Stony Brook University

Developmentally Regulated Transcription Elongation in the *Drosophila* Embryo

A Dissertation Presented

by

Xiaoling Wang

to

The Graduate School

in Parial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

May 2008

Stony Brook University

The Graduate School

Xiaoling Wang

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation

Dr. J. Peter Gergen – Dissertation Advisor
Professor, Department of Biochemistry and Cell Biology

Dr. Gerald H. Thomsen – Chairperson of Defense
Associate Professor, Department of Biochemistry and Cell Biology

Dr. Rolf Sternglanz
Professor, Department of Biochemistry and Cell Biology

Dr. Robert S. Haltiwanger
Professor, Department of Biochemistry and Cell Biology

Dr. Carl W. Anderson
Professor, Biology Department, Brookhaven National Laboratory

Dr. David S. Gilmour
Professor, Department of Biochemistry and Molecular Biology, Penn State University

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Developmentally Regulated Transcription Elongation in the *Drosophila* Embryo

by

Xiaoling Wang

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

2008

For *Drosophila* embryos, gastrulation happens within a 30 minute window, spanning the completion of cellularization and the beginning of germ band extension. During this short time period, the segmentation pattern is established, as cell fates are specified at a single-cell level of resolution. Transcription plays an important role during this critical stage. The simple combinatorial rules for regulation of the *sloppy-paired-1 (slp1)* gene by the pair-rule transcription factors during early embryogenesis offer a unique opportunity to investigate the molecular mechanisms of developmentally regulated transcription. We found that the initial repression of *slp1* in response to pair-rule factors *Runt* and *Fushi-tarazu (Ftz)* does not involve chromatin remodeling, or histone modification. Instead, in *slp1*-repressed cells, RNA-polymerase II initiates transcription at the *slp1* promoter and pauses downstream in a complex that includes negative elongation factor (NELF). Additionally studies indicated that NELF is maternally provided and recruited to gene promoters prior to their transcription. Further functional studies demonstrated that NELF plays a pivotal role in transcriptional regulation during this critical stage of *Drosophila* embryogenesis.

Table of Contents

List of Abbreviations

slp1	sloppy-paired1
ftz	fushi tarazu
wg	wingless
en	engrailed
odd	odd-skipped
eve	even-skipped
opa	odd-paired
hsp	heat shock protein
bro	brother
NGT	nos-Gal4-tubulin
TBP	TATA binding protein
Pol II	RNA polymerase II
CTD	C-terminal domain
NELF	Negative elongation factor
RRM	RNA recognition motif
DRB	5,6-dichloro-1-β-D-ribofuranosylbenzimidazole
DSIF	DRB sensitivity-inducting factor
P-TEFb	Positive transcription elongation factor b
TSS	Transcriptional start site
ChIP	Chromatin Immuno-precipitation
ChIP-chip	Chromatin immuno-precipitation microarray
3C	Chromosome conformation capture
q-RT-PCR	Quantitative reverse transcribed polymerase chain reaction
GLC	Germ line clone
AED	After egg deposition
WT	Wild type
R+F	Runt and Ftz overexpressed

List of Figures

List of Tables

PREFACE

It is not birth, marriage, or death, but gastrulation, which is truly the most important time in your life. LEWIS WOLPERT 1986

For *Drosophila* embryos, gastrulation happens within a 30 minute window, spanning the completion of cellularization and the beginning of germ band extension. During this short time period, the segmentation pattern is established, as cell fates are specified at a single-cell level of resolution. In this thesis, I concluded that the specification of early embryonic cell fates involves the control of transcriptional elongation and found that negative elongation factor NELF has important roles during this critical stage.

Chapter one starts with a brief introduction of *Drosophila* early embryogenesis and the transcription hierarchy involved in segmentation. I will also introduce the genetic background I used for my experiments. At last, I will also describe the steps in transcriptional regulation, how chromatin structure affects transcription, and what methods can be used to address this problem.

In chapter two, I used the *sloppy-paired-1* (*slp1*) gene as a model system to study the process of cell specification. I found that the initial repression of *slp1* in response to Runt and Fushi-tarazu (Ftz) does not involve chromatin remodeling or histone modification. Instead, in *slp1*-repressed cells, *slp1* transcription is initiated and RNA polymerase is paused downstream of the promoter in a complex that includes NELF. The finding that NELF also associates with the promoter regions of *wingless* (*wg*) and *engrailed* (*en*), two other segment polarity genes, strongly suggests that developmentally regulated transcriptional elongation is central to the process of cell fate specification during this stage.

In chapter three, I used biochemical and genetic approaches to investigate the function of NELF during early embryonic development. I found that NELF was maternally provided and recruited to the promoter-proximal region of several segmentation genes prior to their transcriptional activation. Experiments with a dominant negative NELF subunit revealed a functional role for NELF in the segmentation pathway. Additional experiments also suggested that NELF was involved in cell cycle regulation.

Chapter four contains discussions, thoughts, ongoing work and potential future areas of interest. I also included several approaches I tried but which didn't

give me expected or fully developed results. For these experiments, I discuss potential explanations for these results and provide suggestions for future studies.

This thesis is mainly focused on studying the role of transcriptional regulation in the process of cell specification during development. There are many other forms of regulation not touched here, such as microRNA regulation, the spreading of repression at the chromosomal level, differential RNA processing, translational and posttranslational regulation. I believe the reality must be more complicated than the model I propose in this thesis. As Vincent van Gogh said, "*do not quench your inspiration and your imagination, do not become the slave of your model.*"

Chapter 1 General Introduction

Early *Drosophila* development and transcription hierarchy

The *Drosophila* embryo provides a powerful model for investigating transcriptional regulatory mechanisms during development. After egg fertilization, the zygotic nucleus undergoes 8 mitotic divisions within the central portion of the egg with an average of 8 minutes each. The nuclei then migrate to the periphery of the egg, while the mitoses continue at a slower rate. During the ninth division cycle, about five nuclei migrate to the posterior end of the embryo and generate the pole cells that become gametes of the adult. The rest of the nuclei reach the periphery of the embryos at cycle 10 and then undergo four more divisions. Following cycle 13, the oocyte plasma membrane folds inward between the nuclei to form around 6000 single cells. This stage is defined as the cellular blastula stage and happens within 4 hours after fertilization. Transcription starts at cycle 11 and is greatly enhanced at this stage. The development of embryo at this stage can be tracked by the inward growth of the cell membranes and the shape of the cells, which progress from round to square and then to a final elongated shape.

Following the formation of the cellular blastoderm, gastrulation begins. The first movement of gastrulation segregates the presumptive mesoderm, endoderm, and ectoderm. The prospective mesoderm folds inward to produce the ventral furrow and a localized change of cell shapes is associated with an invagination to form the cephalic furrow. These two characters can be clearly seen under the microscope and are readily used to identify gastrula embryos.

After gastrulation, the germ band extends posteriorly and wraps around onto the dorsal surface of the embryo, with the posterior end of the germ band extending to a location immediately behind the head region. The germ band later retracts, placing the presumptive posterior segments at the posterior tip of the embryo. The body segmentation fully forms during these stages and will be kept from the rest of embryogenesis to the larva and adult.

The origins of anterior-posterior polarity of the embryo come from the uneven distribution of maternally provided positional information. As shown in Figure 1-1, the gap genes are then regulated by the maternal effect genes and their localized expression divides the embryo in broad regions. At around cycle 13, the overlapping combinations of gap gene expression will drive the expression of the pair-rule genes into seven stripes. The complementary seven-stripped expression

patterns of two of the pair-rule genes, *even-skipped (eve)* and *fushi tarazu (ftz)* define 14 metameric units referred to as parasegments. Moreover, each row of nuclei within each parasegment expresses a unique combination of pair-rule products. These products will drive the 14-stripe expression of the segment polarity genes in the late blastoderm embryo such that by the onset of gastrulation, cell fates are specified at a single cell level of resolution along the anterior-posterior axis. Table 1-1 lists the major genes affecting segmentation in *Drosophila*.

