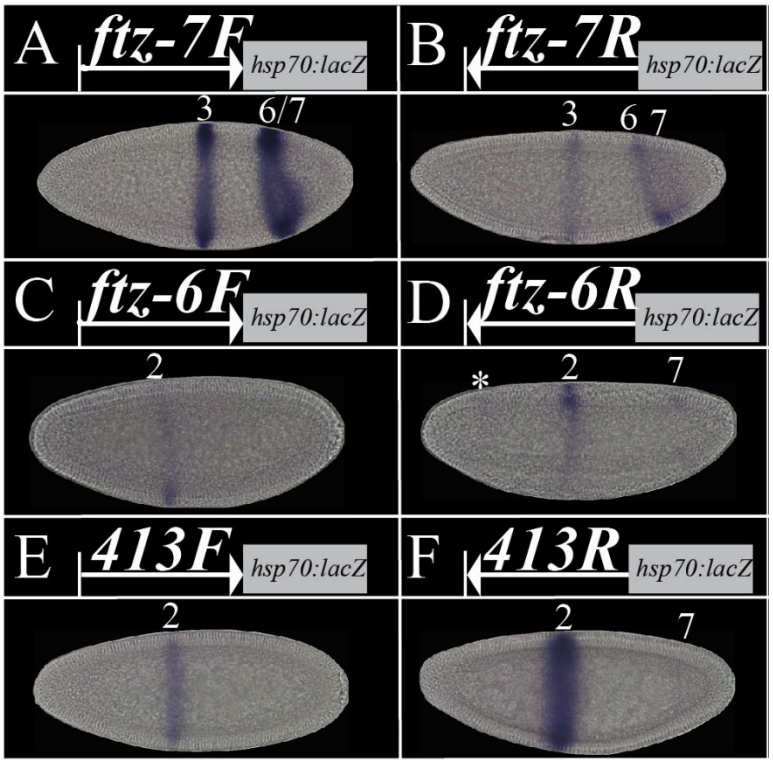


Figure 4-2. Reporter transgenes of upstream stripe-specific *ftz* CREs reflect orientation dependence. Brightfield images of *lacZ* expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization in embryos containing transgenes of the three upstream stripe-specific CREs. Numbers correspond to the presumed *ftz* stripe domains. (A, C, E) The left column shows transgenic embryos with the CRE fused in the forward orientation. (B, D, F) The right column shows transgenic embryos with the CRE fused in the reverse orientation. (A) Strong *lacZ* expression is detected over stripes 3, 6, and 7 of *ftz-7F:lacZ* embryos. Stripes 6 and 7 appear fused (6/7). (B) Clear *lacZ* expression is detected in stripe 6 and faintly over stripes 3 and 7 of *ftz-7R:lacZ* embryos. (C) Weak *lacZ* expression is detected over stripe 2 of

ftz-6F:lacZ embryos. (D) *lacZ* expression is observed over stripes 2 and 7 in *ftz-6R:lacZ* embryos. The stripe 7 expression is weak. A band of expression is detected in the head region, which is not an endogenous *ftz* expression domain (asterisk). (E) Strong expression of *lacZ* is observed over stripe 2 of *ftz413F:lacZ* embryos. (F) Very strong expression is observed over stripe 2 and the surrounding parasegments in *ftz413R:lacZ* embryos. A light band is detected around stripe 7.

4.4.3 One CRE downstream of the *ftz* transcription unit is strongly orientation-dependent

There are two stripe-specific CREs downstream of *ftz*: ftz3'E and ftz+3 (Fig. 4-1). The ftz3'E fragment lies within the original 10 kb rescue fragment. In the forward orientation (*3'E-F:lacZ*), faint transgene expression was detected in stripe 1 and strong expression was observed in stripe 5 (Fig. 4-3A). In the reverse orientation, an exceptionally weak signal was detected in stripe 5, and only when the embryo was slightly overstained (Fig. 4-3B). This is in keeping with unpublished data reporting that transgenes of the ftz3'E fragment were only detected in the forward orientation using the same reasoning given for the ftz413 orientation-dependent effects described above (Alonso, Pick, unpublished). The ftz3'E region shares 1.4kb of genomic sequence with the ftz+3 CREs that is located further downstream from *ftz*. This overlap starts 700 bp from the 5' end of the ftz3'E region, and ftz+3 extends 325 bp past the 3' end of ftz3'E (Schroeder et al., 2011) (Fig. 4-1). Strong expression of *ftz+3F:lacZ* was detected in the stripe 1 and 5 domains (Fig. 4-3C). This is similar to the expression of *ftz+3R:lacZ*, which was detected at weaker levels in the same pattern (Fig. 4-3D). The expression of *ftz+3F:lacZ* more closely resembled the reported expression pattern of this CRE (Schroeder et al., 2011). Though stronger signal was detected from forward fusions of both downstream CREs, *ftz+3F:lacZ* is stronger than *ftz3'E-F:lacZ*. Whereas there is little signal above background detected from the *ftz3'E-R:lacZ* transgenic line, stripes 1

and 5 are clearly detected from *ftz+3R:lacZ*. No signal was observed after gastrulation for any of these stripe-specific transgenes (data not shown).

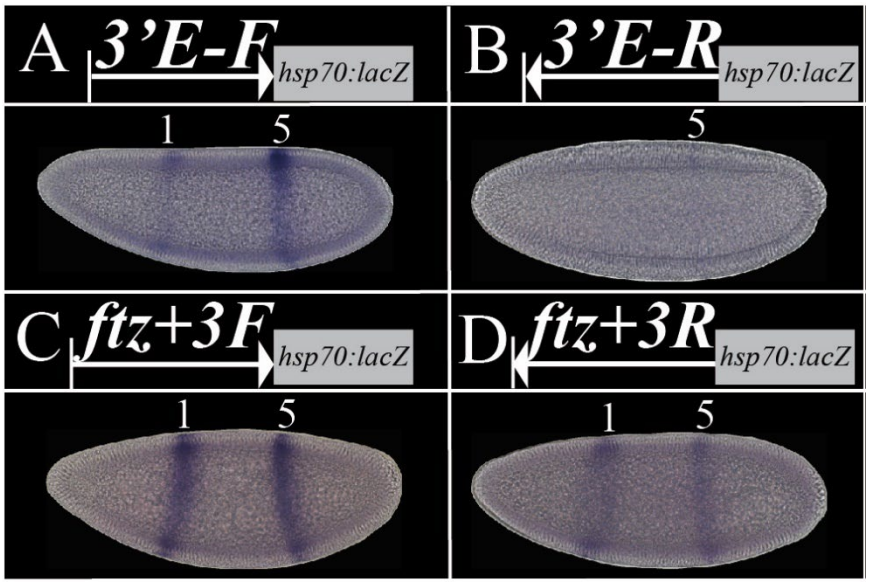


Figure 4-3. Downstream stripe-specific *ftz* CREs are strongly impacted by orientation. Brightfield images of *lacZ* expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization in embryos containing transgenes of the two downstream stripe-specific CREs. Numbers correspond to the presumed *ftz* stripe domain. (A, C) The left column shows the transgenic embryos with the CRE fused in the forward orientation. (B, D) The right column shows the transgenic embryos with the CRE fused in the reverse orientation. (A) Strong *lacZ* expression was detected over stripe 5 in *3'E-F:lacZ* embryos. Weak expression was detected over stripe 1. (B) Low expression of *lacZ* was detected near background levels in *3'E-R:lacZ* embryos. (C) Strong *lacZ* expression was detected over stripes 1 and 5 in *ftz+3F:lacZ* embryos. (D) *lacZ* expression was detected over stripes 1 and 5 in *ftz+3R:lacZ* embryos. Expression from forward CRE fusions was detected at higher levels than reverse CRE fusions for both downstream CREs.

4.4.4 Summary of *ftz* stripe-specific CRE transgene expression

In sum, a 1.6kb genomic fragment approximately seven kilobases upstream of the *ftz* TSS directed *lacZ* expression over stripes 3, 6, and 7 (ftz-7*;* Fig. 4-2A-B). The expression over stripes 6 and 7 appeared fused through the interstripe region. A 1.2kb

fragment approximately six kilobases upstream of the TSS directed *lacZ* expression in stripes 2 and 7 (ftz-6; Fig. 4-2C-D). A genomic fragment contained within the ftz-6 region that is 413 bases directed strong *lacZ* expression in stripe 2 and the surrounding parasegments and weak *lac* expression in stripe 7 (ftz413; Fig. 4-2E-F). On the other side of the *ftz* transcription unit, two fragments of about 2kb were tested. The one immediately downstream of the *ftz* transcription unit and 1.9kb from the TSS directed *lacZ* expression in stripes 1 and 5 that was apparent only when fused in the forward orientation (ftz3'E; Fig. 4-3A-B). The final fragment approximately 3 kb downstream of the *ftz* TSS, which overlaps with ftz3'E, directed *lacZ* expression in stripes 1 + 5 (ftz+3; Fig. 4-3C-D).

A qualitative difference in *lacZ* expression was detected from all five CREs with respect to orientation. This was most striking for the ftz3'E transgenes, for the clear expression pattern that was observed from the forward orientation transgene was near background levels from the reverse orientation transgene (Fig. 4-3A-B). Orientation-dependence was also observed for ftz413; a much stronger signal was found for the reverse orientation fusion than for the forward orientation. Further, an additional stripe (stripe 7) was detected for the reverse orientation, while only stripe 2 was detected from the forward orientation transgene (Fig. 4-2E-F).

4.4.5 Reporter transgenes of *ftz* stripe-specific CREs correspond to endogenous *ftz* stripes

Although these stripe-specific elements were thought to recapitulate the expression of endogenous *ftz* stripes, this has not been explicitly demonstrated. To this end I conducted multiplexed fluorescent *in situ* HCR experiments to compare the expression of the *lacZ* transgenes to the endogenous *ftz* and *eve* striped expression domains. The chromogenic *in*

situ experiments above revealed that the clearest expression signals were detected from one orientation of each stripe-specific CRE. As shown in Fig. 4-4A-B, this is consistent with the HCR results for the *ftz413:lacZ* transgenes. While the forward orientation directed expression of a single stripe over *ftz* stripe 2 with slight overlap to the surrounding *eve* stripes (Fig. 4-4A, purple stripe and number), the reverse orientation directed expression of a stronger and wider stripe that more fully encompasses *eve* stripe 2, *ftz* stripe 2, and *eve* stripe 3 (Fig. 4-4B, purple stripe, number above for *ftz* stripes, red numbers below for *eve* stripes). The expression of *lacZ* was only detected throughout the cellular blastoderm stage and was undetectable at the start of gastrulation for both orientations of the *413:lacZ* transgenes. While expression of *lacZ* was restricted to the dorsal side of the embryo at the start of cellularization, it was detected as a full band around the blastoderm by the end of cellularization (data not shown). From these experiments, it appears that a specific orientation of the CREs in these *lacZ* transgenes can provide a clearer determination in which *ftz* (and *eve*) stripe domains they are expressed. Therefore, only the orientation of the transgene from which a stronger signal was detected from the chromogenic *in situ* experiments was assayed by HCR.

As shown in Fig. 4-4C, *ftz-6R:lacZ* was expressed in a strong stripe overlapping *ftz* stripe 2 and a very weak stripe overlapping *ftz* stripe 7 that was near background levels. Unlike *ftz413-R:lacZ*, stripe 2 expression was confined to the endogenous *ftz* stripe region, and signal was undetectable over *eve* stripes 2 and 3. Ectopic expression was observed as a single, weak band in the head region (asterisk). Expression of *ftz-6R:lacZ* was undetectable at the start of the cellular blastoderm stage, as determined by the degree of cell membrane deposition. Expression over the *ftz* stripe 2 domain became detectable

as a band around the circumference of the blastoderm when a little less than half of the total membrane was developed. Expression over stripe 7 became detectable near the end of cellularization.

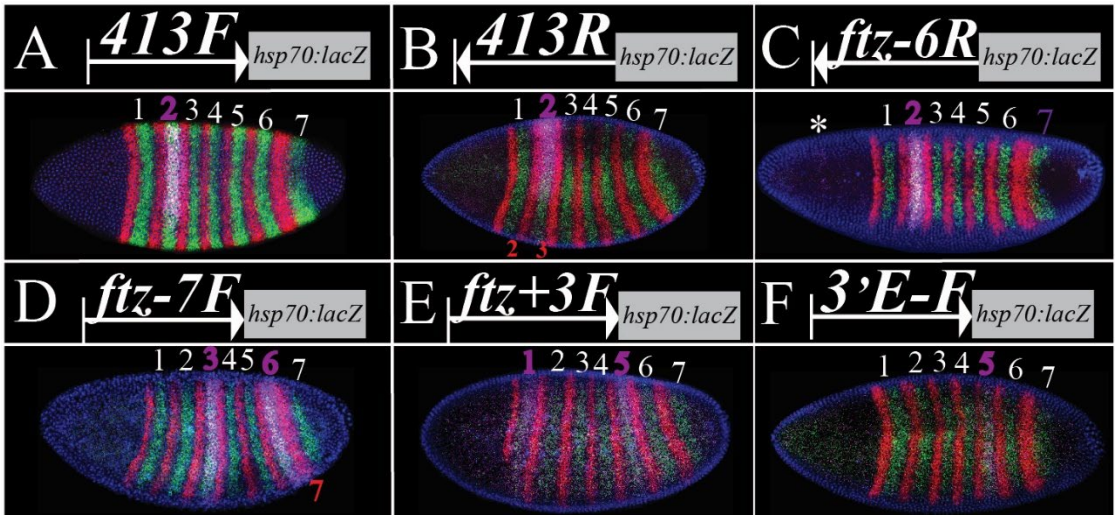


Figure 4-4. Stripe-specific *ftz* CREs direct *lacZ* expression over *ftz* stripe domains. Fluorescent confocal images of *eve*, *ftz*, and *lacZ* expression detected by *in situ* HCR in indicates transgenic reporter lines. Channels indicate the following probes: *eve*, red; *ftz*, green; *lacZ*, purple. Histogram adjusted in ImageJ to enhance qualitative analysis of co-expression domains. White numbers above embryos indicate *ftz* stripes, purple numbers indicate *lacZ* co-expression. Red numbers under embryos indicate *eve* stripe domains with detectable *lacZ* co-expression. (A) In *413F:lacZ* embryos, *lacZ* expression was detected over *ftz* stripe 2. (B) In *413R:lacZ* embryos, *lacZ* expression was detected in a wide stripe over *eve* stripe 2, *ftz* stripe 2, and *eve* stripe 3. (C) In *ftz-6R:lacZ* embryos, strong *lacZ* expression was detected over *ftz* stripe 2. *lacZ* expression was also detected near background levels over *ftz* stripe 7. A band of ectopic expression relative to endogenous *ftz* domains was detected in the head region (asterisk). (D) In *ftz-7F:lacZ* embryos, *lacZ* was detected in a strong stripe over *ftz* stripe 3 and a wide stripe over *ftz* stripe 6 and *eve* stripe 7. (E) In *ftz+3F:lacZ* embryos, *lacZ* was detected over *ftz* stripe 1 and 5. (F) *In ftz3'E-F:lacZ* embryos, *lacZ* was detected at very low levels over *ftz* stripe 5.

ftz-7F:lacZ expression was detected from the HCR *in situ* experiments as a strong stripe over *ftz* stripe 3 and a wide stripe over *ftz* stripe 6 (Fig. 4-4D, purple stripes and numbers). The wide stripe over *ftz* stripe 6 also encompassed *eve* stripe 7 (Fig. 4-4D, red number). Expression of *ftz-7F:lacZ* was undetectable at the start of blastoderm cellularization and became detectable halfway through stage 5. The broad stripe 6 expression resolved into a sharp stripe 6 by the end of cellularization. Signal was detected at the start of cephalic furrow formation but was not maintained through gastrulation. Unlike what was presumed from the colorimetric *in situ* experiments, no expression was detected overlapping *ftz* stripe 7. The presumption of expression over a combined stripe 6/7 domain may have been due to overdevelopment of the signal catalyzed by the chromogenic reaction.

ftz+3F:lacZ was expressed in two stripes overlapping *ftz* stripes 1 and 5 (Fig. 4-4E). These signals were detected at the start of blastoderm cellularization. The fluorescent intensity anticorrelated with the progression of membrane development, suggesting that the contribution of this CRE to the endogenous *ftz* pattern is limited to the start of the cellular blastoderm stage. Though the *lacZ* stripes were approximately the same width as the *ftz* stripes, they appeared to be slightly offset between the *ftz* and *eve* domains by one or two cells.

Weak expression of *ftz3E'-F:lacZ* was detected near background levels overlapping *ftz* stripe 5. This pattern was most clear around the middle of blastoderm cellularization, though at no point was signal detected at high levels. Unlike in the chromogenic *in situ* experiments, no expression was detected overlapping stripe 1. The fluorescent output of HCR is closely correlated to the number of target transcripts present (Choi et al., 2014).

While this does allow for accurate and precise observations of gene expression, it cannot amplify the development of signals over regions of weak expression. This is in contrast to the chromogenic reaction that will continuously develop color until quenched, allowing the observation of low abundance transcripts. These contrasts demonstrate that both HCR and chromogenic *in situ* hybridization experiments can both be valuable to elucidate different properties. In light of this, it is noteworthy that expression of *ftz+3F:lacZ* was markedly stronger than *ftz3'E-F:lacZ*.

This analysis validates previous conclusions – that these genomic fragments recapitulate endogenous *ftz* expression. Thus, stripe-specific elements have been confirmed for *ftz* stripes 1, 2, 3, 5, 6 and 7. Further, the detectable expression of stripe domains from these stripe-specific CRE reporter transgenes occurs in a dynamic order in respect to the progression of blastoderm cellularization. This is reflective of the dynamism by which the seven *ftz* stripes appear. The UPS and the zebra element are the only CREs identified to date that direct expression in *ftz* stripe 4 (Chapters 2 and 3).

4.4.6 Precise CRISPR/Cas9-induced deletions generated the *ftzΔ413* and *ftz413Δftz-6* alleles to test the roles of stripe-specific *ftz* CREs

The substantial difference in stripe 2 intensity when comparing the reporter transgenes of ftz413 and ftz-6 is curious in that ftz413 is embedded within ftz-6 (Fig. 4-1). Because ftz413 is fully contained within the ftz-6 region, I hypothesized that the signal from the ftz-6 transgenes would be stronger than those of ftz413. The greater size of the ftz-6 region could contain more binding sequences for transcriptional activators than the ftz413 region alone. However, the opposite is observed; whereas the ftz413 transgenes shows a strong signal, the ftz-6 transgenes appear to be expressed at relatively

lower levels in refined transcriptional domains. This suggests that the sequences that bind activators for substantial expression in stripe 2 lie within the ftz413 CRE. The surrounding ftz-6 regions likely bind repressors – a mechanism to restrict the expression of *ftz* within a single parasegment and thus prevent ectopic *ftz* expression in the adjacent parasegments.

To test this hypothesis and examine if deleting a strong stripe-specific enhancer causes a detectable difference to endogenous *ftz* expression, CRISPR/Cas9 was utilized to delete ftz413 and generate the *ftzΔ413* allele (Materials and Methods; Fig. 4-5). Briefly, guide RNAs were selected that targeted the termini of the ftz413 region (Fig. 4-5B, black arrowheads). These would direct double-stranded breaks by Cas9 within 10 base pairs of the left terminus and 40 base pairs of the right terminus. The repair template (*pUC19-Δ413)* consisted of homology arms 1.373 kb upstream of ftz413 and 2.158 kb downstream of ftz413 that could be incorporated by the HDR pathway to result in a precise deletion of the ftz413 CRE (Fig. 4-2C). F1 progeny of injected G0 individuals were screened for the deletion with PCR after crosses were established. The primers used in the experiment were MF143 and Dm_UES_2F, which amplify a 786 bp fragment from positive events and a 1198 bp fragment from negative events (Fig. 4-5C-D). Out of 32 fertile G0 lines, 12 produced progeny that tested positive for the precise deletion (37.5%). Specimen from positive lines were used for subsequent rounds of PCR and sent for sequencing to confirm precise removal of the element (Fig. 4-5E). Animals that are homozygous for this $ftz\Delta413^{-/-}$ allele are viable, fertile, progress through all stages of development, and have been maintained as homozygous lines.

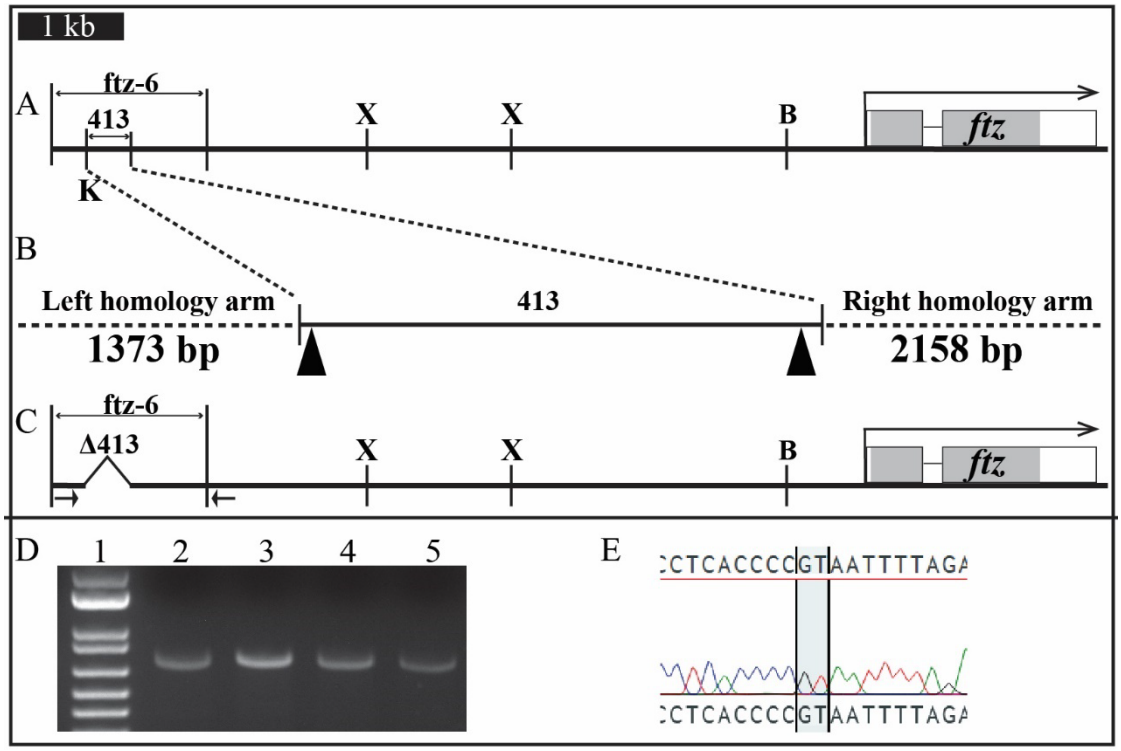


Figure 4-5. Precise deletion of the ftz413 CRE using CRISPR/Cas9 with an HDR template to generate the *ftzΔ413* allele. (A) Schematic of the genomic region upstream relative to the *ftz* transcription unit (*ftz*, black arrow represents transcription from basal promoter, boxes represent exons with CDS in gray, UTR in white, black line represents the one intron). Vertical line labeled with K represent the KpnI site upstream of *ftz*. X stands for XbaI sites. B stands for BalI site. Drawn to scale (scale bar indicates 1 kilobase). Genomic regions of ftz413 and ftz-6 are labeled. (B) Zoom in of the ftz413 fragment with black arrowheads indicating the CRISPR/Cas9 gRNA target sites. Homology arms of the repair template with base length are indicated on either side. (C) Schematic of the *ftzΔ413* allele resulting from the incorporation of *pUC19-ftzΔ413* following CRISPR/Cas9-induced double-stranded breaks. Arrows represent the location of the MF143 and Dm_UES_2F primer binding sites. (D) PCR screen used to identify deletion events with the primers MF143 and Dm_UES_2F to amplify a 786 bp fragment in the absence of the ftz413 region (lanes 2-5). Lane 1: GeneRuler 1kb plus DNA ladder. (E) Sequencing result from a mutant identified from PCR screen. Highlight indicates the junction between the left and right homology arms of the HDR template, indicating a precise deletion.

To test the hypothesis that *ftz* expression in stripe 2 by the ftz-6 CRE is contributed mostly by the nested ftz413 fragment, a similar strategy was used to delete the regions unique to the ftz-6 CRE while maintaining the ftz413 region (Fig. 4-6). Guide RNAs were selected that targeted the termini of the ftz-6 region (Fig. 4-6B, black arrowheads). These would direct double-stranded breaks by Cas9 within 10 base pairs of the left terminus and 20 base pairs of the right terminus. The repair template (*pUC19-ftz413Δftz-6*) consisted of the ftz413 fragment surrounded by homology arms that went 1.117 kb upstream of ftz-6 and 1.586 kb downstream of ftz-6 (Fig. 4-6C). In this way, the whole ftz-6 region could be deleted and the ftz413 CRE then re-incorporated through the HDR pathway using the *pUC19-ftz413Δftz-6* template. F1 progeny of injected G0 flies were screened for the deletion and subsequent HDR incorporation by PCR after crosses were established. The primers used in the experiment were MF21 and MF149, which amplify a 149 bp product from HDR events and 405 bp from negative events (Fig. 4-6C-D). Out of 24 fertile G0 lines, 5 produced positive progeny (20.8%). Specimen from positive lines were used for subsequent rounds of PCR and sent for sequencing to confirm the precise incorporation of the repair template at the left and right junctions (Fig. 4-6E-F, respectively; highlights indicate junctions). Animals that are homozygous for this *ftz413Δftz-6*$^{-/-}$ allele are viable, fertile, progress through all stages of development, and have been maintained as homozygous lines.

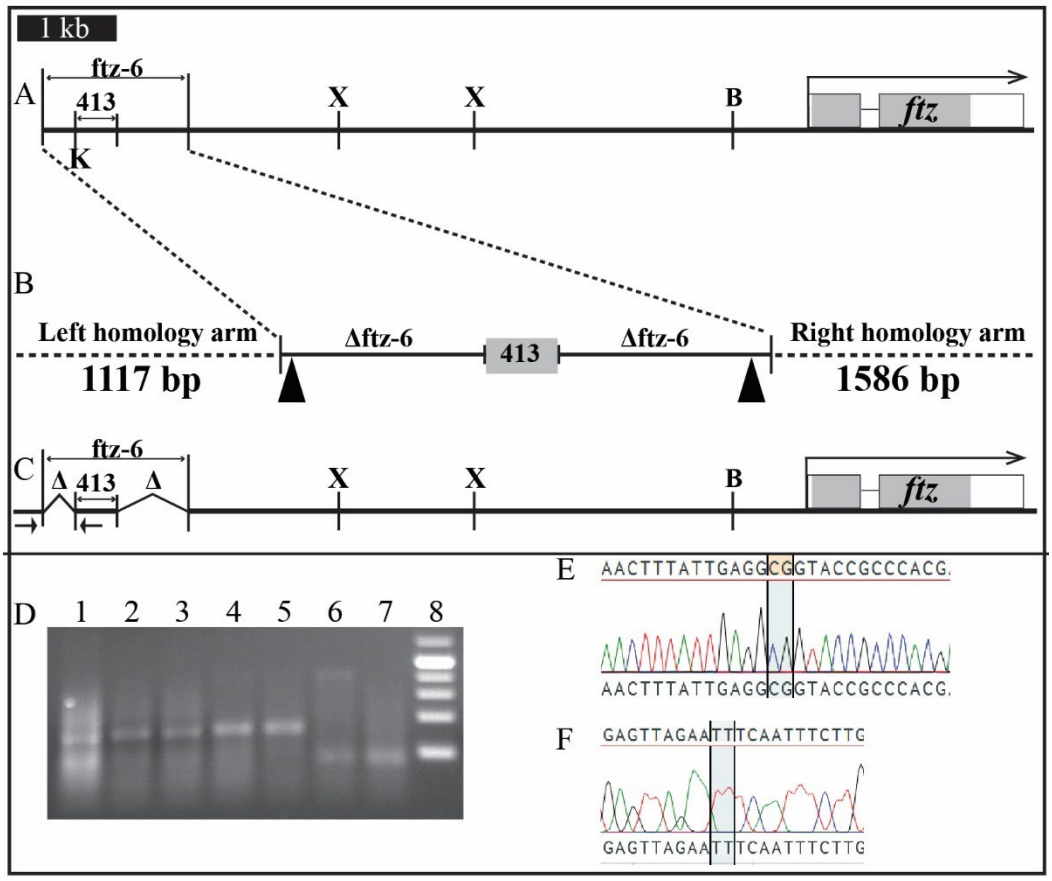


Figure 4-6. Precise deletion of the genomic regions unique to ftz-6 while maintaining the ftz413 CRE using CRISPR/Cas9 and an HDR template generated the *ftz413Δftz-6* allele. (A) Schematic of the genomic region upstream relative to the *ftz* transcription unit. Genomic fragments of ftz413 and ftz-6 are labeled (413 is ftz413). (B) Zoom in of the ftz-6 region indicating where the gRNAs targeted, the regions that were intended for deletion (Δftz-6), and the fragment to be re-incorporated (ftz413) following HDR of the CRISPR/Cas9-induced double-stranded breaks. Homology arms of the repair template with base length are indicated on either side. (C) Schematic of the *ftz413Δftz-6* allele resulting from incorporation of *pUC19-ftz413Δftz-6* following CRISPR/Cas9-induced double-stranded breaks. Arrows represent the location nof the MF21 and MF149 primer binding sites. (D) PCR screen used to identify HDR events with the primers MF21 and MF149, which amplify a 149 bp from positive events (lanes 1-5) and a 405bp fragment from negative events (lanes 6-7). Lane 8 contains GeneRuler 1kb plus DNA ladder. Sequencing result from a mutant identified from PCR screen. Highlight indicates the junction between the left homology arm and ftz413 (E) and between the right homology arm and ftz413 (F).

4.4.7 *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ adults are homozygous viable without apparent segmental defects

Both *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ adults develop with abdominal segments resembling wild type (Fig. 4-7). This appears to be true for the segmentation of the other components of the body plan (head and thorax), as both deletion mutants have fully formed heads and three thoracic segments. Though some mutant adults of either genotype developed with mild segmental aberrations, this was equivalent to the background levels observed for wild type (less than 1% for each; n=2/208 for *ftzΔ413*$^{-/-}$; n=1/240 for *ftz413Δftz-6*$^{-/-}$; n = 1/252 for wild type). However, the *ftz413Δftz-6*$^{-/-}$ females appear to have larger abdomens than wild type (Fig. 4-7D-E). While young wild type female flies have enlarged abdomens for a short duration after eclosion, these mutants have relatively larger abdomens throughout adulthood. Additionally, these *ftz413Δftz-6*$^{-/-}$ mutants appear to lay fewer embryos than wild type, though they are fertile and are maintained as homozygous lines (data not shown).

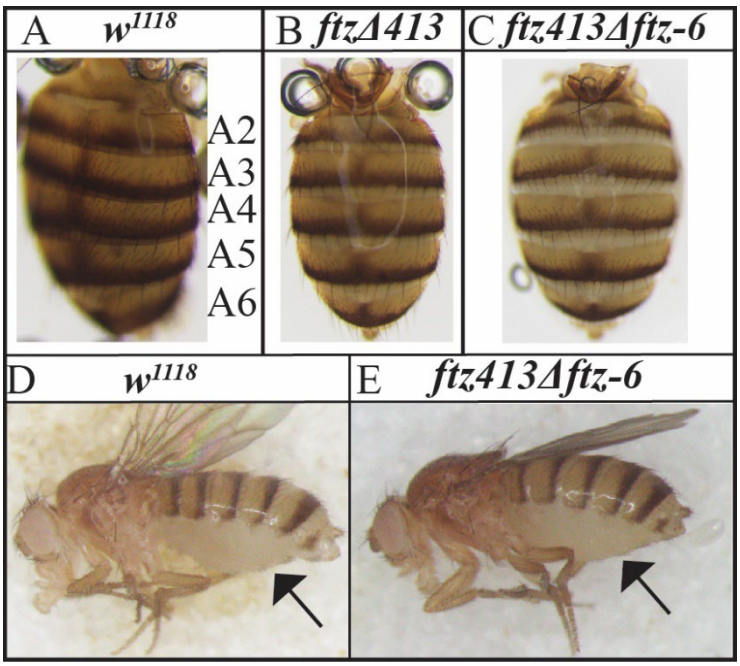


Figure 4-7. Both *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ flies develop to adulthood with wild-type-like segmentation. Brightfield images of adult flies. (A) Abdomens dissected from adult *w*^{1118} homozygotes. Abdominal segments A2-A6 are labeled. (B) Abdomens dissected from adult *ftzΔ413*$^{-/-}$ flies. (C) Abdomens dissected from *ftz413Δftz-6*$^{-/-}$ flies. (D) Lateral images of wild type female adults. (E) Lateral images of *ftz413Δftz-6*$^{-/-}$ female adults. Black arrows point to the region that appears to be larger in *ftz413Δftz-6*$^{-/-}$ females compared to wild type females.

4.4.8 *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ larva develop with wild-type-like cuticle

Though *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ adults are without apparent segmental abnormalities, it is possible that a fraction of mutant embryos failed to develop properly. If these aberrations are lethal, then these individuals would not survive to adulthood. To determine if this phenomenon occurs, larval cuticles were prepared and viewed with darkfield microscopy. Out of 173 wild type larvae, one was missing a segment (Fig. 4-8; 0.5%; n = 1). This suggests that even in a wild type background, a very small number of individuals fail to develop properly. From *ftzΔ413*$^{-/-}$ larvae, it was observed that the majority of individuals developed three thoracic segments and eight abdominal segments (Fig. 4-8B 98.9%; n = 172). Only two *ftzΔ413*$^{-/-}$ larvae developed with abnormalities; one was missing the second thoracic segment (Fig. 4-8C, asterisk; 0.6%; n = 1), and another showed crisscrossing cuticle between the T3 and A1 segments (Fig. 4-8D, asterisk; 0.6%; n = 1). Given that such a small percentage of larva showed any deviation from normal development, it is unlikely that these aberrations can be mechanistically explained by the genotype rather than a failure of the developmental program that is observed even in the wild type background.

The *ftz413Δftz-6*$^{-/-}$ larval cuticles similarly reflected wild type development. The majority of larva developed with all three thoracic and eight abdominal segments (Fig. 4-8 E; 99.4%; n = 311). An individual larva developed with partial fusions between the A1 and A2 segments (Fig. 4-8F, asterisk; 0.3%; n = 1). Another was observed with a slight disruption to the cuticle in the same region (Fig. 4-8G, black arrow points to A1; 0.3%; n = 1). Since the observed deviations from proper segmentation in the ftz*Δ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ larvae occur at the same frequency as in wild type, I conclude that these mutations do not interfere with the development of segmentation.

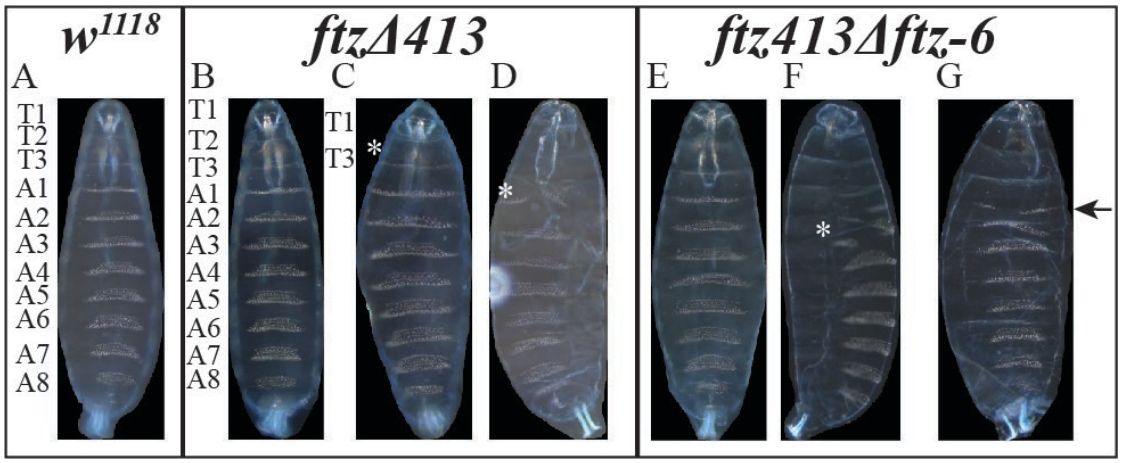


Figure 4-8. *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ larvae develop with wildtype segmentation. (A) Wild type cuticle preparations show three light denticle bands in the thoracic segments (T1-T3) and eight abdominal segments (A1-A8). (B-D) Cuticle patterns observed from ftz*Δ413*$^{-/-}$ larvae. (B) 98.6% of ftz*Δ413*$^{-/-}$ larvae developed with wild-type-like cuticle (n=172). (C) One ftz*Δ413*$^{-/-}$ larva was observed missing T2 segment (asterisk; 0.6%; n=1). (D) One ftz*Δ413*$^{-/-}$ larva had crisscrossing cuticle between T3/A1 (asterisk; 0.6%; n=1). (E) 99.4% of *ftz413Δftz-6*$^{-/-}$ larvae developed with wild-type-like cuticle (n=311). (F) One *ftz413Δftz-6*$^{-/-}$ larva developed with a partial fusion between A1/A2 (asterisk; 0.3%; n=1). (G) One *ftz413Δftz-6*$^{-/-}$ larva developed with a gap in the A1 cuticle (arrow; 0.3%; n=1).

4.4.9 Activation of *ftz* stripes is different in *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ embryos

Since the *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ mutants developed like wild type, I hypothesized that *ftz* expression would be unchanged: seven stripes would be established in a dynamic pattern through the cellular blastoderm stage. However, the order and patterning of *ftz* stripe activation differed in the embryos both mutants from wild type. In wild type, the *ftz* stripes activate in dynamic sets. First, *ftz* is expressed in stripes 1 and 5 (Fig. 4-9i A). Expression is next observed in stripes 2 and 3 (Fig. 4-9i B), and a combined stripe 6/7 begins on the ventral side and progresses dorsally (Fig. 4-9i C). Stripes 6 and 7 then split into separate expression domains while *ftz* expression increases in stripes 1, 2, and 5 (Fig. 4-9i D). Higher abundance of *ftz* transcript is then detected in stripe 3, and stripe 4 expression is detected at low levels (Fig. 4-9i E). Expression in all the stripes increases, but stripe 3 and 4 have relatively lower abundance (Fig. 4-9i F). Stripes 3 and 4 then reach comparable expression levels to other stripes, except 7, which is detected at higher levels (Fig. 4-9i G). As the membrane front reaches the cortex and cellularization is complete, peak expression is reached for all *ftz* stripes (Fig. 4-9i H). All *ftz* stripes remain strong as gastrulation begins (Fig. 4-9i I).

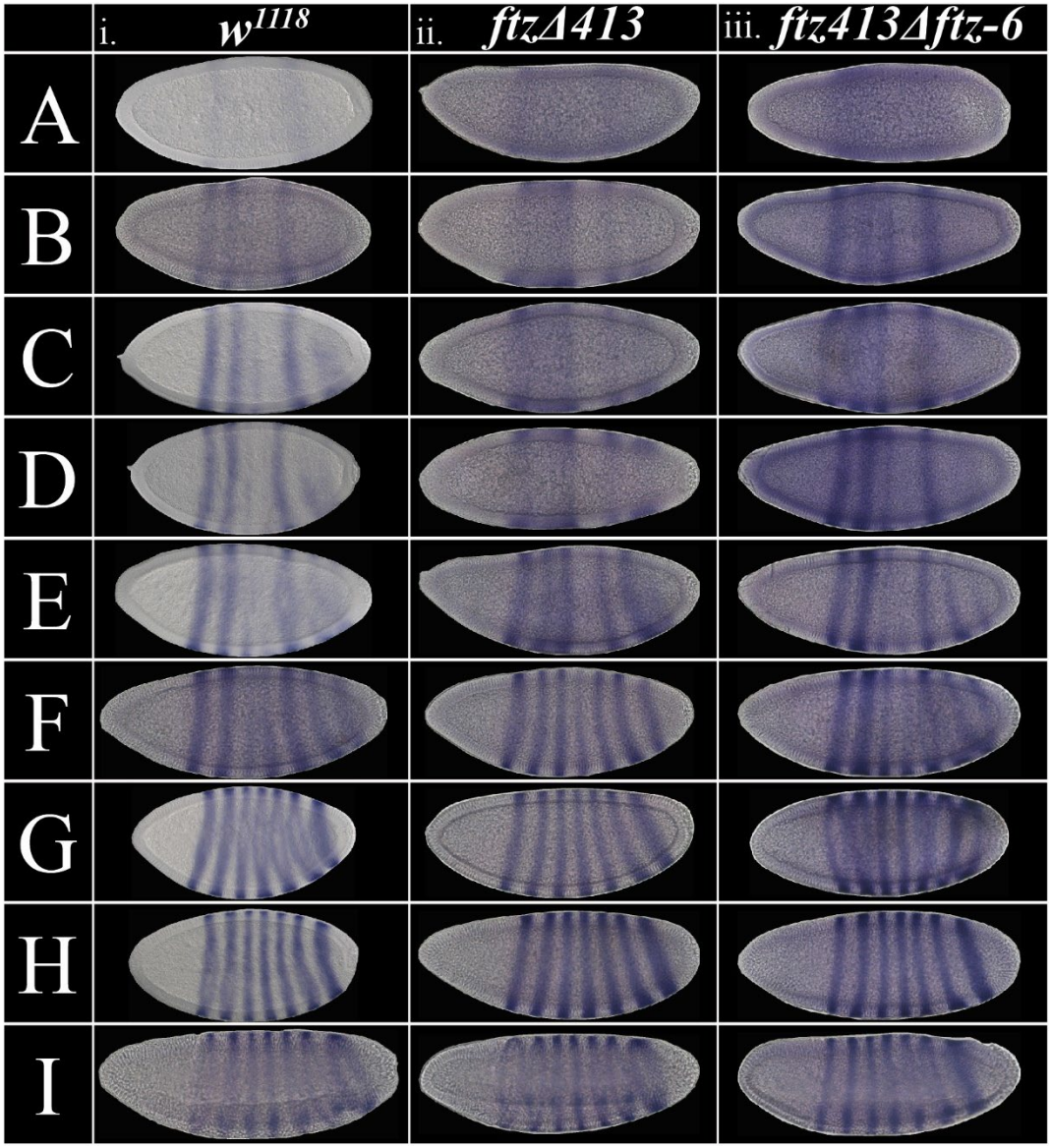


Figure 4-9. The patterning of *ftz* is different in *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ embryos relative to each other and to wild type. Brightfield images of *ftz* expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization in wild type (i), *ftzΔ413*$^{-/-}$ (ii), and *ftz413Δftz-6*$^{-/-}$ (iii) cellular blastoderm embryos. In wild type (i): (i A) *ftz* stripes 1 and 5 are observed first. (i B) Expression initiates in stripes 2 and 3. (i C) The stripe 6/7 domain begins as a single broad band on the ventral side. (i D) The stripe 6 and 7 domains become distinct. Expression increases in stripes 1, 2, 3, and 5. (i E) *ftz* stripe 3 expression increases and stripe 4 expression becomes detectable. (i F) Expression increases in all stripe domains, but remains low in stripes 3 and 4. (i G) The seven-stripe pattern becomes clear. (i H) Peak expression is

reached as cellularization completes. (i H) All *ftz* stripes remain strong as gastrulation begins (i I). In *ftzΔ413*$^{-/-}$ (ii A): *ftz* stripes 1 and 5 initiate. (ii B) A broad band is detected in the anterior of the trunk encompassing the *ftz* stripe 1-3 domains. Expression is detected in stripes 5 and 6. (ii C) Stripe 1 and 3 expression resolves from the anterior band. (ii D) The anterior band subsides as stripes 1 and 3 become resolved. Expression increases slightly in stripes 5 and 6. (ii E) Stripe 2 expression initiates with defined boundaries. Stripe 7 expression initiates as a broad band on the ventral side. (ii F) Stripe 4 expression initiates. (ii G) Peak expression is reached for all *ftz* stripe domains. (ii H) Expression in all stripes remains high as gastrulation begins. In *ftz413Δftz-6*$^{-/-}$ embryos (iii): (A) *ftz* stripes 1 and 5 initiate. (iii B) Stripe 2 and 3 expression initiates. (iii C) Stripe 6 expression initiates as a distinct stripe. (iii D) Stripe 7 expression initiates. Expression in stripe 2 broadens to cover a wider domain. Expression increases in stripes 1 and 5. (iii E) Stripe 4 expression initiates. Expression increases in stripes 6 and 7. (iii F) Expression increases in all seven domains; expression in stripe 2 is very broad. (iii G) the stripe 2 domain retracts and becomes refined. (iii H) Expression of *ftz* is high in all seven stripes as peak levels are reached. (iii I) Expression is maintained in all seven stripes as gastrulation begins.

A variation to the pattern is observed in *ftzΔ413*$^{-/-}$ mutants in the blastoderm stage. As in wild type, *ftz* stripes 1 and 5 are detected first (Fig. 4-9ii A). Next, a broad band of expression ranging from *ftz* stripes 1 to 3 becomes detectable, which contrasts to the wild type pattern where distinct expression of *ftz* stripes 1, 2, and 3 is detected (Fig. 4-9ii B). Another contrast to wild type at the same stage is that *ftz* is detected at lower levels in stripe 5. Additionally, expression in stripe 6 is observed as a band that extends across the dorsal-ventral axis without any detectable broad expression across stripe 7. Next, the stripe 1 and 3 domains start to resolve out of the anterior broad band of expression, and stripes 5 and 6 remain faint (Fig. 4-9ii C). These four stripes (1, 3, 5, and 6) start to become clearer, and the broad band of expression in the anterior of the trunk starts to subside (Fig. 4-9ii D). While *ftz* stripe 2 expression is clearly visible in wild type at this stage, it is not detected in the mutant embryos until the next stage (Fig. 4-9ii E). At this

point, the anterior band is completely resolved into three distinct stripes (stripes 1, 2, and 3). Stripe 5 and 6 expression is also apparent at this stage, and expression of *ftz* stripe 7 begins on the ventral side. The complete seven-stripe pattern arises when *ftz* is then expressed in stripe 4 (Fig. 4-9ii F). Though differences from wild type are observed up to this point (Fig. 4-9A-E), the *ftz* stripe expression is the same as wild type for the rest of the cellular blastoderm stage (Fig. 4-9F-I). All seven stripes become more refined (Fig. 4-9ii G) and *ftz* expression increases until peak *ftz* expression is reached in at the end of blastoderm cellularization (Fig. 4-9ii H). The stripes remain detectable as gastrulation begins (Fig. 4-9ii I).

In summarizing the comparisons of *ftz* stripe dynamism in *ftzΔ413*$^{-/-}$ embryos to the wildtype pattern, three main differences are observed. At the beginning of *ftz* activation, the broad band of expression in the anterior of the trunk is observed in this mutant and not in wild type (Fig. 4-9ii A-C). Next, *ftz* stripe 2 takes much longer to activate than in wild type (Fig. 4-9i B in wild type and 4-9ii E in *ftzΔ413*$^{-/-}$). Finally, *ftz* stripe 7 activates as an individual stripe rather than the broad 6/7 combined stripe that is observed in wild type (Fig. 4-9i C vs. 4-9ii E). Despite these differences, the *ftz* pattern in *ftzΔ413*$^{-/-}$ embryos is the same as wild type at peak expression for both genotypes (Fig. 4-9H).

In *ftz413Δftz-6*$^{-/-}$, yet another order of *ftz* stripe expression is observed. As in wild type, *ftz* stripes 1 and 5 are expressed first (Fig. 4-9iii A). Expression is then detected in stripes 2 and 3 while expression in stripes 1 and 5 increase (Fig. 4-9iii B). Next, Low expression in stripe 6 is observed as an individual stripe (Fig. 4-9iii C). Transcript abundance increases in stripes 1, 2, and 5 while stripe 7 expression begins at very low

levels (Fig. 4-9iii D). Unlike in wild type, the expression of *ftz* in stripes 6 and 7 is slightly offset rather than initiating as a broad band. Another difference is that the expression in the stripe 2 domain is very high relative to the other stripe domains. In the following stage, the expression of *ftz* increases in stripes 6 and 7, and stripe 4 expression becomes detectable (Fig. 4-9iii E). This completes the set of seven stripes. Expression increases in all seven domains, and expression within stripe 2 continues to increase and expand beyond the parasegment (Fig. 4-9iii F). As the seven stripes continue to develop, stripe 2 retracts into a more resolved stripe (Fig. 4-9iii G). Expression of *ftz* reaches peak levels as cellularization is complete (Fig. 4-9iii H). As gastrulation begins, all seven stripes are detected (Fig. 4-9iii I).

In summarizing the comparisons of *ftz* stripe dynamism in *ftz413Δftz-6*$^{-/-}$ embryos to the wildtype pattern, two main differences are observed. Whereas expression in stripes 6 and 7 initiates as a combined stripe that progresses from the ventral side of the embryo in wild type (Fig. 4-9i C), these stripes are detected as distinct domains in *ftz413Δftz-6*$^{-/-}$ embryos (Fig. 4-9iii D-E). A notable difference can be emphasized in regards to the patterning of stripe 2 as well. In wild type, expression of *ftz* in stripe 2 is distinct and refined from the moment it become detectable (Fig. 4-9i B). While it is distinct during the entire *ftz* activation cascade in the *ftz413Δftz-6*$^{-/-}$ mutant embryos, is seems to broaden and become more strongly expressed than in wild type during the majority of the cellular blastoderm (Fig. 4-9iii A-H). Peculiarly, the stripe 2 does become refined and restricted to the expected four cell-wide domain by peak *ftz* expression (Fig. 4-9iii G-H).

Three main differences stand out when comparing *ftz* stripe dynamism in *ftzΔ413*$^{-/-}$ mutants compared to their *ftz413Δftz-6*$^{-/-}$ counterparts. First, there is the broad

band of expression in the anterior of the *ftzΔ413*$^{-/-}$ embryo trunk that extends through the *ftz* stripe 1-3 domains (Fig. 4-9ii B-D). In *ftz413Δftz-6*$^{-/-}$ embryos, stripes 1-3 are distinct once they activate (Fig. 4-9iii B). Second, stripe 2 expression is detected earlier and is initially much wider in *ftz413Δftz-6*$^{-/-}$ embryos than in *ftzΔ413* embryos (Fig. 4-9iii B vs. 4-9ii E). Finally, whereas initiation of *ftz* stripe 7 expression begins as a broad band that resolves in *ftzΔ413*$^{-/-}$ embryos, *ftz* stripe 7 is initially very thin and weak in *ftz413Δftz-6*$^{-/-}$ embryos (Fig. 4-9E).

Despite these differences of *ftz* dynamism throughout the cellular blastoderm, *ftz* expression is identical in both mutants to wild type throughout gastrulation and germband extension (Fig. 4-10). For all three genotypes, the following is observed: expression in all seven stripes remains strong as the embryo transitions from the cellular blastoderm to the germband extension phase (Fig. 4-10A-B). Expression in stripe 1 becomes less detectable around the cephalic furrow (Fig. 4-10C). Stripes 4 and 5 start to lose signal, but they are still clear and present (Fig. 4-10D). As the embryo approaches stage 8, *ftz* expression decreases in all stripes (Fig. 4-10E). Therefore, the effects on *ftz* expression from deleting these stripe-specific CREs seems restricted to *ftz* establishment and normalize to a wildtype pattern before gastrulation begins.

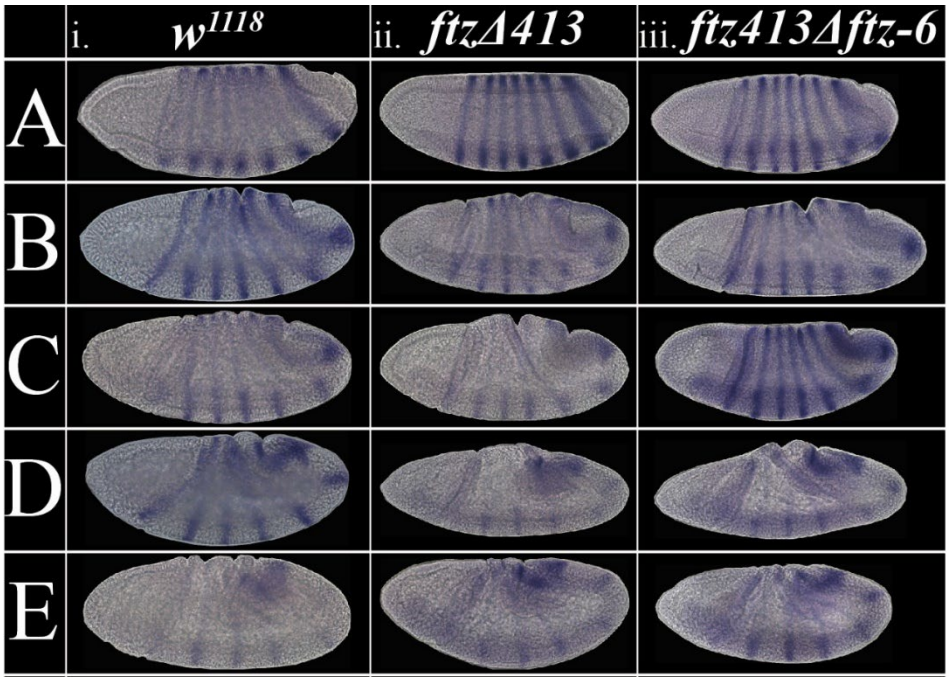


Figure 4-10. *ftz* expression in *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ mutants is indistinguishable from wild type during gastrulation and germband extension. Brightfield images of *ftz* expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization in wild type (i), *ftzΔ413*$^{-/-}$ (ii), and *ftz413Δftz-6*$^{-/-}$ (iii) throughout gastrulation and germband extension. Expression of *ftz* in wild type (i) is identical to that observed in *ftzΔ413*$^{-/-}$(ii) and *ftz413Δftz-6*$^{-/-}$ (iii) embryos. (A-B) Seven stripes are detected from the onset of gastrulation. (C) Expression decreases in all stripe domains, starting with stripe 1. (D) Stripes 4 and 5 start fading. (E) All stripes are present but weaker at the end of gastrulation.

4.4.10 *en* and *slp1* expression patterns of *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ are identical to wild type

Given that *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ animals resemble wild type and that peak *ftz* expression is identical in the mutants compared to wild type, I hypothesized that downstream genes of *ftz* involved in the segmentation pathway would also reflect the wildtype expression pattern. However, the possibility remained that because *ftz* establishment was different in each of these mutants compared to wild type, the

downstream expression may have been perturbed. As described in the previous chapters, *slp1* inter-stripes are repressed by Ftz (Graham et al., 2021; Hang & Gergen, 2017; Prazak et al., 2010). In wild type, *slp1* is expressed in 14 stripes following gastrulation in the germband extension (Fig. 4-11A) and retraction (Fig. 4-11B) stages. Identical patterns to these were observed in both *ftzΔ413*$^{-/-}$ (Fig. 4-11E-F) and *ftz413Δftz-6*$^{-/-}$ (Fig. 4-11I-J) embryos, without the signs of expression domain expansion that were described for the seven-stripe CRE deletions (Chapter 3).

The even stripes of the segmentation polarity gene *en* are activated by Ftz (Ken Howard & Ingham, 1986). Like *slp1*, *en* is expressed in 14 stripe domains throughout the germband extension (Fig. 4-11C) and retraction (Fig. 4-11D) stages. As in wild type, 14 *en* stripes were detected in both the *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ embryos. Taken together with the other phenotypes described of these mutants, it appears that peak *ftz* expression – rather than the order and patterning of early *ftz* stripe establishment – is most consequential for proper regulation of downstream gene expression and subsequent segmentation of the organism.

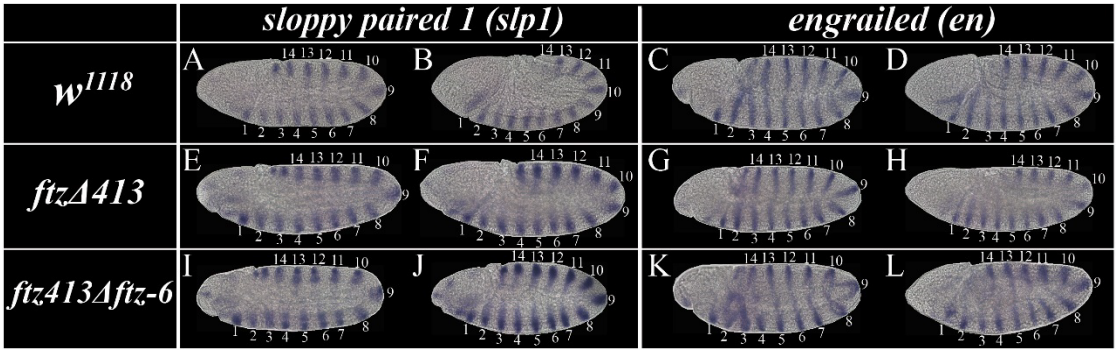


Figure 4-11. Expression of the downstream genes *slp1* and *en* in both *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ embryos is indistinguishable from wild type. Brightfield images of *sloppy paired 1* (*slp1*; left column) and *engrailed* (*en*; right column) expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization throughout stages 7-8 (germband extension) in wild type (A-D), *ftzΔ413*$^{-/-}$ (E-H), and

ftz413Δftz-6$^{-/-}$ (I-L) embryos. (A-B) In wild type embryos, *slp1* is expressed in 14 parasegments. (C-D) In wild type embryos, *en* is expressed in 14 parasegments. (E-F) In *ftzΔ413*$^{-/-}$ embryos, *slp1* is expressed in 14 parasegments. (G-H) In *ftzΔ413*$^{-/-}$embryos, *en* is expressed in 14 parasegments. (I-J) In *ftz413Δftz-6*$^{-/-}$ embryos, *slp1* is expressed in 14 parasegments. (K-L) In *ftz413Δftz-6*$^{-/}$ embryos, *en* is expressed in 14 parasegments.

4.4.11 *Scr* expression is unaffected in *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ embryos

As described in Chapter 3, the *Scr*-specific CRE Scr^{ftz}-XbaI/HindIII (Scr^{ftz}-X/H) is located between the *ftz* UPS and neurogenic CREs (Gindhart et al., 1995). The Scr^{ftz}-X/H CRE directs expression in the anterior mid-gut of stage 13-15 embryos. The regulatory cues of Scr^{ftz}-X/H are specific to *Scr*, and it does not regulate *ftz* transcription. Likewise, the *ftz* CREs do not regulate the transcription of *Scr*. To test if deleting the ftz-6 and ftz413 regions disrupted the specificity of this enhancer-promoter crosstalk, the expression of *ftz* and *Scr* were examined in stage 14 embryos of the *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ mutants with colorimetric *in situ* hybridization using digoxygenin-labeled probes (Fig. 4-12). As in wild type, *Scr* expression was detected in the anterior mid-gut of both deletion mutants. The expression of *ftz* was undetectable in the same domain of both deletion mutants. This suggests that the genomic elements responsible for organizing the specific interaction between this *Scr* CRE and the *Scr* promoter (and not the *ftz* promoter) is not in the ftz-6 genomic region.

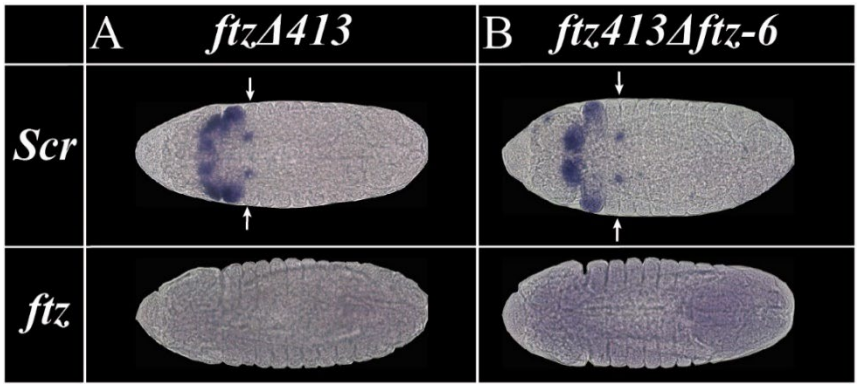


Figure 4-12. *Scr* is detected in the anterior mid-gut of *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ stage 14 mutant embryos. Brightfield images of *Scr* (top panel) and *ftz* (bottom panel) expression using digoxygenin-labeled probes for chromogenic *in situ* hybridization in indicated deletion mutant embryos. (A) Expression of *Scr* and absence of *ftz* in the anterior mid-gut of *ftzΔ413*$^{-/-}$ mutant embryos during stage 14 suggests that the ftz413 genomic region is independent of *Scr-ftz* crosstalk regulation. (B) Expression of *Scr* and absence of *ftz* in the anterior mid-gut of *ftz413Δftz-6*$^{-/-}$ mutant embryos during stage 14 suggests that the ftz-6 genomic region surrounding ftz413 is independent of *Scr-ftz* crosstalk regulation.

4.5 Discussion

The work presented in this chapter has been a comprehensive study of the known stripe-specific CREs of *ftz* (Fig. 4-1). This was conducted with two generalized approaches. First, the sufficiency of these genomic regions in either orientation to direct *lacZ* transgene expression was systematically analyzed (Fig. 4-2 and 4-3). In summary, these experiments revealed that qualitative differences were dependent on the orientation of the CRE. Experiments with these transgenes were then conducted to monitor co-expression in either the *ftz* or *eve* stripe domains (Fig. 4-4). In doing so, I confirmed that most of these genomic regions direct transgene expression specifically in the *ftz* stripe domains. Further, the temporal expression of the transgenes coincided with the dynamic

order of endogenous *ftz* stripe expression. From this, I conclude that these stripe-specific CREs likely function in the native context of the gene to establish *ftz* stripes.

In the second approach, a pair of CREs that share a genomic location (ftz413 and ftz-6) were independently targeted for precise deletion by CRISPR/Cas9 (Fig. 4-5 and 4-6). This was done as a proof of concept to investigate the necessity and role of stripe-specific CREs to endogenous *ftz* expression. Though the order of *ftz* stripe establishment was different in the *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ deletion mutants, the final pattern of *ftz* stripes at peak expression was identical to wild type (Fig. 4-9 and 4-10). Presumably because of this, the downstream expression of two segmentation pathway genes (*slp1* and *en*) in these mutants was detected in a wildtype pattern during the germband extension and retraction stages (Fig. 4-11). The cuticles of these mutant larvae develop with the wild-type-like segmentation (Fig. 4-8), and the adults also resemble wild type (Fig. 4-7). Though the *ftz413Δftz-6*$^{-/-}$ mutant females seem to lay fewer embryos than wild type, they are fertile and easily maintained as a homozygous line. In conclusion, the stripe 2-specific CRE of *ftz* (ftz-6) is redundant and unnecessary for downstream processes, including stripe 2 maintenance and segmentation of the body plan.

4.5.1 Qualitative differences in transgene expression were dependent on orientation

The reporter transgene expression from all five of the CREs examined here showed qualitative orientation-dependent differences. For some of the transgenes, the magnitude of *lacZ* expression was orientation-dependent while the stripe pattern was the same. This was observed for the ftz-7 and ftz+3 transgenes. For both ftz-7 orientations, transgene expression was observed in stripes 3 and 6. While a third stripe appeared to be in stripe 7, the result from the fluorescent *in situ* HCR experiment suggested this was

actually a broad stripe 6 (Fig. 4-4). During *ftz* stripe initiation in wild type embryos, *ftz* expression is detected in a thin stripe 3 and broad stripe that spans the 6/7 domain around the same time (Fig. 4-9B). From these results, I conclude that the ftz-7 CRE likely directs the initial endogenous *ftz* expression within stripes 3 and 6. The contribution from this CRE to transcription within stripe 7, if at all, is likely minor and limited to the initial broad posterior trunk expression.

For both ftz+3 orientations, transgene expression was observed in stripes 1 and 5 in both the chromogenic *in situ* reaction and fluorescent HCR *in situ*. This was in contrast to the near-total orientation dependency of the ftz3'E transgenes. The chromogenic *in situ* experiments detected *ftz3'E-F:lacZ* expression at relatively lower levels in stripe 1 and higher levels in stripe 5. Low expression from the *ftz3'E-R:lacZ* transgene was only detected in stripe 5 around background levels when the chromogenic reaction ran for a longer amount of time, consistent with previous studies on this CRE (Alonso and Pick, unpublished). The two downstream CREs of *ftz* (ftz+3 and ftz3'E) overlap for 1.4kb. Since both *ftz+3F:lacZ* and *ftz3'E-F:lacZ* transgenes are detected in stripes 1 and 5, this shared region likely contains binding sites for *trans*-acting factors that are present in these domains. I will thus refer to this as the regulatory overlap for the discussion. Given that the ftz3'E transgenes reflected nearly total orientation-specificity while the ftz+3 transgenes did not, I conclude that there is a repressive element, such as an insulator or tethering element, in the unique region of the ftz3'E.

4.5.2 A proposed tethering element exists downstream of *ftz*

An insulator in this unique region would block the activity of either downstream CRE from affecting *ftz* expression in the native context of the gene. This does not seem

likely for the following reasons. Moderate *ftz+3F:lacZ* expression was detected in stripes 1 and 5 of early cellular blastoderm embryos. Given that endogenous *ftz* expression is first detected simultaneously in stripes 1 and 5, the initial endogenous expression pattern is probably contributed specifically from the ftz+3 CRE (Fig. 4-9A). This line of reasoning would suggest that the unique region of the ftz3'E contains a tethering element that can facilitate interactions between the *ftz* promoter and the downstream stripe 1+5 CRE (Fig. 4-13). In the context of a transgene, a tethering element could repress expression in an orientation-dependent manner, as follows:

When ftz3'E is fused in the forward orientation, this proposed "unique region tethering element" (URTE) would be placed upstream of the regulatory overlap. Associations of the URTE with other tethering elements around the Phi31C genomic landing site on chromosome 2 would not affect transcription of the forward transgene (Fig. 4-13A). In other words, the URTE activity would be separate from of *lacZ* expression, which is detected in stripes 1 and 5. When fused in the reverse orientation, associations of the URTE with a tethering element elsewhere on chromosome 2 would physically block transcriptional activators bound to the regulatory overlap from transmitting to the *hsp70* promoter (Fig. 4-13B).

In the context of the endogenous *ftz* gene, I propose that this URTE would associate with another tethering element near the *ftz* promoter to bring the ftz+3 CRE in close proximity (Fig. 4-13C). Though a *ftz* promoter-proximal tethering element has yet to be identified, I propose that one may be located in the zebra element based on the disorder of stripe activation observed in $ftz\Delta Z^{-/-}$ mutants (Chapter 2).

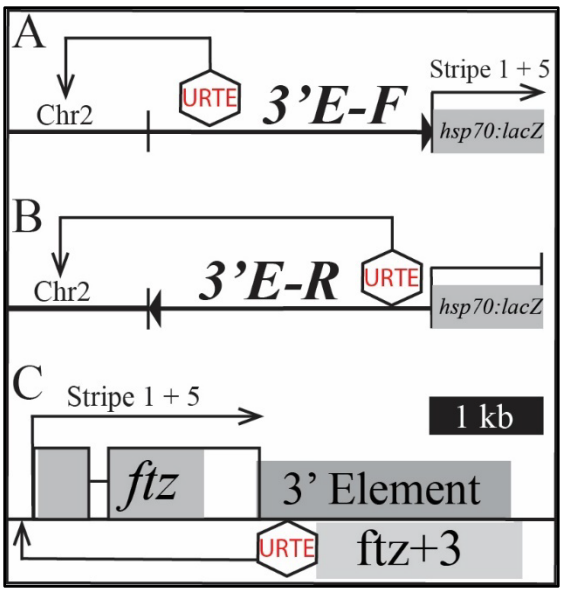


Figure 4-13. Proposed tethering element downstream of *ftz* inhibits reverse transgene expression and facilitates endogenous *ftz* expression in stripes 1 and 5. (A) Schematic of the *ftz3'E-F:lacZ* transgene in the context of the Phi31C docking site on chromosome 2 (Chr2) used in these experiments. The proposed unique region tethering element (URTE) on the 5' end of the *ftz3'E-F:lacZ* transgene associates with another tethering element elsewhere on chromosome 2. The regulatory region of ftz3'E is able to enhance low levels of *lacZ* transcription in the *ftz* stripe 1 and 5 domains. (B) Schematic of the *3'E-R:lacZ* transgene in the context of the same docking site on chromosome 2. The association between the UTRE and a tethering element elsewhere on chromosome 2 intercepts the interaction between ftz3'E and the promoter. Transcription does not occur at detectable levels. (C) Schematic of *ftz* and downstream CREs. Interactions between the UTRE and a putative tethering element upstream of *ftz* brings the downstream stripe 1+5 CRE in close proximity to the promoter, resulting in transcription of *ftz* in stripe 1 and 5.

Note that this type of tethering mechanism has been described for *Scr*, a neighboring gene of *ftz* in the *Antennapedia* gene complex (ANT-C; Batut et al., 2022; Calhoun & Levine, 2003; Li et al., 2015). These studies suggest that the kinetics of the tethering mechanism occur in a dynamic manner that is appropriate for the spatiotemporal regulation of developmental genes (Batut et al., 2022). This contrasts with

insulators, which form stable and more permanent topological associated domains (TADs) that block enhancer-promoter interactions throughout development. Future studies will investigate the role of the URTE in regulating endogenous *ftz* expression.

The presence of tethering elements in the other stripe-specific CREs of *ftz* would further explain the general, qualitative orientation-dependency of the transgenes examined here. If they exist, I hypothesize that the other stripe-specific CREs contain fewer binding sites for tethering element-associated proteins than the ftz3'E region, since the other transgenes did not display the complete orientation-dependence described for the ftz3'E transgenes.

Most of the *ftz* stripe-specific CREs (with exception to ftz413 and ftz3'E) were identified through a computational analysis that identified clusters of maternal and gap gene factor binding sites (Schroeder et al., 2011). Though this analysis was effective in identifying genomic regions that evidently contribute to stripe-specific *ftz* expression (and regions that contribute to the expression of other pair-rule genes as well), it did not search for clusters of tethering element-associated protein binding sites. These putative tethering element-associated protein binding sites around the stripe-specific CREs may not be clustered within the regions identified from that analysis. The ftz+3 CRE, which did not show strong orientation-dependency and was discovered from that study, serves as an example.

Though research on tethering elements is active and much remains to be discovered, they are characterized by H3K4 monomethylation (H3K4me1) and are associated with binding sites for the pioneer factors Zelda (Zld), Grainy head (Grh), and Trithorax-like (Trl) (Batut et al., 2022). Trl is also referred to as the GAGA-associated

factor (GAF). It is notable that several clusters of binding sites for these proteins exist in the promoter-proximal zebra element. A recent publication has presented a Micro-C dataset that maps the transcription-linked chromatin landscape of the early embryo with tight resolution (Batut et al., 2022). By mining this dataset, it could be determined if there is evidence of physical tethering element associations in the *ftz* locus. Functional testing, through more transgene studies or CRISPR/Cas9 deletions of these regions, could test this tethering element hypothesis. Future studies that conduct a combination of bioinformatic analyses and functional experiments would test this hypothesis.

4.5.3 The use of different gene expression detection methods has led to different conclusions

Over the past several decades, different methods have been developed and utilized to monitor gene expression *in vivo*. These methods have drawbacks that result from the indirect nature of their output. In the earliest studies that examined the expression of *lacZ ftz* reporter transgenes, color deposition resulting from a β-galactosidase (β-gal) enzyme reaction was monitored to determine expression domains (Hiromi & Gehring, 1987; Hiromi et al., 1985). This colorimetric assay is dependent upon the stability and activity of the enzyme (Miller, 1972). Thus, instead of monitoring precisely when and where a transcript is expressed, these experiments assayed where catalytically active β-gal protein was located. For example, the original spatiotemporal expression of the *ftz413:lacZ* transgenes used this β-gal activity assay (Pick et al., 1990). Because activity was detected in the stripe 2 domain in *ftz413R:lacZ* transgenic embryos and not in *ftz413F:lacZ* embryos, it was concluded that this CRE is fully orientation dependent. Here, I re-examined the orientation-dependence of these transgenes with *in situ* hybridization using

digoxygenin-labeled probes (Kosman et al., 2004; Tautz & Pfeifle, 1989). An α-digoxygenin antibody conjugated to alkaline phosphatase catalyzed a chromogenic reaction to produce the stain. I believe the reason for the discrepancy between my findings and those previously reported in my PI's postdoctoral work is explained by the relative sensitivity of the assays used.

It should be noted that β-gal is a remarkably stable protein in transgenic organisms, independent of catalytic activity. Immunohistochemistry can detect β-gal for multiple stages of development past the protein's translation (Cohen et al., 1993; Vachon et al., 1992). This is in contrast to the *lacZ* transcript, which decays at a more rapid rate than β-gal. Here, we chose to monitor *lacZ* instead of β-gal to get a more direct readout of transcriptional regulation and to achieve a higher spatiotemporal resolution of transgene expression.

In addition to the chromogenic *in situ* hybridization experiments with digoxygenin-labeled probes, I monitored gene expression with fluorescent *in situ* HCR and confocal microscopy. While both of these assays detect low levels of expression, the fluorescent signal from HCR experiments is limited by the number of target molecules (Choi et al., 2014). This property of HCR makes it an excellent assay for the precision necessary for quantitative measurements (Surkova et al., 2019). However, in comparing my transgene expression experiments, it appears that information can be missed as a result of this tight correlation. Whereas *3'E-F:lacZ* transgene expression was detected in stripes 1 and 5 using the chromogenic *in situ* hybridization method, the transgene was detected at low levels and only in stripe 5 from the HCR experiment using confocal microscopy. The alkaline phosphatase reaction will continuously catalyze a chromogenic

reaction until quenched. My research suggests that this reaction is suitable and ideal for the detection of low abundance transcripts, assuming that substantial background staining does not occur. In conclusion, many assays exist for monitoring gene expression, and the specific properties and desired parameters (such as sensitivity, accuracy, and precision) should be carefully considered when designing experiments.