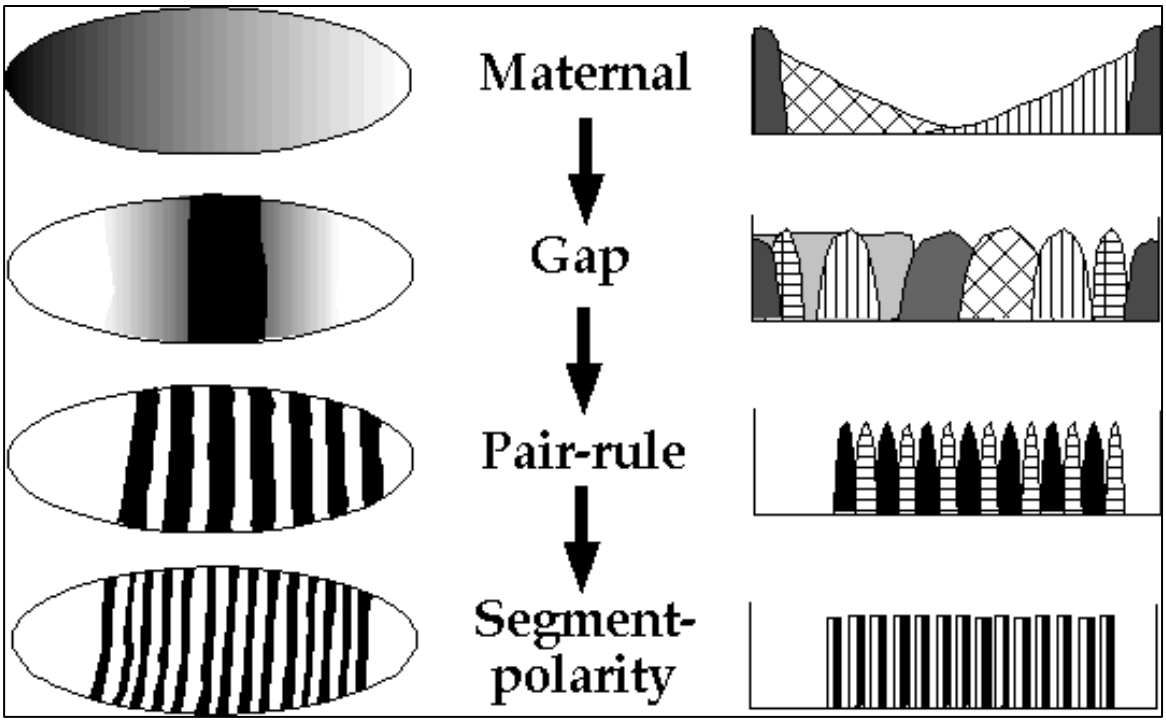


Figure 1-1 ***Drosophila*** **segmentation gene hierarchy** Maternal genes establish initial positional gradients to regulate the expression patterns of Gap genes. Differing concentrations of the gap proteins drive the periodic expression of Pair-rule genes. The pair-rule proteins then enable the expression of 14-stripe Segment-polarity genes to establish the *Drosophila* body segments.

Category		Category		Category	
Gap genes	*Kruppel(Kr)* *knirps (kni)* *hunchback (hb)* *giant (gt)* *tailless (tll)* *huckebein (hkb)* *buttonhead (btd)* *empty spiracles (ems)* *orthodenticle (otd)*	**Pair-rule genes Primary**	*hairy (h)* *even-skipped (eve)* *runt (run)*	**Segment polarity genes**	*sloppy paired (slp)* *engrailed (en)* *wingless (wg)* *cubitus interruptus (ci)* *hedgehog (hh)* *fused (fu)* *armadillo (arm)* *patched (ptc)* *gooseberry (gsb)* *pangolin (pan)*
		Pair-rule gene Secondary	*fushi tarazu (ftz)* *odd-paired (opa)* *odd-skipped (odd)* *paired (prd)*		

Table 1-1 Major genes affecting segmentation pattern in *Drosophila* (Modified from Scott F. Gilbert, Developmental Biology Sixth Edition, Table 9.2)

Quantitative analysis of gene function in the *Drosophila* embryo

Our laboratory previously developed a system that allowed for genetically controlled ectopic expression of any gene of interest under normal physiological conditions in the early *Drosophila* embryo (Tracey et al. 2000). As shown in Figure 1-2, females have a transgene call NGT (nos-GAL4-tubulin). Under the control of the nanos promoter, the yeast transcription factor GAL4 can be expressed specifically during oogenesis. The 3' UTR from the *Tub84B* mRNA uniformly distributes maternally expressed GAL4 throughout the embryo. When these NGT females are crossed to males containing an UAS transgene X, gene X will be activated when transcription starts in embryo. The amount of gene X product will continue to increase until several hours later, when the embryos are at full germ band extension. Different NGT lines have been established that produce different levels of maternal GAL4 activity. Presumably, these differences are due to the sites of transgene insertion. Similarly, different UAS transgenes also have different strengths. By using different cross combinations, we are able to study the effects of ectopic gene expression over a wide range. Importantly, this system can be quantitatively calibrated by crossing NGT females to *UAS-LacZ* expressing embryos and measuring β-galactosidase activity.

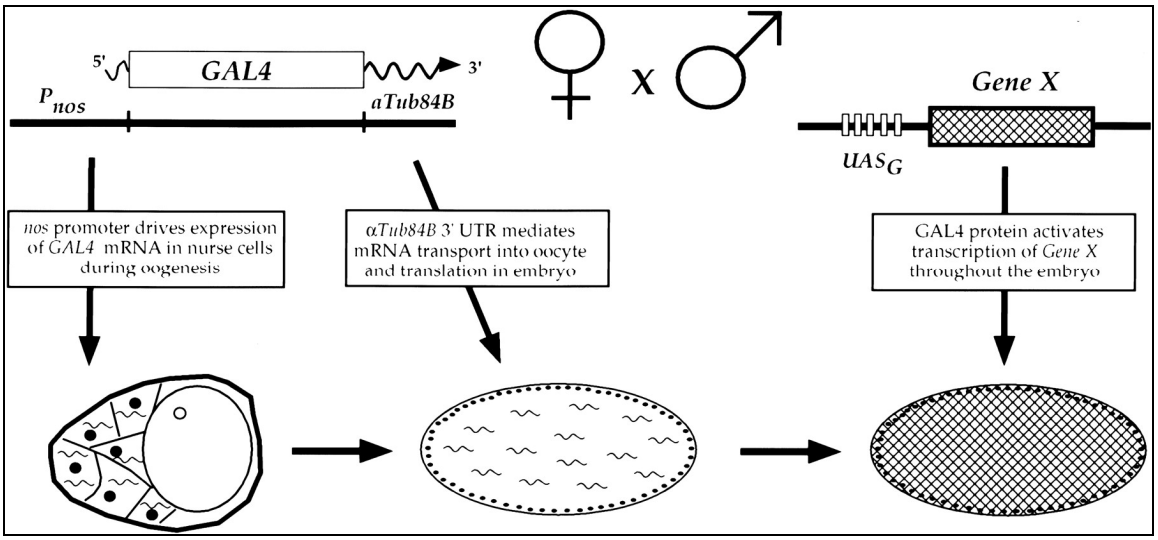


Figure 1-2 GAL4-dependent transcriptional regulation in early *Drosophila* embryos (Adapted from Tracey et al, Genetics 154:273, Figure 1) GAL4-dependent transcriptional regulation is achieved in early embryos that contain maternally expressed GAL4 mRNA. The nos promoter is used to drive expression of the GAL4 mRNA specifically during oogenesis. Inclusion of the 3' untranslated region of a maternally expressed tubulin mRNA allows for uniform deposition of the GAL4 mRNA in the developing oocyte. *GAL4*-dependent transgenes can be inherited either from the male in the cross (as shown) or from the female.

Runt-dependent *sloppy paired 1 (slp1)* regulation

The pair-rule transcription factor Runt is the founding member of the Runx family of developmental regulators (Kagoshima et al. 1993). All Runx family proteins contain a highly conserved 128 amino acid motif, referred to as the Runt domain. This domain functions in DNA binding, ATP binding, and protein-protein interactions (Coffman 2003). From sea urchins to mammals, Runx proteins are involved in regulating cell proliferation and differentiation. Mutations in *runx* genes are linked to a number of diseases, such as acute leukemia, cleidocranial dysphasia, and gastric cancer. An important property of Runt as well as other genes involved in cell differentiation is that they can function both as transcriptional activators and repressors, depending on the target gene and developmental context (Coffman 2003). But how these opposing functions can be achieved by one protein remains unclear.

During the past several years, our lab identified the segment polarity gene *sloppy paired 1 (slp1)* as a good model to study the mechanism of Runt-dependent regulation during early embryo development (Swantek and Gergen 2004). *slp1* is a sensitive target of Runt, Eve, and Ftz. Figure 1-3 outlines the combinatorial regulation for *slp1* in the segmentated region of the body. Runt and Opa together in the absence of Ftz activate the odd stripes of *slp1*. Eve alone can repress *slp1* in odd-numbered parasegments. Runt and Ftz together repress *slp1* in even parasegments. These relatively simple rules and the sensitive response of *slp1* make it possible for us to manipulate *slp1* transcription genetically.