4.5.4 Analysis of maternal effect and gap gene product predicted binding sites in the stripe 2 and 7 CREs of *ftz* and *eve*

The expression domains of *ftz* and *eve* are interdigitated and offset from each other (Peter A. Lawrence et al., 1987). Can the model proposed to explain *eve* stripe 2 (see Introduction) explain how such definitive boundaries are formed between the *ftz* and *eve* domains? The *ftz* stripe-specific CREs ftz413 and ftz-6 direct expression in domains that are adjacent to *eve* stripes 2, 3, and 7 (Fig. 4-14A). A mid-sagittal scan of the cellularized blastoderm shows that there are distinct cells that express *ftz* and *eve* (Fig. 4-14B). As summarized in the Intro, the *eve* stripe 2 or *eve* stripe 3/7 CRE model is based on differential binding of Gt, Bcd, Kni, Hb, and Kr. Accordingly, I compared the location of predicted consensus binding sites for Gt, Bcd, Kni, Hb, and Kr in four CREs that direct adjacent expression patterns: the *eve* stripe 2 CRE, the *eve* stripe 3/7 CRE, ftz413, and ftz-6 (Fig. 4-14C).

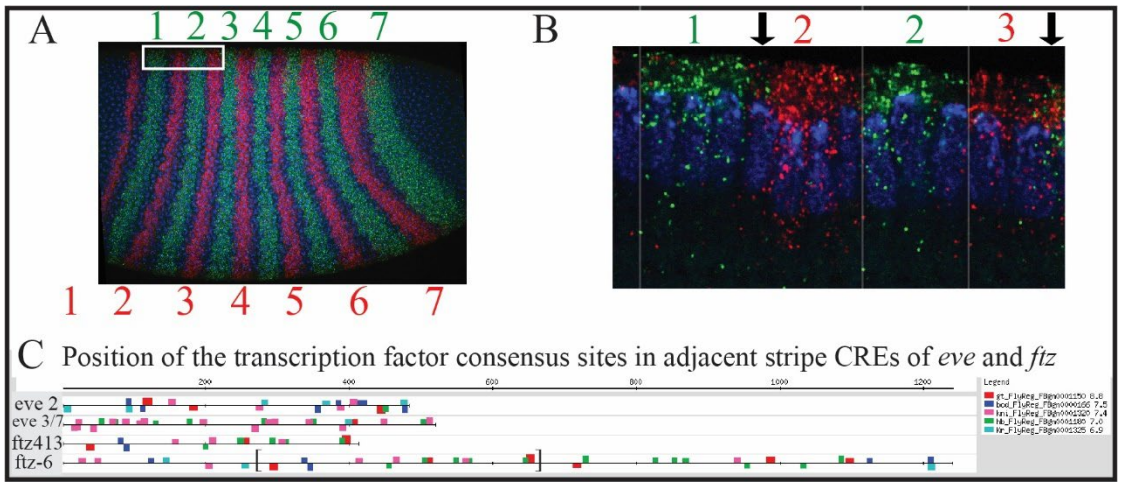


Figure 4-14. Distribution of binding sites within the stripe 2 and 7 CREs of *eve* and *ftz*. (A) Fluorescent *in situ* HCR confocal image of a fully cellularized wild type blastoderm stage embryo. *ftz* (green) and *eve* (red) are expressed in adjacent parasegment registers. Green numbers above the embryo label the *ftz* stripes, and the red numbers below label the *eve* stripes. Hoechst (blue) used as a counterstain. White box shows region of digital zoom. (B) Mid-sagittal scan of dorsal blastoderm layer. Opaque lines were drawn between cells of distinct *ftz* and *eve* expression domains (approximately three to four cells wide). Arrows indicate cells at borders in which both *ftz* and *eve* were detected. Histograms in ImageJ were automatically adjusted. (C) Schematic of consensus binding sites for Giant (Gt, red), Bicoid (Bcd, blue), Knirps (Kni, pink), Hunchback (Hb, green), and Kruppel (Kr, turquoise) in the *eve* stripe 2, *eve* stripe 3+7, ftz413, and ftz-6 CREs. Analysis conducted with the RSAT motif search tool with parameters adjusted to closely replicate the reported *eve* stripe 2 CRE analysis (Arnosti et al., 1996; Nguyen et al., 2018; Small et al., 1992). Brackets in the ftz-6 analysis indicate the position of ftz413 within this region. Each CRE contains a different combination of consensus binding sites. In the *eve* stripe 2 CRE, the following sites were found: 3x Gt, 6x Bcd, 4x Kni, 1x Hb, and 5x Kr. In the *eve* stripe 3+7 CRE: 1x Gt, 0x Bcd, 16x Kni, 6x Hb, and 1x Kr. In the ftz413 CRE: 3x Gt, 2x Bcd, 3x Kni, 5x Hb, and 0x Kr. In the ftz-6 CRE: 6x Gt, 5x Bcd, 7x Kni, 12x Hb, and 3x Kr.

My analysis predicted three Gt, six Bcd, four Kni, one Hb, and five Kr binding sites in the *eve* stripe 2 CRE. One Gt, zero Bcd, 16 Kni, six Hb, and one Kr binding sites were predicted in the *eve* stripe 3/7 CRE. Three Gt, two Bcd, three Kni, five Hb, and zero

Kr binding sites were predicted in the ftz413 region. Finally, six Gt, five Bcd, seven Kni, 12 Hb, and three Kr sites were identified in ftz-6. It is worth emphasizing that the combination of binding sites predicted from this analysis is similar but not identical to the studies conducted on the *eve* CREs (Small et al., 1992; Small et al., 1996).

When focusing on the ftz-6 and *eve* stripe 2 CREs, the predicted binding sites give a general idea of how *ftz* stripe 2 is posterior to *eve* stripe 2. There are more Gt binding sites predicted in ftz-6 compared to the *eve* stripe 2 CRE. This suggests Gt may be repressing the anterior boundary of this *ftz* stripe. There are fewer Kr binding sites predicted in ftz-6 compared to the *eve* stripe 2 CRE. This would position the stripe further along the Kr gradient, suggesting a lower repression on the posterior compared to the *eve* stripe 2 CRE. There is one less Bcd site predicted, but there is also less Bcd protein in the *ftz* stripe 2 domain. This would suggest strong activation by Bcd does not explain expression of *ftz* in the stripe 2 domain, though it helps to explain why this CRE does not direct expression in a more anterior domain. There are fewer predicted sites for Kr in ftz-6, which would suggest a higher concentration of Kr would repress ftz-6 compared to the *eve* stripe 2 CRE. This is also consistent with the relatively more posterior location of *ftz* stripe 2. The relatively high number of predicted Kni sites in ftz-6 may explain why this CRE does not contribute to stripes between 2 and 7.

The final comparison is between the Hb sites, which becomes very confusing with the contradictory regulatory logic of *eve* stripe 2 versus stripe 3/7. However, if the *eve* stripe 2 logic that Hb functions as an activator (or counter-repressor) is followed, the high number of predicted Hb sites in ftz-6 would suggest that Hb is the primary activator of *ftz* stripe 2. Because Hb is also present at lower levels in the posterior, the low level of

transgene expression in the *ftz* stripe 7 domain is consistent with this hypothesis. I cannot logically compare the predicted binding sites in ftz-6 and the *eve* stripe 3/7 CRE in regards to stripe 7 regulation, for the proposed Hb activation of *ftz* in stripes 2 and 7 is contradictory to the Hb repression of *eve* around stripe 7, for which there is evidence.

The *ftz413:lacZ* reporters are expressed in a wider domain around *ftz* stripe 2 as compared to the refined expression of the *ftz-6:lacZ* reporters (Fig. 4-3 and 4-4). The ftz413 fragment is nested within the ftz-6 CRE. Whereas I had hypothesized that ftz413 consists mostly of binding sites for transcriptional activators, this predictive analysis rather suggests that there are more binding sites for repressors in the rest of ftz-6. There are as many predicted sites for Gt in ftz413 as the *eve* stripe 2 CRE. This could explain the anterior expansion of transgene expression over *eve* stripe 2, for these two CREs would be repressed at the same concentrations of Gt. There are no predicted Kr sites in ftz413, so the posterior boundary would not be repressed. Since the posterior expansion of the *ftz413R:lacZ* transgene extended to the *eve* stripe 3 domain and no further, this prediction is probably not accurate. Alternatively, the activators that bind to this CRE may only be present in sufficient concentrations within the expanded transcriptional domain (*eve* stripe 2 + *ftz* stripe 2 + *eve* stripe 3). There are fewer predicted Hb sites in ftz413 compared to ftz-6. Since the Hb gradient declines between 30-50% egg length, there may not be enough protein to effectively bind to this CRE beyond the *eve* stripe 3 domain (the boundary of *ftz413R:lacZ* stripe expansion). This is in assuming that Hb functions as an activator for this CRE.

A major assumption of the discussion above is that predicted binding sites are meaningful. The studies on the *eve* stripe 2 and 3/7 CREs confirmed direct binding by

biochemical experiments *in vitro*. Similar biochemical and ChIP analysis of these predicted sites on *ftz* stripe regulation would be needed to test my hypotheses. However, on the whole, the predicted sites do potentially explain why *ftz* stripe 2 is posterior to *eve* stripe 2, purely on the basis of maternal and gap gene product predicted binding sites.

4.5.5 Deleting the *ftz* stripe 2 CRE had a negligible impact on gene expression and development

It is interesting to further compare the binding site analysis to the altered order of endogenous *ftz* stripe expression in the *ftzΔ413*$^{-/-}$ and *ftz413Δftz-6*$^{-/-}$ mutants (Fig. 4-9). Because such a high level of expression was detected from the *ftz413:lacZ* transgenes, I had hypothesized that *ftz* expression within this domain would be substantially reduced in the *ftzΔ413*$^{-/-}$ deletion mutants. Instead, I was surprised by two observations. First, though the initiation of expression in *ftz* stripe 2 was delayed compared to wild type (Fig. 4-9A-D), the final pattern of *ftz* expression looked wild-type-like: all seven stripes were present, all within their expected boundaries, and at levels that qualitatively resembled wild type (Fig. 4-9H). Second, there appeared to be a loss of repression between *ftz* stripes 1 through 3. Whereas discrete stripes become detectable in wild type throughout cellularization, a wide band of expression was observed in the *ftzΔ413*$^{-/-}$ mutants up until stripe 2 became detectable (Fig. 4-9A-E). The only *trans*-factors that are in this region are the activators Hb and Bcd and the repressor Gt. In the absence of the ftz413 region, the remaining ftz-6 CRE would contain seven predicted sites for Hb, three predicted sites for Bcd, and three predicted sites for Gt. I find it difficult to precisely understand this phenotype, but based on the *eve* stripe 2 model, it appears that the loss of Gt repression combined with continued Hb/Bcd activation results in a low level of expression between

ftz stripes 1 through 3. In summary, I conclude that stripe initiation was aberrant in the *ftzΔ413*$^{-/-}$ mutants mostly due to loss of repressor binding sites, though the loss of activator sites contributed to lower levels of expression early on.

The initiation of *ftz* stripe expression in *ftz413Δftz-6*$^{-/-}$ mutants was different from wild type as well. Based on the expression of the ftz-6 transgenes, I hypothesized that *ftz* would be expressed beyond the normal boundaries of stripe 2, and *ftz* expression in stripe 7 would be slightly reduced. As with the *ftzΔ413*$^{-/-}$mutants, my observations partially supported this hypothesis. The expression of *ftz* was expanded beyond stripe 2 through almost all of the cellular blastoderm stage but eventually normalized to the wild-type-like pattern (Fig. 4-9D-G). Expression of *ftz* stripe 7 appeared slightly impacted, though this was a minor effect that quickly normalized to the wildtype pattern (Fig. 4-9E-F). If the loss of Gt, Hb, and Bcd binding sites caused this phenotype, I hypothesize it would follow the same logic I described above for the *ftz413:lacZ* transgene expression. This is because the *ftz413Δftz-6* allele is essentially a modification of the ftz-6 CRE into ftz413.

Aside from the alterations to *ftz* stripe establishment, these deletion mutants otherwise appeared like wild type. Expression of *ftz* by the end of the cellular blastoderm stage and through gastrulation resembled the wildtype pattern. Downstream genes of *ftz* do not appear to be affected by the aberrant stripe initiation. Mutant larvae and animals resemble wild type, and these mutants are easily maintained as homozygotes. In conclusion, the *ftz* stripe 2 CRE is unnecessary, and its removal is inconsequential to the development of *Drosophila melanogaster*.

Chapter 5: Conclusions and Future Directions

Here, I have comprehensively examined *ftz* CREs to further understand their qualitative and quantitative contributions towards transcriptional pattern establishment in *Drosophila melanogaster*. The experiments utilized to study the qualitative contributions were carefully controlled at the molecular level. All of the isolated genomic regions were fused in both orientations into the same restriction site of the same plasmid. These reporter genes were then integrated into the same Phi31C docking site on chromosome 2. To test the necessity of four different regulatory regions towards *ftz* function, the CRISPR/Cas9 system was utilized to create precise deletions. In total, my analyses have been conducted with 36 reporter transgenic lines and seven CRISPR/Cas9-induced CRE deletion mutants. The generation and analysis of the *ftzΔZ* deletion mutant was done in collaboration with another member of the lab, Dr. Patricia Graham, and a previous undergraduate student, Abhigya Giri. This collaborative work, presented here with minor additions as Chapter 2, has been peer-reviewed and published in the journal *G3: Genes, Genomes, Genetics* (Graham et al., 2021). Dr. Graham and I are co-first authors of that article.

There is substantial redundancy to the *cis*-regulation of *ftz* stripe establishment and maintenance. Expression of *ftz* was significantly reduced in stripe 4 in both the $ftz\Delta Z^{-/-}$ and $ftz\Delta UPS^{-/-}$ embryos – however, the seven-stripe pattern was qualitatively present. In the $ftz\Delta UPS^{7S-/-}$ embryos, *ftz* expression was not maintained through gastrulation. The same result was observed from homozygous ftz^{9H34} mutants, supporting previous conclusions that this element is autoregulatory. However, this autoregulatory

function is evidently redundant: the majority of *ftzΔUPS*$^{7S-/-}$ animals appear wild-type-like, undergo all stages of development, and are fertile. The *ftzΔZ*$^{-/-}$ mutants were also homozygous viable and fertile, though these adults often developed segmental abnormalities of varying degrees. In *ftzΔ413*$^{-/-}$ embryos, *ftz* stripe 2 is established and refined by the end of the cellular blastoderm stage despite altered stripe initiation. These animals are also homozygous viable and fertile, and the adults appear completely wild-type-like.

Out of the seven deletion mutants presented in this work, only the *ftzΔUPS*^{7S}*ΔZ* allele was homozygous lethal. Taken together with the findings that stripe 4 was undetectable above background levels in the *ftzΔUPS*^{7S}*ΔZ*$^{-/-}$ mutants but present at low levels in the *ftzΔZ*$^{-/-}$ mutants, I conclude three things: 1) the UPS^{7S} and the zebra element are the only seven-stripe CREs of *ftz*. 2) The stripe-specific CREs alone are insufficient for *ftz* function and survival. 3) The UPS^{7S} contributes to low stripe 4 expression independent of its role in autoregulation. Future experiments need to verify this by evaluating if the *ftzUPS*^{7S}*:lacZ* transgene is expressed in stripe 4 in a *ftz* mutant background. I hypothesize that this would occur at a low level that reflects the approximately 15-25% reduced *ftz* stripe 4 expression observed in the *ftzΔZ*$^{-/-}$ mutant background, which only received contributions to this stripe from the UPS^{7S}. Though previous studies have investigated this and failed to detect Ftz-independent expression, the methods utilized then were insufficient for detecting low levels of expression.

Regardless, it appears that there is a threshold of *ftz* abundance in each stripe required to convert the analog information of gene expression into the functional, digital information of segmentation. In comparing the HCR data calculations from each of the

seven-stripe deletion mutants, I determine them to be: under 45% for stripe 1, around 15-25% for stripe 4, around 27% for stripe 5, and between 15-32% for stripe 7. Further determinations or tighter resolution will require careful dissection of the CREs, perhaps by removing other *ftz* CREs from these sensitized backgrounds (*ftzΔUPS*^{7S}, *ftzΔZ*, or *ftzΔUPS*^{7S}*ΔZ*, from which segmentation formation can be observed in larval cuticles).

Future experiments could determine if the seven-stripe CREs alone (along with the neurogenic element) are sufficient for *ftz* function and survival. However, there is already early evidence that this would be the case. One of the original rescue constructs of *ftz* contained only the upstream genomic regions including the zebra element, neurogenic element, and the original UPS (Hiromi et al., 1985). At the time, it was not known that the downstream fragment directs expression in stripes 1 and 5. Aside from the piece of the ftz-6/ftz413 CRE located on the 5' end of the UPS, this rescue construct only contained the seven-stripe CREs and it successfully rescued the *ftz* mutants.

That so much redundancy has evolved for the establishment of *ftz* stripes is striking. Previous research suggested that the *ftz* CREs function in a manner that is non-additive (Yu & Pick, 1995). The expression seen from reporter zebra and UPS transgenes does not equate to simple additions of the pattern conferred from either individually. For example, a strong anterior band of expression is observed from the *zebra:lacZ* transgenes. This is not an endogenous *ftz* domain. When combined with the UPS, this anterior expression is dramatically reduced, as demonstrated from the *7SN:lacZ* transgenes. In the *ftz413Δftz-6*$^{-/-}$ embryos, the initial *ftz* stripe 2 expression domain is both more intense and wider than it is in wild type. However, just before the end of cellularization, this stripe is refined and becomes wild-type-like. These mutants develop into adults that are viable and

fertile. Interestingly, *ftz* stripe 2 is significantly expanded in *ftzΔUPS^{7S}ΔZ$^{-/-}$* mutants at the end of the cellular blastoderm. Thus, it appears the UPS^{7S} and zebra elements together play a role in stripe maintenance and pattern resolution. I hypothesize that the *trans*-acting factors that are directly involved in this inter-stripe repression and late cellular blastoderm refinement of *ftz* stripes, or at least *ftz* stripe 2 specifically, bind to the seven-stripe CREs.

It may be that stripe regulatory redundancy has evolved as a means to buffer against environmental stresses or mutations. Though most of the CRE deletion mutants presented here are homozygous viable, the flies are not entirely healthy. *ftzΔZ$^{-/-}$* mutants frequently are missing a segment, have a relatively shorter lifespan, and seem to lay fewer embryos than wild type. *ftz413Δftz-6$^{-/-}$* mutants also seem to lay fewer embryos than wild type and the females specifically seem to have a shorter lifespan, but the segmentation appears wild-type-like. *ftzΔUPS$^{-/-}$* mutants are seemingly and surprisingly as fit as wild type, though fitness or lifespan have not been carefully analyzed. However, they occasionally develop with segmental fusions.

A UPS-like CRE has been reported and functionally studied from *Drosophila hydei* (*D. hydei*) (Jost et al., 1995). Like the *Drosophila melanogaster* (*D. mel*) UPS, the *D. hydei* UPS directs expression in seven-stripes throughout gastrulation. Functional studies in the mosquito *Anopheles stephensi* have demonstrated that the segmentation gene network is highly conserved among *Diptera* (Jarvela et al., 2020). However, the *ftz* stripe establishment and formation of segments is fundamentally different in *Anopheles* compared to *Drosophila*. The *ftz* stripes of *Anopheles* are established one-by-one from anterior to posterior. Preliminary experiments by others in the lab have attempted to

identify stripe-specific *ftz* CREs in *Anopheles* on the basis of sequence identity. Though several putative CREs were identified, none of them direct *ftz*-like transgene expression when *Drosophila melanogaster* was used as a heterologous system. Future work could investigate the evolution of *ftz* stripe *cis*-regulation, as its expression changed from an ancestral homeotic pattern, to a pair-rule striped pattern in holometabolous insects (Heffer et al., 2013).

Much of the work outlined in this thesis has been a re-evaluation of previous findings. Though my experiments were carefully conducted, the possibility remained that all of this work would validate the previous studies without contributing new information. To be honest, this was at some level expected, and I have been surprised, humbled, and grateful that novel findings from this research can benefit the fields of developmental genetics and transcriptional *cis*-regulation.

Appendices

Appendix A: List of oligonucleotides used in this Thesis

Primer	Sequence	Purpose	Notes
CFDN-Seq	GGGCTTTGAGTGTGTGTAGAC	Colony PCR and sequencing	
MF1	GCTCTAGAGGTACCGCCCACGAGCC	ftz413 cloning and UPS^0 cloning	
MF2	GCTCTAGAATTCTAACTCAATTTAAGTCATAAAAATAACCAAAGCATTG	ftz413 cloning and colony PCR	
MF3	GCTCTAGATGCCGGTGGTCGTAGGAG	ftz-6 cloning	
MF4	GCTCTAGATGTAATGGCAGTAATGTTTCGC	ftz-6 cloning	
MF20	TCTAGAGGATCCCCGGGTACGCCCGCTGATTTTTCAAATG	ΔUPS HDR template	
MF21	TATCGCTTTACGACCGGAGG	ftz413Δftz-6 CRISPR screen and colony PCR	
MF22	GGTTCGGATCGCAGGCAATC	Sequencing	
MF24	GGACCCATGACACGGTGC	Sequencing	
MF25	ACATTCGCAGTTCAACATCTTAATC	Δ413 HDR template, sequencing, and colony PCR	
MF40	CAGTGAATTCGAGCTCGGTACGTGTACCCATTGACAATTTTACAAAG	ΔUPS, Δ413, and ftz413Δftz-6 HDR templates	
MF41	CGGCTCGAGGGTACCTCTAGAATGTTTCTCGTTCGAGGTTG	ftz-7 cloning	
MF42	TCTATTTATACTCCGGCGCTCTAGACCGGCAGCCTCAATAAAG	ftz-7 cloning	
MF43	TCTATTTATACTCCGGCGCTCTAGAATGTTTCTCGTTCGAGGTTG	ftz-7 cloning	
MF44	CGGCTCGAGGGTACCTCTAGACCGGCAGCCTCAATAAAG	ftz-7 cloning	
MF45	CGGCTCGAGGGTACCTCTAGACTATTTATTTAAATATACGAGTAAAGTAAATCGATCGAA	ftz3' element cloning	
MF46	TCTATTTATACTCCGGCGCTCTAGACTGCCGGCTAAGTGGACAC	ftz3' element cloning	
MF47	TCTATTTATACTCCGGCGCTCTAGACTATTTATTTAAATATACGAGTAAAGTAAATCGATCGAA	ftz3' element cloning	
MF48	CGGCTCGAGGGTACCTCTAGACTGCCGGCTAAGTGGACAC	ftz3' element cloning	
MF63	CGGCTCGAGGGTACCTCTAGACCGCTGCACGTTCCTCTATT	ftz+3 cloning	
MF64	TCTATTTATACTCCGGCGCTCTAGAGCCTGTCGTGTTCCTTCC	ftz+3 cloning	
MF65	TCTATTTATACTCCGGCGCTCTAGACCGCTGCACGTTCCTCTATT	ftz+3 cloning	
MF66	CGGCTCGAGGGTACCTCTAGAGCCTGTCGT	ftz+3 cloning	

	GTTCCTTCC		
MF94	GGGCGGTACCGCCTCAATAAAGTTTTATTAGGACC	ftz413Δftz-6 HDR template	
MF95	GAGGCGGTACCGCCCACGAGCC	ftz413Δftz-6 HDR template	
MF96	CACAAGAAATTGAAATTCTAACTCAATTTAAGTCATAAAAATAACCAAAGCATTG	ftz413Δftz-6 HDR template and colony PCR	
MF97	TTATGACTTAAATTGAGTTAGAATTTCAATTTCTTGTGAGACGGGG	ftz413Δftz-6 HDR template and sequencing	
MF108	CCGACGACTCCTCCGGAGGTGCACCAGCCGGGAATCGA	UPS sgRNA PCR	
MF109	ACCTCCGGAGGAGTCGTCGGGTTTTAGAGCTAGAAATAGCA	UPS sgRNA PCR	
MF110	TGACGATCGACAATTGTGGCGCACCAGCCGGGAATCGA	ebony sgRNA PCR	
MF112	TTATTTTAACTTGCTATTTCTAGCTCTAAAACTGACGATCGACAATTGTGGC	ebony sgRNA PCR	
MF113	GATCGGATCGAGTCGTTAAGTC	Sequencing	
MF114	GCTCTAGATTCAATTTCTTGTGAGACGGG	UPS^{7S} cloning	
MF115	GCTCTAGAATGTCTAGATAAATGAAGTAATCTTCAG	UPS^{7S} and UPS^{0} cloning	
MF120	TCGATTCCCGGCCGATGCATTCTTGTGAGACGGGGGCGTGTTTTAGAGCTAGAAATAGCA	UPS sgRNA PCR	
MF123	CTTGATGAATTTGCCAGATGCTGTAATGGCAGTAATGTTTCGC	ΔUPS HDR Template	
MF124	AAACATTACTGCCATTACAGCATCTGGCAAATTCATCAAGATG	ΔUPS HDR Template	
MF125	ACCTAGTAGGATTACGAACAAG	Seq	
MF127	AAAGTGCCCACAGATGCC	Seq	
MF128	CGGCTCGAGGGTACCTTTCAATTTCTTGTGAGACGGGG	7SN-Z	
MF130	TCTATTTATACTCCGGCGCTTTCAATTTCTTGTGAGACGGGG	7SN-Z	
MF131	CGGCTCGAGGGTACCTAGAGAGCCCTCGGCACAC	7SN	
MF132	TCTATTTATACTCCGGCGCTCCAAGCAAGTGCGCAACT	7SN-Z	
MF133	CGGCTCGAGGGTACCTCCAAGCAAGTGCGCAACT	7SN-Z	
MF135	TCTATTTATACTCCGGCGCTCTAGAGGTACCGCCCACGAGCC	UPS^{0}	
MF136	CGGCTCGAGGGTACCTCTAGACATCTGGCAAATTCATCAA	Scr^{ftz}-XbaI/HindIII	
MF137	TCTATTTATACTCCGGCGCTCTAGAAAGCTTCAAAAGTCTGTTATAAATGG	Scr^{ftz}-XbaI/HindIII	
MF138	TCGATTCCCGGCCGATGCACTGTCTTCCCGGCTCGTGGGGTTTTAGAGCTAGAAATAGCA	413 sgRNA PCR	
MF139	AAGCATTGTCTAATGCTATTGCACCAGCCGGGAATCGA	413 sgRNA PCR	
MF140	AATAGCATTAGACAATGCTTGTTTTAGAGC	413 sgRNA PCR	

	TAGAAATAGCA		
MF141	GTTAAATAAAATATTTTTCAATTAAAATCT AAAATTACGGGGTGAGGGCAGAAAC	Δ413 HDR template	
MF142	GTAATTTTAGATTTTAATTGAAAAATATTT TATTTAAC	Δ413 HDR template and sequencing	
MF143	CTGTAATGGCAGTAATGTTTC	Δ413 HDR template and colony PCR	
MF144	TCTAGAGGATCCCCGGGTACTCTAGATAAA TGAAGTAATCTTCAGTTTC	ftz413Δftz-6 HDR template	
MF145	TCGATTCCCGGCCGATGCAAGCAAAACTCC TACGACCACGTTTTAGAGCTAGAAATAGCA	ftz-6 sgRNA PCR	
MF146	TTTGCGAAACATTACTGCCAGCACCAGCCG GGAATCGA	ftz-6 sgRNA PCR	
MF147	TGGCAGTAATGTTTCGCAAAGTTTTAGAGC TAGAAATAGCA	ftz-6 sgRNA PCR	
MF149	GCCGGAATTGTTCCTAATCCCC	ftz413Δftz-6 CRISPR screen	
MF152	GCTCTAGATAATACGACTCACTATAGGGAG ATTGTCGTTTCGTCTCGCC	Scr probe cloning	
MF153	CCCGGAATTCATTCGCAAACTCACCAGCG	ftz probe PCR	
MF154	GCTCTAGATAATACGACTCACTATAGGGAG AATGCCAAAGTCTCCTCGATG	ftz probe PCR	
MF155	CCCGGAATTCACTTGAACGGAGGTGAACG	Scr probe cloning	
lacZprobe1	CCAAGAAGAAGCGCAAGGTG	lacZ probe synthesis	From Alys Jarvela
lacZprobe2 T7	TAATACGACTCACTATAGGGAGACTCATCG ATAATTTCACCGCCG	lacZ probe synthesis	From Alys Jarvela
ColonyPCR NandZ	TCGTGGCCAAGCAAGTGCGC	Sequencing	From Patricia Graham
Dm_UES_2 F	TAGGCCTCCTTTGGTAGCCTC	Colony PCR	From Patricia Graham
Dm_UES_2 R	GCTTAAGGACCCATGACACGG	UPS^0 cloning	From Patricia Graham
Dm_UES_3 F	CCGTGTCATGGGTCCTTAAGC	UPS^0 cloning	From Patricia Graham
Dm_UES_3 R	GTTACGGACGCTCCAACTGAG	UPS^0 cloning and sequencing	From Patricia Graham
Dm ftz zebra gRNA1	TCAGGAATACGCGTGGCCGA	zebra sgRNA cloning	From Patricia Graham
Dm ftz zebra gRNA2	CGCAGCAGGTAGGCACCGTA	zebra sgRNA cloning	From Patricia Graham
zebra1	ATTTTGGAAGTGCGTTTGTTG	ΔZ PCR Screen	From Patricia Graham
zebra2	TGACAGCTGACGAGGATTTCT	ΔZ PCR Screen	From Patricia Graham

DmFtz_Neuro1F	TCAAGATGATTTCCCCAGCGC	Sequencing	From Patricia Graham
DmFtz_Neurogenic2.01F	CAAGGCAAAGTGCGGTATTT	7SN cloning and sequencing	From Patricia Graham
DmFtz_Neurogenic2.01R	TGTGAAATCCAATCGTTCGA	Sequencing	From Patricia Graham
DmFtz_Neurogenic2.02F	TAAAGTTGCCTGGGTTCTGG	7SN cloning and sequencing	From Patricia Graham
DmFtz_UESSeq_Inner1F	CGCGCATTCTCGTCATTGTT	Sequencing	From Patricia Graham
DmFtz_UESSeq_Inner2F	TTTATTCCCAGCGGGTCGTC	7SN cloning and colony PCR	From Patricia Graham
DmFtz_UESSeq1F	TCGAAAAATACTCCCGTGAACT	Sequencing	From Patricia Graham
DmFtz_UESSeq2F	GTCGGGGAATTTACACCAATC	Sequencing	From Patricia Graham
DmFtz_UESSeq2R	AATACCGCACTTTGCCTTGTA	7SN cloning and colony PCR	From Patricia Graham
DmUES3F	CCGTGTCATGGGTCCTTAAGC	Sequencing	From Patricia Graham
DmUES3R	GTTACGGACGCTCCAACTGAG	Sequencing	From Patricia Graham
Dm Zebra fullL	TCCAAAACCGACGCATGAAGT	ΔZ PCR Screen	From Patricia Graham
Dm_ZebraFullR	CCAGAGTATGCGGATGGGGTA	ΔZ PCR Screen and sequencing	From Patricia Graham
DmZebraFullL2	GAGCTCCGGATTCGCATTCTA	ΔZ HDR template and sequencing	From Patricia Graham
Dm Zebra fullR2	ATCGATCGCTGAGAACCCATC	ΔZ HDR template	From Patricia Graham
Dm_zebra homolRS	CACACACTCTCTCCTAGGCCAAGCAAGTGCGCAACTCCC	ΔZ HDR template	From Patricia Graham
zebra RS homol3	AGAGAGTGTGTGCCTAGGGCTATAACGGCGAGCGTGTGC	ΔZ HDR template	From Patricia Graham
Zebra RShomology2	GATTTTGCTATATATGCAGGA	ΔZ-GAGA HDR template	From Patricia Graham
NeuroSeq1	GACCTCGAACAGAAGCCAAG	Colony PCR	From

			Patricia Graham
DZ1	GCTCTAGACATCTGGCAAATTC	Scr^{ftz}-$XbaI^{2}$ cloning	From Daniel Zheng
DZ3	GCTCTAGAAAGCTTCAAAAGTCTGTTATAAATGG	Scr^{ftz}-$XbaI^{2}$ cloning	From Daniel Zheng
M13F	GTAAAACGACGGCCAGT	Sequencing	Genewiz sequencing primer
M13R	CAGGAAACAGCTATGAC	Sequencing	Genewiz sequencing primer
SeqF	AAATGCTTGGATTTCACTGG	Colony PCR and sequencing	From general lab stocks
SeqR	GTTTGAATTGAATTGTCGCTCC	Colony PCR and sequencing	From general lab stocks

Appendix B: Table of *in situ* probes and plasmid sequences.