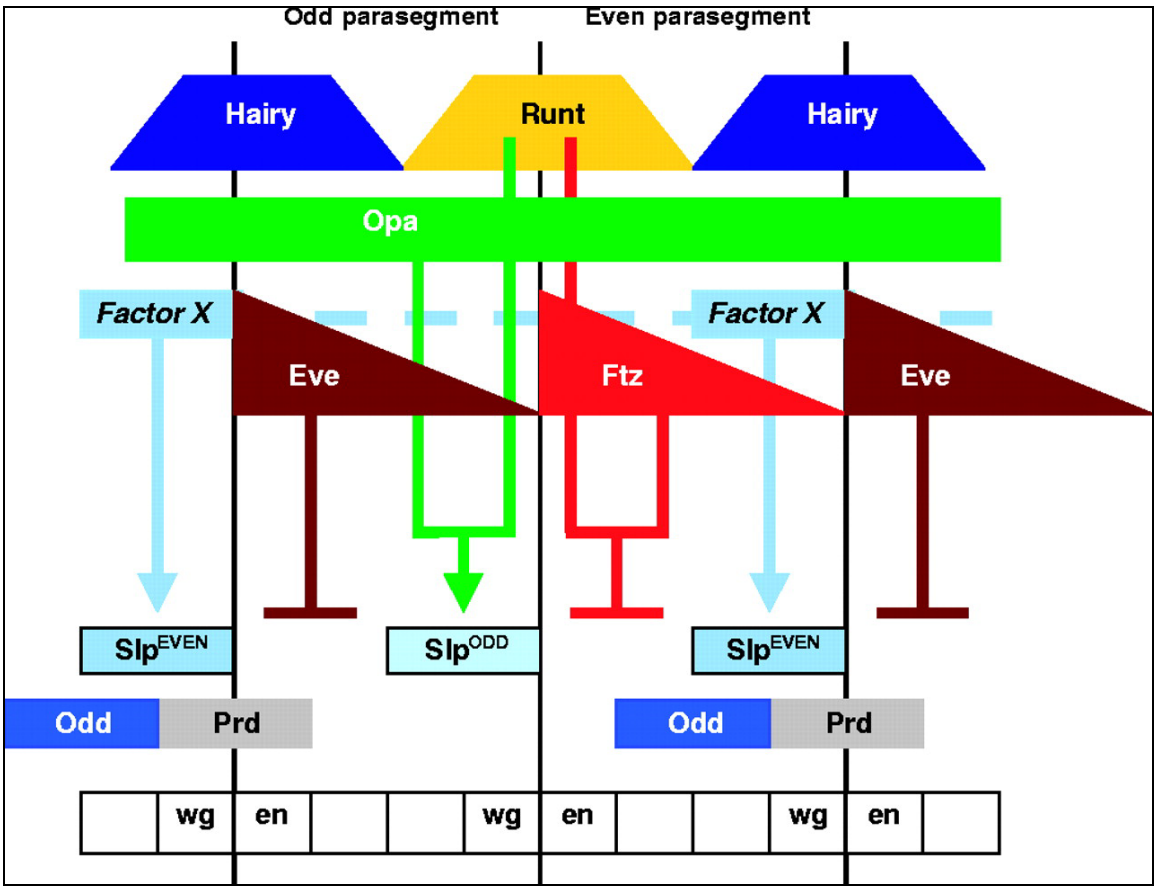


Figure 1-3 ***slp1* regulation by pair-rule transcription factors** (Adapted from Swantek and Gergen, Development 131:2281, Figure 6) The activation and repression of *slp1* in different cells within the pre-segmental region of the blastoderm embryo is explained by pair-rule genes. The combinatorial regulation by Eve, Runt, Opa and Ftz accounts for all aspects of *slp1* regulation, except for activation in even-numbered parasegments. The minimal spatial domain of activity of a Factor X that is proposed to be responsible for this aspect of *slp1* expression is depicted in pale blue. Factor X-dependent activation may also contribute to the expanded *slp1* expression obtained by transient elimination of *eve* or *runt.* The possibility that Factor X is active in other cells within the pre-segmental region is indicated by the broken blue line. The strip of cells along the anteroposterior axis drawn at the bottom of this diagram shows the relationship of *wg* and *en* expression in each parasegment to that of *slp1* and the pair-rule transcription factors.

Transcription regulation

The appropriate regulation of mRNA transcription is crucial to the proper function, cell differentiation, and development of complex organisms. In brief, mRNAs are synthesized by the action of many general and gene specific transcription factors in a cooperative manner. Although it sounds simple, the entire transcriptional machinery is not controlled through a simple on/off switch at the promoter. Instead, it can be regulated at many different levels. In this section I will briefly describe the RNA polymerase (Pol) II involved transcription regulation. Transcription by RNA Pol I (ribosomal RNA genes) or RNA Pol III (small nuclear and transfer RNA genes) will not be discussed here or in the thesis.

Steps in early transcription

The general transcriptional machinery for RNA Pol II and a common transcription pathway are shared by thousands of genes in eukaryotic cells. We can use a simplified step-by-step model to describe the process of producing an elongationally competent Pol II complex (Lis 1998). First, tight-packed chromatin structure can create an early barrier to prevent the recruitment of transcription factors to a gene locus. Some sequence-specific DNA binding factors as well as chromatin remodeling machinery are needed to open the chromatin. When the chromatin is open, the TFIID complex (including TBP the TATA binding protein) will be recruited to the promoter region. Based on the crystal structure, TBP tightly bound to the TATA element creates a unique and huge DNA distortion (Kim et al. 1993). The severe change in the path of the DNA may increase the proximity of binding by other proteins and disfavor nucleosome packaging to prevent chromatin repression (Orphanides et al. 1996). Other factors and Pol II are then recruited via an ordered assembly (Buratowski et al. 1989) or as a holoenzyme that contains many of the general transcription factors, mediator, and the core Pol II (Thompson et al. 1993; Kim et al. 1994). The Pol II complex then opens the downstream DNA and forms the first phosphodiester bond to initiate transcription. Pol II at this stage can synthesize a short transcript but may be paused for some reason. Specific activators may be recruited later to push the Pol II transcriptional machinery to pass the pausing barrier and finish transcription.

Transcription repression

Transcriptional repression plays important roles in the precise control of gene expression. The repression of transcription can be produced by the neutralization of a positively acting factor, or by direct repression of the basal transcriptional complex as well as by alteration of chromatin structure, or the inhibition of transcriptional elongation (Maldonado et al. 1999; Courey and Jia 2001).

Even for the same gene, different mechanisms are used to achieve transcriptional repression during development. By studying Runt-dependent *engrailed* repression, our laboratory found that this repression is a two-step process and that the establishment and maintenance of repression are distinct: establishment requires Tramtrack but does not require high-affinity DNA binding by Runt, and maintenance requires Groucho, dCtBP, Rdp3 and DNA-binding by Runt (Wheeler et al. 2002).

The separation of establishment and maintenance of repression can also be observed in other repression pathways. For example, the patterns of *Drosophila* Hox transcription set up by transiently expressed factors during development are subsequently maintained in a very stable and heritable manner by chromatin modifications mediated by the polycomb group (PcG) proteins (Brock and van Lohuizen 2001).

Pol II pausing and transcription elongation

The pausing of Pol II and regulation of transcription at an elongation step has been shown in several genes such as *Drosophila hsp70, hsp26, hsp27, βTub*, human *c-myc*, *c-fos*, *Junb*, *Igk*, and the *human immunodeficiency virus (HIV)* (Fujinaga et al. 1998; Lis 1998; Raschke et al. 1999; Aiyar et al. 2004). Among these genes, the *Drosophila* hsp70 promoter is an extensively studied example of regulated transcriptional elongation (Rougvie and Lis 1988; Boehm et al. 2003; Lebedeva et al. 2005). Under normal growth conditions, Pol II initiates transcription but pauses in the region 20 to 40 nucleotides downstream. In response to heat shock, a DNA sequence-specific activator, heat shock factor (HSF) rapidly trimerizes and binds to heat shock loci (Westwood et al. 1991). This immediately triggers transcription. Within 70 seconds, the first wave of Pol II is able to be detected beyond the paused region of *hsp70*. Within 3 minutes the amount of Pol II on *hsp70* is near its fully induced level (O'Brien and Lis 1993). Many factors are involved in this process. Figure 1-4 is a summary of the regulation of *hsp70*.

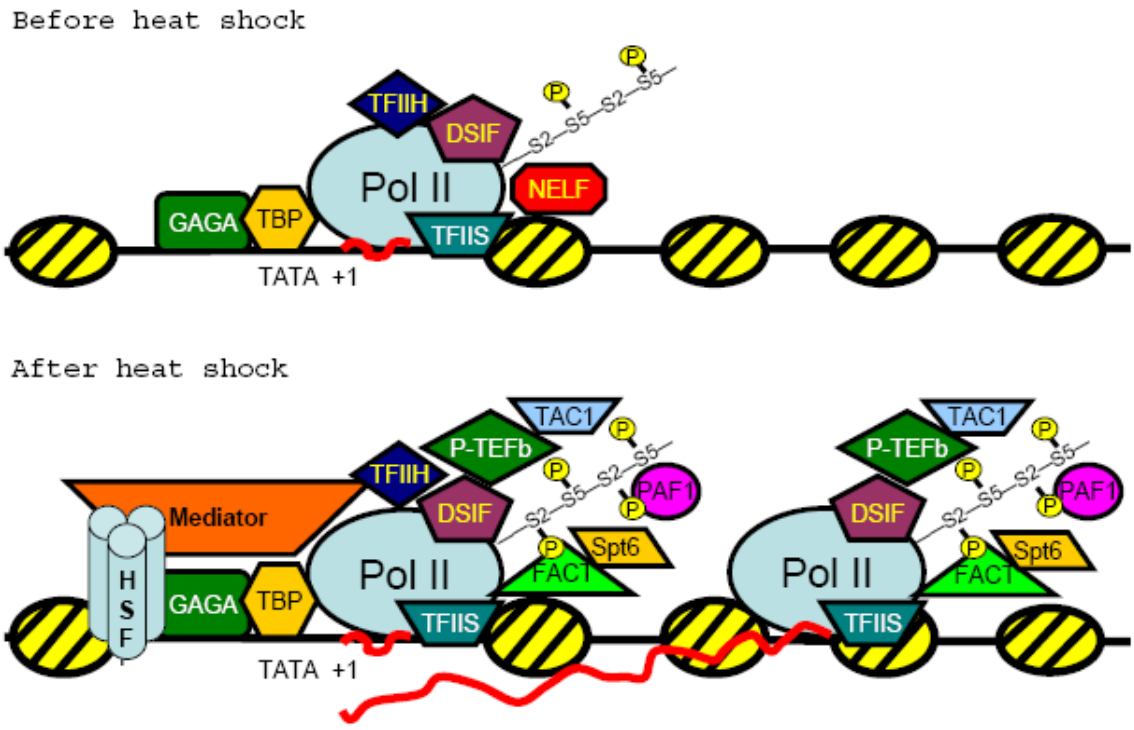


Figure 1-4 Regulation of transcription elongation: the hsp70 paradigm The *Drosophila* hsp70a promoter is an extensively studied example of regulated transcriptional elongation. Before heat shock, chromatin is open; GAGA and TBP bind to promoter. Pol II initiates transcription and pauses immediately downstream from the promoter with Serine 5 phosphorylation on the CTD. DSIF and NELF are thought to help Pol II pausing. TFIIS will facilitate efficient release of Pol II from the pause site. Upon heat shock, trimerized HSF immediately binds upstream of the promoter and recruits Mediator, P-TEFb, TAC1, FACT, Spt6 and PAF1. NELF dissociates from the Pol II complex. The Pol II CTD is phosphorylated at serine 2 and transcription continues to the end.

Recent work by Dr. Frank Pugh pointed out that the +1 nucleosome is very likely to be the barrier to inhibit elongation by Pol II. His group sequenced nucleosomes containing H3 and the histone variant H2A.Z with a median 4-base-pair resolution from 0 to 12 hour old *Drosophila* embryos. By comparing Pol II chromatin immuno-precipitation microarray (ChIP-chip) data with histone distribution data, they found that in very large populations, paused Pol II sits right next to the +1 nucleosome and even pushed it 10bp downstream (Personal communication). But how repressors work together to set the nucleosome barrier or how activators remove the barrier still remains unknown.

Recently, several genome-wide Pol II ChIP-chip assays in human cells and *Drosophila* S2 cells identified subsets of genes having Pol II pausing with no RNA

expression (Kim et al. 2005; Lee et al. 2006; Muse et al. 2007). More interestingly, ChIP-chip assay with *Drosophila* 2-4 hour old embryos shows that around 12% of genes have stalled Pol II at the promoter proximal region. Further analysis shows that these genes are highly enriched for functions in development, including neurogenesis, ectoderm development and muscle differentiation (Zeitlinger et al. 2007). This strongly suggests that regulation of transcriptional elongation is a potentially widespread strategy in response to either stimuli or rapid temporal and spatial changes during development.

NELF (Negative elongation factor) and DSIF (DRB sensitivity-inducing factor) are two protein factors that act together to regulate transcription elongation by Pol II. They were first discovered from HeLa cell nuclear extracts. With the addition of a transcription inhibitor DRB (5,6-dichloro-1-β-D-ribofuranosylbenzimidazole), DSIF and NELF can cause the pausing of Pol II *in vitro* (Wada et al. 1998; Yamaguchi et al. 1999). Human DSIF, as well as *Drosophila* DSIF, contains two subunits, Spt4 and Spt5, respectively. Human NELF contains 4 subunits, NELFA, NELFB, NELF-D and NELF-E. *Drosophila* NELF also contains four subunits with significant sequence similarity to the four subunits of human NELF (Wu et al. 2005). Immuno-depletion of either protein from *in vitro* transcription reactions or ablation of NELF-E from salivary glands with RNAi reduces the level of paused polymerase in the promoter proximal region of hsp70 (Wu *et al.* 2003; Wu *et al.* 2005).

This negative transcription effect caused by DSIF and NELF can be alleviated by P-TEFb (positive transcription elongation factor) (Ptashne and Gann 1997; Wada et al. 1998). P-TEFb is a kinase that phosphorylates the C-terminal domain (CTD) of Pol II (Marshall et al. 1996) P-TEFb also phosphorylates the human Spt5 protein (Kim and Sharp 2001), making DSIF function as a positive elongation factor (Yamada et al. 2006). P-TEFb can also phosphorylate human NELF-E at several serine residues. This modification reduces the RNA binding ability of NELF-E and removes NELF from the paused Pol II complex. As a consequence, transcription is no longer repressed and productive elongation of transcription ensues (Fujinaga et al. 2004). *Drosophila* P-TEFb contains two subunits, Cdk9 and CycT. Figure 1-5 is a summary of promoter-proximal pausing and the factors involved in the escape to productive elongation (Saunders et al. 2006).

Interestingly, Although DSIF and P-TEFb have homologs in eukaryotes ranging from yeast to human, no homologs of the four subunits of NELF are identified in model organisms such as yeast or *Caenorhabditis elegans* (Narita et al. 2003). Thus, the regulatory potential provided by NELF could be restricted to a subset of eukaryotes.