Name	Sequence
ftz probe chromogenic	TGCCAAAGTCTCCTCGATGTGCGACCAATTGAAATCGCCGGCTCCATTCGGG GCTGTCACAATTCGATGATTGATCTCCTGGCTGACAGCTGACGAGGATTTCT CCCCCGGCGATTGGGGTGGCGTGCTGATGCCTTCGAGGGGCGGCAGAGAGG TGGGCGTGGTTGGCGGAGGGGTGGCAAAGTCGCCATTCTTCAGCTTCTGCG TCTGCGACTGGGGCGAGTAGACGTCCAAGTAGTCGACATCCTCGGAGGCGC TGGGCGTGGGCACAGTCACGTACTCTTGGTCGTAGCTGGGAGCGGGGCTGG CGGTGACCTTGGTGCTTACGGCGGGAGCCTTCTTCACCTGCTCGACGGTTGT GTAGAAATAGTCGGGGGTGTACTTCAGCTTCTTCACGGGATTGGTGAGCAG AGCCCTCAGTGTGCTGGGTCGCTCCTCCACGGCGGCGATGATGGAAGCAGC ATCATCTTCGGCCTTGCGCTTGGTGGCCTTGGGCGGCGGGGTGGTGGGTTGC ACGGGCGGTACAGTCTGGGTGGTCACCTGCTCCTGATTGTTGTAGTAGTAGC AGCTCTCCGAGTAACTCTCCTGGGGATAGTAGCCCTGATAATTGGAGGTGTT CTGATAGTAGGCATTGCTGCCTGAATTATCGTAGTAGGTGGGCGGCAGGCT GTGGGGGTGATACATGTTGTACATGTTCATGTTGTCGGCGTAGCTGTAGTGG CTCTGGCTGTTTGTGGTGGCCATATCGGATGTGTATTGCTAGATTTCTTCTCT AACTCTGCGATGTGCACGCAACGCTGGTGAGTTTGCGAAT
lacZ probe chromogenic	CTCATCGATAATTTCACCGCCGAAAGGCGCGGTGCCGCTGGCGACCTGCGT TTCACCCTGCCATAAAGAAACTGTTACCCGTAGGTAGTCACGCAACTCGCC GCACATCTGAACTTCAGCCTCCAGTACAGCGCGGCTGAAATCATCATTAAA GCGAGTGGCAACATGGAAATCGCTGATTTGTGTAGTCGGTTTATGCAGCAA CGAGACGTCACGGAAAATGCCGCTCATCCGCCACATATCCTGATCTTCCAG ATAACTGCCGTCACTCCAGCGCAGCACCATCACCGCGAGGCGGTTTTCTCCG GCGCGTAAAAATGCGCTCAGGTCAAATTCAGACGGCAAACGACTGTCCTGG CCGTAACCGACCCAGCGCCCGTTGCACCACAGATGAAACGCCGAGTTAACG CCATCAAAAATAATTCGCGTCTGGCCTTCCTGTAGCCAGCTTTCATCAACAT TAAATGTGAGCGAGTAACAACCCGTCGGATTCTCCGTGGGAACAAACGGCG GATTGACCGTAATGGGATAGGTCACGTTGGTGTAGATGGGCGCATCGTAAC CGTGCATCTGCCAGTTTGAGGGGACGACGACAGTATCGGCCTCAGGAAGAT CGCACTCCAGCCAGCTTTCCGGCACCGCTTCTGGTGCCGGAAACCAGGCAA AGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGC GGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGC GATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGA CGGGATCTTTCACCTTGCGCTTCTTCTTGG
sloppy paired 1 probe chromogenic (template from W. Ray Anderson)	GAGGGCGTTGTAGCTGTATGGCGGTTTCTTGGTATCGCTGCCGGCGGTCATT TTCTGTTTCTTGCTGGGATTGCCATCCTCGCTCTCCTGGTCCTCGTCCAGCTG ATCCTCCAGCTCATCGTCGAACTCCACGTCGAGCTTGTCATTGTTTTGCGAC GATAGACTCTCGGCGGCACTGCTCATGGGTGTGCTGGTCCTGCTGGGGCTGT CGAAATCCTCCGAAGCCGAGAGTTCGCCATCGGAATTGGAGTACGGATGAA CGTACTGATGATGATGATGCACCGGTTCGGTTTTGATGGGTTGAGTTGCGGT GTTGATGGGTTTCTTGGCCAAGATGGCATCGATGGAGAAGTTGCTCTTGAAT TCCATTTCACTTTTCACCATTTTTCTGATTCGGATATCTTTAGATTTTTGTTTT CTTTGGTTTCTGATTCTCACACACGACTTGGGATCGCTTGAGATGTTTGCTCT TGAGATGCTAGCTCT

engrailed probe from Alys Cheatle Jarvela (FlyBase clone LD16125)	TTTTTTTTTTTTTTTTTTTTAATTTTTTTCCATAATTCTTTTATTTTGTTTTTAA TTTTCTTTTCATGAACTTGCTTTAGCACAAACATTTCGTTGTTATTTTCTTAG GTGCATTTGCCAGCTATTAGTTGCTCGTTAACATTTCGCTTAGAAACTAAAA TACTCTTTCATAGCAATCCTTAAGTGTGTAAACTAGGAATTAAGCTAGGGAA CTGTTAAGTTACTTAATTTACAGAGCGGTTGCAAGCGCGTAATTTATAGGGA GTAATAACTATAGATAGTGTTTTGTGTAGGAATAAACGTCTGAAGGTCGAA GCTTTTGAAGTGCGGCGATAGCGATAGGCTTTGTGTCGGGTGATTACGATA ATACTTTCGGCTATCCAGGGGTGGTTAAATAAATACGCCAACAACTTACAG CTCGAATTCGCTACGGATGGGTCTTACTCTAGGTTAGGATTATACGTATATA TATATATATTTTTAGAGATATATACAAAAACATATGTACGCTACAATTATAT CACATGAAAACCATTTGTACGCCCTCGAATATTGCACGCCCCCGTCAGAAG GAGTAACCCCTTATGGGTAACCATTGGTCAGCGCTTAGGGGATCTGCCCGTT CATGCGCATCTCGAGCTCCTCCTCCTCCTTGGTCAGCGGCACGGTGGTGTGG TTGTACAATCCCTGGGCCATCAGCTGCAGTGCCAGCGGATTTTTGGAGCCCG TCGACTTCTTGATCTTGGCCCGCTTGTTCTGGAACCAGATCTTGATCTGCGC CTCGTTCAGGCCCAACTCGCTGCTCAGCTGCTGGCGTCTCCGCTCGGTCAGA TAGCGATTCTCGTTGAACTCCCGCTTAAGGCGGGCCAACTGCTCGCTGGAG AACGCGGTGCGTGGACGCTTCTCGTCGTTGGTCTTGTCCTTTGGCTGTTTGG GGCGGCGGTAGCGGGGTCCTGAGCTGGGACGATCGCTGTAGCGGGTGCAGT ACACCCAGGCGGGCCACATCTCGTTCTTGCCGCCCTCTGTGGTGGTGGATCC CGTCTCGGAACGCGTGTCATCCGAGGACTCCATTCCGGAATCTCGGCTAAC GGCGGATGGCGGCGGGATTGGCTGCGGCTGGGGACTGCTGGCATTGACTCC GGATCCACTGGCTCCCAGTCGGCTACTGGGCGAGGAGTTCAAGGAACTGGC CGCCGAGGAGCAACCCGACGAGGAGGAGCTCGAGCTGGTGGCCACGGAAG AGGTGCTGCTAATGGTCGAGGCATTGGAGGCGGGTGGCGGACTGGCCAAGC TGCTGGCACTGCTACTCAGCGGAGGTTGGGTCATCGTAGGGGCAGCAGGTT GTCCGATCTGCGAGACCGCCTTGCACAGGGATCCCAGAGCAGATTTCTCTG GATTGGCATCGCTCATCTTGGAGCTCATGGGCGGCGGGATGACGGCATTGG CTGCACTCCTGCGCAGCCGACTCTGCACGATTTCCTCGTGTATCCTGGGATA GGCAGCCGGATTGAAGCAGTTAAGGAAGTTGGCCCGTTCCAGCATCATGGC CGCAGTAGCGGCTGCGGCGGCAGCCTGCTGTTGCCGGCTAAACTCCAGCAG ATCCACTCTTGTAAAGGCGGAGGGCGTGGCAGTCTGGGAACGACTCGCCTC GAAGGGGCGGAATATGCTGGCCTGGTTCTCCATCGACTTGCCCGGCTTCTGG ACATCTCCGAAACGATCGCTCAGGATGTTGGAGATGGAAAAGGCCAGCGAG GGCTTGGCCGTCGACTGCTGCTGGTGGGCGGGTGGTGGCAGGCGTCCGCCC GCCGACGTATCATCCACATCCACATCAATGTCGTCATCCTCCTCCGCCTCCT CCTCCTCGTTGGTGGTGTGCGTCTGATTGTGGAAACTCATGTCACCGATCAC GGAGTCGCTTTCCTCCTCCTTGATGGTCAGCGGCGTGGACGAGGCGGGCGT GGAGCAGGATGCGGGCGATCCACTGCCACTGAGGCGCTGCTGCAGTGCAGC AGCATGAGCGTGGGCGGCCGCAGCAGCAGCTGCAGCAGCGGCGGCTGCAG CGGCATCGAAGGCCATTGCCGGATGGTGGAAGACACCGGCGGCCAGTTGCT GTTGGTGCAACTGCTGCAGTTGCTGCAGCTGGTGGAGGTGCTGCATTTGCTG CTGCTGTTGCTGCTGCTGCTGTTGCTGGTGGTGAAGATGCTGCATTTGTAGG GTAATGGGGCTGGGCGCTGACTGTGGACTGCAGCGATCCTCCAGGGCCATT GGTTTCGACTTGGTGCTAAATCGCAACTCGCTTAAGACACTTGTCACTGGGT CACTTGACACTGAGCCACTGATTCTTCTGATTGCTCAGAATCTTCTGATTCTC TTCTCTGTTTGGTGTTTTTTGTTTGGCTTATGCTTTCGTGTTTGATTTGCAACT GTTGTTCCGACACGCACGTCTGTTCACATCGAG
Scr probe chromogenic	ATGTCGTTTCGTCTCGCCATTGGCATTCACAGTACTCGTCCCGAGATGCACT CGCTTCATCCACGGATATATCTGCGGCGGGTTCTTCTTTCCATTGCCGCTGTT CTGGCTACTGCCCGCCTCATTGCCGCTATCGCTCTCGGAGTCCGAATCCCCG CTCCCGGGTGATTCTGCTGCGGATTCATCTGCTGCGGATAACAGGAGGCCA GCGAGTTGACGAACTGGTACGAGGACATCGCAAAGCAGTCGGGATCCATTT TAAAGCCAGGGGTCGTTGTCGTGGATTGGAGTAATTTTAGTTAGCCGACGG C

HCR *ftz* probe CDS submitted to Molecular Instruments	ATTCGCAAACTCACCAGCGTTGCGTGCACATCGCAGAGTTAGAGAAGAAAT CTAGCAATACACATCCGATATGGCCACCACAAACAGCCAGAGCCACTACAG CTACGCCGACAACATGAACATGTACAACATGTATCACCCCCACAGCCTGCC GCCCACCTACTACGATAATTCAGGCAGCAATGCCTACTATCAGAACACCTC CAATTATCAGGGCTACTATCCCCAGGAGAGTTACTCGGAGAGCTGCTACTA CTACAACAATCAGGAGCAGGTGACCACCCAGACTGTACCGCCCGTGCAACC CACCACCCCGCCGCCCAAGGCCACCAAGCGCAAGGCCGAAGATGATGCTGC TTCCATCATCGCCGCCGTGGAGGAGCGACCCAGCACACTGAGGGCTCTGCT CACCAATCCCGTGAAGAAGCTGAAGTACACCCCCGACTATTTCTACACAAC CGTCGAGCAGGTGAAGAAGGCTCCCGCCGTAAGCACCAAGGTCACCGCCAG CCCCGCTCCCAGCTACGACCAAGAGTACGTGACTGTGCCCACGCCCAGCGC CTCCGAGGATGTCGACTACTTGGACGTCTACTCGCCCCAGTCGCAGACGCA GAAGCTGAAGAATGGCGACTTTGCCACCCCTCCGCCAACCACGCCCACCTC TCTGCCGCCCCTCGAAGGCATCAGCACGCCACCCCAATCGCCGGGGGAGAA ATCCTCGTCAGCTGTCAGCCAGGAGATCAATCATCGAATTGTGACAGCCCC GAATGGAGCCGGCGATTTCAATTGGTCGCACATCGAGGAGACTTTGGCATC AGATTGCAAAGACTACAGCTCCCCGGAGCACTGTGGTGCCGGCTACACCGC GATGCTGCCGCCACTGGAGGCCACAAGCACCGCCACCACCGGGGCACCATC GGTGCCAGTGCCCATGTACCACCACCACCAAACCACCGCCGCCTACCCCGC TTACAGCCACAGTCACAGTCATGGTTATGGCCTGCTCAATGATTACCCTCAG CAGCAGACCCACCAGCAGTACGATGCCTACCCGCAGCAGTACCAACATCAG TGCAGCTACCAGCAACATCCACAGGACCTCTACCATCTGTCTTGAGGTCCGG CGATGCTCAGTTACTCTCTTCCCCAGAGCGGAACCGAAAGCCGTACCGCCA CGAAACCGAAGCGCACTTCTCTCGACCATTTGTAGGTGACACGCAAATGAC ACAGCCGAGAACGAAGCTGCGACGCGATGAGTTGCACAGTAGAGGGCGCA CTCCCTACGGTGCCCAGGACATTTTGGGCACAAGGACGAGTGCGCAAGTGC AGAAGGCAGAGGCAAAAGAGGCAGCGCAAACAGAAAAGGAGCCTTGCTGC GCGCGGAACCCAGTGGCTGGCCATGATGGGTTCTCAGCGATCGATTAGCTG CGGCCAAACACAAGCCCAAAACACTCAGCTGGGAGTGATAATGGCCAAGA GACTTGGAGACTGACACACATGTTTTTGTACATATAGTAGTTAAGATATTCC TATCATAGAATTCTATTTATTTAAATATACGAGTAAAGTAAATCGATCGAAT TTAAAC
HCR *eve* probe CDS submitted to Molecular Instruments	AATGCCTATCCAGTCCGGATAACTCCTTGAACGGCAGCCGCGGCTCGGAGA TTCCCGCCGACCCGTCGGTACGCCGCTATCGCACCGCCTTCACCCGTGACCA GCTGGGTCGCTTGGAGAAGGAGTTCTACAAGGAGAACTACGTGTCCCGTCC CCGTCGCTGCGAACTGGCCGCCGTCAGAGGATCGCCGTCGCCTGGCCCTAC GCAGCCGTCTACTCCGATCCCGCCTTCGCCGCCTCCATCCTCCAGGCCGCCG CCAACAGCGTGGGCATGCCCTATCCGCCCTACGCCCCCGCTGCTGCCGCCGC TGCTGCCGCCGCCGCTGCCGTGGCCACCAATCCGATGATGGCCACCGGAAT GCCCCCGATGGGCATGCCCCAGATGCCCACAATGCAGATGCCCGGACACTC GGGACATGCCGGCCATCCATCGCCCTACGGACAGTACCGCTACACGCCCTA CCACATCCCCGCCCGCCCGGCGCCGCCACATCCCGCTGGTCCTCATATGCAT CATCCGCACATGATGGGATCCAGCGCAACGGGATCGTCGTACTCCGCCGGT GCCGCCGGCCTTTTGGGCGCTCTGCCCTCCGCCACCTGCTATACCGGACTGG GTGTGGGTGTGCCCAAGACCCAGACGCCGCCGCTGGATCTGCAGTCGTCGT CATCGCCGCACTCCTCCACGCTGTCGCTCTCGCCAGTGGGATCCGATCACGC CAAGGTGTTCGACCGCAGTCCAGTGGCTCAATCCGCTCCATCAGTTCCTGCT CCCGCTCCACTGACCACCACCAGCCCGCTGCCCGCTCCAGGCCTCCTGATGC CCAGTGCCAAGCGGCCTGCCTCCGACATGTCGCCGCCGCCGACGACAACTG TGATTGCGGAGCCCAAGCCGAAGCTCTTCAAGCCCTACAAGACTGAGGCGT AAGCCCGCGATCCACACACACTCTCTCCCCCCCCCCCATGCTCCCCCAAAAG ATTGTACAAACTAGTCTTAGTCAGCCTCATCTATTTATTCCCGAAGATTGTA CAGATTGTAGAGTAGCTAATTGTAGTCATAATTAAGGCGCAAAATCAAATT AAGAAATAAATGCGAAAATAACATTG

HCR *lacZ* probe submitted to Molecular Instruments	ATGCCTGCTATTGGAATCCCCAAGAAGAAGCGCAAGGTGAAAGATCCCGTC GTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTCCTTCCCAACAGTTGCG CAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGT GCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGT CCCCTCAAACTGGCAGATGCACGGTTACGATGCGCCCATCTACACCAACGT GACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACG GGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCC AGACGCGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAA CGGGCGCTGGGTCGGTTACGGCCAGGACAGTCGTTTGCCGTCTGAATTTGA CCTGAGCGCATTTTTACGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCT GCGCTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAG CGGCATTTTCCGTGACGTCTCGTTGCTGCATAAACCGACTACACAAATCAGC GATTTCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGCTGTACTGG AGGCTGAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGGGTAACAG TTTCTTTATGGCAGGGTGAAACGCAGGTCGCCAGCGGCACCGCGCCTTTCG GCGGTGAAATTATCGATGAGCGTGGTGGTTATGCCGATCGCGTCACACTAC GTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCT ATCGTGCGGTGGTTGAACTGCACACCGCCGACGGCACGCTGATTGAAGCAG AAGCCTGCGATGGTATGTGGTGGATGAAGCCAATATTGAAACCCACGGCAT GGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAG CGAACGCGTAACGCGAATGGTGCAGCGCGATCGTAATCACCCGAGTGTGAT CATCTGGTCGCTGGGGAATGAATCAGGCCACGGCGCTAATCACGACGCGCT GTATCGCTGGATCAAATCTGTCGATCCTTCCCGCCCGGTGCAGTATGAAGGC GGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGATGTACGCGCGC GTGGATGAAGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAA TGGCTTTCGCTACCTGGAGAGACGCGCCCGCTGATCCTTTGCGAATACGCCC ACGCGATGGGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCGTTTC GTCAGTATCCCCGTTTACAGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTC GCTGATTAAATATGATGAAAACGGCAACCCGTGGTCGGCTTACGGCGGTGA TTTTGGCGATACGCCGAACGATCGCCAGTTCTGTATGAACGGTCTGGTCTTT GCCGACCGCACGCCGCATCCAGCGCTGACGGAAGCAAAACACCAGCAGCA GTTTTTCCAGTTCCGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATAC CTGTTCCGTCATAGCGATAACGAGCTCCTGCACTGGATGGTGGCGCTGGAT GGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGT AAACAGTTGATTGAACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAA CTCTGGCTCACAGTACGCGTAGTGCAACCGAACGCGACCGCATGGTCAGAA GCCGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGT GTGACGCTCCCCGCCGCGTCCCACGCCATCCCGCATCTGACCACCAGCGAA ATGGATTTTTGCATCGAGCTGGGTAATAAGCGTTGGCAATTTAACCGCCAGT CAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACTGCTGACGC CGCTGCGCGATCAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAA GTGAAGCGACCCGCATTGACCCTAACGCCTGGGTCGAACGCTGGAAGGCGG CGGGCCATTACCAGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACAC TTGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGGCAGCATCAGGGGA AAACCTTATTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGTCAAA TGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGGCGC GGATTGGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGC TCGGATTAGGGCCGCAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTT TGACCGCTGGGATCTGCCATTGTCAGACATGTATACCCCGTACGTCTTCCCG AGCGAAAACGGTCTGCGCTGCGGGACGCGCGAATTGAATTATGGCCCACAC CAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAA CTGATGGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACATGG CTGAATATCGACGGTTTCCATATGGGGATTGGTGGCGACGACTCCTGGAGC CCGTCAGTATCGGCGGAATTACAGCTGAGCGCCGGTCGCTACCATTACCAG TTGGTCTGGTGTCAAAAATAA

ftzΔZ-GAGA HDR template	CGATATGTTAGGGGTGTTCAACGGGGGTGGCTAATAAAATTTATCCAGGCC GCGTAAGCCGATCAGAAAGTCATGTAAACACAAAAGTGGCTGAGAGATGA TTGTTTTTTACCAGTCAATTTAGCTGAGCTGATTTCAGGTTATACGCGAACC TAAGATGGAAGTAGATTTTTAGATTACCAGAAATAACCTTCAGTTTCTAAAA GAATGAGAAATAGTAAATCGTAAAGCATGTGGCATTCTTACCCCTTACTGA AATGAATGTAACAGGTAGAAGGCAGCAATTGCAGTGCTACAAAGTATATAT ATTCTTGATCCATATAAGTATAATTATAAGAGCTATAATGTTATAATTAAAC ATGAAGATCCTACGCTGTACAAGTTTTGTGCCACAAATTTAACTCAAGATTA CCTCTCACAAAGGATACATTACCATTTGGCCATTAAAAATTACAGCATCCAT AGACAACCTACTTAAAATTTATAAAAACTTGAGGGTTATATCACAGAGTTA CCGAAAAAAAGCGTAAAAGCTTTATATTCTCAACAATATTATGCTATTAAA ATATTGCTGGTTTTCTGCTGTTATAGAATCATTTTTAAAAGTATAACGTAAA AAATAAAATAAACTAGTATTCATTTGAAAATTCAGCGGGCATATAATTTAT ATCATATTTTTAAAATTTAGGCAAAGGATGTTTGCATAAAGTTTTTACTGTT TACTAGTCATTTTGGAAGTGCGTTTGTTGGTTTTTAGGCAAATACCGGGCAC AGGAGTGAGTTTGGGAATCGGGAGTTGCGCACTTGCTTGGCCTAGGGATTT TGCTATATATGCAGGATCTGCCGCAGGACCAGCTCATTCGCAAACTCACCA GCGTTGCGTGCACATCGCAGAGTTAGAGAAGAAATCTAGCAATACACATCC GATATGGCCACCACAAACAGCCAGAGCCACTACAGCTACGCCGACAACATG AACATGTACAACATGTATCACCCCCACAGCCTGCCGCCCACCTACTACGAT AACTCAGGCAGCAATGCCTACTATCAGAACACCTCCAATTATCAGGGCTAC TATCCCCAGGAGAGTTACTCGGAGAGCTGCTACTACTACAACAATCAGGAG CAGGTGACCACCCAGACTGTACCGCCCGTGCAACCCACCACCCCGCCGCCC AAGGCCACCAAGCGCAAGGCCGAAGATGATGCTGCTTCCATCATCGCCGCC GTGGAGGAGCGACCCAGCACACTGAGGGCTCTGCTCACCAACCCCGTGAAG AAGCTGAAGTACACCCCCGACTATTTCTACACCACCGTCGAGCAGGTGAAG AAGGCTCCCGCCGTAACCACCAAGGTCACCGCCAGCCCCGCTCCCAGCTAC GACCAAGAGTACGTGACTGTGCCCACGCCCAGCGCCTCCGAGGATGTCGAC TACTTGGACGTCTACTCGCCCCAGTCGCAGACGCAGAAGCTGAAGAATGGC GACTTTGCCACCCCTCCGCCAACCACGCCCACCTCTCTGCCGCCCCTCGAAG GCATCAGCACGCCACCCCAATCGCCGGGGGAGAAATCCTCGTCAGCTGTCA GCCAGGAGATCAATCATCGAATTGTGACAGCCCCGAATGGAGCCGGCGATT TCAATTGGTCGCACATCGAGGAGACTTTGGCATCAGGTAGGCATCACACAC GATTAACAACCCCTAAAAATACACTTTGAAAATATTGAAAATATGTTTTTGT ATACATTTTTGATATTTTCAAACAATACGCAGTTATAAAACTCATTAGCTAA CCCATTTTTTCTTTGCTTATGCTTACAGATTGCAAAGACTCGAAACGCACCC GTCAGACGTACACCCGCTACCAGACCCTGGAGCTCGAGAAGGAGTTCCACT TCAATAGATACATCACCCGGCGTCGTCGCATCGATATCGCCAAGCCC
zebra-40 transgene	TCTAGACCACGAGGGCAAACAAAAAGCGCAAACACGCGACCCTCGGCCAC GCGTATTCCTGATCCCAGGGATCGGACGTAATGTTATCCTTTGGCCGCCCAG TGCCACGAAATAAATTCGGAGGGAAAGGGCATCGGGTTCCGGGAACAACT GGCAGCCAGTCTTCGGTGTTTTGCGCGCTGGCAAAAATCCAGAGAAATTTTT AGGGAACCATAAACGGGCCGGGGAAAAAGCCTCTGCGCCGAAGGAACGTT TTCAGCAACAGTTTACAGTTTTTATGTCTTTATGATTATTGCAATTAGAGGG AGATCGGCTGAGAGTCGCGCCCTCTCGCTCTGCGCACCTCATAGGTAGGCA CCTCATGGCCGTAATTACTGCAGCACCGTCTCAAGGTCGCCGAGTAGGAGA AGCGCGCGGGCGGATAAATCGCGATGATAATGGGCGCGATGGGTAGGTAA TAAGCCGCGCAGCAGGTAGGCACCGTACGGATAAAGTTGCCAGGACCTCGG ATAACTTCCCCTCTCCGTGCCTGCAAGGACATTTCGCCGGAGGGGTGGCTGC GAACAGCAGGCGGCAAAGTGTCATGCGCAGGGATATTTATGCGCTATAACG GCGAGCGTGTGCCGAGGGCTCTCTTCTAGA

zebra-73 transgene	TCTAGACCACGAGGGCAAACAAAAAGCGCAAACACGCGACCCTCGGCCAC GCGTATTCCTGATCCCAGGGATCGGACGTAATGTTATCCTTTGGCCGCCCAG TGCCACGAAATAAATTCGGAGGGAAAGGGCATCGGGTTCCGGGAACAACT GGCAGCCAGTCTTCGGTGTTTTGCGCGCTGGCAAAAATCCAGAGAAATTTTT AGGGAACCATAAACGGGCCGGGGAAAAAGCCTCTGCGCCGAAGGAACGTT TTCAGCAACAGTTTACAGTTTTTATGTCTTTATGATTATTGCAATTAGAGGG AGATCGGCTGAGAGTCGCGCCCTCTCGCTCTGCGCACCTCATAGGTAGGCA CCTCATGGCCGTAATTACTGCAGCACCGTCTCAAGGTCGCCGAGTAGGAGA AGCGCGCGGGCGGATAAATCGCGATGATAATGGGCGCGATGGGTAGGTAA TAAGCCGCGCAGCAGGTAGGCACCGTACGGATAAAGTTGCCAGGACCTCGG ATAACTTCCCCTCTCCGTGCCTGCAAGGACATTTCGCCGGAGGGGTGGCTGC GAACAGCAGGCGGCAAAGTGTCATGCGCAGGGATATTTATGCTCTAGA
UPS^0 transgene	GTACCGCCCACGAGCCGGGAAGACAGTCAATAAAACGTCACCTTCACCTCT TACACAATATTTCACAACGTTCTTGCAAATCCGGGGATTAGGAACAATTCCG GCGTTCTCAAAAGGTTTTGGGCATCATGCCCTTCCGTAGTCGCCGTTTGCTC TAATTTAAATTACCATTCTGCATTGTTTTGGAATGTTTCCTAAAAACTGACCT AATTTACGCCACACACTTTTCGTTACAAGTCTTTCGTTTTTATGACCCAAAA ATGGTAAAAGTTTCCGCGAATTTAATCATTTTTATGATGTGAACGATTTTTT GAATAGATGTTGCGAATTCCGAAATCGGGAAATAGAAATATGATTAAAATA GCATTAGaCAATGCTTTGGTTATTTTTATGACTTAAATTGAGTTAGAATGTAA TTTTAGATTTTAATTGAAAAATATTTTATTTAACATATAGTAACTAAGAATT TTTTTTATTAATAATTATATATGTATATATATATATTATATAAGTATATTATT ATATATATATTTTATTTTATTATTATTATATTATATTATTAAATTTTTTTAGCT TTTCTTATAATTTTAATATTTTTATATATTTTTTTTTTGCAAAAGATAAATTA ACAAACCAATATTTCCGCTTCACAAGCAATCCAAGCCCCGATAATAAAATG TTCTATATTTCTTcCTAAAAATATATACTATTATTAATTATGTTTTATTACTAT ATATGTATTTATATTTGCAATGATTATTGTTCAATCGAAGAAAAAATATTGC ATGCAATACTTTGGGGTGTTCTTAGTTTTTCTGAATATATTTTTTATTTATTTT ACTTTAGTTTTGTTATCTTAACAAATAATCTTAATTTATAATATAGTTATAAA CAAATGTTCAATCGTTTTTTCCAGACAAAATGTCTTTTACATTCGCAGTTCA ACATCTTAATCCTTTTGCGAAACATTACTGCCATTACAGTTCAATTTCTTGTG AGACGGGGGCGTTGGCTAAAGGAAAGTCTGTTTTGGGTAAACAGAGCGAA AATCCCTTTCGCCGCCATTTCCATCCATTTCCACCTCTGATTTCTCTGATTTC CACGGACGAATCAAGGCCTACCTGGGAAGGGAAGACGCGTGTGTCTCGCAT TAGACTGATGTTCTGTTGAACGGGACGGGAAATGGAAATGCGTGGCCGAAT CGAGGACGAGGCTCGCGCTACTCGACTTGGACCGACTTGAGTCGTCTGTTTA TGGATAGTAAATGCACCGTGTCATGGGTCCTTAAGCGTCTTAATTGCCCGCG TATCCTTCGCCGAGAAGCAGCGCAATTGGTTTCCAAGGCGGAAATATTATT GGCACATCATAAAACTATAAAACAGGCCGCAGGGGGCCCCCTTCCTCCTTC CTTCGTTTTCCTGTCCAGTTCAGTTGAGCTCTTCACCTCGGATTGCGAGCTCT GAGATTCAAACTGAGACCATCCAATGCCATCAGCGAATTCGCCGAATTTTC GAGCAGCGCGCGCGACTTCGATTCCCCGGCGTACGACTCGTCAGCGGAATC GCATTGACAGCCTACGCCGAACTCCCCGGAGACCTCCGTTCCGATCATTGTT AATGGACGTTATCCTTATTAGATGTTGATGTCCCACGATGTCGAATAGCTAC GGGACATACATATCTACATATATAAAAGATATGCCCATGTTAGGTATTAAG ATCTGATTCAGTTCATGTTCAGCTTCATTATATTATAGATTAAAATTATGTG AATGCTTTTATTTCTATTTCCGTTTCATCAGTTTGCATTTTATATTCCAAATT GTATAATCCTTTAAATTATTAAAATTGTTATAGTGCACGGACCTTCGAGCCT GCTTAAAGAATCGAAACTTTAAATTACTACACAAGGTACAAGGTATTAGGA GAATTAACATAGTTTCACATGACATTATTTATCAATAGTTTTTAACTGATGA CACTCTGACATTTTTCAAAAAGCCATCAATCTTTGAGCTCAGTTGGAGCGTC CGTAACAGAAGGTGTTAACGAGGTCCTAGCTGAGATTTAAACAATTCGAAA TTTATATCTAGGGGCCATAAGCAGTACGATGCGATGATGCACACAGTTCAC ATTTCATTCAGCTTACGGGGTCATCATTGCTATTTAAAAGCGGCAGGACAAT TCGAATAAGTGTGACAGGAGCAATTACAGCCTTATCCTGATGTCTTCGATGT CAACACACCAGTGCTCAAGACATCGCAGGCAATTCAAGGATATGTAGGACG

	CACAGGACCTCGAACAGAAGCCAAGGACACAGGCGACGCGTGACGCATTG GGAAAATATTTGTACAAAACATTGAGGATATCTCAAAAGTAGCACAGCGTT TCGGCATCCTTGACTTTGATTGGTGCCGATCCAAGGACGAAACGCTATATCC ACCCATTGAGAtGACCTCCGGAGGAGTCGTCGGTGGGCCTCCGATCGAAGGT CACCCAGAAATGCATCCTGTCTGGCTTTCTCTTCAaTtTATAGTGAATTTTAG AAACTGAAGATTACTTCATTTAT
UPS[7S] transgene	TTCAATTTCTTGTGAGACGGGGGCGTTGGCTAAAGGAAAGTCTGTTTTGGGT AAACAGAGCGAAAATCCCTTTCGCCGCCATTTCCATCCATTTCCACCTCTGA TTTCTCTGATTTCCACGGACGAATCAAGGCCTACCTGGGAAGGGAAGACGC GTGTGTCTCGCATTAGACTGATGTTCTGTTGAACGGGACGGGAAATGGAAA TGCGTGGCCGAATCGAGGACGAGGCTCGCGCTACTCGACTTGGACCGACTT GAGTCGTCTGTTTATGGATAGTAAATGCACCGTGTCATGGGTCCTTAAGCGT CTTAATTGCCCGCGTATCCTTCGCCGAGAAGCAGCGCAATTGGTTTCCAAGG CGGAAATATTATTGGCACATCATAAAACTATAAAACAGGCCGCAGGGGGCC CCCTTCCTCCTTCCTTCGTTTTCCTGTCCAGTTCAGTTGAGCTCTTCACCTCG GATTGCGAGCTCTGAGATTCAAACTGAGACCATCCAATGCCATCAGCGAAT TCGCCGAATTTTCGAGCAGCGCGCGCGACTTCGATTCCCCGGCGTACGACTC GTCAGCGGAATCGCATTGACAGCCTACGCCGAACTCCCCGGAGACCTCCGT TCCGATCATTGTTAATGGACGTTATCCTTATTAGATGTTGATGTCCCACGAT GTCGAATAGCTACGGGACATACATATCTACATATATAAAAGATATGCCCAT GTTAGGTATTAAGATCTGATTCAGTTCATGTTCAGCTTCATTATATTATAGA TTAAAATTATGTGAATGCTTTTATTTCTATTTCCGTTTCATCAGTTTGCATTT TATATTCCAAATTGTATAATCCTTTAAATTATTAAAATTGTTATAGTGCACG GACCTTCGAGCCTGCTTAAAGAATCGAAACTTTAAATTACTACACAAGGTA CAAGGTATTAGGAGAATTAACATAGTTTCACATGACATTATTTATCAATAGT TTTTAACTGATGACACTCTGACATTTTTCAAAAAGCCATCAATCTTTGAGCT CAGTTGGAGCGTCCGTAACAGAAGGTGTTAACGAGGTCCTAGCTGAGATTT AAACAATTCGAAATTTATATCTAGGGGCCATAAGCAGTACGATGCGATGAT GCACACAGTTCACATTTCATTCAGCTTACGGGGTCATCATTGCtattTAAAAGC GGCAGGACAATTCGAATAAGTGTGACAGGAGCAATTACAGCCTTATCCTGA TGTCTTCGATGTCAACACACCAGTGCTCAAGACATCGCAGGCAATTCAAGG ATATGTAGGACGCACAGGACCTCGAACAGAAGCCAAGGACACAGGCGACG CGTGACGCATTGGGAAAATATTTGTACAAAACATTGAGGATATCTCAAAAG TAGCACAGCGTTTCGGCATCCTTGACTTTGATTGGTGCCGATCCAAGGACGA AACGCTATATCCACCCATTGAGAtGACCTCCGGAGGAGTCGTCGGTGGGCCT CCGATCGAAGGTCACCCAGAAATGCATCCTGTCTGGCTTTCTCTTCAaTtTAT AGTGAATTTTAGAAACTGAAGATTACTTCATTTAT
pCFD5-ΔUPS	GGGCTTTGAGTGTGTGTAGACATCAAGCATCGGTGGTTCAGTGGTAGAATG CTCGCCTGCCACGCGGGCGGCCCGGGTTCGATTCCCGGCCGATGCATTCTTG TGAGACGGGGGCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTA GTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTAACAAAGCACC AGTGGTCTAGTGGTAGAATAGTACCCTGCCACGGTACAGACCCGGGTTCGA TTCCCGGCTGGTGCACCTCCGGAGGAGTCGTCGGGTTTTAGAGCTAGAAAT AGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTAACAAAGCACCAGTGGTCTAGTGGTAGAATAGTACCCTGCCA CGGTACAGACCCGGGTTCGATTCCCGGCTGGTGCGCCACAATTGTCGATCGT CAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAA CTTGAAAAAGTGGCACCGAGTCGGTGC

pUC19-ΔUPS	GTGTACCCATTGACAATTTTACAAAGAATACACTCCTACAATATCACTGAAC AGTTGGATAATAAAAAAACTCAGAATGTTTTTTTTtAAATGAGTGTTCATAA TAAACTTGAACGCAGTCTTGCCGAGTTTTTTAAGCGACAATATGTCTATTAA ACATATTGATGTAGTGgCATTTAAAAATTATTTGTGCTATTAtTTTTTTAAAT AAGTACTCCTTTTTAGCTTTCATAGTTCCCGAGATCACATCGTTCATATGGA CAGACaGATTGACAGATAAGCCCAGATTGAATTGGCCAGTAATCCCGATAA AGAATACGTATTCACTTTCATCCTTTGATACATACTTTTCATACTTTATACCC TTTTACTCTACGAGGAACGGGAGTAATTGTTGGAAATGTTTATTCACCACTC CCAGTTATTGCATTTATACAAATTATTACTTGACTAGATTTTATAATCAAAA TATTTAATTAAAATCCtCCGGTCATTATGTCCTGCTCTACATATAATTTAGAC CTCCCATCGGGTgTTGACTGTCCGTAGCTGTACTATCATTATCCTCCTAAGGC TGTCAAGTTATCTACGCCGCTTGAGGCCATTGTCAgTGTCTTCGCCGGAACT GCTACAAAAAATTGTGACAATTTACACTGATCTCTGGAACCTGGTCCACAG ACCCGTCCTTGACTGCCATAAAGTCAGTTCCACTAAAAATCGTCGAAAAAT ACTCCCGTGAACTAGAAAAAATAAAGCACACAAAAAATGTGCCTCAGTTCT AGTTTAGTTTCAAGAAATTCTCTGTTCGCTCGGAAGGAAAAAATGGAAACT AGAATACGCGGGACAATAAAAAAACGATGCGCTGTAAAAACGTTTTAGACC GAATTCACCTTTGTGTTTCGCGCATTCTCGTCATTGTTAAAAAAAAATCaAG CAtTTTTTTGGGTCGCGACACAATTTTTTGCCGTAGTGTCTCCAGGATTCCTC TAGATCCTGGTCTCGTTCTCAGGCCGTGGCCTCGTTTTCCATTTAAAATGCG CGACGCACAAAAGTTTTTACACGAATTTATGTTTTTATCGCTTTACGACCGG AGGTCCTAATAAAACTTTATTGAGGCTGCCGGTGGTCGTAGGAGTTTTGCTC GAGTTTCTTACCGTAGTAGGCCTCCTTTGGTAGCCTCAAAACTCAAAACTGT AGCAgGGTCAGGTTCCTCGGGATAAGCGCGTATCCTGCGAATTAAGCCAGC CCAAGGACGAGCCGGGTTATGCAAGTCTATGCAACGTATGGGTTTGCATAA TAGAGGTAAATGTTTGTTTAACTGGGAACTGCTCTAGTTAGCATTGTCTTGA TGGCATTGTTTCTGCCCTCACCCCGGTACCGCCCACGAGCCGGGAAGACAG TCAATAAAACGTCACCTTCACCTCTTACACAATATTTCACAACGTTCTTGCA AATCCGGGGATTAGGAACAATTCCGGCGTTCTCAAAAGGTTTTGGGCATCA TGCCCTTCCGTAGTCGCCGTTTGCTCTAATTTAAATTACCATTCTGCATTGTT TTGGAATGTTTCCTAAAAACTGACCTAATTTACGCCACACACTTTTCGTTAC AAGTCTTTCGTTTTTATGACCCAAAAATGGTAAAAGTTTCCGCGAATTTAAT CATTTTTATGATGTGAACGATTTTTTGAATAGATGTTGCGAATTCCGAAATC GGGAAATAGAAATATGATTAAAATAGCATTAGaCAATGCTTTGGTTATTTTT ATGACTTAAATTGAGTTAGAATGTAATTTTAGATTTTAATTGAAAAATATTT TATTTAACATATAGTAACTAAGAATTTTTTTTTATTAATAATTATATATGTATA TATATATATTATATAAGTATATTATTATATATATATTTTATTTTATTATTATT ATATTATATTATTAAATTTTTTTAGCTTTTCTTATAATTTTAATATTTTTATAT ATTTTTTTTTTGCAAAAGATAAATTAACAAACCAATATTTCCGCTTCACAAG CAATCCAAGCCCCGATAATAAAATGTTCTATATTTCTTcCTAAAAATATATA CTATTATTAATTATGTTTTATTACTATATATGTATTTATATTTGCAATGATTA TTGTTCAATCGAAGAAAAAATATTGCATGCAATACTTTGGGGTGTTCTTAGT TTTTCTGAATATATTTTTTATTTATTTTACTTTAGTTTTGTTATCTTAACAAAT AATCTTAATTTATAATATAGTTATAAACAAATGTTCAATCGTTTTTTCCAGA CAAAATGTCTTTTACATTCGCAGTTCAACATCTTAATCCTTTTGCGAAACAT TACTGCCATTACAGCATCTGGCAAATTCATCAAGATGATTTCCCCAGCGCAG ACTTAACGACTCGATCCGATCTGCGTGCAGTTTGATTGTTAACTCAAACAAA TCCaATAACGCAAAAGAAATTTTGAATAATTtTCGAGCATCGCTACTAATTA GCCGCCGATTGAGGCAATCAAATTGATTGCCTGCGATCCGAACCAAAGAAA ACCTCACGCTGGGATCCCCGAACCCACCATAGACCGATTGACGCTGATACT