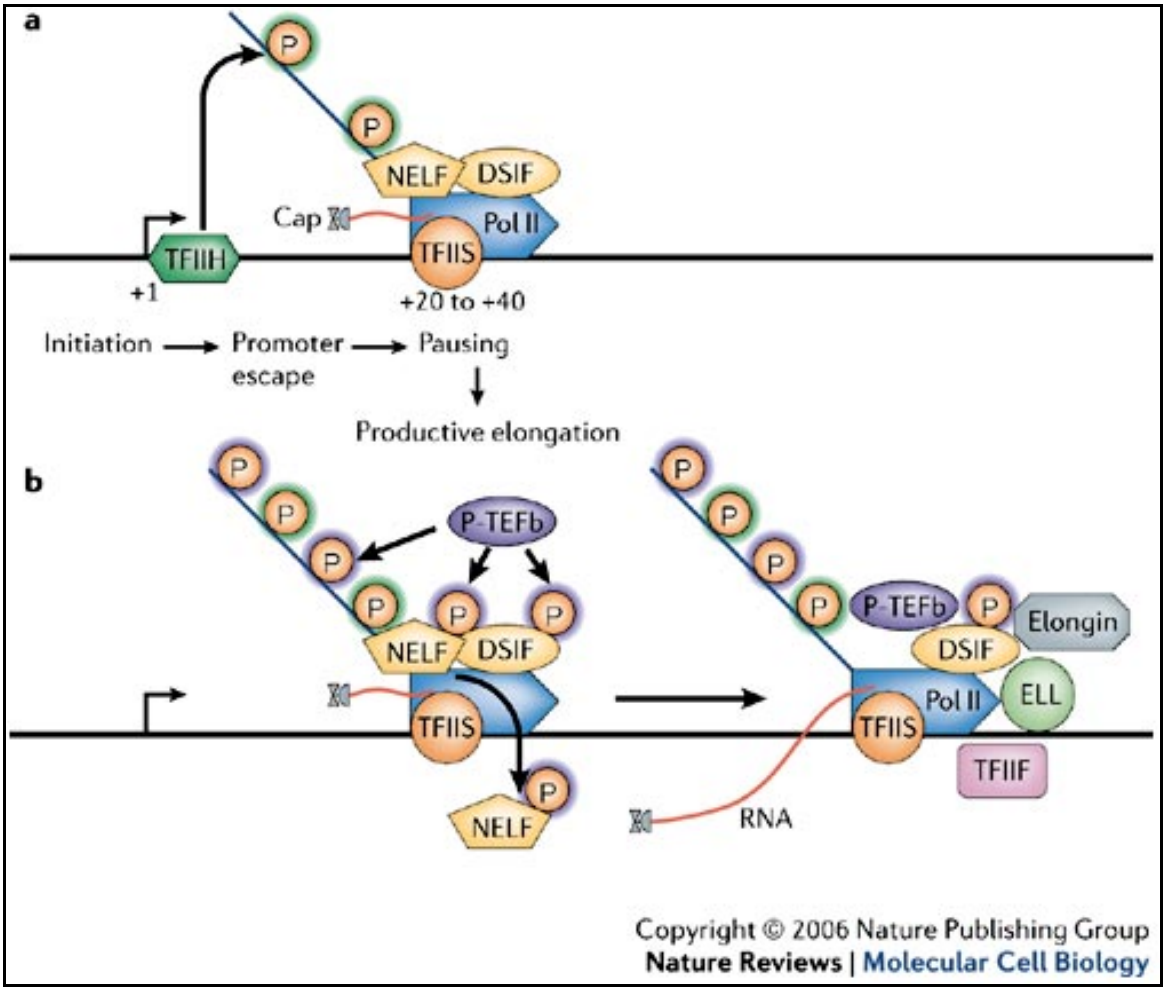


Figure 1-5 Promoter-proximal pausing and escape to productive elongation (Adapted from Saunders *et al. Nature Reviews Molecular Cell Biology* **7**, 557–567 (August 2006) **a.** TFIIH-mediated phosphorylation of Ser5 of the CTD of Pol II occurs on pre-initiation complex formation or before promoter-proximal pausing. DSIF and NELF probably facilitate Pol II pausing in the promoter-proximal region, and TFIIS also associates with the paused polymerase. TFIIS stimulates the intrinsic RNA-cleavage activity of Pol II to create a new RNA 3′-OH in the Pol II active site after backtracking of the polymerase. Capping enzyme associates with the Ser5-phosphorylated CTD and with Spt5, and the nascent RNA becomes capped during this first stage of elongation. The RNA cap is formed by addition of a methylated guanosine to the 5′ end of the RNA through the action of capping enzyme and an RNA (guanine-7) methyltransferase. **b**. P-TEFb-mediated phosphorylation of DSIF, NELF and Ser2 of the Pol II CTD stimulates productive elongation, and the capping enzyme might contribute to this process by counteracting the negative effects of DSIF and NELF. TFIIS facilitates efficient release of Pol II from the pause site by aiding the escape of backtracked transcription complexes. NELF dissociates from the transcription complex and DSIF, TFIIS and P-TEFb track with Pol II along the gene. TFIIF, ELL, and elongin, which stimulate Pol II elongation activity, might also associate with the elongation complex.

Role of chromatin structure in gene regulation

DNA in eukaryotic cells is packaged into a structure known as chromatin. The fundamental unit of this structure is the nucleosome. A nucleosome has 147bp of DNA wrapped approximately twice around a unit of eight histone molecules (two each of histones H2A, H2B, H3 and H4). Each nucleosome is bound by linker histones (most commonly H1) and separated by an average of 29-43 bp of linker DNA (Kornberg and Lorch 1999; Woodcock 2006; Woodcock et al. 2006). Nucleosomes disassemble and reassemble quickly to affect the access of transcription factors to their appropriate binding sites. In general, active or potentially active genes have a loose nucleosome structure, and inactive DNA has a tight nucleosome structure. (Figure 1-6) Also, research in past years reveals several histone variants (H2A.Z, H3.3) and many types of histone modifications consisting of phosphorylation, methylation, acetylation and ubiquitination (Berger 2002). These all may contribute to the establishment and maintenance of “open” or “closed” chromatin structure.

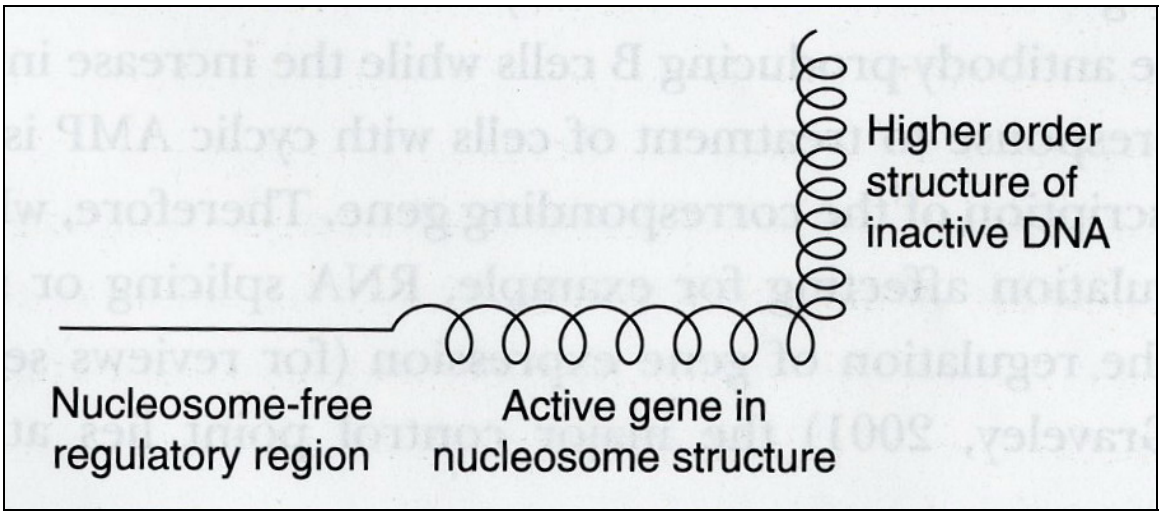


Figure 1-6 Levels of chromatin structure in active or inactive DNA. (Adapted from Eukaryotic transcription Factors, Fourth Edition, Figure 1.1) Chromatin structure is compacted tightly in genes which are not transcriptionally active or about to become active. In contrast, active or potentially active genes exist in the simple nucleosomal structure. Moreover, in the regulatory regions of these genes nucleosomes are either removed altogether or undergo a structural alteration which facilitates the binding of specific transcription factors to their binding sites in these regions.

Histone modifications are tightly linked to transcription regulation. For example, histone acetylation at lysine residues is generally correlated to transcription activation. Histone arginine methylation is involved in gene activation. Histone H3 lysine 9 (K9) tri-methylation is well known to be a transcription repression or silencing marker, while histone H3 K4 tri-methylation marks active euchromatin. Phosphorylation of Ser-10 of histone H3 functions both in transcriptional activation and in chromosome condensation during mitosis (Berger 2002; Kouzarides 2002; Peterson and Laniel 2004). The many different modifications of histones are referred to as a histone code.

Several methods are commonly used to study chromatin structure and the role of the histone code in transcription. The DNase I hypersensitive site assay (DHS) is often used to detect open regions of chromatin. DNA packaged into condensed chromatin is unavailable to the transcription initiation complex and also resistant to digestion by DNase I. When a gene is activated and the chromatin is remodeled into a transcriptionally permissive configuration, the DNA also can show an increased sensitivity to digestion with DNase I.

The *in vivo* permanganate footprinting assay can provide a detailed look at transcriptionally active regions at single base pair resolution. When the transcriptional machinery is recruited to DNA, it melts the DNA helix to form a transcription bubble. Permanganate oxidizes the C5-C6 double bond in thymines and renders the DNA backbone sensitive to cleavage by piperidine. This cleavage can later be detected by Ligation-mediated PCR. Recently, the method has been successfully used in *Drosophila* embryos to study developmentally regulated genes (Wang *et al.* 2007; Zeitlinger *et al.* 2007).

Chromosome conformation capture (3C) is a relative new technique to investigate physical and spatial interactions between regulatory regions. Cells are fixed with formaldehyde which forms DNA-protein and protein-protein cross-links between regions of the genome in close proximity. Subsequent restriction enzyme digestion and intra-molecular ligation produces novel junctions between restriction fragment ends that are in proximity in the nucleus. These novel ligation products can be detected by PCR (Dekker et al. 2002). This method makes it possible to virtually draw the interactions between promoters and enhancers and observe potential changes of chromatin structure that correlate with changes of gene expression during development (Lanzuolo et al. 2007).

Chromatin immuno-precipitation (ChIP) has been widely used in studying protein-DNA interactions and mapping histone modifications or histone variants. Detected by quantitative PCR, this method can detect as small as a 20% change in the level of association. Moreover, when combined with microarrays, this

method can map protein binding or histone code changes across the entire genome (known as ChIP-chip). Mark Biggin at Lawrence Berkeley National Laboratory has mapped the binding of many transcriptional factors in *Drosophila* embryo, which provided us extremely rich and useful information to study transcription regulation during development (Li et al. 2008). Analysis of multiple co-existing histone modifications on a single locus can also be achieved by ChIP-sequencing, at a very high resolution (Albert et al. 2007).

Chapter 2 Transcription elongation controls cell fate specification in the *Drosophila* embryo

This chapter has been adapted from a manuscript with a same title published in *Genes and Development* 21:1031-1036 (2007). The majority of this work is done by me and Dr. Gergen, with the help of Chanhyo Lee and David S. Gilmour on permanganate footprinting. (Figure 2-4)

Abstract

The simple combinatorial rules for regulation of the *sloppy-paired-1* (*slp1*) gene by the pair-rule transcription factors during early *Drosophila* embryogenesis offer a unique opportunity to investigate the molecular mechanisms of developmentally regulated transcription repression. We find that the initial repression of *slp1* in response to Runt and Fushi-tarazu (Ftz) does not involve chromatin remodeling, or histone modification. Chromatin immunoprecipitation and *in vivo* footprinting experiments indicate RNA polymerase II (Pol II) initiates transcription in *slp1*-repressed cells and pauses downstream from the promoter in a complex that includes the negative elongation factor NELF. The finding that NELF also associates with the promoter regions of *wingless* (*wg*) and *engrailed* (*en*), two other pivotal targets of the pair-rule transcription factors, strongly suggests that developmentally regulated transcriptional elongation is central to the process of cell fate specification during this critical stage of embryonic development.

Introduction

The *Drosophila* embryo provides a powerful model for investigating the regulatory mechanisms of cell fate specification during development. Gradients of maternally provided positional information are translated into differential programs of gene expression at a single-cell level of resolution within the 6000 cells that comprise the cellular blastoderm stage embryo at 3 h of development (Ingham 1988). On the antero–posterior axis, this initial specification is revealed by the segmentally repeated expression of several segment-polarity genes (Akam 1987). These metameric expression patterns are generated in response to positional cues provided by the DNA-binding transcription factors encoded by the pair-rule segmentation genes (Howard 1990). One key component of segment-polarity gene regulation is context-dependent activation and repression by the pair-rule transcription factor Runt, the founding member of the Runx family of developmental regulators (Coffman 2003; Levanon and Groner 2004). Recent studies reveal that Runt activity is modulated by the Ftz homeodomain protein and further identify the *slp1* gene as an attractive model for investigating the molecular mechanisms of Runt-dependent transcriptional regulation (Swantek and Gergen 2004).

The segmentally repeated expression of *slp1* initiates during the late cellular blastoderm stage and becomes fully apparent by the onset of gastrulation (~3.5 h of development). The combinatorial rules that generate this pattern differ in odd- and even-numbered parasegments. Expression in the posterior half of odd parasegments requires the pair-rule transcription factors Runt and Odd-paired (Opa). In contrast, in the adjacent cells that comprise the anterior half of the even parasegments, the presence of Ftz converts Runt from an activator to a repressor of *slp1* expression (Swantek and Gergen 2004). Normally, these combinatorial rules only apply within the presegmental region of the embryo. However, ectopic coexpression of Runt and Opa in embryos that are mutant for Ftz leads to *slp1* activation in all somatic blastoderm cells, whereas ectopic coexpression of both Runt and Ftz gives *slp1* repression throughout the embryo (Figure 2-1A). Thus, all late blastoderm stage nuclei can adopt *slp1* expression states characteristic of posterior odd-numbered or anterior even-numbered parasegments in response to these pair-rule transcription factors. The uniform repression of *slp1* in response to Runt and Ftz is of particular interest, as it is straightforward to generate staged populations of these embryos that can be used to investigate the *in vivo* biochemistry of this repression. Indeed, in matched embryo collections processed 3–4 h AED (after egg deposition), quantitative RT–PCR reveals 40-fold less *slp1* mRNA in embryos that are ectopically coexpressing Runt and Ftz compared with

wild-type embryos. Our goal in this study was to identify the molecular changes at the *slp1* locus that account for this 40-fold difference in expression.

Results and Discussion

We began by using DNase I hypersensitivity to probe the chromatin structure of the *slp1* locus. These assays revealed the presence of a DNase I-hypersensitive site near to the 5'-end of the *slp1* transcription unit (Figure 2-2A). Chromatin immunoprecipitation (ChIP) experiments with antiserum against histone H3 provide an explanation for this DNase I hypersensitivity. There is significantly reduced association of H3 with the *slp1* promoter region compared with both the structural gene as well as sequences upstream of the promoter (Figure 2-2B,C). These observations strongly suggest that the promoter region is nucleosome free. Importantly, matched collections of wild-type and Runt + Ftz (R+F) embryos show both the same pattern of DNase I hypersensitivity (Figure 2-2A) and histone H3 association (Figure 2-2C). These results indicate that the 40-fold decrease in mRNA expression in *slp1*-repressed embryos is not due to gross changes in the accessibility of the *slp1* promoter region.

Histone acetylation and deacetylation are important for transcriptional regulation with a general correlation between histone acetylation and active transcription (Roth et al. 2001; Schubeler et al. 2004). Indeed, prior work from our laboratory demonstrated that the *Rpd3* histone deacetylase is important for maintaining the Runt-dependent repression of the segment-polarity gene *en* (Wheeler et al. 2002). ChIP experiments reveal no significant difference in the H3 acetylation pattern of *slp1* chromatin from wild-type versus R+F embryos (Figure 2-2D). Although we detect no differences in H3 or Ac-H3 association that correlate with *slp1* repression, there are interesting differences in the H3 acetylation levels at different genomic locations. The *slp1* structural gene (primer pairs 6 and 7) shows stronger Ac-H3 association than the upstream region (primer pair 2) (Figure 2-2D). This difference is not observed for the association of H3 with these same intervals (Figure 2-2C), suggesting that H3 acetylation marks genomic regions that are permissive for transcription. The relative levels of H3 and Ac-H3 association with *Brother* (*Bro*), a gene that is not transcribed in the early embryo (as measured by RT–PCR) (data not shown), provide additional evidence for this trend. Although H3 association with the *Bro* gene is greater than for any region of the *slp1* locus, the level of Ac-H3 association with *Bro* is lower than for any region of *slp1* (Figure 2-2C,D). Based on the observation that we can detect differences in H3 and Ac-H3 association that correlate with transcriptional potential and yet detect no differences between wild-type and *slp1*-repressed embryos, we conclude that H3 acetylation plays a negligible role in the establishment of *slp1* repression.

The above observations led us to characterize the interactions of the transcriptional machinery with *slp1*. Association of the TATA-box-binding protein (TBP) is a first step in assembly of the transcriptional machinery on a promoter (Green 2000). As expected, TBP association is detected with a promoter-proximal interval centered 6 base pairs (bp) upstream of the *slp1* transcript initiation site in chromatin from wild-type embryos (Figure 2-3A). A weaker signal is detected for an interval within the 5' untranslated region (UTR), centered 124 bp downstream from the start site, whereas all other intervals give background level signals. We find very similar levels of TBP association in chromatin from R+F embryos (Figure 2-3A). More surprising is the finding that there is almost no difference in the level of Pol II association with the *slp1* promoter-proximal interval in chromatin from wild-type and R+F embryos (Figure 2-3B). Pol II is also associated with the *slp1* structural gene in wild-type embryos, but at lower levels than at the promoter (Figure 2-3B). In contrast, Pol II association with the *slp1* structural gene is markedly reduced in R+F embryos and near to background levels for regions downstream from the 5'-UTR (Figure 2-3B). Based on these results, we conclude that promoter recruitment of Pol II is not blocked in *slp1*-repressed embryos. We further characterized *slp1*-associated Pol II using an antibody that recognizes the Phospho-Ser-5 form of the heptad repeats that comprise the C-terminal domain (CTD) of the largest Pol II subunit. Phospho-Ser-5 modification of the CTD is associated with transcription initiation (Komarnitsky et al. 2000; Sims et al. 2004; Morris et al. 2005). This antiserum also gives the strongest signals with the *slp1* promoter-proximal interval in wild-type chromatin, and this signal is not reduced in chromatin from R+F embryos (Figure 2-3C). This result indicates that *slp1* repression occurs at a step downstream from transcription initiation.