	ATATATAGACGGTGTATCGGAATAGTCGGGGAATTTACACCAATCTCGAGT TGTTTGTTTTGTTTGTTGATAATTTTTCTCGGCCCTTTACTATTTTATTCCCAG CGGGTCGTCGCCTGTTTCACTGTGATAATGAGTGTGATATACTCGATTGTTT TTATCACCCCTTCCCCAATTAcGTGGTATTTGGTATATAATGAGAAAACCGC CGGAGGAGATCAGCAGCAACTTCAAAGGAAACTGGCCCCAGCATGTGATCA TCGCAGTTTTTCTCAAGCTAGAAAACCAAACAAATCCCGAGACTAACTGAC ACTGTGGTATCCTATTGgGTCTTATGTAATGCCCCACTGTTGGCAGGTGCTG CATTGGCCCATTTCTAATAGGCTTTTGAAGTTTCAAATATTAATCGTCTGCA GGCAGCAGGATCACCAATCGATAGGCCAATAAAAAACAATACCCTGTGTCA TCTGTGTATACAAGGCAAAGTGCGGTATTTTGGATATATGTAAAAGCCAGC TTTACCgCTATTGGCAGGTGCTGTATTGGGCCATTTATAACAGACTTTTGAA GCTTCAAATATTAATGTGCTGCAATCAGAAGAATCGCCAACCGATAACCTA ATAAAAAGATAATGCCTTATTATCTGCGTAAACCAGACATATTACGAATTA GTTTGACATCTCTAGACTTATATAAATGTTTAAATGGTTTGAATTACTATAAa TACTATAATAGgCATATGGAGAACTTATGTGCATAATTTAATCCAAAAAATG TTACGGCGACAACTCAATAATTAATAATGTTTGTGGATAATAGCATACAACT TTTTGGCAAtAAATGACAGTTTAGTTTCCTAATGATCTTGGTCACTTTGGATT CGCTTGACCCACGGCAGCTGATAACGCTGCTCCGGGAAAAATAACGATGAG ATTGACTgCAACAGCAGCCAGAATATATCTGTGGGCTGGGAATTAACACATT GGCCTCCTCCAGGCGCGTGCCCGTTTCACCTTCCCGTGCATTTATGAAAAGT CTGCCGGACCGACCCACCGAACCGTCCGCTGTCTAGCTCTATACGAGTAGA TGTGTATCTGGCTGTTAACCCCTCTGGAATTGCCAAAAATGTGCAACGGCAT CAGCGAAAGAACGGCAGGAAAGATGGTCGGAATCGGGGAGGGGTTCAAAC TCCAAAGGATATAAGGGACATGGGACATGCAGCTAGTCGGGCACAATGGG ATGGTCCCGATTCACTGTCAAGGTAATTTATTAATTCGACAATAAAATCGCA CGAGTTCTTGGGACGAGCTCCGGATTCGCATTCTATGTTGGGTTTGGGTGCG GACTTGCGATGGGGTGTGTTGTGGTCGGACACTCGCATGTGGGCCTTACAA CTCAATTAAAGTTGgCTGGGTTCTGGATCTGAGCTCGCTGCATTTCCGTGGA GAACCGGGTGGGCAAAGTATTATATGACCATGTGCTAGGTTTGGATTAAGC TTAACAGAGGAATTTGTGAAATCCAATCGTTCGAAAGTGAAAAAGTTAAGT GCCCTTATGCCACGTAAACCCATCATTAAATATTTACAAAAACAAACCATTC ATAAACAGCTAGACAAGTTAGTTTAAAATCTTAGGCCATGTTCATGTAAAA TAATTCAACTAATTAATCTTTTTAAGGTATAATTCAATAAGTATATATCTAA ATTCAAAGACCTAGTCAAAAGTAGTAGCCCTCGCCCTTATTTATTGCGTTAA AAAAGGATCCTAACAATATTTAAATAAAATGAGATGAAATTATTTTAGTAA ATGATTCGGTATCGAAAGAATTTTATTTTTGGGGAGCTGTTTTTCCTTTCTTT GATTGATAGATATAAACATAAGTTATCGGATtGTTTTATATCCTACTCTACAC CCTTATTGTTAAGTTTACTGCAAAACCAGAAAGGAGGCTCACATTCACTGGC AATATAGCAATCACAGGACATCGGAGAGATGCGAACTGCAGTCTCTTCTGT CAAACTGAAAAACGCCATAAATTTTGTGGAACAATGGAATCTGCCAGAGTA TGCGGATGGGGTAAGGGTAGAATTCGCACATGCACACATATGTCAGGAGTA TCGATATGTTAGGGGTGTTCAACGGGGGTGGCTAATAAAATTTATCCAGGC CGCGTAAGCCGATCAGAAAGTCATGTAAACACAAAAGTGGCTGAGAGATG ATTGTTTTTTACCAGTCAATTTAGCTGAGCTGATTTCAGGTTATACGCGAAC CTAAGATGGAAGTAGATTTTTAGATTACCAGAAATAACCTTCAGTTTCTAAA AGAATGAGAAATAGTAAATCGTaAAGCaTGTGGCATTCTTACCCCTTACTGA AATGAATGTAACAGGTAGAAGGCAGCAATTGCAGTGCTACAAAGTATATAT ATTCTTGATCCATATAAGTATAAtTATAAGAGCTATAATGTTATAATTAAAC ATGAAGATCCTACGCTGTaCAAGTTTTGTGCCACAAATTTAACTCAAGATTA CCtCTCACAAAGGATACATTACCATTTGGCCATTAAAAATTACAGCATCCAT AGACAACCTACTTAAAATTTATAAAAACTTGAGGGTTATATCACAGAGTTA CCGAAAAAAAGCGTAAAAGCTTTATATTCTCAACAATATTATGCTATTAAA ATATTGCTGGTTTTCTGCTGTTATAGAATCATTTTTAAAAGTATAACGTAAA AAATAAAATAAACTAGTATTCATTTGAAAAtTCAGCGGGC

7SN-F transgene (highlight indicates region missing in 7SN-R)	TTCAATTTCTTGTGAGACGGGGGCGTTGGCTAAAGGAAAGTCTGTTTTGGGT AAACAGAGCGAAAATCCCTTTCGCCGCCATTTCCATCCATTTCCACCTCTGA TTTCTCTGATTTCCACGGACGAATCAAGGCCTACCTGGGAAGGGAAGACGC GTGTGTCTCGCATTAGACTGATGTTCTGTTGAACGGGACGGGAAATGGAAA TGCGTGGCCGAATCGAGGACGAGGCTCGCGCTACTCGACTTGGACCGACTT GAGTCGTCTGTTTATGGATAGTAAATGCACCGTGTCATGGGTCCTTAAGCGT CTTAATTGCCCGCGTATCCTTCGCCGAGAAGCAGCGCAATTGGTTTCCAAGG CGGAAATATTATTGGCACATCATAAAACTATAAAACAGGCCGCAGGGGGCC CCCTTCCTCCTTCCTTCGTTTTCCTGTCCAGTTCAGTTGAGCTCTTCACCTCG GATTGCGAGCTCTGAGATTCAAACTGAGACCATCCAATGCCATCAGCGAAT TCGCCGAATTTTCGAGCAGCGCGCGACTTCGATTCCCCGGCGTACGACTC GTCAGCGGAATCGCATTGACAGCCTACGCCGAACTCCCCGGAGACCTCCGT TCCGATCATTGTTAATGGACGTTATCCTTATTAGATGTTGATGTCCCACGAT GTCGAATAGCTACGGGACATACATATCTACATATATAAAAGATATGCCCAT GTTAGGTATTAAGATCTGATTCAGTTCATGTTCAGCTTCATTATATTATAGA TTAAAATTATGTGAATGCTTTTATTTCTATTTCCGTTTCATCAGTTTGCATTT TATATTCCAAATTGTATAATCCTTTAAATTATTAAAATTGTTATAGTGCACG GACCTTCGAGCCTGCTTAAAGAATCGAAACTTTAAATTACTACACAAGGTA CAAGGTATTAGGAGAATTAACATAGTTTCACATGACATTATTTATCAATAGT TTTTAACTGATGACACTCTGACATTTTTCAAAAAGCCATCAATCTTTGAGCT CAGTTGGAGCGTCCGTAACAGAAGGTGTTAACGAGGTCCTAGCTGAGATTT AAACAATTCGAAATTTATATCTAGGGGCCATAAGCAGTACGATGCGATGAT GCACACAGTTCACATTTCATTCAGCTTACGGGGTCATCATTGCtattTAAAAGC GGCAGGACAATTCGAATAAGTGTGACAGGAGCAATTACAGCCTTATCCTGA TGTCTTCGATGTCAACACACCAGTGCTCAAGACATCGCAGGCAATTCAAGG ATATGTAGGACGCACAGGACCTCGAACAGAAGCCAAGGACACAGGCGACG CGTGACGCATTGGGAAAATATTTGTACAAAACATTGAGGATATCTCAAAAG TAGCACAGCGTTTCGGCATCCTTGACTTTGATTGGTGCCGATCCAAGGACGA AACGCTATATCCACCCATTGAGAtGACCTCCGGAGGAGTCGTCGGTGGGCCT CCGATCGAAGGTCACCCAGAAATGCATCCTGTCTGGCTTTCTCTTCAaTtTAT AGTGAATTTTAGAAACTGAAGATTACTTCATTTATCTAGACATCTGGCAAAT TCATCAAGATGATTTCCCCAGCGCAGACTTAACGACTCGATCCGATCTGCGT GCAGTTTGATTGTTAACTCAAACAAATCCaATAACGCAAAAGAAATTTTGAA TAATTtTCGAGCATCGCTACTAATTAGCCGCCGATTGAGGCAATCAAATTGA TTGCCTGCGATCCGAACCAAAGAAAACCTCACGCTGGGATCCCCGAACCCA CCATAGACCGATTGACGCTGATACTATATATAGACGGTGTATCGGAATAGT CGGGGAATTTACACCAATCTCGAGTTGTTTGTTTTGTTTGTTGATAATTTTTC TCGGCCCTTTACTATTTTATTCCCAGCGGGTCGTCGCCTGTTTCACTGTGATA ATGAGTGTGATATACTCGATTGTTTTTATCACCCCTTCCCCAATTAcGTGGTA TTTGGTATATAATGAGAAAACCGCCGGAGGAGATCAGCAGCAACTTCAAAG GAAACTGGCCCCAGCATGTGATCATCGCAGTTTTTCTCAAGCTAGAAAACC AAACAAATCCCGAGACTAACTGACACTGTGGTATCCTATTGgGTCTTATGTA ATGCCCCACTGTTGGCAGGTGCTGCATTGGCCCATTTCTAATAGGCTTTTGA AGTTTCAAATATTAATCGTCTGCAGGCAGCAGGATCACCAATCGATAGGCC AATAAAAAACAATACCCTGTGTCATCTGTGTATACAAGGCAAAGTGCGGTA TTTTGGATATATGTAAAAGCCAGCTTTACCgCTATTGGCAGGTGCTGTATTG GGCCATTTATAACAGACTTTTGAAGCTTCAAATATTAATGTGCTGCAATCAG AAGAATCGCCAACCGATAACCTAATAAAAAGATAATGCCTTATTATCTGCG TAAACCAGACATATTACGAATTAGTTTGACATCTCTAGACTTATATAAATGT TTAAATGGTTTGAATTACTATAAaTACTATAATAGgCATATGGAGAACTTAT GTGCATAATTTAATCCAAAAAATGTTACGGCGACAACTCAATAATTAATAA TGTTTGTGGATAATAGCATACAACTTTTTGGCAAtAAATGACAGTTTAGTTTC CTAATGATCTTGGTCACTTTGGATTCGCTTGACCCACGGCAGCTGATAACGC TGCTCCGGGAAAAATAACGATGAGATTGACTgCAACAGCAGCCAGAATATA

	TCTGTGGGCTGGGAATTAACACATTGGCCTCCTCCAGGCGCGTGCCCGTTTC ACCTTCCCGTGCATTTATGAAAAGTCTGCCGGACCGACCCACCGAACCGTCC GCTGTCTAGCTCTATACGAGTAGATGTGTATCTGGCTGTTAACCCCTCTGGA ATTGCCAAAAATGTGCAACGGCATCAGCGAAAGAACGGCAGGAAAGATGG TCGGAATCGGGGAGGGGTTCAAACTCCAAAGGATATAAGGGACATGGGAC ATGCAGCTAGTCGGGCACAATGGGATGGTCCCGATTCACTGTCAAGGTAAT TTATTAATTCGACAATAAAATCGCACGAGTTCTTGGGACGAGCTCCGGATTC GCATTCTATGTTGGGTTTGGGTGCGGACTTGCGATGGGGTGTGTTGTGGTCG GACACTCGCATGTGGGCCTTACAACTCAATTAAAGTTGCCTGGGTTCTGGAT CTGAGCTCGCTGCATTTCCGTGGAGAACCGGGTGGGCAAAGTATTATATGA CCATGTGCTAGGTTTGGATTAAGCTTAACAGAGGAATTTGTGAAATCCAATC GTTCGAAAGTGAAAAAGTTAAGTGCCCTTATGCCACGTAAACCCATCATTA AATATTTACAAAAACAAACCATTCATAAACAGCTAGACAAGTTAGTTTAAA ATCTTAGGCCATGTTCATGTAAAATAATTCAACTAATTAATCTTTTTAAGGT ATAATTCAATAAGTATATATCTAAATTCAAAGACCTAGTCAAAAGTAGTAG CCCTaGCCCTTATTTATTGCGTTAAAAAAGGATCCTAACAATATTTAAATAA AATGAGATGAAATTATTTTAGTAcATGATTCGGTATCGAAAGAATTTTATTT TTGGGGAGCTGTTTTTCCTTTCTTTGATTGATAtATATAAACATAAGTTATCG GATGGTTTTATATCCTACTCTACACCCTTATTGTTAAGTTTACTGCAAAACC AGAAAGGAGGCTCACATTCACTGGCAATATAGCAATCACAGGACATCGGAG AGATGCGAACTGCAGTCTCTTCTGTCAAACTGAAAAACGCCATAAATTTTGT GGAACAATGGAATCTGCCAGAGTATGCGGATGGGGTAAGGGTAGAATTCGC ACATGCACACATATGTCAGGAGTATCGATATGTTAGGGGTGTTCAACGGGG GTGGCTAATAAAATTTATCCAGGCCGCGTAAGCCGATCAGAAAGTCATGTA AACACAAAAGTGGCTGAGAGATGATTGTTTTTTACCAGTCAATTTAGCTGA GCTGATTTCAGGTTATACGCGAACCTAAGATGGAAGTAGATTTTTAGATTAC CAGAAATAACCTTCAGTTTCTAAAAGAATGAGAAATAGTAAATCGTGAAGC CTGTGGCATTCTTACCCCTTACTGAAATGAATGTAACAGGTAGAAGGCAGC AATTGCAGTGCTACAAAGTATATATATTCTTGATCCATATAAGTATAAGTAT AAGAGCTATAATGTTATAATTAAACATGAAGATCCTACGCTGTGCAAGTTTT GTGCCACAAATTTAACTCAAGATTACCACTCACAAAGGATACATTACCATTT GGCCATTAAAAATTACAGCATCCATAGACAACCTACTTAAAATTTATAAAA ACTTGAGGGTTATATCACAGAGTTACCGAAAAAAAGCGTAAAAGCTTTATA TTCTCAACAATATTATGCTATTAAAATATTGCTGGTTTTCTGCTGTTATAGAA TCATTTTTAAAAGTgTAACGTAAAAAATAAAATAAAACTAGTATTaATTTGA AAAATCAGCGGGCATATAATTTATATCATATTTTTAAAATTTAGGCAAAGG ATGTTTGCATAAAGTTTTTACTGTTTACTAGTCATTTTGGAAGTGCGTTTGTT GGTTTTTAGGCAAATACCGGGCACAGGAGTGAGTTTGGGAATCGGGAGTTG CGCACTTGCTTGGCCACGAGGGCAAACAAAAAGCGCAAACACGCGACCCTC GGCCACGCGTATTCCTGATCCCAGGGATCGGACGTAATGTTATCCTTTGGCC GCCCAGTGCCACGtAATAAATTCGGAGGGAAAGGGCATCGGGTTCCGGGAA CAACTGGCAGCCAGTCTTCGGTGTTTTGCGCGCTGGCAAAAATCCAGAGAA ATTTTTAGGGAACCATAAACGGGCCGGGGAAAAAGCCTCTGCgCCGAAGGA ACGTTTTCAGCAACAGTTTACAGTTTTTATGTCTTTATGATTATTGCAATTAG AGGGAGATCGGCTGAGAGTCGCGCCCTCTCGCTCTGCGCACCTCATAGGTA GGCACCTCATGGCCGTAATTACTGCAGCACCGTCTCAAGGTCGCCGAGTAG GAGAAGCGCGCGGGCGGATAAATCGCGATGATAATGGGCGCGATGGGTAG GTAATAAGCCGCGCAGCAGGTAGGCACCGTACGGATAAAGTTGCCAGGACC TCGGATAACTTCCCCTCTCCGTGCCTGCAAGGACATTTCGCCGGAGGGGTGG CTGCGAACAGCAGGCGGCAAAGTGTCATGCGCAGGGATATTTATGCGCTAT AACGcCGAGCGTGTGCCGAGGGCTCTCT

7SN-Z-F transgene (highlight indicates region missing in 7SN-R)	TTCAATTTCTTGTGAGACGGGGGCGTTGGCTAAAGGAAAGTCTGTTTTGGGT AAACAGAGCGAAAATCCCTTTCGCCGCCATTTCCATCCATTTCCACCTCTGA TTTCTCTGATTTCCACGGACGAATCAAGGCCTACCTGGGAAGGGAAGACGC GTGTGTCTCGCATTAGACTGATGTTCTGTTGAACGGGACGGGAAATGGAAA TGCGTGGCCGAATCGAGGACGAGGCTCGCGCTACTCGACTTGGACCGACTT GAGTCGTCTGTTTATGGATAGTAAATGCACCGTGTCATGGGTCCTTAAGCGT CTTAATTGCCCGCGTATCCTTCGCCGAGAAGCAGCGCAATTGGTTTCCAAGG CGGAAATATTATTGGCACATCATAAAACTATAAAACAGGCCGCAGGGGGCC CCCTTCCTCCTTCCTTCGTTTTCCTGTCCAGTTCAGTTGAGCTCTTCACCTCG GATTGCGAGCTCTGAGATTCAAACTGAGACCATCCAATGCCATCAGCGAAT TCGCCGAATTTTCGAGCAGCGCGCGACTTCGATTCCCCGGCGTACGACTC GTCAGCGGAATCGCATTGACAGCCTACGCCGAACTCCCCGGAGACCTCCGT TCCGATCATTGTTAATGGACGTTATCCTTATTAGATGTTGATGTCCCACGAT GTCGAATAGCTACGGGACATACATATCTACATATATAAAAGATATGCCCAT GTTAGGTATTAAGATCTGATTCAGTTCATGTTCAGCTTCATTATATTATAGA TTAAAATTATGTGAATGCTTTTATTTCTATTTCCGTTTCATCAGTTTGCATTT TATATTCCAAATTGTATAATCCTTTAAATTATTAAAATTGTTATAGTGCACG GACCTTCGAGCCTGCTTAAAGAATCGAAACTTTAAATTACTACACAAGGTA CAAGGTATTAGGAGAATTAACATAGTTTCACATGACATTATTTATCAATAGT TTTTAACTGATGACACTCTGACATTTTTCAAAAAGCCATCAATCTTTGAGCT CAGTTGGAGCGTCCGTAACAGAAGGTGTTAACGAGGTCCTAGCTGAGATTT AAACAATTCGAAATTTATATCTAGGGGCCATAAGCAGTACGATGCGATGAT GCACACAGTTCACATTTCATTCAGCTTACGGGGTCATCATTGCtattTAAAAGC GGCAGGACAATTCGAATAAGTGTGACAGGAGCAATTACAGCCTTATCCTGA TGTCTTCGATGTCAACACACCAGTGCTCAAGACATCGCAGGCAATTCAAGG ATATGTAGGACGCACAGGACCTCGAACAGAAGCCAAGGACACAGGCGACG CGTGACGCATTGGGAAAATATTTGTACAAAACATTGAGGATATCTCAAAAG TAGCACAGCGTTTCGGCATCCTTGACTTTGATTGGTGCCGATCCAAGGACGA AACGCTATATCCACCCATTGAGAtGACCTCCGGAGGAGTCGTCGGTGGGCCT CCGATCGAAGGTCACCCAGAAATGCATCCTGTCTGGCTTTCTCTTCAaTtTAT AGTGAATTTTAGAAACTGAAGATTACTTCATTTATCTAGACATCTGGCAAAT TCATCAAGATGATTTCCCCAGCGCAGACTTAACGACTCGATCCGATCTGCGT GCAGTTTGATTGTTAACTCAAACAAATCCaATAACGCAAAAGAAATTTTGAA TAATTtTCGAGCATCGCTACTAATTAGCCGCCGATTGAGGCAATCAAATTGA TTGCCTGCGATCCGAACCAAAGAAAACCTCACGCTGGGATCCCCGAACCCA CCATAGACCGATTGACGCTGATACTATATATAGACGGTGTATCGGAATAGT CGGGGAATTTACACCAATCTCGAGTTGTTTGTTTTGTTTGTTGATAATTTTTC TCGGCCCTTTACTATTTTATTCCCAGCGGGTCGTCGCCTGTTTCACTGTGATA ATGAGTGTGATATACTCGATTGTTTTTATCACCCCTTCCCCAATTAcGTGGTA TTTGGTATATAATGAGAAAACCGCCGGAGGAGATCAGCAGCAACTTCAAAG GAAACTGGCCCCAGCATGTGATCATCGCAGTTTTTCTCAAGCTAGAAAACC AAACAAATCCCGAGACTAACTGACACTGTGGTATCCTATTGgGTCTTATGTA ATGCCCCACTGTTGGCAGGTGCTGCATTGGCCCATTTCTAATAGGCTTTTGA AGTTTCAAATATTAATCGTCTGCAGGCAGCAGGATCACCAATCGATAGGCC AATAAAAAACAATACCCTGTGTCATCTGTGTATACAAGGCAAAGTGCGGTA TTTTGGATATATGTAAAAGCCAGCTTTACCgCTATTGGCAGGTGCTGTATTG GGCCATTTATAACAGACTTTTGAAGCTTCAAATATTAATGTGCTGCAATCAG AAGAATCGCCAACCGATAACCTAATAAAAAGATAATGCCTTATTATCTGCG TAAACCAGACATATTACGAATTAGTTTGACATCTCTAGACTTATATAAATGT

	TTAAATGGTTTGAATTACTATAAaTACTATAATAGgCATATGGAGAACTTAT GTGCATAATTTAATCCAAAAAATGTTACGGCGACAACTCAATAATTAATAA TGTTTGTGGATAATAGCATACAACTTTTTGGCAAtAAATGACAGTTTAGTTTC CTAATGATCTTGGTCACTTTGGATTCGCTTGACCCACGGCAGCTGATAACGC TGCTCCGGGAAAAATAACGATGAGATTGACTgCAACAGCAGCCAGAATATA TCTGTGGGCTGGGAATTAACACATTGGCCTCCTCCAGGCGCGTGCCCGTTTC ACCTTCCCGTGCATTTATGAAAAGTCTGCCGGACCGACCCACCGAACCGTCC GCTGTCTAGCTCTATACGAGTAGATGTGTATCTGGCTGTTAACCCCTCTGGA ATTGCCAAAAATGTGCAACGGCATCAGCGAAAGAACGGCAGGAAAGATGG TCGGAATCGGGGAGGGGTTCAAACTCCAAAGGATATAAGGGACATGGGAC ATGCAGCTAGTCGGGCACAATGGGATGGTCCCGATTCACTGTCAAGGTAAT TTATTAATTCGACAATAAAATCGCACGAGTTCTTGGGACGAGCTCCGGATTC GCATTCTATGTTGGGTTTGGGTGCGGACTTGCGATGGGGTGTGTTGTGGTCG GACACTCGCATGTGGGCCTTACAACTCAATTAAAGTTGCCTGGGTTCTGGAT CTGAGCTCGCTGCATTTCCGTGGAGAACCGGGTGGGCAAAGTATTATATGA CCATGTGCTAGGTTTGGATTAAGCTTAACAGAGGAATTTGTGAAATCCAATC GTTCGAAAGTGAAAAAGTTAAGTGCCCTTATGCCACGTAAACCCATCATTA AATATTTACAAAAACAAACCATTCATAAACAGCTAGACAAGTTAGTTTAAA ATCTTAGGCCATGTTCATGTAAAATAATTCAACTAATTAATCTTTTTAAGGT ATAATTCAATAAGTATATATCTAAATTCAAAGACCTAGTCAAAAGTAGTAG CCCTaGCCCTTATTTATTGCGTTAAAAAAGGATCCTAACAATATTTAAATAA AATGAGATGAAATTATTTTAGTAcATGATTCGGTATCGAAAGAATTTTATTT TTGGGGAGCTGTTTTTCCTTTCTTTGATTGATAtATATAAACATAAGTTATCG GATGGTTTTATATCCTACTCTACACCCTTATTGTTAAGTTTACTGCAAAACC AGAAAGGAGGCTCACATTCACTGGCAATATAGCAATCACAGGACATCGGAG AGATGCGAACTGCAGTCTCTTCTGTCAAACTGAAAAACGCCATAAATTTTGT GGAACAATGGAATCTGCCAGAGTATGCGGATGGGGTAAGGGTAGAATTCGC ACATGCACACATATGTCAGGAGTATCGATATGTTAGGGGTGTTCAACGGGG GTGGCTAATAAAATTTATCCAGGCCGCGTAAGCCGATCAGAAAGTCATGTA AACACAAAAGTGGCTGAGAGATGATTGTTTTTTACCAGTCAATTTAGCTGA GCTGATTTCAGGTTATACGCGAACCTAAGATGGAAGTAGATTTTTAGATTAC CAGAAATAACCTTCAGTTTCTAAAAGAATGAGAAATAGTAAATCGTGAAGC CTGTGGCATTCTTACCCCTTACTGAAATGAATGTAACAGGTAGAAGGCAGC AATTGCAGTGCTACAAAGTATATATATTCTTGATCCATATAAGTATAAGTAT AAGAGCTATAATGTTATAATTAAACATGAAGATCCTACGCTGTGCAAGTTTT GTGCCACAAATTTAACTCAAGATTACCACTCACAAAGGATACATTACCATTT GGCCATTAAAAATTACAGCATCCATAGACAACCTACTTAAAATTTATAAAA ACTTGAGGGTTATATCACAGAGTTACCGAAAAAAAGCGTAAAAGCTTTATA TTCTCAACAATATTATGCTATTAAAATATTGCTGGTTTTCTGCTGTTATAGAA TCATTTTTAAAAGTgTAACGTAAAAAATAAAATAAAACTAGTATTaATTTGA AAAATCAGCGGGCATATAATTTATATCATATTTTTAAAATTTAGGCAAAGG ATGTTTGCATAAAGTTTTTACTGTTTACTAGTCATTTTGGAAGTGCGTTTGTT GGTTTTTAGGCAAATACCGGGCACAGGAGTGAGTTTGGGAATCGGGAGTTG CGCACTTGCTTGGCCACGAGGGCAAACAAAAAGCGCAAACACGCGACCCTC GGCCACGCGTATTCCTGATCCCAGGGATCGGACGTAATGTTATCCTTTGGCC GCCCAGTGCCACGtAATAAATTCGGAGGGAAAGGGCATCGGGTTCCGGGAA CAACTGGCAGCCAGTCTTCGGTGTTTTGCGCGCTGGCAAAAATCCAGAGAA ATTTTTAGGGAACCATAAACGGGCCGGGGAAAAAGCCTCTGCgCCGAAGGA ACGTTTTCAGCAACAGTTTACAGTTTTTATGTCTTTATGATTATTGCAATTAG AGGGAGATCGGCTGAGAGTCGCGCCCTCTCGCTCTGCGCACCTCATAGGTA GGCACCTCATGGCCGTAATTACTGCAGCACCGTCTCAAGGTCGCCGAGTAG GAGAAGCGCGCGGGCGGATAAATCGCGATGATAATGGGCGCGATGGGTAG GTAATAAGCCGCGCAGCAGGTAGGCACCGTACGGATAAAGTTGCCAGGACC TCGGATAACTTCCCCTCTCCGTGCCTGCAAGGACATTTCGCCGGAGGGGTGG CTGCGAACAGCAGGCGGCAAAGTGTCATGCGCAGGGATATTTATGCGCTAT AACGcCGAGCGTGTGCCGAGGGCTCTCT

7SN-R transgene	AGAGAGCCCTCGGCACACGCTCGgCGTTATAGCGCATAAATATCCCTGCGC ATGACACTTTGCCGCCTGCTGTTCGCAGCCACCCCTCCGGCGAAATGTCCTT GCAGGCACGGAGAGGGGAAGTTATCCGAGGTCCTGGCAACTTTATCCGTAC GGTGCCTACCTGCTGCGCGGCTTATTACCTACCCATCGCGCCCATTATCATC GCGATTTATCCGCCCGCGCGCTTCTCCTACTCGGCGACCTTGAGACGGTGCT GCAGTAATTACGGCCATGAGGTGCCTACCTATGAGGTGCGCAGAGCGAGAG GGCGCGACTCTCAGCCGATCTCCCTCTAATTGCAATAATCATAAAGACATA AAAACTGTAAACTGTTGCTGAAAACGTTCCTTCGGcGCAGAGGCTTTTTCCC CGGCCCGTTTATGGTTCCCTAAAAATTTCTCTGGATTTTTGCCAGCGCGCAA AACACCGAAGACTGGCTGCCAGTTGTTCCCGGAACCCGATGCCCTTTCCCTC CGAATTTATTaCGTGGCACTGGGCGGCCAAAGGATAACATTACGTCCGATCC CTGGGATCAGGAATACGCGTGGCCGAGGGTCGCGTGTTTGCGCTTTTTGTTT GCCCTCGTGGCCAAGCAAGTGCGCAACTCCCGATTCCCAAACTCACTCCTGT GCCCGGTATTTGCCTAAAAACCAACAAACGCACTTCCAAAATGACTAGTAA ACAGTAAAAACTTTATGCAAACATCCTTTGCCTAAATTTTAAAAATATGATA TAAATTATATGCCCGCTGATTTTTCAAATtAATACTAGTTTTATTTTATTTTTT ACGTTAcACTTTTAAAAATGATTCTATAACAGCAGAAAACCAGCAATATTTT AATAGCATAATATTGTTGAGAATATAAAGCTTTTACGCTTTTTTTCGGTAAC TCTGTGATATAACCCTCAAGTTTTTATAAATTTTAAGTAGGTTGTCTATGGA TGCTGTAATTTTTAATGGCCAAATGGTAATGTATCCTTTGTGAGTGGTAATC TTGAGTTAAATTTGTGGCACAAAACTTGCACAGCGTAGGATCTTCATGTTTA ATTATAACATTATAGCTCTTATACTTATACTTATATGGATCAAGAATATATA TACTTTGTAGCACTGCAATTGCTGCCTTCTACCTGTTACATTCATTTCAGTAA GGGGTAAGAATGCCACAGGCTTCACGATTTACTATTTCTCATTCTTTTAGAA ACTGAAGGTTATTTCTGGTAATCTAAAAATCTACTTCCATCTTAGGTTCGCG TATAACCTGAAATCAGCTCAGCTAAATTGACTGGTAAAAAACAATCATCTC TCAGCCACTTTTGTGTTTACATGACTTTCTGATCGGCTTACGCGGCCTGGAT AAATTTTATTAGCCACCCCCGTTGAACACCCCTAACATATCGATACTCCTGA CATATGTGTGCATGTGCGAATTCTACCCTTACCCCATCCGCATACTCTGGCA GATTCCATTGTTCCACAAAATTTATGGCGTTTTTCAGTTTGACAGAAGAGAC TGCAGTTCGCATCTCTCCGATGTCCTGTGATTGCTATATTGCCAGTGAATGT GAGCCTCCTTTCTGGTTTTGCAGTAAACTTAACAATAAGGGTGTAGAGTAGG ATATAAAACCATCCGATAACTTATGTTTATATaTATCAATCAAAGAAAGGAA AAACAGCTCCCCAAAAATAAAATTCTTTCGATACCGAATCATgTACTAAAAT AATTTCATCTCATTTTATTTAAATATTGTTAGGATCCTTTTTTAACGCAATAA ATAAGGGCtAGGGCTACTACTTTTGACTAGGTCTTTGAATTTAGATATATACT TATTGAATTATACCTTAAAAAGATTAATTAGTTGAATTATTTTACATGAACA TGGCCTAAGATTTTAAACTAACTTGTCTAGCTGTTTATGAATGGTTTGTTTTT GGCACTTAACTTTTTCACTTTCGAACGATTGGATTTCACAAATTCCTCTGTTA AGCTTAATCCAAACCTAGCACATGGTCATATAATACTTTGCCCACCCGGTTC TCCACGGAAATGCAGCGAGCTCAGATCCAGAACCCAGGCAACTTTAATTGA GTTGTAAGGCCCACATGCGAGTGTCCGACCACAACACACCCCATCGCAAGT CCGCACCCAAACCCAACATAGAATGCGAATCCGGAGCTCGTCCCAAGAACT CGTGCGATTTTATTGTCGAATTAATAAATTACCTTGACAGTGAATCGGGACC ATCCCATTGTGCCCGACTAGCTGCATGTCCCATGTCCCTTATATCCTTTGGA GTTTGAACCCCTCCCCGATTCCGACCATCTTTCCTGCCGTTCTTTCGCTGATG CCGTTGCACATTTTTGGCAATTCCAGAGGGGTTAACAGCCAGATACACATCT ACTCGTATAGAGCTAGACAGCGGACGGTTCGGTGGGTCGGTCCGGCAGACT TTTCATAAATGCACGGGAAGGTGAAACaGGCACGCGCCTGGAGGAGGCCAA TGTGTTAATTCCCAGCCCACAGATATATTCTGGCTGCTGTTGAAGTCAATCT

	CATCGTTATTTTTCCCGGAGCAGCGTTATCAGCTGCCGTGGGTCAAGCGAAT CCAAAGTGACCAAGATCATTAGGAAACTAAACTGTCATTTCTTGCCAAAAA GTTGTATGCTATTATCCACAAACATTATTAATTATTGAGTTGTCGCCGTAAC ATTTTTTGGATTAAATTATGCACATAAGTTCTCCATATGcCTATTATAGTAtTT ATAGTAATTCAAACCATTTAAACATTTATATAAGTCTAGAGATGTCAAACTA ATTCGTAATATGTCTGGTTTACGCAGATAATAAGGCATTATCTTTTTATTAG GTTATCGGTTGGCGATTCTTCTGATTGCAGCACATTAATATTTGAAGCTTCA AAAGTCTGTTATAAATGGCCCAATACAGCACCTGCCAATAGcGGTAAAGCT GGCTTTTACATATATCCAAAATACCGCACTTTGCCTTGTATACACAGATGAC ACAGGGTATTGTTTTTTATTGGCCTATCGATTGGTGATCCTGCTGCCTGCAG ACGATTAATATTTGAAACTTCAAAAGCCTATTAGAAATGGGCCAATGCAGC ACCTGCCAACAGTGGGGCATTACATAAGACcCAATAGGATACCACAGTGTC AGTTAGTCTCGGGATTTGTTTGGTTTTCTAGCTTGAGAAAAACTGCGATGAT CACATGCTGGGGCCAGTTTCCTTTGAAGTTGCTGCTGATCTCCTCCGGCGGT TTTCTCATTATATACCAAATACCACgTAATTGGGGAAGGGGTGATAAAAACA ATCGAGTATATCACACTCATTATCACAGTGAAACAGGCGACGACCCGCTGG GAATAAAATAGTAAAGGGCCGAGAAAAATTATCAACAAACAAAACAAACA ACTCGAGATTGGTGTAAATTCCCCGACTATTCCGATACACCGTCTATATATA GTATCAGCGTCAATCGGTCTATGGTGGGTTCGGGGATCCCAGCGTGAGGTTT TCTTTGGTTCGGATCGCAGGCAATCAATTTGATTGCCTCAATCGGCGGCTAA TTAGTAGCGATGCTCGAaAATTATTCAAAATTTCTTTTGCGTTATtGGATTTG TTTGAGTTAACAATCAAACTGCACGCAGATCGGATCGAGTCGTTAAGTCTG CGCTGGGGAAATCATCTTGATGAATTTGCCAGATGTCTAGATAAATGAAGT AATCTTCAGTTTCTAAAATTCACTATAaAtTGAAGAGAAAGCCAGACAGGAT GCATTTCTGGGTGACCTTCGATCGGAGGCCCACCGACGACTCCTCCGGAGG TCATCTCAATGGGTGGATATAGCGTTTCGTCCTTGGATCGGCACCAATCAAA GTCAAGGATGCCGAAACGCTGTGCTACTTTTGAGATATCCTCAATGTTTTGT ACAAATATTTTCCCAATGCGTCACGCGTCGCCTGTGTCCTTGGCTTCTGTTC GAGGTCCTGTGCGTCCTACATATCCTTGAATTGCCTGCGATGTCTTGAGCAC TGGTGTGTTGACATCGAAGACATCAGGATAAGGCTGTAATTGCTCCTGTCAC ACTTATTCGAATTGTCCTGCCGCTTTtaaaTAGCAATGATGACCCCGTAAGCTG AATGAAATGTGAACTGTGTGCATCATCGCATCGTACTGCTTATGGCCCCTAG ATATAAATTTCGAATTGTTTAAATCTCAGCTAGGACCTCGTTAACACCTTCT GTTACGGACGCTCCAACTGAGCTCAAAGATTGATGGCTTTTTGAAAAATGTC AGAGTGTCATCAGTTAAAAACTATTGATAAATAATGTCATGTGAAACTATG TTAATTCTCCTAATACCTTGTACCTTGTGTAGTAATTTAAAGTTTCGATTCTT TAAGCAGGCTCGAAGGTCCGTGCACTATAACAATTTTAATAATTTAAAGGA TTATACAATTTGGAATATAAAATGCAAACTGATGAAACGGAAATAGAAATA AAAGCATTCACATAATTTTAATCTATAATATAATGAAGCTGAACATGAACT GAATCAGATCTTAATACCTAACATGGGCATATCTTTTATATATGTAGATATG TATGTCCCGTAGCTATTCGACATCGTGGGACATCAACATCTAATAAGGATA ACGTCCATTAACAATGATCGGAACGGAGGTCTCCGGGGAGTTCGGCGTAGG CTGTCAATGCGATTCCGCTGACGAGTCGTACGCCGGGGAATCGAAGTCGCG CGCGCTGCTCGAAAATTCGGCGAATTCGCTGATGGCATTGGATGGTCTCAGT TTGAATCTCAGAGCTCGCAATCCGAGGTGAAGAGCTCAACTGAACTGGACA GGAAAACGAAGGAAGGAGGAAGGGGGCCCCCTGCGGCCTGTTTTATAGTTT TATGATGTGCCAATAATATTTCCGCCTTGGAAACCAATTGCGCTGCTTCTCG GCGAAGGATACGCGGGCAATTAAGACGCTTAAGGACCCATGACACGGTGC ATTTACTATCCATAAACAGACGACTCAAGTCGGTCCAAGTCGAGTAGCGCG AGCCTCGTCCTCGATTCGGCCACGCATTTCCATTTCCCGTCCCGTTCAACAG AACATCAGTCTAATGCGAGACACACGCGTCTTCCCTTCCCAGGTAGGCCTTG ATTCGTCCGTGGAAATCAGAGAAATCAGAGGTGGAAATGGATGGAAATGG CGGCGAAAGGGATTTTCGCTCTGTTTACCCAAAACAGACTTTCCTTTAGCCA ACGCCCCCGTCTCACAAGAAATTGAA