The *Drosophila hsp70a* promoter is an extensively studied example of regulated transcriptional elongation (Rougvie and Lis 1988; Boehm et al. 2003; Lebedeva et al. 2005). Pol II initiates transcription at the *hsp70a* promoter, and then, in the absence of a heat shock, pauses immediately downstream from the promoter. All somatic cells in 3–4-h-AED embryos are capable of activating the *hsp70a* gene, and as expected, Phospho-Ser-5-modified Pol II is readily detected on the *hsp70a* promoter in our chromatin preparations from non-heat-shocked embryos (Figure 2-3C). The paused Pol II complex on the *hsp70a* promoter is also readily detected using permanganate footprinting due to the increased sensitivity of thymine residues in single-stranded regions (Weber et al. 1997). We used this same technique to carry out footprinting studies on the *slp1* promoter region. The results reveal strong hyperreactivity of thymine residues at +15, +28, +30, +38, and +50 downstream from the transcription start site in blastoderm stage embryos (Figure 2-4). This interval is similar, though perhaps somewhat larger than the interval detected for *hsp70a*, within which the most prominent increases in

reactivity are at residues +22 and +30 (Weber et al. 1997). The pattern of reactivity on *slp1* is extremely similar in both wild-type and *slp1*-repressed embryos, indicating that the hyperreactivity is not due to active transcription of the *slp1* gene. Importantly, this pattern is not observed in nuclei from *Drosophila* tissue culture cells (Figure 2-4). Thus, unlike *hsp70a*, the footprint on the *slp1* 5'-UTR is developmentally regulated.

The negative elongation factor NELF is thought to play a key role in establishing the paused Pol II complex on the *hsp70a* promoter (Wu et al. 2003). Indeed, NELF association provides a marker for the paused complex as heat-shock-induced transcriptional elongation involves release of NELF (Wu et al. 2003; Wu et al. 2005). In agreement with the results of our footprinting studies, ChIP experiments reveal the NELF-D and NELF-E subunits are associated with the *slp1* promoter region in chromatin from wild-type embryos (Figure 2-5A), but not in chromatin from *Drosophila* tissue culture cells (data not shown). Strong signals are obtained in chromatin from embryos with both the promoter-proximal and 5'-UTR intervals, whereas background level signals are obtained with other intervals of the *slp1* locus. It is notable that the promoter-proximal signal is less than or equal to the signal detected for the 5'-UTR interval. This pattern of association contrasts that obtained with TBP, which shows a threefold stronger signal with the promoter-proximal primer pair (Figure 2-5B,D). Our interpretation of these association patterns is that NELF is bound downstream from the *slp1* transcription start site, presumably as a component of the paused Pol II complex. Consistent with this interpretation, we find a similar differential pattern of TBP and NELF association with promoter-proximal and 5'-UTR intervals of *hsp70a* (Figure 2-5C,D). These results strongly suggest that NELF plays a key role in regulating *slp1* elongation in the blastoderm-stage *Drosophila* embryo.

The initial indications that *slp1* expression was regulated at a step downstream from transcription initiation came from ChIP experiments on chromatin from a homogeneous population of embryos that uniformly repress *slp1.* Localized association of NELF in a region downstream from the transcription start site is a hallmark of promoter-proximal pausing. Importantly, this association provides a method for detecting paused Pol II complexes in chromatin from embryos that contain a mixture of cells, some of which are expressing full-length mRNA transcripts. We used ChIP assays to determine whether NELF associates with the promoter regions of *wg* and *en*, two pivotal segment-polarity gene targets of the pair-rule transcription factors. The results reveal specific association of NELF with the promoter-proximal and 5'-UTR regions of both genes in 3–4-h-AED embryos (Figure 2-5E,F). Furthermore, the differential association pattern of TBP and NELF with these two intervals indicates that NELF is localized to a region

immediately downstream from the initiation sites for both genes (Figure 2-5D). These findings indicate that regulation of transcriptional elongation is likely to be central in generating the initial patterns of segment-polarity gene expression in the *Drosophila* embryo.

Regulation of transcriptional elongation has been described for several genes in addition to the *Drosophila* heat-shock genes, including human *c-myc*, *c-myb*, *c-fos*, *junB*, and *p21* (Bender et al. 1987; Pinaud and Mirkovitch 1998; Chen and Sytkowski 2001; Aida et al. 2006; Gomes et al. 2006). A feature shared by these previously characterized examples is rapid induction of gene expression in response to external stimuli. The initial establishment of segment-polarity gene-expression patterns in response to the pair-rule transcription factors occurs within a relatively brief developmental window of ~30 min, spanning the completion of cellularization and the beginning of germ band extension. The temporal advantages offered by regulating these genes at a transcriptional elongation step as compared with chromatin remodeling and/or Pol II initiation complex assembly may be essential for the timely establishment of differing gene expression programs during cell fate specification in the *Drosophila* blastoderm embryo. The observations that Pol II molecules are enriched at the 5'-ends of a number of genes (Gariglio et al. 1981; Rougvie and Lis 1988), coupled with findings that defects in transcriptional elongation factors produce specific developmental defects (Guo et al. 2000; Jennings et al. 2004), strongly suggest that regulation of transcriptional elongation is a hitherto overlooked, but potentially widespread strategy for controlling gene expression during development.

Materials and methods

Drosophila strains and transgenes

The *Drosophila* strain *y w[67c23]* that was used to develop different transgenic lines was used as the wild-type control for all in situ hybridization and biochemical experiments. Uniform ectopic expression of different pair-rule genes was obtained using the *NGT* (*nanos-GAL4-tubulin*) maternal GAL4 expression system (Tracey et al. 2000). The genetic cross to express Runt and Opa in embryos mutant for *ftz* was between females homozygous for the *NGT[40]* GAL4-driver and heterozygous for the *ftz[11]* mutation and males homozygous for both the *UAS-runt[15]*and *UAS-opa[14]* transgenes and heterozygous for *ftz[11]*. Ectopic coexpression of Runt and Ftz was obtained by crossing females homozygous for both the *NGT[40]* and *NGT[A]* GAL4 drivers with males homozygous for both the *UAS-runt[15]* and *UAS-ftz[263]* transgenes. These different *NGT*, *UAS-runt*, *UAS-opa*, and *UAS-ftz* transgenes have all been described previously (Swantek and Gergen 2004).

In situ hybridization and RT–PCR

In situ hybridization was carried out as described previously (Swantek and Gergen 2004) using the digoxigenin-labeled riboprobe for *slp1* described in Wheeler et al. (2002). RNA used for PCR after reverse-transcription was isolated from homogenates of ~200 embryos from appropriately staged collections. RNA was extracted using the High Pure RNA isolation Kit (Roche). cDNA was synthesized using the 1st Strand cDNA synthesis Kit (Roche) programmed with 1 ug of RNA. Quantitative PCR was conducted using a Light Cycler (Roche) with a primer pair centered 504 bp downstream from the *slp1* promoter (primer pair 6 in Figure 2-2B). The *slp1* RT–PCR signal obtained with different RNA samples was normalized using the RT–PCR signal for *rp49*.

DNase I-hypersensitive site assay

Nuclei isolated from ~250 mg of staged *Drosophila* embryos (Cartwright et al. 1999) were resuspended in DNase I digestion buffer, aliquoted into microfuge tubes, and incubated with 0, 2, 4, or 8 U of DNase I for 3 min on ice. Digestion was stopped by addition of EDTA to 15 mM and SDS to 0.5%. DNA was purified by phenol/chloroform extraction and ethanol precipitation, and resuspended in water. Approximately 5 ug of DNA samples digested with BstXI were run on 1% agarose gels and transferred to nylon membranes for Southern blot hybridization. The *slp1* riboprobe used for these experiments is similar to the one used for in situ

hybridization, although it was not subjected to partial hydrolysis prior to use for hybridization. DNA size standards were detected on Southern blots using a digoxigenin-labeled DNA probe.