7SN-Z-R transgene (highlight indicates region missing in 7SN-R)	CCAAGCAAGTGCGCAACTCCCGATTCCCAAACTCACTCCTGTGCCCGGTATT TGCCTAAAAACCAACAAACGCACTTCCAAAATGACTAGTAAACAGTAAAAA CTTTATGCAAACATCCTTTGCCTAAATTTTAAAAATATGATATAAATTATAT GCCCGCTGATTTTTCAAATGAATACTAGTTTTATTTTATTTTTTACGTTATAC TTTTAAAAATGATTCTATAACAGCAGAAAACCAGCAATATTTTAATAGCAT AATATTGTTGAGAATATAAAGCTTTTACGCTTTTTTTCGGTAACTCTGTGAT ATAACCCTCAAGTTTTTATAAATTTTAAGTAGGTTGTCTATGGATGCTGTAA TTTTTAATGGCCAAATGGTAATGTATCCTTTGTGAGTGGTAATCTTGAGTTA AATTTGTGGCACAAAACTTGCACAGCGTAGGATCTTCATGTTTAATTATAAC ATTATAGCTCTTATACTTATACTTATATGGATCAAGAATATATATACTTTGT AGCACTGCAATTGCTGCCTTCTACCTGTTACATTCATTTCAGTAAGGGGTAA GAATGCCACAGGCTTCACGATTTACTATTTCTCATTCTTTTAGAAACTGAAG GTTATTTCTGGTAATCTAAAAATCTACTTCCATCTTAGGTTCGCGTATAACCT GAAATCAGCTCAGCTAAATTGACTGGTAAAAAACAATCATCTCTCAGCCAC TTTTGTGTTTACATGACTTTCTGATCGGCTTACGCGGCCTGGATAAATTTTAT TAGCCACCCCCGTTGAACACCCCTAACATATCGATACTCCTGACATATGTGT GCATGTGCGAATTCTACCCTTACCCCATCCGCATACTCTGGCAGATTCCATT GTTCCACAAAATTTATGGCGTTTTTCAGTTTGACAGAAGAGACTGCAGTTCG CATCTCTCCGATGTCCTGTGATTGCTATATTGCCAGTGAATGTGAGCCTCCT TTCTGGTTTTGCAGTAAACTTAACAATAAGGGTGTAGAGTAGGATATAAAA CaATCCGATAACTTATGTTTATATCTATCAATCAAAGAAAGGAAAAACAGCT CCCCAAAAATAAAATTCTTTCGATACCGAATCATTTACTAAAATAATTTCAT CTCATTTTATTTAAATATTGTTAGGATCCTTTTTTAACGCAATAAATAAGGG CGAGGGCTACTACTTTTGACTAGGTCTTTGAATTTAGATATATACTTATTGA ATTATACCTTAAAAAGATTAATTAGTTGAATTATTTTACATGAACATGGCCT AAGATTTTAAACTAACTTGTCTAGCTGTTTATGAATGGTTTGTTTTTGTAAAT ATTTAATGATGGGTTTACGTGGCATAAGGGCACTTAACTTTTTCACTTTCGA ACGATTGGATTTCACAAATTCCTCTGTTAAGCTTAATCCAAACCTAGCACAT GGTCATATAATACTTTGCCCACCCGGTTCTCCACGGAAATGCAGCGAGCTCA GATCCAGAACCCAGGCAACTTTAATTGAGTTGTAAGGCCCACATGCGAGTG TCCGACCACAACACACCCCATCGCAAGTCCGCACCCAAACCCAACATAGAA TGCGAATCCGGAGCTCGTCCCAAGAACTCGTGCGATTTTATTGTCGAATTAA TAAATTACCTTGACAGTGAATCGGGACCATCCCATTGTGCCCGACTAGCTGC ATGTCCCATGTCCCTTATATCCTTTGGAGTTTGAACCCCTCCCCGATTCCGAC CATCTTTCCTGCCGTTCTTTCGCTGATGCCGTTGCACATTTTTGGCAATTCCA GAGGGGTTAACAGCCAGATACACATCTACTCGTATAGAGCTAGACAGCGGA CGGTTCGGTGGGTCGGTCCGGCAGACTTTTCATAAATGCACGGGAAGGTGA AACGGGCACGCGCCTGGAGGAGGCCAATGTGTTAATTCCCAGCCCACAGAT ATATTCTGGCTGCTGTTGcAGTCAATCTCATCGTTATTTTTCCCGGAGCAGCG TTATCAGCTGCCGTGGGTCAAGCGAATCCAAAGTGACCAAGATCATTAGGA AACTAAACTGTCATTTaTTGCCAAAAAGTTGTATGCTATTATCCACAAACAT TATTAATTATTGAGTTGTCGCCGTAACATTTTTTGGATTAAATTATGCACAT AAGTTCTCCATATGcCTATTATAGTAtTTATAGTAATTCAAACCATTTAAACA TTTATATAAGTCTAGAGATGTCAAACTAATTCGTAATATGTCTGGTTTACGC AGATAATAAGGCATTATCTTTTTATTAGGTTATCGGTTGGCGATTCTTCTGA TTGCAGCACATTAATATTTGAAGCTTCAAAAGTCTGTTATAAATGGCCCAAT ACAGCACCTGCCAATAGcGGTAAAGCTGGCTTTTACATATATCCAAAATACC GCACTTTGCCTTGTATACACAGATGACACAGGGTATTGTTTTTTATTGGCCT ATCGATTGGTGATCCTGCTGCCTGCAGACGATTAATATTTGAAACTTCAAAA GCCTATTAGAAATGGGCCAATGCAGCACCTGCCAACAGTGGGGCATTACAT AAGACcCAATAGGATACCACAGTGTCAGTTAGTCTCGGGATTTGTTTGGTTT TCTAGCTTGAGAAAAACTGCGATGATCACATGCTGGGGCCAGTTTCCTTTGA AGTTGCTGCTGATCTCCTCCGGCGGTTTTCTCATTATATACCAAATACCACgT AATTGGGGAAGGGGTGATAAAAACAATCGAGTATATCACACTCATTATCAC AGTGAAACAGGCGACGACCCGCTGGGAATAAAATAGTAAAGGGCCGAGAA AAATTATCAACAAACAAAACAAACAACTCGAGATTGGTGTAAATTCCCCGA

	CTATTCCGATACACCGTCTATATATAGTATCAGCGTCAATCGGTCTATGGTG GGTTCGGGGATCCCAGCGTGAGGTTTTCTTTGGTTCGGATCGCAGGCAATCA ATTTGATTGCCTCAATCGGCGGCTAATTAGTAGCGATGCTCGAaAATTATTC AAAATTTCTTTTGCGTTATtGGATTTGTTTGAGTTAACAATCAAACTGCACGC AGATCGGATCGAGTCGTTAAGTCTGCGCTGGGGAAATCATCTTGATGAATTT GCCAGATGTCTAGATAAATGAAGTAATCTTCAGTTTCTAAAATTCACTATAa AtTGAAGAGAAAGCCAGACAGGATGCATTTCTGGGTGACCTTCGATCGGAG GCCCACCGACGACTCCTCCGGAGGTCATCTCAATGGGTGGATATAGCGTTTC GTCCTTGGATCGGCACCAATCAAAGTCAAGGATGCCGAAACGCTGTGCTAC TTTTGAGATATCCTCAATGTTTTGTACAAATATTTTCCCAATGCGTCACGCGT CGCCTGTGTCCTTGGCTTCTGTTCGAGGTCCTGTGCGTCCTACATATCCTTGA ATTGCCTGCGATGTCTTGAGCACTGGTGTGTTGACATCGAAGACATCAGGAT AAGGCTGTAATTGCTCCTGTCACACTTATTCGAATTGTCCTGCCGCTTTtaaaT AGCAATGATGACCCCGTAAGCTGAATGAAATGTGAACTGTGTGCATCATCG CATCGTACTGCTTATGGCCCCTAGATATAAATTTCGAATTGTTTAAATCTCA GCTAGGACCTCGTTAACACCTTCTGTTACGGACGCTCCAACTGAGCTCAAAG ATTGATGGCTTTTTGAAAAATGTCAGAGTGTCATCAGTTAAAAACTATTGAT AAATAATGTCATGTGAAACTATGTTAATTCTCCTAATACCTTGTACCTTGTG TAGTAATTTAAAGTTTCGATTCTTTAAGCAGGCTCGAAGGTCCGTGCACTAT AACAATTTTAATAATTTAAAGGATTATACAATTTGGAATATAAAATGCAAA CTGATGAAACGGAAATAGAAATAAAAGCATTCACATAATTTTAATCTATAA TATAATGAAGCTGAACATGAACTGAATCAGATCTTAATACCTAACATGGGC ATATCTTTTATATATGTAGATATGTATGTCCCGTAGCTATTCGACATCGTGG GACATCAACATCTAATAAGGATAACGTCCATTAACAATGATCGGAACGGAG GTCTCCGGGGAGTTCGGCGTAGGCTGTCAATGCGATTCCGCTGACGAGTCG TACGCCGGGGAATCGAAGTCGCGCGCGCTGCTCGAAAATTCGGCGAATTCG CTGATGGCATTGGATGGTCTCAGTTTGAATCTCAGAGCTCGCAATCCGAGGT GAAGAGCTCAACTGAACTGGACAGGAAAACGAAGGAAGGAGGAAGGGGG CCCCCTGCGGCCTGTTTTATAGTTTTATGATGTGCCAATAATATTTCCGCCTT GGAAACCAATTGCGCTGCTTCTCGGCGAAGGATACGCGGGCAATTAAGACG CTTAAGGACCCATGACACGGTGCATTTACTATCCATAAACAGACGACTCAA GTCGGTCCAAGTCGAGTAGCGCGAGCCTCGTCCTCGATTCGGCCACGCATTT CCATTTCCCGTCCCGTTCAACAGAACATCAGTCTAATGCGAGACACACGCGT CTTCCCTTCCCAGGTAGGCCTTGATTCGTCCGTGGAAATCAGAGAAATCAGA GGTGGAAATGGATGGAAATGGCGGCGAAAGGGATTTTCGCTCTGTTTACCC AAAACAGACTTTCCTTTAGCCAACGCCCCCGTCTCACAAGAAATTGAA
Scr^{ftz}-XbaI^{2} transgene	CTAGACATCTGGCAAATTCATCAAGATGATTTCCCCAGCGCAGACTTAACG ACTCGATCCGATCTGCGTGCAGTTTGATTGTTAACTCAAACAAATCCAATAA CGCAAAAGAAATTTTGAATAATTtTCGAGCATCGCTACTAATTAGCCGCCGA TTGAGGCAATCAAATTGATTGCCTGCGATCCGAACCAAAGAAAACCTCACG CTGGGATCCCCGAACCCACCATAGACCGATTGACGCTGATACTATATATAG ACGGTGTATCGGAATAGTCGGGGAATTTACACCAATCTCGAGTTGTTTGTTT TGTTTGTTGATAATTTTTCTCGGCCCTTTACTATTTTATTCCCAGCGGGTCGT CGCCTGTTTCACTGTGATAATGAGTGTGATATACTCGATTGTTTTTATCACCC CTTCCCCAATTAcGTGGTATTTGGTATATAATGAGAAAACCGCCGGAGGAGA TCAGCAGCAACTTCAAAGGAAACTGGCCCCAGCATGTGATCATCGCAGTTT TTCTCAAGCTAGAAAACCAAACAAATCCCGAGACTAACTGACACTGTGGTA TCCTATTGgGTCTTATGTAATGCCCCACTGTTGGCAGGTGCTGCATTGGCCC ATTTCTAATAGGCTTTTGAAGTTTCAAATATTAATCGTCTGCAGGCAGCAGG ATCACCAATCGATAGGCCAATAAAAAACAATACCCTGTGTCATCTGTGTAT ACAAGGCAAAGTGCGGTATTTTGGATATATGTAAAAGCCAGCTTTACCGCT ATTGGCAGGTGCTGTATTGGGCCATTTATAACAGACTTTTGAAGCTTCAAAT ATTAATGTGCTGCAATCAGAAGAATCGCCAACCGATAACCTAATAAAAAGA TAATGCCTTATTATCTGCGTAAACCAGACATATTACGAATTAGTTTGACATC T

Scr^{ftz}-XbaI/HindIII transgene	ACATCTGGCAAATTCATCAAGATGATTTCCCCAGCGCAGACTTAACGACTC GATCCGATCTGCGTGCAGTTTGATTGTTAACTCAAACAAATCCaATAACGCA AAAGAAATTTTGAATAATTtTCGAGCATCGCTACTAATTAGCCGCCGATTGA GGCAATCAAATTGATTGCCTGCGATCCGAACCAAAGAAAACCTCACGCTGG GATCCCCGAACCCACCATAGACCGATTGACGCTGATACTATATATAGACGG TGTATCGGAATAGTCGGGGAATTTACACCAATCTCGAGTTGTTTGTTTTGTT TGTTGATAATTTTTCTCGGCCCTTTACTATTTTATTCCCAGCGGGTCGTCGCC TGTTTCACTGTGATAATGAGTGTGATATACTCGATTGTTTTTATCACCCCTTC CCCAATTAcGTGGTATTTGGTATATAATGAGAAAACCGCCGGAGGAGATCA GCAGCAACTTCAAAGGAAACTGGCCCCAGCATGTGATCATCGCAGTTTTTCT CAAGCTAGAAAACCAAACAAATCCCGAGACTAACTGACACTGTGGTATCCT ATTGgGTCTTATGTAATGCCCCACTGTTGGCAGGTGCTGCATTGGCCCATTTC TAATAGGCTTTTGAAGTTTCAAATATTAATCGTCTGCAGGCAGCAGGATCAC CAATCGATAGGCCAATAAAAAACAATACCCTGTGTCATCTGTGTATACAAG GCAAAGTGCGGTATTTTGGATATATGTAAAAGCCAGCTTTACCgCTATTGGC AGGTGCTGTATTGGGCCATTTATAACAGACTTTTGAAGCTT
ftz-7 transgene	ATGTTTCTCGTTCGAGGTTGGAGTTGAATTCGGAGTCGGAGTTGGCTGGCTG GCTGCCTTGCTGGTGAAGGGATTGGCCTTCGTTTCACCCATCACCGCCAGAA AATTTCCCGGGCTGATTAGCATGAGGCGTAACAGTTTTTAATTGCAGCCACA TTTCGCCGAGCTCCGAATGTGCAGCAACAAAAGCCTGGCAAACACATAATG ATAGCAATACTTATCTTGCCATCAGAAAGTtGAAAAAGTTTAAGCACTTTTT CAATTTTTTTGTAACAACCACCAATGCTTCTATGCGACAGCTTTCCCTCTACT GTGACTTCTTTAGATGTTTTCaGCTGGATTGATATATTTTATTTCGTTTAAAA GATTGTTAAAAGAAAACacAACTAatAaTGtTTTTCATACGATACATTCCATTTT ATTGTAATTTTGTGTACCCATTGACAATTTTACAAAGAATACACTCCTACAA TATCACTGAACAGTTGGATAATAAAAAAACTCAGAATGTTTTTTTTtAAATG AGTGTTCATAATAAACTTGAACGCAGTCTTGCCGAGTTTTTTAAGCGACAAT ATGTCTATTAAACATATTGATGTAGTGgCATTTAAAAATTATTTGTGCTATTA tTTTTTTAAATAAGTACTCCTTTTTAGCTTTCATAGTTCCCGAGATCACATCG TTCATATGGACAGACaGATTGACAGATAAGCCCAGATTGAATTGGCCAGTA ATCCCGATAAAGAATACGTATTCACTTTCATCCTTTGATACATACTTTTCAT ACTTTATACCCTTTTACTCTACGAGGAACGGGAGTAATTGTTGGAAATGTTT ATTCACCACTCCCAGTTATTGCATTTATACAAATTATTACTTGACTAGATTTT ATAATCAAAATATTTAATTAAAATCCtCCGGTCATTATGTCCTGCTCTACATA TAATTTAGACCTCCCATCGGGTgTTGACTGTCCGTAGCTGTACTATCATTATC CTCCTAAGGCTGTCAAGTTATCTACGCCGCTTGAGGCCATTGTCAgTGTCTTC GCCGGAACTGCTACAAAAAATTGTGACAATTTACACTGATCTCTGGAACCT GGTCCACAGACCCGTCCTTGACTGCCATAAAGTCAGTTCCACTAAAAATCGT CGAAAAATACTCCCGTGAACTAGAAAAAATAAAGCACACAAAAAATGTGC CTCAGTTCTAGTTTAGTTTCAAGAAATTCTCTGTTCGCTCGGAAGGAAAAAA TGGAAACTAGAATACGCGGGACAATAAAAAAACGATGCGCTGTAAAAACG TTTTAGACCGAATTCACCTTTGTGTTTCGCGCATTCTCGTCATTGTTAAAAAA AATCAAGCATTTTTTTGGGTCGCGACACAATTTTTTGCCGTAGTGTCTCCAG GATTCCTCTAGATCCTGGTCTCGTTCTCAGGCCGTGGCCTCGTTTTCCATTTA AAATGCGCGACGCACAAAAGTTTTTACACGAATTTATGTTTTTATCGCTTTA CGACCGGAGGTCCTAATAAAACTTTATTGAGGCTGCCGG

ftz-6 transgene	TGCCGGTGGTCGTAGGAGTTTTGCTCGAGTTTCTTACCGTAGTAGGCCTCCT TTGGTAGCCTCAAAACTCAAAACTGTAGCAgGGTCAGGTTCCTCGGGATAA GCGCGTATCCTGCGAATTAAGCCAGCCCAAGGACGAGCCGGGTTATGCAAG TCTATGCAACGTATGGGTTTGCATAATAGAGGTAAATGTTTGTTTAACTGGG AACTGCTCTAGTTAGCATTGTCTTGATGGCATTGTTTCTGCCCTCACCCCGGT ACCGCCCACGAGCCGGGAAGACAGTCAATAAAACGTCACCTTCACCTCTTA CACAATATTTCACAACGTTCTTGCAAATCCGGGGATTAGGAACAATTCCGG CGTTCTCAAAAGGTTTTGGGCATCATGCCCTTCCGTAGTCGCCGTTTGCTCT AATTTAAATTACCATTCTGCATTGTTTTGGAATGTTTCCTAAAAACTGACCT AATTTACGCCACACACTTTTCGTTACAAGTCTTTCGTTTTTATGACCCAAAA ATGGTAAAAGTTTCCGCGAATTTAATCATTTTTATGATGTGAACGATTTTTT GAATAGATGTTGCGAATTCCGAAATCGGGAAATAGAAATATGATTAAAATA GCATTAGaCAATGCTTTGGTTATTTTTATGACTTAAATTGAGTTAGAATGTAA TTTTAGATTTTAATTGAAAAATATTTTATTTAACATATAGTAACTAAGAATT TTTTTTATTAATAATTATATATGTATATATATATATTATATAAGTATATTATT ATATATATATTTTATTTTATTATTATTATATTATATTATTAAATTTTTTTAGCT TTTCTTATAATTTTAATATTTTTATATATTTTTTTTTTGCAAAAGATAAATTA ACAAACCAATATTTCCGCTTCACAAGCAATCCAAGCCCCGATAATAAAATG TTCTATATTTCTTcCTAAAAATATATACTATTATTAATTATGTTTTATTACTAT ATATGTATTTATATTTGCAATGATTATTGTTCAATCGAAGAAAAAATATTGC ATGCAATACTTTGGGGTGTTCTTAGTTTTTCTGAATATATTTTTTATTTATTTT ACTTTAGTTTTGTTATCTTAACAAATAATCTTAATTTATAATATAGTTATAAA CAAATGTTCAATCGTTTTTTCCAGACAAAATGTCTTTTACATTCGCAGTTCA ACATCTTAATCCTTTTGCGAAACATTACTGCCATTACAT
ftz413 transgene	GGTACCGCCCACGAGCCGGGAAGACAGTCAATAAAACGTCACCTTCACCTC TTACACAATATTTCACAACGTTCTTGCAAATCCGGGGATTAGGAACAATTCC GGCGTTCTCAAAAGGTTTTGGGCATCATGCCCTTCCGTAGTCGCCGTTTGCT CTAATTTAAATTACCATTCTGCATTGTTTTGGAATGTTTCCTAAAAACTGAC CTAATTTACGCCACACACTTTTCGTTACAAGTCTTTCGTTTTTATGACCCAAA AATGGTAAAAGTTTCCGCGAATTTAATCATTTTTATGATGTGAACGATTTTT TGAATAGATGTTGCGAATTCCGAAATCGGGAAATAGAAATATGATTAAAAT AGCATTAGaCAATGCTTTGGTTATTTTTATGACTTAAATTGAGTTAGAAT

ftz 3' element transgene	CTATTTATTTAAATATACGAGTAAAGTAAATCGATCGAATTTAAACAAATCA AGTTGAACATTCATTTGGCAATTTGTGAAGAAGAGTCTTGGGCATGCTGCA ATTTGACTGCTTTAAAATTTTAAATTTATAGGCCGTGGGCCGCGTATGTGGA ATACATTTCATATGTATATGTGTTGAAATAcAATTAAATGCCTTTCAATGAT AACTACTCAATAAATTCTTAAAATAGTTTGGATCAACGTATGGTATTTGTCT AACTACTTTCACCGAAAACATATCACAAGGCAAAGTTTACCTTCAGCCTATT TaCATACCTTATATATGCATATAATACATATATATGATAAAAATGGAGTTTA GGAGTTAAGGAGTTTAcATTATTTAAGTATAAGCCTAATATTTAATTGTTCG TTGAGAAGCTTAAGAGATCTCTTGGTTTACTTTCTGATGAAgTGCACATGGT TGAAGCATTACTTTCGAACTTATACGAAACGCAAACGATTTAATGTTGAGC ACGAATCGTACAAATTCGAGCAGCTGCATTTTGTCGCTTCAGTCCCCCTCAT CCCTGACCCATTGCTgTCTCCCGGATTTTCTATTAAATGCACTCTTTTCGCCA GAGAAAATGTCACATTTTGGTCTGGCTTCGGGGCATATCTACCACCGCATCC CTGCTCCCTTCCTCCCTCCGaCGCTGCACGTTCCTCTATTGAAGTGAGACATT GATTGGTAATTTTTCATTGCACATCCGTGACAGTTATGGGTAACGCAACGCA AAAGGAAAAGCCCGGTGCGGAATCGGATTCGGAATCAGAATCAATATCAA AGGCAAAGGGGATCCATGTGCAGTCGAATATATCCCGGCACCATTTATCCG ACTCATTAGACAAAAGTGGTCGGTGGGTGAGTTCGAAGTGGCATCTGTGGG CACTTTAGTGTTAGGTGCGAATTGGTGGCATTGGTTAAGCTTAGTGGTGATA TGTATCTATGCCGCTGCGAAAGAAATTTTTTGTCTGATTTAATGATTTAGAT GGCTTCAATGTGCTACTTAATTGATTTGGAAACCATGTACATTTTGTTCAAA TATTTTATAAAATAAATATTTAACAATGAACAaACCTTCAATAGATCAGATT CTATCCTACtCTTCTAGAATaTAGTTATGACAGACAAATTAGTTAACAGATAT AAATCAAAAGACCGAAAACGGAAGAATCCCTTTGCAAATTCGAATTTGTaG CATTTTTGCAATTATATAATTAGGCCATAAATATATAAAtGATGGTGGTTGG GTAAAAAAGCCCTACAGTGACATACATATCTACAATACAAACTACTTGGCA AATACAGTATGCGTCAGAAATATAAGTCAACGAAGTATTTATATTTTCTAAA TAGTCAACTTTTATTAAACCCGTTACGCGTATTAGAAAAGGGTATACTTGAT CCGTTGAAAAGTATGTAACAGCTAAAAGTTTTTAAAATAAAAGTTTTTTTTC CGTATAAGCTAGGATTCCAGAGATGCACACACACAATTAATACGTTGAAAT TTCTTGTTCGTAATCCTACTAGGTGATAGTCGAAAAACTCGACAATACCATT CCCTCTAGTTGTTTTGCAAACATCAACCAAGCGGATTTTACTGGTTCCTTTG ACTCCTTAAAGTTATCAATTAGTCAACATTAAACAAACGAAATTGCAAGGA CATTTACCCGTTTCTcAAAAGTCAACATCTCTTTTCATACGTTTCCTTGGAAG TTAAAGAAAAACTCAAGAAACTGGTAAAATATTTACGTGTAAATTGCGAAA GAGTTACGTTAACTCTTCCTGTGTCCTTCTGCGAAACGAAGTCTTTTTTATCC TTTTTATTCGCATTCTGCTTTATTAGAGGTCCTCGGAGTCGGTTCTTAACTCA TTCGCGAGCCATAAACCCCGAGATTCGCGGAAAGGGTATTCGTTGTGGGCG GAATCCTTTTGCAAACCCGGTTTCAGTCAAGGCTGCAAAGTCAATCCTTTAG CATAATCCTTTAAATAGCCAAAATAAAGCCCGCCTTTGTCAAAACAAAGCG ATTCGCTGGCAGAACTCGTATTTGCGAATAGGCGTGCCATTGTGTCCACTTA GCCGGCAG

ftz+3 transgene	CCGCTGCACGTTCCTCTATTGAAGTGAGACATTGATTGGTAATTTTTCATTG CACATCCGTGACAGTTATGGGTAACGCAACGCAAAAGGAAAAGCCCGGTGC GGAATCGGATTCGGAATCAGAATCAATATCAAAGGCAAAGGGGATCCATGT GCAGTCGAATATATCCCGGCACCATTTATCCGACTCATTAGACAAAAGTGG TCGGTGGGTGAGTTCGAAGTGGCATCTGTGGGCACTTTAGTGTTAGGTGCG AATTGGTGGCATTGGTTAAGCTTAGTGGTGATATGTATCTATGCCGCTGCGA AAGAAATTTTTTGTCTGATTTAATGATTTAGATGGCTTCAATGTGCTACTTA ATTGATTTGGAAACCATGTACATTTTGTTCAAATATTTTATAAAATAAATAT TTAACAATGAACAaACCTTCAATAGATCAGATTCTATCCTACtCTTCTAGAAT aTAGTTATGACAGACAAATTAGTTAACAGATATAAATCAAAAGACCGAAAA CGGAAGAATCCCTTTGCAAATTCGAATTTGTaGCATTTTTGCAATTATATAAT TAGGCCATAAATATATAAAtGATGGTGGTTGGGTAAAAAAGCCCTACAGTG ACATACATATCTACAATACAAACTACTTGGCAAATACAGTATGCGTCAGAA ATATAAGTCAACGAAGTATTTATATTTTCTAAATAGTCAACTTTTATTAAAC CCGTTACGCGTATTAGAAAAGGGTATACTTGATCCGTTGAAAAGTATGTAA CAGCTAAAAGTTTTTAAAATAAAAGTTTTTTTTCCGTATAAGCTAGGATTCC AGAGATGCACACACACAATTAATACGTTGAAATTTCTTGTTCGTAATCCTAC TAGGTGATAGTCGAAAAACTCGACAATACCATTCCCTCTAGTTGTTTTGCAA ACATCAACCAAGCGGATTTTACTGGTTCCTTTGACTCCTTAAAGTTATCAAT TAGTCAACATTAAACAAACGAAATTGCAAGGACATTTACCCGTTTCTcAAAA GTCAACATCTCTTTTCATACGTTTCCTTGGAAGTTAAAGAAAAACTCAAGAA ACTGGTAAAATATTTACGTGTAAATTGCGAAAGAGTTACGTTAACTCTTCCT GTGTCCTTCTGCGAAACGAAGTCTTTTTTATCCTTTTTATTCGCATTCTGCTT TATTAGAGGTCCTCGGAGTCGGTTCTTAACTCATTCGCGAGCCATAAACCCC GAGATTCGCGGAAAGGGTATTCGTTGTGGGCGGAATCCTTTTGCAAACCCG GTTTCAGTCAAGGCTGCAAAGTCAATCCTTTAGCATAATCCTTTAAATAGCC AAAATAAAGCCCGCCTTTGTCAAAACAAAGCGATTCGCTGGCAGAACTCGT ATTTGCGAATAGGCGTGCCATTGTGTCCACTTAGCCGGCAGGTACCGCCTCC TCAAAACCCTCGGTCGCGGAGGACTCCTTGGTCTGCATCTCTATCCCTTCCA CCAGCCTCAGCCATAACGATCGTAAAGTTATCCTTTCGCAAACACGAGCATt CGGTTACCGCCTTAAGCTGGCGAAGGAGTTACAAAGTCATAAAAACAAAAC ACCTTTGTCCGgGCAGCATTGTCAGTTGGCGAATTATCCGTTCAGTCGCCGC TGCCGAATCCTTAATCCGTGCCGTTTTTATAGCCAGTAAGTAGATGGAGATA GTAGGAAATTCGCGATGGGCTCGTCCCTGGCGAAAGGGGAAGGAACACGA CAGGC
pCFD5-Δ413	CTGTCTTCCCGGCTCGTGGGGTTTTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTAACAA AGCACCAGTGGTCTAGTGGTAGAATAGTACCCTGCCACGGTACAGACCCGG GTTCGATTCCCGGCTGGTGCAATAGCATTAGACAATGCTTGTTTTAGAGCTA GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGC ACCGAGTCGGTGCTAACAAAGCACCAGTGGTCTAGTGGTAGAATAGTACCC TGCCACGGTACAGACCCGGGTTCGATTCCCGGCTGGTGCGCCACAATTGTC GATCGTCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGT TATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

413 HDR template	GTGTACCCATTGACAATTTTACAAAGAATACACTCCTACAATATCACTGAAC AGTTGGATAATAAAAAAACTCAGAATGTTTTTTTTtAAATGAGTGTTCATAA TAAACTTGAACGCAGTCTTGCCGAGTTTTTTAAGCGACAATATGTCTATTAA ACATATTGATGTAGTGgCATTTAAAAATTATTTGTGCTATTAtTTTTTTAAAT AAGTACTCCTTTTTAGCTTTCATAGTTCCCGAGATCACATCGTTCATATGGA CAGACaGATTGACAGATAAGCCCAGATTGAATTGGCCAGTAATCCCGATAA AGAATACGTATTCACTTTCATCCTTTGATACATACTTTTCATACTTTATACCC TTTTACTCTACGAGGAACGGGAGTAATTGTTGGAAATGTTTATTCACCACTC CCAGTTATTGCATTTATACAAATTATTACTTGACTAGATTTTATAATCAAAA TATTTAATTAAAATCCtCCGGTCATTATGTCCTGCTCTACATATAATTTAGAC CTCCCATCGGGTgTTGACTGTCCGTAGCTGTACTATCATTATCCTCCTAAGGC TGTCAAGTTATCTACGCCGCTTGAGGCCATTGTCAgTGTCTTCGCCGGAACT GCTACAAAAAATTGTGACAATTTACACTGATCTCTGGAACCTGGTCCACAG ACCCGTCCTTGACTGCCATAAAGTCAGTTCCACTAAAAATCGTCGAAAAAT ACTCCCGTGAACTAGAAAAAATAAAGCACACAAAAAATGTGCCTCAGTTCT AGTTTAGTTTCAAGAAATTCTCTGTTCGCTCGGAAGGAAAAAATGGAAACT AGAATACGCGGGACAATAAAAAAACGATGCGCTGTAAAAACGTTTTAGACC GAATTCACCTTTGTGTTTCGCGCATTCTCGTCATTGTTAAAAAAAAtcaagcatTT TTTTGGGTCGCGACACAATTTTTTGCCGTAGTGTCTCCAGGATTCCTCTAGA TCCTGGTCTCGTTCTCAGGCCGTGGCCTCGTTTTCCATTTAAAATGCGCGAC GCACAAAAGTTTTTACACGAATTTATGTTTTTATCGCTTTACGACCGGAGGT CCTAATAAAACTTTATTGAGGCTGCCGGTGGTCGTAGGAGTTTTGCTCGAGT TTCTTACCGTAGTAGGCCTCCTTTGGTAGCCTCAAAACTCAAAACTGTAGCA gGGTCAGGTTCCTCGGGATAAGCGCGTATCCTGCGAATTAAGCCAGCCCAA GGACGAGCCGGGTTATGCAAGTCTATGCAACGTATGGGTTTGCATAATAGA GGTAAATGTTTGTTTAACTGGGAACTGCTCTAGTTAGCATTGTCTTGATGGC ATTGTTTCTGCCCTCACCCCGTAATTTTAGATTTTAATTGAAAAATATTTTAT TTAACATATAGTAACTAAGAATTTTTTTTATTAATAATTATATATGTATATAT ATATATTATATAAGTATATTATTATATATATATTTTATTTTATTATTATTATA TTATATTATTAAATTTTTTTAGCTTTTCTTATAATTTTAATATTTTTATATATT TTTTTTTTGCAAAAGATAAATTAACAAACCAATATTTCCGCTTCACAAGCAA TCCAAGCCCCGATAATAAAATGTTCTATATTTCTTcCTAAAAATATATACTAT TATTAATTATGTTTTATTACTATATATGTATTTATATTTGCAATGATTATTGT TCAATCGAAGAAAAAATATTGCATGCAATACTTTGGGGTGTTCTTAGTTTTT CTGAATATATTTTTTATTTATTTTACTTTAGTTTTGTTATCTTAACAAATAAT

	CTTAATTTATAATATAGTTATAAACAAATGTTCAATCGTTTTTTCCAGACAA AATGTCTTTTACATTCGCAGTTCAACATCTTAATCCTTTTGCGAAACATTACT GCCATTACAGTTCAATTTCTTGTGAGACGGGGGCGTTGGCTAAAGGAAAGT CTGTTTTGGGTAAACAGAGCGAAAATCCCTTTCGCCGCCATTTCCATCCATT TCCACCTCTGATTTCTCTGATTTCCACGGACGAATCAAGGCCTACCTGGGAA GGGAAGACGCGTGTGTCTCGCATTAGACTGATGTTCTGTTGAACGGGACGG GAAATGGAAATGCGTGGCCGAATCGAGGACGAGGCTCGCGCTACTCGACTT GGACCGACTTGAGTCGTCTGTTTATGGATAGTAAATGCACCGTGTCATGGGT CCTTAAGCGTCTTAATTGCCCGCGTATCCTTCGCCGAGAAGCAGCGCAATTG GTTTCCAAGGCGGAAATATTATTGGCACATCATAAAACTATAAAACAGGCC GCAGGGGGCCCCCTTCCTCCTTCCTTCGTTTTCCTGTCCAGTTCAGTTGAGCT CTTCACCTCGGATTGCGAGCTCTGAGATTCAAACTGAGACCATCCAATGCCA TCAGCGAATTCGCCGAATTTTCGAGCAGCGCGCGCGACTTCGATTCCCCGGC GTACGACTCGTCAGCGGAATCGCATTGACAGCCTACGCCGAACTCCCCGGA GACCTCCGTTCCGATCATTGTTAATGGACGTTATCCTTATTAGATGTTGATG TCCCACGATGTCGAATAGCTACGGGACATACATATCTACATATATAAAAGA TATGCCCATGTTAGGTATTAAGATCTGATTCAGTTCATGTTCAGCTTCATTAT ATTATAGATTAAAATTATGTGAATGCTTTTATTTCTATTTCCGTTTCATCAGT TTGCATTTTATATTCCAAATTGTATAATCCTTTAAATTATTAAAATTGTTATA GTGCACGGACCTTCGAGCCTGCTTAAAGAATCGAAACTTTAAATTACTACA CAAGGTACAAGGTATTAGGAGAATTAACATAGTTTCACATGACATTATTTAT CAATAGTTTTTAACTGATGACACTCTGACATTTTTCAAAAAGCCATCAATCT TTGAGCTCAGTTGGAGCGTCCGTAACAGAAGGTGTTAACGAGGTCCTAGCT GAGATTTAAACAATTCGAAATTTATATCTAGGGGCCATAAGCAGTACGATG CGATGATGCACACAGTTCACATTTCATTCAGCTTACGGGGTCATCATTGCtatt TAAAAGCGGCAGGACAATTCGAATAAGTGTGACAGGAGCAATTACAGCCTT ATCCTGATGTCTTCGATGTCAACACACCAGTGCTCAAGACATCGCAGGCAA TTCAAGGATATGTAGGACGCACAGGACCTCGAACAGAAGCCAAGGACACA GGCGACGCGTGACGCATTGGGAAAATATTTGTACAAAACATTGAGGATATC TCAAAAGTAGCACAGCGTTTCGGCATCCTTGACTTTGATTGGTGCCGATCCA AGGACGAAACGCTATATCCACCCATTGAGAtGACCTCCGGAGGAGTCGTCG GTGGGCCTCCGATCGAAGGTCACCCAGAAATGCATCCTGTCTGGCTTTCTCT TCAaTtTATAGTGAATTTTAGAAACTGAAGATTACTTCATTTATCTAGA
pCFD5-Δftz-6	AGCAAAACTCCTACGACCACGTTTTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTAACAA AGCACCAGTGGTCTAGTGGTAGAATAGTACCCTGCCACGGTACAGACCCGG GTTCGATTCCCGGCTGGTGCTGGCAGTAATGTTTCGCAAAGTTTTAGAGCTA GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGC ACCGAGTCGGTGCTAACAAAGCACCAGTGGTCTAGTGGTAGAATAGTACCC TGCCACGGTACAGACCCGGGTTCGATTCCCGGCTGGTGCGCCACAATTGTC GATCGTCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGT TATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
pUC19-413Δftz-6	GTGTACCCATTGACAATTTTACAAAGAATACACTCCTACAATATCACTGAAC AGTTGGATAATAAAAAAACTCAGAATGTTTTTTTTtAAATGAGTGTTCATAA TAAACTTGAACGCAGTCTTGCCGAGTTTTTTAAGCGACAATATGTCTATTAA ACATATTGATGTAGTGgCATTTAAAAATTATTTGTGCTATTAtTTTTTTAAAT AAGTACTCCTTTTTAGCTTTCATAGTTCCCGAGATCACATCGTTCATATGGA CAGACaGATTGACAGATAAGCCCAGATTGAATTGGCCAGTAATCCCGATAA AGAATACGTATTCACTTTCATCCTTTGATACATACTTTTCATACTTTATACCC TTTTACTCTACGAGGAACGGGAGTAATTGTTGGAAATGTTTATTCACCACTC CCAGTTATTGCATTTATACAAATTATTACTTGACTAGATTTTATAATCAAAA TATTTAATTAAAATCCtCCGGTCATTATGTCCTGCTCTACATATAATTTAGAC CTCCCATCGGGTgTTGACTGTCCGTAGCTGTACTATCATTATCCTCCTAAGGC TGTCAAGTTATCTACGCCGCTTGAGGCCATTGTCAgTGTCTTCGCCGGAACT

	GCTACAAAAAATTGTGACAATTTACACTGATCTCTGGAACCTGGTCCACAG ACCCGTCCTTGACTGCCATAAAGTCAGTTCCACTAAAAATCGTCGAAAAAT ACTCCCGTGAACTAGAAAAAATAAAGCACACAAAAAATGTGCCTCAGTTCT AGTTTAGTTTCAAGAAATTCTCTGTTCGCTCGGAAGGAAAAAATGGAAACT AGAATACGCGGGACAATAAAAAAACGATGCGCTGTAAAAACGTTTTAGACC GAATTCACCTTTGTGTTTCGCGCATTCTCGTCATTGTTAAAAAAAAtcaagcatTT TTTTGGGTCGCGACACAATTTTTTGCCGTAGTGTCTCCAGGATTCCTCTAGA TCCTGGTCTCGTTCTCAGGCCGTGGCCTCGTTTTCCATTTAAAATGCGCGAC GCACAAAAGTTTTTACACGAATTTATGTTTTTATCGCTTTACGACCGGAGGT CCTAATAAAACTTTATTGAGGCGGTACCGCCCACGAGCCGGGAAGACAGTC AATAAAACGTCACCTTCACCTCTTACACAATATTTCACAACGTTCTTGCAAA TCCGGGGATTAGGAACAATTCCGGCGTTCTCAAAAGGTTTTGGGCATCATG CCCTTCCGTAGTCGCCGTTTGCTCTAATTTAAATTACCATTCTGCATTGTTTT GGAATGTTTCCTAAAAACTGACCTAATTTACGCCACACACTTTTCGTTACAA GTCTTTCGTTTTTATGACCCAAAAATGGTAAAAGTTTCCGCGAATTTAATCA TTTTTATGATGTGAACGATTTTTTGAATAGATGTTGCGAATTCCGAAATCGG GAAATAGAAATATGATTAAAATAGCATTAGaCAATGCTTTGGTTATTTTTAT GACTTAAATTGAGTTAGAATTTCAATTTCTTGTGAGACGGGGGCGTTGGCTA AAGGAAAGTCTGTTTTGGGTAAACAGAGCGAAAATCCCTTTCGCCGCCATT TCCATCCATTTCCACCTCTGATTTCTCTGATTTCCACGGACGAATCAAGGCC TACCTGGGAAGGGAAGACGCGTGTGTCTCGCATTAGACTGATGTTCTGTTG AACGGGACGGGAAATGGAAATGCGTGGCCGAATCGAGGACGAGGCTCGCG CTACTCGACTTGGACCGACTTGAGTCGTCTGTTTATGGATAGTAAATGCACC GTGTCATGGGTCCTTAAGCGTCTTAATTGCCCGCGTATCCTTCGCCGAGAAG CAGCGCAATTGGTTTCCAAGGCGGAAATATTATTGGCACATCATAAAACTA TAAAACAGGCCGCAGGGGGCCCCCTTCCTCCTTCCTTCGTTTTCCTGTCCAG TTCAGTTGAGCTCTTCACCTCGGATTGCGAGCTCTGAGATTCAAACTGAGAC CATCCAATGCCATCAGCGAATTCGCCGAATTTTCGAGCAGCGCGCGCGACT TCGATTCCCCGGCGTACGACTCGTCAGCGGAATCGCATTGACAGCCTACGC CGAACTCCCCGGAGACCTCCGTTCCGATCATTGTTAATGGACGTTATCCTTA TTAGATGTTGATGTCCCACGATGTCGAATAGCTACGGGACATACATATCTAC ATATATAAAAGATATGCCCATGTTAGGTATTAAGATCTGATTCAGTTCATGT TCAGCTTCATTATATTATAGATTAAAATTATGTGAATGCTTTTATTTCTATTT CCGTTTCATCAGTTTGCATTTTATATTCCAAATTGTATAATCCTTTAAATTAT TAAAATTGTTATAGTGCACGGACCTTCGAGCCTGCTTAAAGAATCGAAACTT TAAATTACTACACAAGGTACAAGGTATTAGGAGAATTAACATAGTTTCACA TGACATTATTTATCAATAGTTTTTAACTGATGACACTCTGACATTTTTCAAA AAGCCATCAATCTTTGAGCTCAGTTGGAGCGTCCGTAACAGAAGGTGTTAA CGAGGTCCTAGCTGAGATTTAAACAATTCGAAATTTATATCTAGGGGCCAT AAGCAGTACGATGCGATGATGCACACAGTTCACATTTCATTCAGCTTACGG GGTCATCATTGCtattTAAAAGCGGCAGGACAATTCGAATAAGTGTGACAGGA GCAATTACAGCCTTATCCTGATGTCTTCGATGTCAACACACCAGTGCTCAAG ACATCGCAGGCAATTCAAGGATATGTAGGACGCACAGGACCTCGAACAGA AGCCAAGGACACAGGCGACGCGTGACGCATTGGGAAAATATTTGTACAAA ACATTGAGGATATCTCAAAAGTAGCACAGCGTTTCGGCATCCTTGACTTTGA TTGGTGCCGATCCAAGGACGAAACGCTATATCCACCCATTGAGAtGACCTCC GGAGGAGTCGTCGGTGGGCCTCCGATCGAAGGTCACCCAGAAATGCATCCT GTCTGGCTTTCTCTTCAaTtTATAGTGAATTTTAGAAACTGAAGATTACTTCA TTTATCTAGA

Appendix C: Polymorphisms in the *ftz* locus of *nos-Cas9* flies.

Sequences surrounding the zebra element in the *nos-Cas9* line were compared to the *ftz* region of *Drosophila melanogaster* chromosome 3 (AC001653). Features indicated: zebra region (bold, no underline), gRNA sequences (purple, bold and underlined), transcription and translation start sites (green, bold and underlined) and stop codon (dark red, bold and underlined). Several polymorphisms were observed (bright red, bold, and underlined). Three of these occur in the coding sequence, but only one (orange, bold and underlined) results in an amino acid change of a serine to a threonine, which is conservative and unlikely to impact function. This sequencing was assembled by Patricia Graham.

>Sequenced *nosCas9* reads

TTTTATATCCTACTCTACACCCTTATTGTTAAGTTTACTGCAAAACCAGAAAG
GAGGCTCACATTCACTGGCAATATAGCAATCACAGGACATCGGAGAGATGCG
AACTGCAGTCTCTTCTGTCAAACTGAAAAACGCCATAAATTTTGTGGAACAAT
GGAATCTGCCAGAGTATGCGGATGGGGTAAGGGTAGAATTCGCACATGCACA
CATATGTCAGGAGTATCGATATGTTAGGGGTGTTCAACGGGGGTGGCTAATA
AAATTTATCCAGGCCGCGTAAGCCGATCAGAAAGTCATGTAAACACAAAAGT
GGCTGAGAGATGATTGTTTTTTACCAGTCAATTTAGCTGAGCTGATTTCAGGT
TATACGCGAACCTAAGATGGAAGTAGATTTTTAGATTACCAGAAATAACCTT
CAGTTTCTAAAAGAATGAGAAATAGTAAATCGTAAAGCATGTGGCATTCTTA
CCCCTTACTGAAATGAATGTAACAGGTAGAAGGCAGCAATTGCAGTGCTACA
AAGTATATATATTCTTGATCCATATAAGTATAATTATAAGAGCTATAATGTTA
TAATTAAACATGAAGATCCTACGCTGTACAAGTTTTGTGCCACAAATTTAACT
CAAGATTACCTCTCACAAAGGATACATTACCATTTGGCCATTAAAAATTACA
GCATCCATAGACAACCTACTTAAAATTTATAAAAACTTGAGGGTTATATCAC
AGAGTTACCGAAAAAAAGCGTAAAAGCTTTATATTCTCAACAATATTATGCT
ATTAAAATATTGCTGGTTTTCTGCTGTTATAGAATCATTTTTAAAAGTATAAC
GTAAAAAATAAAATAAACTAGTATTCATTTGAAAATTCAGCGGGCATATAAT
TTATATCATATTTTTAAAATTTAGGCAAAGGATGTTTGCATAAAGTTTTTACT
GTTTACTAGTCATTTTGGAAGTGCGTTTGTTGGTTTTTAGGCAAATACCGGGC
ACAGGAGTGAGTTTGGGAATCGGGAGTTGCGCACTTGCTTGGCCACGAG**GGC**
AAACAAAAAGCGCAAACACGCGACCCTCGGCCACGCGTATTCCTGATCC
CAGGGATCGGACGTAATGTTATCCTTTGGCCGCCCAGTGCCACGAAATA
AATTCGGAGGGAAAGGGCATCGGGTTCCGGGAACAACTGGCAGCCAGT
CTTCGGTGTTTTGCGCGCTGGCAAAAATCCAGAGAAATTTTTAGGGAAC

CATAAACGGGCCGGGGAAAAAGCCTCTGCGCCGAAGGAACGTTTTCAGC
AACAGTTTACAGTTTTTATGTCTTTATGATTATTGCAATTAGAGGGAGAT
CGGCTGAGAGTCGCGCCCTCTCGCTCTGCGCACCTCATAGGTAGGCACC
TCATGGCCGTAATTACTGCAGCACCGTCTCAAGGTCGCCGAGTAGGAGA
AGCGCGCGGGCGGATAAATCGCGATGATAATGGGCGCGATGGGTAGGT
AATAAGCCGCGCAGCAGGTAGGCACCGTACGGATAAAGTTGCCAGGACC
TCGGATAACTTCCCCTCTCCGTGCCTGCAAGGACATTTCGCCGGAGGGG
TGGCTGCGAACAGCAGGCGGCAAAGTGTCATGCGCAGGGATATTTATGC
GCTATAACGGCGAGCGTGTGCCGAGGGCTCTCTGATTTTGCTATATATG
CAGGATCTGCCGCAGGACCAGCTCATTCGCAAACTCACCAGCGTTGCGT
GCACATCGCAGAGTTAGAGAAGAAATCTAGCAATACACATCCGATATGG
CCACCACAAACAGCCAGAGCCACTACAGCTACGCCGACAACATGAACAT
GTACAACATGTATCACCCCCACAGCCTGCCGCCCACCTACTACGATAACTCA
GGCAGCAATGCCTACTATCAGAACACCTCCAATTATCAGGGCTACTATCCCC
AGGAGAGTTACTCGGAGAGCTGCTACTACTACAACAATCAGGAGCAGGTGAC
CACCCAGACTGTACCGCCCGTGCAACCCACCACCCCGCCGCCCAAGGCCACC
AAGCGCAAGGCCGAAGATGATGCTGCTTCCATCATCGCCGCCGTGGAGGAGC
GACCCAGCACACTGAGGGCTCTGCTCACCAACCCCGTGAAGAAGCTGAAGTA
CACCCCCGACTATTTCTACACCACCGTCGAGCAGGTGAAGAAGGCTCCCGCC
GTAACCACCAAGGTCACCGCCAGCCCCGCTCCCAGCTACGACCAAGAGTACG
TGACTGTGCCCACGCCCAGCGCCTCCGAGGATGTCGACTACTTGGACGTCTA
CTCGCCCCAGTCGCAGACGCAGAAGCTGAAGAATGGCGACTTTGCCACCCCT
CCGCCAACCACGCCCACCTCTCTGCCGCCCCTCGAAGGCATCAGCACGCCAC
CCCAATCGCCGGGGGAGAAATCCTCGTCAGCTGTCAGCCAGGAGATCAATCA
TCGAATTGTGACAGCCCCGAATGGAGCCGGCGATTTCAATTGGTCGCACATC
GAGGAGACTTTGGCATCAGGTAGGCATCACACACGATTAACAACCCCTAAAA
ATACACTTTGAAAATATTGAAAATATGTTTTTGTATACATTTTTGATATTTTCA
AACAATACGCAGTTATAAAACTCATTAGCTAACCCATTTTTTCTTTGCTTATG
CTTACAGATTGCAAAGACTCGAAACGCACCCGTCAGACGTACACCCGCTACC
AGACCCTGGAGCTCGAGAAGGAGTTCCACTTCAATAGATACATCACCCGGCG
TCGTCGCATCGATATCGCCAATGCCCTGAGCCTGAGCGAAAGGCAGATCAAG
ATCTGGTTCCAAAACCGACGCATGAAGTCGAAGAAGGATCGCACGCTGGACA
GCTCCCCGGAGCACTGTGGTGCCGGCTACACCGCGATGCTGCCGCCACTGGA
GGCCACAAGCACCGCCACCACCGGGGCACCATCGGTGCCAGTGCCCATGTAC
CACCACCACCAAACCACCGCCGCCTACCCCGCTTACAGCCACAGTCACAGTC
ATGGTTATGGCCTGCTCAATGATTACCCTCAGCAGCAGACCCACCAGCAGTA
CGATGCCTACCCGCAGCAGTACCAACATCAGTGCAGCTACCAGCAACATCCA
CAGGACCTCTACCATCTGTCTTGAGGTCCGGCGATGCTCAGTTACTCTCTTCC
CCAGAGCGGAACCGAAAGCCGTACCGCCACGAAACCGAAGCGCACTTCTCTC
GACCATTTGTAGGTGACACGCAAATGACACAGCCGAGAACGAAGCTGCGAC
GCGATGAGTTGCACAGTAGAGGGCGCACTCCCTACGGTGCCCAGGACATTTT
GGGCACAAGGACGAGTGCGCAAGTGCAGAAGGCAGAGGCAAAAGAGGCAG
CGCAAACAGAAAAGGAGCCTGC

Appendix D: Confocal microscopy image acquisition, processing, and data visualization.

To allow for quantitative analysis, all samples were imaged in the same session and with the same confocal microscope parameters for gain and offset. These calibrations were made following previously published specifications for quantitative purposes (Surkova, Myasnikova, Janssens, et al., 2008; Surkova et al., 2013, 2019; Surkova, Myasnikova, Kozlov, et al., 2008). Briefly, the gain for *ftz*- and *eve*-detecting channels were set to be just below pixel saturation of the strongest observed sample stripe. Because the same gain and offset were used for all samples, the raw fluorescent intensity output could be extracted and used for comparative quantitative data analysis.

Fiji ImageJ software (Schindelin et al., 2012) was used to separate image stacks from LPA AF .lif and .czi files into discreet grayscale TIF files to be used as input for the ProStack software (Kozlov et al., 2007). Briefly, a previously published pipeline for ProStack allows for automated image processing designed for the quantification of fluorescent *Drosophila* embryo images (Surkova, Myasnikova, Janssens, et al., 2008). This automated pipeline detects the shape of the embryo and location of stained nuclei within it, rotates the embryo to be anterior left/dorsal up, crops the image, and detects fluorescent signals from each nuclear position. These nuclear positions are referred to as centroids, and each one is assigned an X and Y coordinate. The fluorescent signal intensity from each channel of each centroid is organized as a column delineated output file. This file can be opened in Excel or by other means, such as R, for further processing.

All data processing for visualization and statistical analyses of the ProStack output file were conducted in Rstudio (Team, 2022). Fluorescent data from centroids in the central 10% (45%-55%) of the dorsal-ventral axis were used for one-dimensional

(1D) analysis along the anterior-posterior (A-P) axis as previously described (Surkova, Myasnikova, Janssens, et al., 2008; Surkova et al., 2013, 2019). This allowed each centroid to be characterized by its X coordinate along the A-P axis, described as percent body length. These are the datapoints used in the HCR line plots shown here, and stripe domains can be clearly visualized with this 1D analysis.

It was previously described that the background fluorescent noise along this 1D analysis takes on a parabolic shape along the A-P axis (Surkova et al., 2013). Background was removed from each channel and from each embryo with the following process: centroid data points below the stripe domains were used to create a quadratic model that was fit to the background noise. Each centroid data point was subtracted by the corresponding predicted background point from the model, resulting in the removal of the parabolic background noise. Datapoints that became slightly negative due to being less than the model were converted to values of zero.

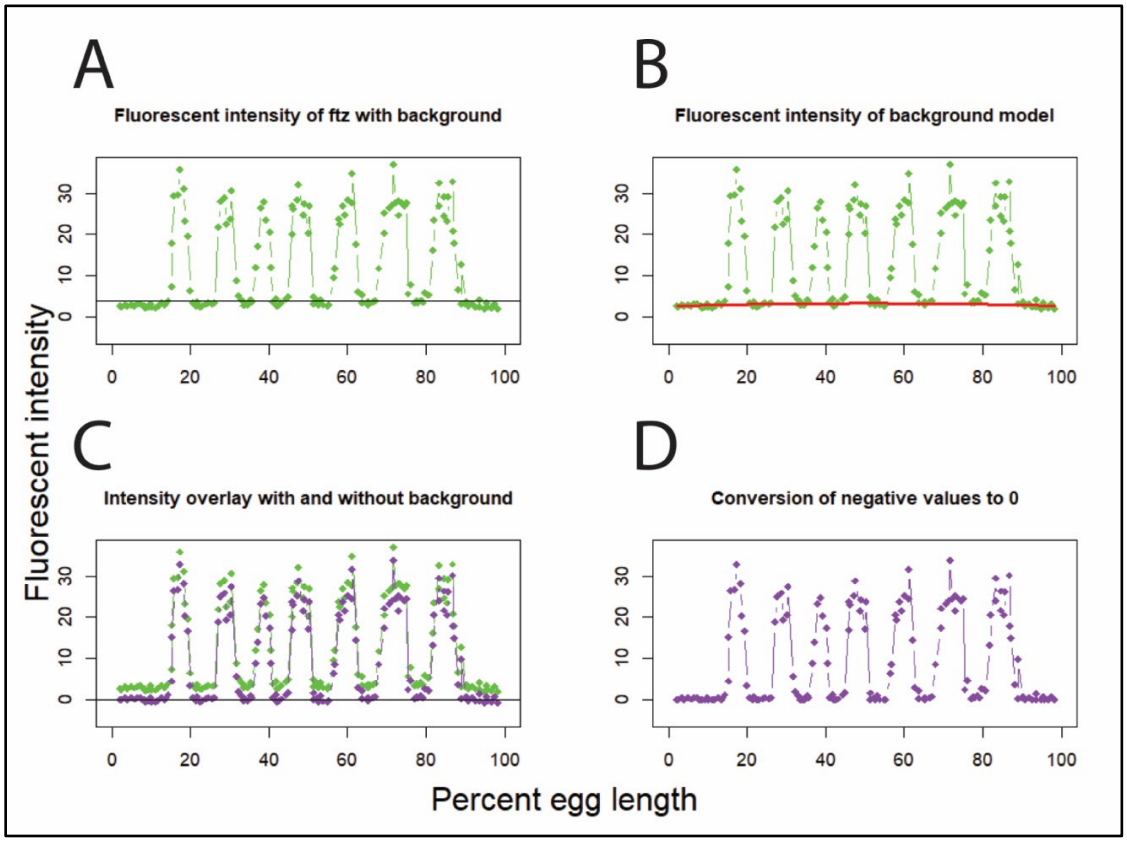
Stripe domain boundaries of *ftz* were determined by first identifying the minimal fluorescent signal (threshold) of the stripe centroids. A script was written which analyzes each centroid from 0% to 100% body length; the left boundary of each domain was determined as the first centroid between a data point below the threshold and a data point above the threshold. Conversely, the right boundary of each domain was determined as the first centroid between a data point above the threshold and a data point below the threshold. This automated processing was manually curated to accurately determine the boundaries of each stripe in the 1D analysis.

The intensity of each *ftz* stripe was determined by integrating between each of the identified boundaries. This was calculated with the Area Under the Curve (AUC)

function with the ‘trapezoid’ method in the ‘DescTools’ package in R (Signorell {Andri et mult. al.}, 2022). Because this calculation accounts for both the amplitude and width of the stripe, it portrays the intensity from this 1D analysis, generating the value to be used for the quantitative stripe intensity comparisons between embryos of different genotypes.

Due to the highly dynamic nature of pair-rule gene expression during the cleavage cycle 14A interphase, embryos that were nearing complete cellularization and thus near peak expression were used to quantify *ftz* stripe domain expression levels. These embryos were selected from the confocal images on the basis of nuclear elongation, plasma membrane deposition, and relative max values of *eve* stripes. Selecting embryos in the same stage of cycle14A in this manner helped reduce how this stage’s dynamism could confound analysis. Previously published *eve* stripe dynamics convey a pattern of relative stripe intensity that correlates with the embryo’s developmental timepoint within cycle 14A (Surkova et al., 2019). By measuring the maximal *eve* intensity values for each stripe and determining the degree of membrane depositions, embryos were temporally categorized.

Appendix E: Example of processing and removal of fluorescent background from one-dimensional analysis.



Plots produced from centroids in central 45-55% of dorsal-ventral axis. Green: *ftz* signal. (A) Plot of *ftz* signal from central strip with background present. Note the parabolic curve of background. Black line represents fluorescent intensity threshold; all points below threshold are used in background quadratic model calculations. (B) Red line represents the fit parabolic model representing the background. (C) Purple: difference between *ftz* with background (green) and background model (red line from B). Black line drawn at 0 to highlight datapoints that became negative (all negative points are between 0 and -1). (D) All negative data points converted to 0. This results in the processed dataset used for calculating stripe integrals.

Appendix F: Partial sequence of the deletion region of the *ftzΔZp* allele.

PCR products of the region around the zebra element were prepared from single ftzΔZp homozygous flies and sequenced to examine the deleted region. Features of interest are indicated with text that is colored, bold and underlined. Indel (pink); portions of the 3' gRNA remaining (-276 to -261 in diagrams, blue), transcription start site, green (+1 in diagrams) and translation start codon. The 5' region of the zebra CRE that remains stops 13 bases before the 5' gRNA (at base -637 in diagrams). This sequencing and analysis were done by Patricia Graham.

>Sequenced_ftzΔZp_reads

GGAGTATCGATATGTTAGGGGTGTTCAACGGGGGTGGCTAATAAAATTTATC
CAGGCCGCGTAAGCCGATCAGAAAGTCATGTAAACACAAAAGTGGCTGAGA
GATGATTGTTTTTTACCAGTCAATTTAGCTGAGCTGATTTCAGGTTATACGCG
AACCTAAGATGGAAGTAGATTTTTAGATTACCAGAAATAACCTTCAGTTTCTA
AAAGAATGAGAAATAGTAAATCGTAAAGCATGTGGCATTCTTACCCCTTACT
GAAATGAATGTAACAGGTAGAAGGCAGCAATTGCAGTGCTACAAAGTATATA
TATTCTTGATCCATATAAGTATAATTATAAGAGCTATAATGTTATAATTAAAC
ATGAAGATCCTACGCTGTACAAGTTTTGTGCCACAAATTTAACTCAAGATTAC
CTCTCACAAAGGATACATTACCATTTGGCCATTAAAAATTACAGCATCCATAG
ACAACCTACTTAAAATTTATAAAAACTTGAGGGTTATATCACAGAGTTACCG
AAAAAAAGCGTAAAAGCTTTATATTCTCAACAATATTATGCTATTAAAATATT
GCTGGTTTTCTGCTGTTATAGAATCATTTTTAAAAGTATAACGTAAAAAATAA
AATAAACTAGTATTCATTTGAAAATTCAGCGGGCATATAATTTATATCATATT
TTTAAAATTTANGCAAAGGATGTTTGCATAAAGTTTTTACTGTTTACTAGTCA
TTTTGGAAGTGCGTTTGTTGGTTTTTANGCAAATACCGGGCACAGGAGTGAGT
TTGGGAATCGGGAGTTGCGCACTTGCTTGGCCACGAGGGCAAACAAAAAGCG
CAAACACGCGA**ACAAA**CAGGTAGGCACCGTACGGATAAAGTTGCCAGGAC
CTCGGATAACTTCCCCTCTCCGTGCCTGCAAGGACATTTCGCCGGAGGGGTG
GCTGCGAACAGCAGGCGGCAAAGTGTCATGCGCAGGGATATTTATGCGCTAT
AACGGCGAGCGTGTGCCGAGGGCTCTCTGATTTTGCTATATATGCAGGATCTG
CCGCAGGACCAGCTC**A**TTCGCAAACTCACCAGCGTTGCGTGCACATCGCAGA
GTTAGAGAAGAAATCTAGCAATACACATCCGAT**ATG**GCCACCACAAACAGC
CAGAGCCACTACAGCTACGCCGACAACATGAACATGTACAACATGTATCACC
CCCACAGCCTGCCGCCCACCTACTACGATAACTCAGGCAGCAATGCCTACTA
TCAGAACACCTCCAATTATCAGGGCTACTATCCCCAGGAGAGTTACTCGGAG
AGCTGCTACTACTACAACAATCAGGAGCAGGTGACCACCCAGACTGTACCGC
CCGTGCAACCCACCACCCCGCCGCCCAAGGCCACCAAGCGCAAGGCCGAAG
ATGATGCTGCTTCCATCATCGCCGCCGTGGAGGAGCGACCCAGCACACTGAG

GGCTCTGCTCACCAACCCCGTGAAGAAGCTGAAGTACACCCCCGACTATTTC
TACACCACCGTCGAGCAGGTGAAGAAGGCTCCCGCCGTAACCACCAAGGTCA
CCGCCAGCCCCGCTCCCAGCTACGACCAAGAGTACGTGACTGTGCCCACGCC
CAGCGCCTCCGAGGATGTCGACTACTTGGACGTCTACTCGCCCCAGTCGCAG
ACGCAGAAGCTGAAGAATGGCGACTTTGCCACCCCTCCGCCAACCACGCCCA
CCTCTCTGCCGCCCCTCGAAGGCATCAGCACGCCACCCCAATCGCCGGGGGA
GAAATCCTCGTCAGCTGTCAGCCAGGAGATCAATCATCGAATTGTGACAGCC
CCGAATGGAGCCGGCGATTTCAATTGGTCGCACATCGAGGAGACTTTGGCAT
CAGGTAGGCATCACACACGATTAACAACCCCTAAAAATACACTTTGAAAATA
TTGAAAATATGTTTTTGTATACATTTTTGATATTTTCAAACAATACGCAGTTAT
AAAACTCATTAGCTAACCCATTTTTTCTTTGCTTATGCTTACAGATTGCAAAG
ACTCGAAACGCACCCGTCAGACGTACACCCGCTACCAGACCCTGGAGCTCGA
GAAGGAGTTCCACTTCAATAGATACATCACCCGGCGTCGTCGCATCGATATC
GCCAATGCCC

Appendix G: Partial sequence of the deletion region of the *ftzΔZ* allele.

PCR products of the region around the zebra element were prepared from single *ftzΔZ* homozygous flies and sequenced to examine the deleted region. Features of interest are indicated with text that is colored, bold and underlined: gold, AvrII site; green the transcription start site (+1 in diagrams) and the translation start codon. This sequencing and analysis were done by Patricia Graham.

>Sequenced ftzΔZ reads

TATGTCNGGAGTATCGATATGTTAGGGGTGTTCAACGGGGGTGGCTAATAAA
ATTTATCCAGGCCGCGTAAGCCGATCAGAAAGTCATGTAAACACAAAAGTGG
CTGAGAGATGATTGTTTTTTACCAGTCAATTTAGCTGAGCTGATTTCAGGTTA
TACGCGAACCTAAGATGGAAGTAGATTTTTAGATTACCAGAAATAACCTTCA
GTTTCTAAAAGAATGAGAAATAGTAAATCGTAAAGCATGTGGCATTCTTACC
CCTTACTGAAATGAATGTAACAGGTAGAAGGCAGCAATTGCAGTGCTACAAA
GTATATATATTCTTGATCCATATAAGTATAATTATAAGAGCTATAATGTTATA
ATTAAACATGAAGATCCTACGCTGTACAAGTTTTGTGCCACAAATTTAACTCA
AGATTACCTCTCACAAAGGATACATTACCATTTGGCCATTAAAAATTACAGC
ATCCATAGACAACCTACTTAAAATTTATAAAAACTTGAGGGTTATATCACAG
AGTTACCGAAAAAAAGCGTAAAAGCTTTATATTCTCAACAATATTATGCTATT
AAAATATTGCTGGTTTTCTGCTGTTATAGAATCATTTTTAAAAGTATAACGTA
AAAAATAAAATAAACTAGTATTCATTTGAAAATTCAGCGGGCATATAATTTA
TATCATATTTTTAAAATTTAGGCAAAGGATGTTTGCATAAAGTTTTTACTGTTT
ACTAGTCATTTTGGAAGTGCGTTTGTTGGTTTTTAGGCAAATACCGGGCACAG
GAGTGAGTTTGGGAATCGGGAGTTGCGCACTTGCTTGG**CCTAGG**GCTATAAC
GGCGAGCGTGTGCCGAGGGCTCTCTGATTTTGCTATATATGCAGGATCTGCCG
CAGGACCAGCTC**A**TTCGCAAACTCACCAGCGTTGCGTGCACATCGCAGAGTT
AGAGAAGAAATCTAGCAATACACATCCGAT**ATG**GCCACCACAAACAGCCAG
AGCCACTACAGCTACGCCGACAACATGAACATGTACAACATGTATCACCCCC
ACAGCCTGCCGCCCACCTACTACGATAACTCAGGCAGCAATGCCTACTATCA
GAACACCTCCAATTATCAGGGCTACTATCCCCAGGAGAGTTACTCGGAGAGC
TGCTACTACTACAACAATCAGGAGCAGGTGACCACCCAGACTGTACCGCCCG
TGCAACCCACCACCCCGCCGCCCAAGGCCACCAAGCGCAAGGCCGAAGATG
ATGCTGCTTCCATCATCGCCGCCGTGGAGGAGCGACCCAGCACACTGAGGGC
TCTGCTCACCAACCCCGTGAAGAAGCTGAAGTACACCCCCGACTATTTCTAC
ACCACCGTCGAGCAGGTGAAGAAGGCTCCCGCCGTAACCACCAAGGTCACCG
CCAGCCCCGCTCCCAGCTACGACCAAGAGTACGTGACTGTGCCCACGCCCAG
CGCCTCCGAGGATGTCGACTACTTGGACGTCTACTCGCCCCAGTCGCAGACG
CAGAAGCTGAAGAATGGCGACTTTGCCACCCCTCCGCCAACCACGCCCACCT
CTCTGCCGCCCCTCGAAGGCATCAGCACGCCACCCCAATCGCCGGGGGAGAA
ATCCTCGTCAGCTGTCAGCCAGGAGATCAATCATCGAATTGTGACAGCCCCG

AATGGAGCCGGCGATTTCAATTGGTCGCACATCGAGGAGACTTTGGCATCAG
GTAGGCATCACACACGATTAACAACCCCTAAAAATACACTTTGAAAATATTG
AAAATATGTTTTTGTATACATTTTTGATATTTTCAAACAATACGCAGTTATAA
AACTCATTAGCTAACCCATTTTTTCTTTGCTTATGCTTACAGATTGCAAAGACT
CGAAACGCACCCGTCAGACGTACACCCGCTACCAGACCCTGGAGCTCGAGAA
GGAGTTCCACTTCAATAGATACATCACCCGGCGTCGTCGCATCGATATCGCC
AATGCCC

References

(DGT), T. F. C. and the R. P. and C. (2014). A promoter-level mammalian expression atlas. *Nature*, *507*(7493), 462–470. https://doi.org/10.1038/nature13182

Amano, T., Sagai, T., Tanabe, H., Mizushina, Y., Nakazawa, H., & Shiroishi, T. (2009). Chromosomal dynamics at the Shh locus: Limb bud-specific differential regulation of competence and active transcription. *Developmental Cell*, *16*(1), 47–57. https://doi.org/10.1016/j.devcel.2008.11.011

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., … Sandelin, A. (2014). An atlas of active enhancers across human cell types and tissues. *Nature*, *507*(7493), 455–461. https://doi.org/10.1038/nature12787

Arnosti, D. N., Barolo, S., Levine, M., & Small, S. (1996). The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. *Development*, *122*(1), 205–214. https://doi.org/10.1242/dev.122.1.205

Asturias, F. J., Jiang, Y. W., Myers, L. C., Gustafsson, C. M., & Kornberg, R. D. (1999). Conserved structures of mediator and RNA polymerase II holoenzyme. *Science*, *283*(5404), 985–987. https://doi.org/10.1126/science.283.5404.985

Banerji, J., Rusconi, S., & Schaffner, W. (1981). Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell*, *27*(2 PART 1), 299–308. https://doi.org/10.1016/0092-8674(81)90413-X

Batut, P. J., Bing, X. Y., Sisco, Z., Raimundo, J., Levo, M., & Levine, M. S. (2022). Genome organization controls transcriptional dynamics during development. *Science*, *375*(6580), 566–570. https://doi.org/10.1126/science.abi7178

Bazett-Jones, D. P., Leblanc, B., Herfort, M., & Moss, T. (1994). Short-range DNA looping by the Xenopus HMG-box transcription factor, xUBF. *Science*, *264*(5162), 1134–1137. https://doi.org/10.1126/science.8178172

Bernard, V., Brunaud, V., & Lecharny, A. (2010). TC-motifs at the TATA-box expected position in plant genes: a novel class of motifs involved in the transcription regulation. *BMC Genomics*, *11*(1), 166. https://doi.org/10.1186/1471-2164-11-166

Berghammer AJ, Klingler M, Wimmer EA. 1999. A universal marker for transgenic insects. *Nature* **402**: 370-371.

Boija, A., Klein, I. A., Sabari, B. R., Dall'Agnese, A., Coffey, E. L., Zamudio, A. V., … Young, R. A. (2018). Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell*, *175*(7), 1842-1855.e16. https://doi.org/10.1016/j.cell.2018.10.042

Boyle M, DiNardo S. 1995. Specification, migration and assembly of the somatic cells of the Drosophila gonad. *Development* **121**: 1815-1825.

Burke, T. W., & Kadonaga, J. T. (1997). The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAF II 60 of Drosophila. *Genes & Development*, *11*(22), 3020–3031. https://doi.org/10.1101/gad.11.22.3020

Burke, T. W., Willy, P. J., Kutach, A. K., Butler, J. E. F., & Kadonaga, J. T. (1998). The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box. *Cold Spring Harbor Symposia on Quantitative Biology*, *63*, 75–82. https://doi.org/10.1101/sqb.1998.63.75

Calhoun, V. C., & Levine, M. (2003). Long-range enhancer-promoter interactions in the

Scr-Antp interval of the Drosophila Antennapedia complex. *Proceedings of the National Academy of Sciences*, *100*(17), 9878–9883. https://doi.org/10.1073/pnas.1233791100

Calhoun, V. C., Stathopoulos, A., & Levine, M. (2002). Promoter-proximal tethering elements regulate enhancer-promoter specificity in the Drosophila Antennapedia complex. *Proceedings of the National Academy of Sciences*, *99*(14), 9243–9247. https://doi.org/10.1073/pnas.142291299

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., … Hayashizaki, Y. (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nature Genetics*, *38*(6), 626–635. https://doi.org/10.1038/ng1789

Carroll, S. B., & Vavra, S. H. (1989). The zygotic control of Drosophila pair-rule gene expression. II. Spatial repression by gap and pair-rule gene products. *Development*, *107*(3), 673–683. https://doi.org/10.1242/dev.107.3.673

Carroll, Sean B. (1990). Zebra patterns in fly embryos: Activation of stripes or repression of interstripes? *Cell*, *60*(1), 9–16. https://doi.org/10.1016/0092-8674(90)90711-M

Carroll, Sean B., & Scott, M. P. (1985). Localization of the fushi tarazu protein during Drosophila embryogenesis. *Cell*, *43*(1), 47–57. https://doi.org/10.1016/0092-8674(85)90011-X

Carroll, Sean B., & Scott, M. P. (1986). Zygotically active genes that affect the spatial expression of the fushi tarazu segmentation gene during early Drosophila embryogenesis. *Cell*, *45*(1), 113–126. https://doi.org/10.1016/0092-8674(86)90543-X

Cheatle Jarvela, A. M., Trelstad, C. S., & Pick, L. (2020). Regulatory gene function handoff allows essential gene loss in mosquitoes. *Communications Biology*, *3*(1). https://doi.org/10.1038/s42003-020-01203-w

Chen, D., & Lei, E. P. (2019). Function and regulation of chromatin insulators in dynamic genome organization. *Current Opinion in Cell Biology*, *58*, 61–68. https://doi.org/10.1016/j.ceb.2019.02.001

Chen, J., Zhang, Z., Li, L., Chen, B.-C., Revyakin, A., Hajj, B., … Liu, Z. (2014). Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell*, *156*(6), 1274–1285. https://doi.org/10.1016/j.cell.2014.01.062

Chen, R. A.-J., Down, T. A., Stempor, P., Chen, Q. B., Egelhofer, T. A., Hillier, L. W., … Ahringer, J. (2013). The landscape of RNA polymerase II transcription initiation in C. elegans reveals promoter and enhancer architectures. *Genome Research*, *23*(8), 1339–1347. https://doi.org/10.1101/gr.153668.112

Chetverina D, Erokhin M, Schedl P. 2021. GAGA factor: a multifunctional pioneering chromatin protein. *Cell Mol Life Sci*.

Choi, H. M. T., Beck, V. A., & Pierce, N. A. (2014). Next-generation in situ hybridization chain reaction: Higher gain, lower cost, greater durability. *ACS Nano*, *8*(5), 4284–4294. https://doi.org/10.1021/nn405717p

Choi, H. M. T., Chang, J. Y., Trinh, L. A., Padilla, J. E., Fraser, S. E., & Pierce, N. A. (2010). Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature Biotechnology*, *28*(11), 1208–1212. https://doi.org/10.1038/nbt.1692

Clark E, Akam M. 2016. Odd-paired controls frequency doubling in Drosophila segmentation by altering the pair-rule gene regulatory network. *Elife* **5**.

Cohen, B., Simcox, A. A., & Cohen, S. M. (1993). Allocation of the thoracic imaginal primordia in the Drosophila embryo. *Development*, *117*(2), 597–608. https://doi.org/10.1242/dev.117.2.597

Danino, Y. M., Even, D., Ideses, D., & Juven-Gershon, T. (2015). The core promoter: At the heart of gene expression. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1849*(8), 1116–1131. https://doi.org/10.1016/j.bbagrm.2015.04.003

de Mendiburu, F. (2021). agricolae: Statistical Procedures for Agricultural Research. Retrieved from https://cran.r-project.org/package=agricolae

Dearolf CR, Topol J, Parker CS. 1989a. The *caudal* gene product is a direct activator of *fushi tarazu* transcription during *Drosophila* embryogenesis. *Nature* **341**: 340-343.

Dearolf CR, Topol J, Parker CS. 1989b. Transcriptional control of *Drosophila fushi tarazu* zebra stripe expression. *Genes & Development* **3**: 384-398.

Delker RK, Ranade V, Loker R, Voutev R, Mann RS. 2019. Low affinity binding sites in an activating CRM mediate negative autoregulation of the Drosophila Hox gene Ultrabithorax. *PLoS Genet* **15**: e1008444.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P. D., … Blobel, G. A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell*, *149*(6), 1233–1244. https://doi.org/10.1016/j.cell.2012.03.051

DiNardo, S., & O'Farrell, P. H. (1987). Establishment and refinement of segmental pattern in the Drosophila embryo: spatial control of engrailed expression by pair-rule genes. *Genes & Development*, *1*(10), 1212–1225. https://doi.org/10.1101/gad.1.10.1212

DiNardo, Stephen, Kuner, J. M., Theis, J., & O'Farrell, P. H. (1985). Development of embryonic pattern in D. melanogaster as revealed by accumulation of the nuclear engrailed protein. *Cell*, *43*(1), 59–69. https://doi.org/10.1016/0092-8674(85)90012-1

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., … Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, *485*(7398), 376–380. https://doi.org/10.1038/nature11082

Doe, C. Q., Hiromi, Y., Gehring, W. J., & Goodman, C. S. (1988). Expression and function of the segmentation gene fushi tarazu during Drosophila neurogenesis. *Science*, *239*(4836), 170–175.

Dos Santos, G., Schroeder, A. J., Goodman, J. L., Strelets, V. B., Crosby, M. A., Thurmond, J., … Grumbling, G. (2015). FlyBase: Introduction of the Drosophila melanogaster Release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Research*, *43*(D1), D690–D697. https://doi.org/10.1093/nar/gku1099

Eagen, K. P., Aiden, E. L., & Kornberg, R. D. (2017). Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(33), 8764–8769. https://doi.org/10.1073/pnas.1701291114

Edgar BA, Weir MP, Schubiger G, Kornberg T. 1986. Repression and turnover pattern *fushi tarazu* RNA in the early *Drosophila* embryo. *Cell* **47**: 747-754.

El-Sharnouby, S., Fischer, B., Magbanua, J. P., Umans, B., Flower, R., Choo, S. W., …

White, R. (2017). Regions of very low H3K27me3 partition the Drosophila genome into topological domains. *PLoS ONE*, *12*(3), 1–23. https://doi.org/10.1371/journal.pone.0172725

Farley, E. K., Olson, K. M., Zhang, W., Brandt, A. J., Rokhsar, D. S., & Levine, M. S. (2015). Suboptimization of developmental enhancers. *Science*, *350*(6258), 325–328. https://doi.org/10.1126/science.aac6948

Farley, E. K., Olson, K. M., Zhang, W., Rokhsar, D. S., & Levine, M. S. (2016). Syntax compensates for poor binding sites to encode tissue specificity of developmental enhancers. *Proceedings of the National Academy of Sciences*, *113*(23), 6508–6513. https://doi.org/10.1073/pnas.1605085113

Fischer, M. D., Mgboji, E., & Liu, Z. (2018). Pyrite cloning: a single tube and programmed reaction cloning with restriction enzymes. *Plant Methods*, *14*(1), 91. https://doi.org/10.1186/s13007-018-0359-7

Fjose, A., Mcginnis, W. J., & Gehring, W. J. (1985). Isolation of a homoeo box-containing gene from the engrailed region of Drosophila and the spatial distribution of its transcripts. *Nature*, *313*, 284–289.

Flynn, R. A., Almada, A. E., Zamudio, J. R., & Sharp, P. A. (2011). Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. *Proceedings of the National Academy of Sciences*, *108*(26), 10460–10465. https://doi.org/10.1073/pnas.1106630108

Frasch, M., & Levine, M. (1987). Complementary patterns of even-skipped and fushi tarazu expression involve their differential regulation by a common set of segmentation genes in Drosophila. *Genes & Development*, *1*(9), 981–995. https://doi.org/10.1101/gad.1.9.981

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2016). Formation of chromosomal domains by loop extrusion. *Cell Reports*, *15*(9), 2038–2049. https://doi.org/10.1016/j.celrep.2016.04.085

Fujioka M, Yusibova GL, Patel NH, Brown SJ, Jaynes JB. 2002. The repressor activity of Even-skipped is highly conserved, and is sufficient to activate engrailed and to regulate both the spacing and stability of parasegment boundaries. *Development* **129**: 4411-4421.

Gambetta, M. C., & Furlong, E. E. M. (2018). The insulator protein CTCF is required for correct hox gene expression, but not for embryonic development in Drosophila. *Genetics*, *210*(1), 129–136. https://doi.org/10.1534/genetics.118.301350

Gindhart, J. G., King, A. N., & Kaufman, T. C. (1995). Characterization of the cis-Regulatory Region of the Drosophila Homeotic Gene Sex combs reduced, *795*.

Gloor GB, Engels WR. 1992. Single fly DNA preps for PCR. *Dros Inform Ser* **71**: 148-149.

Goto, T., Macdonald, P., & Maniatis, T. (1989). Early and late periodic patterns of even skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell*, *57*(3), 413–422. https://doi.org/10.1016/0092-8674(89)90916-1

Graham, P. L., Fischer, M. D., Giri, A., & Pick, L. (2021). The fushi tarazu zebra element is not required for Drosophila viability or fertility . *G3 Genes|Genomes|Genetics*. https://doi.org/10.1093/g3journal/jkab300

Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison,

M. M., … O'connor-Giles, K. M. (2013). Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics*, *194*(4), 1029–1035. https://doi.org/10.1534/genetics.113.152710

Grossniklaus, U., Pearson, R. K., & Gehring, W. J. (1992). The Drosophila sloppy paired locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. *Genes and Development*, *6*(6), 1030–1051. https://doi.org/10.1101/gad.6.6.1030

Haberle, V., Li, N., Hadzhiev, Y., Plessy, C., Previti, C., Nepal, C., … Lenhard, B. (2014). Two independent transcription initiation codes overlap on vertebrate core promoters. *Nature*, *507*(7492), 381–385. https://doi.org/10.1038/nature12974

Haberle, V., & Stark, A. (2018). Eukaryotic core promoters and the functional basis of transcription initiation. *Nature Reviews Molecular Cell Biology*, *19*(10), 621–637. https://doi.org/10.1038/s41580-018-0028-8

Hafen, E., Kuroiwa, A., & Gehring, W. J. (1984). Spatial distribution of transcripts from the segmentation gene fushi tarazu during Drosophila embryonic development. *Cell*, *37*(3), 833–841. https://doi.org/10.1016/0092-8674(84)90418-5

Han, W, Yu, Y., Altan, N., & Pick, L. (1993). Multiple proteins interact with the fushi tarazu proximal enhancer. *Molecular and Cellular Biology*, *13*(9), 5549–5559. https://doi.org/10.1128/MCB.13.9.5549

Han, W, Yu, Y., Su, K., Kohanski, R. a, & Pick, L. (1998). A binding site for multiple transcriptional activators in the fushi tarazu proximal enhancer is essential for gene expression in vivo. *Molecular and Cellular Biology*, *18*(6), 3384–3394. https://doi.org/10.1128/mcb.18.6.3384

Hang, S., & Gergen, J. P. (2017). Different modes of enhancer-specific regulation by Runt and Even-skipped during Drosophila segmentation. *Molecular Biology of the Cell*, *28*(5), 681–691. https://doi.org/10.1091/mbc.E16-09-0630

Hare, E. E., Peterson, B. K., Iyer, V. N., Meier, R., & Eisen, M. B. (2008). Sepsid even-skipped enhancers are functionally conserved in Drosophila despite lack of sequence conservation. *PLoS Genetics*, *4*(6). https://doi.org/10.1371/journal.pgen.1000106

Heffer, A., Xiang, J., & Pick, L. (2013). Variation and constraint in Hox gene evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(6), 2211–2216. https://doi.org/10.1073/pnas.1210847110

Heitz, E. (1928). Das Heterochromatin der Moose. *Jahrb Wiss Botanik*, (69), 762–818.

Hiromi, Y., & Gehring, W. J. (1987). Regulation and function of the Drosophila segmentation gene fushi tarazu. *Cell*, *50*(6), 963–974. https://doi.org/10.1016/0092-8674(87)90523-X

Hiromi, Y., Kuroiwa, A., & Gehring, W. J. (1985). Control Elements of the Drosophila Segmentation Gene fushi tarazu, *43*(December 1965), 603–613. https://doi.org/10.1016/0092-8674(85)90232-6

Hong JW, Hendrix DA, Levine MS. 2008. Shadow enhancers as a source of evolutionary novelty. *Science* **321**: 1314.

Hoskins, R. A., Landolin, J. M., Brown, J. B., Sandler, J. E., Takahashi, H., Lassmann, T., … Celniker, S. E. (2011). Genome-wide analysis of promoter architecture in Drosophila melanogaster. *Genome Research*, *21*(2), 182–192. https://doi.org/10.1101/gr.112466.110

Housden, B. E., Valvezan, A. J., Kelley, C., Sopko, R., Hu, Y., Roesel, C., … Perrimon,

N. (2015). Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi. *Science Signaling*, *8*(393), 1–10. https://doi.org/10.1126/scisignal.aab3729

Howard, K., Ingham, P., & Rushlow, C. (1988). Region-specific alleles of the Drosophila segmentation gene hairy. *Genes & Development*, *2*(8), 1037–1046. https://doi.org/10.1101/gad.2.8.1037

Howard, K. R., & Struhl, G. (1990). Decoding positional information: Regulation of the pair-rule gene hairy. *Development*, *110*(4), 1223–1231. https://doi.org/10.1242/dev.110.4.1223

Howard, Ken, & Ingham, P. (1986). Regulatory interactions between the segmentation genes fushi tarazu, hairy, and engrailed in the Drosophila blastoderm. *Cell*, *44*(6), 949–957. https://doi.org/10.1016/0092-8674(86)90018-8

Hsieh, T.-H. S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., & Rando, O. J. (2015). Mapping nucleosome resolution chromosome folding in yeast by Micro-C. *Cell*, *162*(1), 108–119. https://doi.org/10.1016/j.cell.2015.05.048

Hug, C. B., Grimaldi, A. G., Kruse, K., & Vaquerizas, J. M. (2017). Chromatin Architecture emerges during zygotic genome activation independent of transcription. *Cell*, *169*(2), 216-228.e19. https://doi.org/10.1016/j.cell.2017.03.024

Ingham, P., & Gergen, P. (1988). lnteractions between the pair-rule genes runt, hairy, even-skipped and fushi tarazu and the establishment of periodic pattern in the Drosophila embryo. *Development*, (104(Supplement)), 51–60.

Jaeger, J. (2011). The gap gene network. *Cellular and Molecular Life Sciences*, *68*(2), 243–274. https://doi.org/10.1007/s00018-010-0536-y

Johnston, D. S., & Nüsslein-Volhard, C. (1992). The origin of pattern and polarity in the Drosophila embryo. *Cell*, *68*(2), 201–219. https://doi.org/10.1016/0092-8674(92)90466-P

Jost, D., Carrivain, P., Cavalli, G., & Vaillant, C. (2014). Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Research*, *42*(15), 9553–9561. https://doi.org/10.1093/nar/gku698

Jost, W., Yu, Y., Pick, L., Preiss, A., & Maier, D. (1995). Structure and regulation of the fushi tarazu gene from Drosophila hydei. *Roux's Archives of Developmental Biology*, *205*(3–4), 160–170. https://doi.org/10.1007/BF00357762

Judd J, Duarte FM, Lis JT. 2021. Pioneer-like factor GAF cooperates with PBAP (SWI/SNF) and NURF (ISWI) to regulate transcription. *Genes Dev* **35**: 147-156.

Jurgens, G., Nüsslein-Volhard, C., Wieschaus, E., & Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster II. Zygotic loci on the third chromosome. *Roux's Archives of Developmental Biology*, *193*, 283–295. https://doi.org/10.3989/revmetalm.2004.v40.i3.265

Kane, N. S., Vora, M., Varre, K. J., & Padgett, R. W. (2017). Efficient screening of CRISPR/Cas9-induced events in Drosophila using a Co-CRISPR strategy. *G3: Genes, Genomes, Genetics*, *7*(1), 87–93. https://doi.org/10.1534/g3.116.036723

Karch F, Weiffenbach B, Peifer M, Bender W, Duncan I, Celniker S, Crosby M, Lewis EB. 1985. The abdominal region of the bithorax complex. *Cell* **43**: 81-96.

Klingler, M., Soong, J., Butler, B., & Gergen, J. P. (1996). Disperse versus compact elements for the regulation of runt stripes in Drosophila. *Developmental Biology*, *177*(1), 73–84. https://doi.org/10.1006/dbio.1996.0146

Knipple, D. C., Seifert, E., Rosenberg, U. B., Preiss, A., & Jäckle, H. (1985). Spatial and temporal patterns of krüppel gene expression in early drosophila embryos. *Nature*, *317*(6032), 40–44. https://doi.org/10.1038/317040a0

Kornberg, T., Sidén, I., O'Farrell, P., & Simon, M. (1985). The engrailed locus of drosophila: In situ localization of transcripts reveals compartment-specific expression. *Cell*, *40*(1), 45–53. https://doi.org/10.1016/0092-8674(85)90307-1

Kosman, D., Mizutani, C. M., Lemons, D., Cox, W. G., McGinnis, W., & Bier, E. (2004). Multiplex detection of RNA expression in Drosophila embryos. *Science*, *305*(5685), 846. https://doi.org/10.1126/science.1099247

Kozlov, K., Pisarev, A., Matveeva, A., Kaandorp, J., & Samsonova, M. (2007). Image processing package ProStack for quantification of biological images. In *Proceedings of the 4th International Symposium on Networks in Bioinformatics (ISNB)* (p. 204).

Kulkarni, M. M., & Arnosti, D. N. (2003). Information display by transcriptional enhancers. *Development*, *130*(26), 6569–6575. https://doi.org/10.1242/dev.00890

Kutach, A. K., & Kadonaga, J. T. (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. *Molecular and Cellular Biology*, *20*(13), 4754–4764. https://doi.org/10.1128/MCB.20.13.4754-4764.2000

Kyrchanova, O., & Georgiev, P. (2021). Mechanisms of enhancer-promoter interactions in higher eukaryotes. *International Journal of Molecular Sciences*, *22*(2), 671. https://doi.org/10.3390/ijms22020671

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., & Valen, E. (2019). CHOPCHOP v3: Expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research*, *47*(W1), W171–W174. https://doi.org/10.1093/nar/gkz365

Lawrence, P. (1992). *The making of a fly: the genetics of animal design*. Oxford [England]: Blackwell Science.

Lawrence, P A, & Johnston, P. (1989). Pattern formation in the Drosophila embryo: allocation of cells to parasegments by even-skipped and fushi tarazu. *Development (Cambridge, England)*, *105*(4), 761–767. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2598812

Lawrence, Peter A., Johnston, P., Macdonald, P., & Struhl, G. (1987). Borders of parasegments in Drosophila embryos are delimited by the fushi tarazu and even-skipped genes. *Nature*, *328*(6129), 440–442. https://doi.org/10.1038/328440a0

Li, M., Ma, Z., Liu, J. K., Roy, S., Patel, S. K., Lane, D. C., & Cai, H. N. (2015). An Organizational Hub of Developmentally Regulated Chromatin Loops in the Drosophila Antennapedia Complex. *Molecular and Cellular Biology*, *35*(23), 4018–4029. https://doi.org/10.1128/mcb.00663-15

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., … Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*(5950), 289–293. https://doi.org/10.1126/science.1181369

Lifton, R. P., Goldberg, M. L., Karp, R. W., & Hogness, D. S. (1978). The organization of the histone genes in Drosophila melanogaster: functional and evolutionary implications. *Cold Spring Harbor Symposia on Quantitative Biology*, *42*, 1047–1051. https://doi.org/10.1101/SQB.1978.042.01.105

Macdonald, P. M., Ingham, P., & Struhl, G. (1986). Isolation, structure, and expression of even-skipped: A second pair-rule gene of Drosophila containing a homeo box. *Cell*, *47*(5), 721–734. https://doi.org/10.1016/0092-8674(86)90515-5

Magrassi, L., & Lawrence, P. A. (1988). The pattern of cell death in fushi tarazu, a segmentation gene of Drosophila. *Development*, *104*(3), 447–451. https://doi.org/10.1242/dev.104.3.447

Medina-Rivera, A., Defrance, M., Sand, O., Herrmann, C., Castro-Mondragon, J. A., Delerce, J., … van Helden, J. (2015). RSAT 2015: Regulatory sequence analysis tools. *Nucleic Acids Research*, *43*(W1), W50–W56. https://doi.org/10.1093/nar/gkv362

Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I. E., Males, M., Viales, R. R., & Furlong, E. E. M. (2018). The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes and Development*, *32*(1), 42–57. https://doi.org/10.1101/gad.308619.117

Miller, J. H. (1972). Experiments in Molecular Genetics. *Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.*, 352–355.

Mohler, J., Eldon, E. D., & Pirrotta, V. (1989). A novel spatial transcription pattern associated with the segmentation gene, giant, of Drosophila. *EMBO Journal*, *8*(5), 1539–1548. https://doi.org/10.1002/j.1460-2075.1989.tb03538.x

Moon, H., Filippova, G., Loukinov, D., Pugacheva, E., Chen, Q., Smith, S. T., … Lobanenkov, V. (2005). CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Reports*, *6*(2), 165–170. https://doi.org/10.1038/sj.embor.7400334

Naito, Y., Hino, K., Bono, H., & Ui-Tei, K. (2015). CRISPRdirect: Software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*, *31*(7), 1120–1123. https://doi.org/10.1093/bioinformatics/btu743

Newlon, C. S., & Theis, J. F. (1993). The structure and function of yeast ARS elements. *Current Opinion in Genetics & Development*, *3*(5), 752–758. https://doi.org/10.1016/S0959-437X(05)80094-2

Nguyen, N. T. T., Contreras-Moreira, B., Castro-Mondragon, J. A., Santana-Garcia, W., Ossio, R., Robles-Espinoza, C. D., … Thomas-Chollier, M. (2018). RSAT 2018: regulatory sequence analysis tools 20th anniversary. *Nucleic Acids Research*, *46*(W1), W209–W214. https://doi.org/10.1093/nar/gky317

Nüsslein-Volhard, C., & Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *October*, *287*(October), 795–801.

Ohtsuki, S., Levine, M., & Cai, H. N. (1998). Different core promoters possess distinct regulatory activities in the Drosophila embryo. *Genes and Development*, *12*(4), 547–556. https://doi.org/10.1101/gad.12.4.547

Pankratz, M. J., Seifert, E., Gerwin, N., Billi, B., Nauber, U., & Jäckle, H. (1990). Gradients of Krüppel and knirps gene products direct pair-rule gene stripe patterning in the posterior region of the drosophila embryo. *Cell*, *61*(2), 309–317. https://doi.org/10.1016/0092-8674(90)90811-R

Perrimon N, Mahowald AP. 1986. *l(1) hopscotch*, a larval-pupal zygotic lethal with a specific maternal effect on segmentation in *Drosophila*. *Developmental Biology* **118**: 28-41.

Perrimon N. 2020. https://www.flyrnai.org/evaluateCrispr/input.

Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T.-T. B., Misra, S., Murphy, C., … Rubin, G. M. (2008). Tools for neuroanatomy and neurogenetics in Drosophila. *Proceedings of the National Academy of Sciences*, *105*(28), 9715–9720. https://doi.org/10.1073/pnas.0803697105

Pick, L. (1998). Segmentation : Painting Stripes, *10*(May), 1–10.

Pick, L., Schier, A., Affolter, M., Schmidt-Glenewinkel, T., & Gehring, W. J. (1990). Analysis of the ftz upstream element: Germ layer-specific enhancers are independently autoregulated. *Genes and Development*, *4*(7), 1224–1239. https://doi.org/10.1101/gad.4.7.1224

Port, F., & Bullock, S. L. (2016). Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. *Nature Methods*, *13*(10), 852–854. https://doi.org/10.1038/nmeth.3972

Porto, M. S., Pinheiro, M. P. N., Batista, V. G. L., dos Santos, R. C., de Albuquerque Melo Filho, P., & de Lima, L. M. (2014). Plant Promoters: An Approach of Structure and Function. *Molecular Biotechnology*, *56*(1), 38–49. https://doi.org/10.1007/s12033-013-9713-1

Prazak, L., Fujioka, M., & Gergen, J. P. (2010). Non-additive interactions involving two distinct elements mediate sloppy-paired regulation by pair-rule transcription factors. *Developmental Biology*, *344*(2), 1048–1059. https://doi.org/10.1016/j.ydbio.2010.04.026

Ramakrishnan, A.-B., Chen, L., Burby, P. E., & Cadigan, K. M. (2021). Wnt target enhancer regulation by a CDX/TCF transcription factor collective and a novel DNA motif. *Nucleic Acids Research*, *49*(15), 8625–8641. https://doi.org/10.1093/nar/gkab657

Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., … Aiden, E. L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, *159*(7), 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021

Rhee, H. S., & Pugh, B. F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell*, *147*(6), 1408–1419. https://doi.org/10.1016/j.cell.2011.11.013

Riddihough, G., & Ish-Horowicz, D. (1991). Individual stripe regulatory elements in the Drosophila hairy promoter respond to maternal, gap, and pair-rule genes. *Genes and Development*, *5*(5), 840–854. https://doi.org/10.1101/gad.5.5.840

Rowley, M. J., Nichols, M. H., Lyu, X., Ando-Kuri, M., Rivera, I. S. M., Hermetz, K., … Corces, V. G. (2017). Evolutionarily conserved principles predict 3D chromatin organization. *Molecular Cell*, *67*(5), 837-852.e7. https://doi.org/10.1016/j.molcel.2017.07.022

Royo, J. L., Bessa, J., Hidalgo, C., Fernández-Miñán, A., Tena, J. J., Roncero, Y., … Casares, F. (2012). Identification and analysis of conserved cis-regulatory regions of the MEIS1 gene. *PLoS ONE*, *7*(3), e33617. https://doi.org/10.1371/journal.pone.0033617

Rubin, G. M., & Spradling, A. C. (1982). Genetic transformation of Drosophila with transposable element vectors. *Science*, *218*(4570), 348–353. https://doi.org/10.1126/science.6289436

Rubin, G. M., & Spradling, A. C. (1983). Vectors for P element-mediated gene transfer in

Drosophila. *Nucleic Acids Research*, *11*(18), 6341–6351. https://doi.org/10.1093/nar/11.18.6341

Schaffner, W. (2015). Enhancers, enhancers - from their discovery to today's universe of transcription enhancers. *Biological Chemistry*, *396*(4), 311–327. https://doi.org/10.1515/hsz-2014-0303

Schier AF, Gehring WJ. 1992. Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene. *Nature* **356**: 804-807.

Schier AF, Gehring WJ. 1993a. Analysis of a *fushi tarazu* autoregulatory element: multiple sequence elements contribute to enhancer activity. *EMBO J* **12**: 1111-1119.

Schier AF, Gehring WJ. 1993b. Functional specificity of the homeodomain protein fushi tarazu: The role of DNA-binding specificity. *Proc Natl Acad Sci USA* **90**: 1450-1454.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., … Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Schroeder MD, Pearce M, Fak J, Fan H, Unnerstall U, Emberly E, Rajewsky N, Siggia ED, Gaul U. 2004. Transcriptional control in the segmentation gene network of Drosophila. *PLoS Biol* **2**: E271.

Schroeder, M. D., Greer, C., & Gaul, U. (2011). How to make stripes: deciphering the transition from non-periodic to periodic patterns in Drosophila segmentation. *Development*, *138*(14), 3067–3078. https://doi.org/10.1242/dev.062141

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., … Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell*, *148*(3), 458–472. https://doi.org/10.1016/j.cell.2012.01.010

Shrinivas, K., Sabari, B. R., Coffey, E. L., Klein, I. A., Boija, A., Zamudio, A. V., … Chakraborty, A. K. (2019). Enhancer features that drive formation of transcriptional condensates. *Molecular Cell*, *75*(3), 549-561.e7. https://doi.org/10.1016/j.molcel.2019.07.009

Signorell {Andri et mult. al.}. (2022). {DescTools}: Tools for descriptive statistics. Retrieved from https://cran.r-project.org/package=DescTools

Small, S., Blair, A., & Levine, M. (1992). Regulation of even-skipped stripe 2 in the Drosophila embryo. *EMBO Journal*, *11*(11), 4047–4057. https://doi.org/10.1002/j.1460-2075.1992.tb05498.x

Small, Stephen, & Arnosti, D. N. (2020). Transcriptional enhancers in Drosophila. *Genetics*, *216*(1), 1–26. https://doi.org/10.1534/genetics.120.301370

Small, Stephen, Blair, A., & Levine, M. (1996). Regulation of two pair-rule stripes by a single enhancer in the Drosophila embryo. *Developmental Biology*, *175*(2), 314–324. https://doi.org/10.1006/dbio.1996.0117

Small, Stephen, Kraut, R., Hoey, T., Warrior, R., & Levine, M. (1991). Transcriptional regulation of a pair-rule stripe in Drosophila. *Genes and Development*, *5*(5), 827–839. https://doi.org/10.1101/gad.5.5.827

Spitz, F., & Furlong, E. E. M. (2012). Transcription factors: From enhancer binding to developmental control. *Nature Reviews Genetics*, *13*(9), 613–626. https://doi.org/10.1038/nrg3207

Spradling, A. C., & Rubin, G. M. (1982). Transposition of cloned P elements into Drosophila germ line chromosomes. *Science*, *218*(4570), 341–347. https://doi.org/10.1126/science.6289435

Štanojević, D., Hoey, T., & Levine, M. (1989). Sequence-specific DNA-binding activities of the gap proteins encoded by hunchback and Krüppel in Drosophila. *Nature*, *341*(6240), 331–335. https://doi.org/10.1038/341331a0

Stanojevic, D., Small, S., & Levine, M. (1991). Regulation of a segmentation stripe by overlapping activators and repressors in the Drosophila embryo. *Science*, *254*(5036), 1385–1387. https://doi.org/10.1126/science.1683715

Stormo, G. D. (2000). DNA binding sites: Representation and discovery. *Bioinformatics*, *16*(1), 16–23. https://doi.org/10.1093/bioinformatics/16.1.16

Struhl G. 1985. Near-reciprocal phenotypes caused by inactivation or indiscriminate expression of the *Drosophila* segmentation gene *ftz*. **318**: 677-680.

Su, W., Jackson, S., Tjian, R., & Echols, H. (1991). DNA looping between sites for transcriptional activation: Self-association of DNA-bound Sp1. *Genes and Development*, *5*(5), 820–826. https://doi.org/10.1101/gad.5.5.820

Surkova, S., Myasnikova, E., Janssens, H., Kozlov, K. N., Samsonova, A. A., Reinitz, J., & Samsonova, M. (2008). Pipeline for acquisition of quantitative data on segmentation gene expression from confocal images, *2*(2), 58–66.

Surkova, S., Myasnikova, E., Kozlov, K. N., Pisarev, A., Reinitz, J., & Samsonova, M. (2013). Quantitative imaging of gene expression in drosophila embryos. *Cold Spring Harbor Protocols*, *8*(6), 488–497. https://doi.org/10.1101/pdb.top075101

Surkova, S., Myasnikova, E. M., Kozlov, K. N., Samsonova, A. A., Reinitz, J., & Samsonova, M. G. (2008). Methods for acquisition of quantitative from confocal images of gene expression in situ. *Tsitologiya*, *50*(4), 352–369. https://doi.org/10.1134/S1990519X08020156

Surkova, S., Sokolkova, A., Kozlov, K., Nuzhdin, S. V., & Samsonova, M. (2019). Quantitative analysis reveals genotype- and domain- specific differences between mRNA and protein expression of segmentation genes in Drosophila. *Developmental Biology*, *448*(1), 48–58. https://doi.org/10.1016/j.ydbio.2019.01.006

Tautz, D. (1988). Regulation of the Drosophila segmentation gene hunchback by two maternal morphogenetic centres. *Nature*, *332*(6161), 281–284. https://doi.org/10.1038/332281a0

Tautz, D., & Pfeifle, C. (1989). A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback. *Chromosoma*, *98*(2), 81–85. https://doi.org/10.1007/BF00291041

Team, R. C. (2022). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Tillberg, P. W., Chen, F., Piatkevich, K. D., Zhao, Y., Yu, C.-C., English, B. P., … Boyden, E. S. (2016). Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nature Biotechnology*, *34*(9), 987–992. https://doi.org/10.1038/nbt.3625

Tolhuis, B., Palstra, R.-J., Splinter, E., Grosveld, F., & de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active β-globin locus. *Molecular Cell*,

10(6), 1453–1465. https://doi.org/10.1016/S1097-2765(02)00781-5

Topol T, Dearolf CR, Prakash K, Parker CS. 1991. Synthetic oligonucleotides recreate *Drosophila fushi tarazu* zebra-stripes expression. *Genes & Dev* **5**: 855-867.

Tsai, A., Alves, M. R. P., & Crocker, J. (2019). Multi-enhancer transcriptional hubs confer phenotypic robustness. *ELife*, *8*. https://doi.org/10.7554/eLife.45325

Tsai, A., Muthusamy, A. K., Alves, M. R. P., Lavis, L. D., Singer, R. H., Stern, D. L., & Crocker, J. (2017). Nuclear microenvironments modulate transcription from low-affinity enhancers. *ELife*, *6*, 1–18. https://doi.org/10.7554/eLife.28975

Tunnacliffe, E., & Chubb, J. R. (2020). What Is a Transcriptional Burst? *Trends in Genetics*, *36*(4), 288–297. https://doi.org/10.1016/j.tig.2020.01.003

Vachon, G., Cohen, B., Pfeifle, C., McGuffin, M. E., Botas, J., & Cohen, S. M. (1992). Homeotic genes of the bithorax complex repress limb development in the abdomen of the Drosophila embryo through the target gene Distal-less. *Cell*, *71*(3), 437–450. https://doi.org/10.1016/0092-8674(92)90513-C

van Steensel, B., & Furlong, E. E. M. (2019). The role of transcription in shaping the spatial organization of the genome. *Nature Reviews Molecular Cell Biology*, *20*(6), 327–337. https://doi.org/10.1038/s41580-019-0114-6

Verger, A., Monté, D., & Villeret, V. (2019). Twenty years of Mediator complex structural studies. *Biochemical Society Transactions*, *47*(1), 399–410. https://doi.org/10.1042/BST20180608

Vincent, B. J., Staller, M. V., Lopez-Rivera, F., Bragdon, M. D. J., Pym, E. C. G., Biette, K. M., … DePace, A. H. (2018). Hunchback is counter-repressed to regulate even-skipped stripe 2 expression in Drosophila embryos. *PLoS Genetics*, *14*(9), 1–21. https://doi.org/10.1371/journal.pgen.1007644

Wieschaus, E., & Nusslein-Volhard, C. (2016). The Heidelberg screen for pattern mutants of drosophila: A personal account. *Annual Review OfCell and Developmental Biology*. https://doi.org/10.1146/annurev-cellbio-113015-023138

Yu Y, Li W, Su K, Han W, Yussa M, Perrimon N, Pick L. 1997. The nuclear hormone receptor FTZ-F1 is a cofactor for the *Drosophila* homeodomain protein Ftz. *Nature* **385**: 552-555.

Yu, Y., & Pick, L. (1995). Non-periodic cues generate seven ftz stripes in the Drosophila embryo. *Mechanisms of Development*, *50*(2–3), 163–175. https://doi.org/10.1016/0925-4773(94)00333-I

Zhang, X., Odom, D. T., Koo, S.-H., Conkright, M. D., Canettieri, G., Best, J., … Montminy, M. (2005). Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. *Proceedings of the National Academy of Sciences*, *102*(12), 4459–4464. https://doi.org/10.1073/pnas.0501076102

Zhao, H., Friedman, R. D., & Fournier, R. E. K. (2007). The locus control region activates serpin gene expression through recruitment of liver-specific transcription factors and RNA polymerase II. *Molecular and Cellular Biology*, *27*(15), 5286–5295. https://doi.org/10.1128/MCB.00176-07