ChIP

Staged embryo collections were dechorionated, fixed with 2% formaldehyde for 15 min at room temperature, washed, and snap-frozen at –80°C for storage (Orlando et al. 1998). Approximately 100 mg of embryos were homogenized for 1 min in 10 mM EDTA and 50 mM Tris (pH 8.1). After addition of SDS to a final concentration of 1% and incubation on ice for 10 min, glass beads (150–200 µm) were added and the homogenates were sonicated to give sheared chromatin preparations with an average DNA size of 300–400 bp. Chromatin preparations from the equivalent of ~ 25 mg of embryos were used for each immunoprecipitation using the Chromatin Immunoprecipitation Assay Kit (Upstate Biotechnology) with the following antisera: anti-rabbit IgG (Sigma), anti-mouse IgG (Sigma), anti-mouse IgM (Sigma), anti-Histone H3 (Abcam), anti-acetyl-Histone H3 (Upstate Biotechnology), anti-TBP (Lebedeva et al. 2005), anti-RNA pol II 8WG16 (Covance), anti-RNA pol II H14 (Covance), and anti-NELF-D and anti-NELF-E (Wu et al. 2003). The relative amount of immunoprecipitated DNA was quantified using real-time PCR (Roche Light-Cycler) using primer pairs with a T_m in the range of 59°C–66°C that generate products between 150 and 211 bp. Primer sequences are available on request. The percent precipitation values that are reported were calculated using a dilution series of input chromatin to determine the relative efficiency for each primer pair. PCR was done multiple times on samples from at least two independent immunoprecipitation assays using at least two different chromatin preparations.

Permanganate footprinting

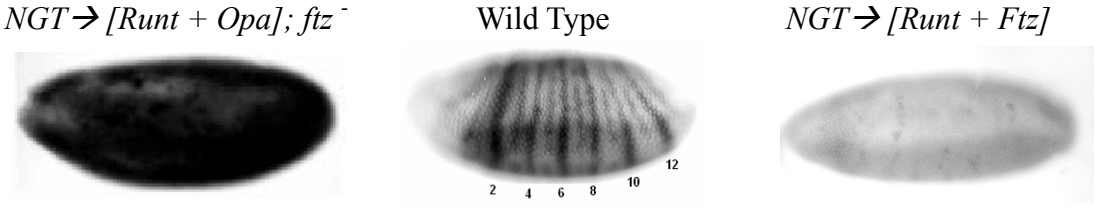
Permanganate footprinting on *Drosophila* S2 cells was performed as described previously using 10 mM $KMnO_4$ for a 1-min incubation (Weber et al. 1997). For embryos, collections of 3–4-h-AED embryos (10 µL, corresponding to ~10^3 embryos or 10^7 genomes) were dechorionated, transferred into 150 µL of ice-cold dissection buffer (130 mM NaCl, 5 mM NaCl, 5 mM KCl, 1.5 mM CaCl2), and briefly homogenized prior to addition of 50 µL of 40 mM $KMnO_4$. Reactions were stopped after 1 min by addition of 200 µL of stop solution (20 mM Tris at pH 7.5, 20 mM NaCl, 40 mM EDTA, 1% SDS, 0.8 M β-mercaptoethanol) to the lysate. Nested primers for detecting cleavage within the *slp1* promoter and 5'-UTR interval by Ligation-Mediated PCR span the interval from 161 to 199 nucleotides (nt) downstream from the transcription start site. The sequences and annealing temperature for PCR are as follows: *slp1* primer-1,

5'-GTTTTGATGGGTTGAGTTG-3' (51°C); *slp1* primer-2 5'-GGGT TGAGTTGCGGTGT-3' (55°C); *slp1* primer-3, 5'-GCGGTGTTGAT GGGTTTCTT-3' (58°C).

Acknowledgments

We thank Steven Naymagon and Jocelin Abraham for technical assistance, and Achim Breiling, Ping Hu, Nouria Hernandez, Ann Sutton, and Rolf Sternglanz for technical advice in developing ChIP protocols for early *Drosophila* embryos. We thank James Kadonaga for providing antisera to TBP. Comments on the manuscript were provided by Carl Anderson, Michael Frohman, Maurice Kernan, and Rolf Sternglanz. This work was supported by research grants from the NSF (MCB 0344486, to J.P.G) and the NIH (GM 47477 to D.S.G).

A



B

Genotype / Signal	WT 0~2h	WT 3~4h	*NGT→ [Runt+Ftz]* 3~4h	WT 4~5h	*NGT→ [Runt+Ftz]* 4~5h
RT (-) control	0	0	0	0	0
slp1 mRNA amount	3%	100%	2.5%	103%	5%

Figure 2-1 Manipulation of *slp1* expression by pair-rule transcription factors. **(*A*)** Expression of *slp1* mRNA in gastrula-stage embryos as revealed by in situ hybridization. (*Middle* panel) The wild-type pattern shows 14 distinct stripes in the presegmental region with stronger expression in even-numbered parasegments. The embryos on the *left* and *right* demonstrate the response of *slp1* to ectopic expression of different pair-rule transcription factors using the *NGT* (*nanos-GAL4-tubulin*) expression system (Tracey et al. 2000). (*Left* panel) *NGT*-driven expression of Runt and Opa in embryos that are mutant for *ftz* leads to activation throughout the presegmental region as well as in the anterior head region. (*Right* panel) In contrast, *NGT*-driven coexpression of Runt and Ftz leads to uniform *slp1* repression. **(*B*)** Quantitative RT–PCR was used to determine the differences in *slp1* expression in staged collections of wild-type (WT) and Runt + Ftz (R+F) coexpressing embryos.

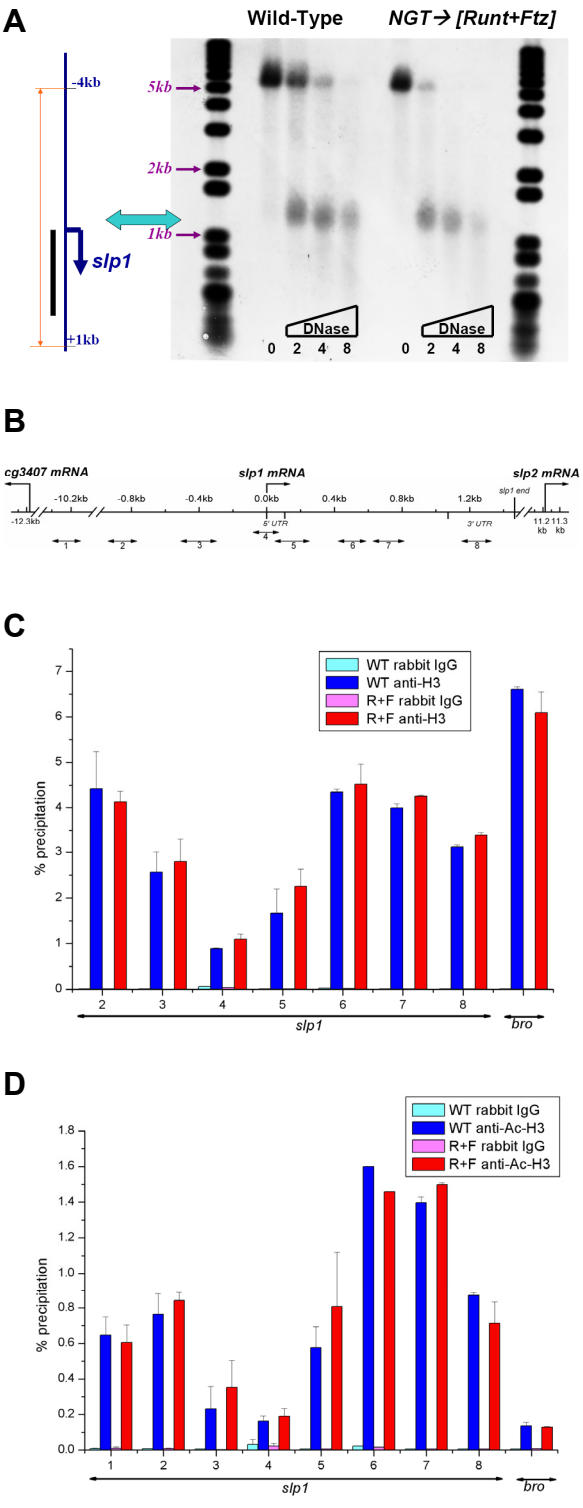


Figure 2-2 *slp1* repression is established without alterations in chromatin structure. **(*A*)** DNase I digestion reveals the presence of a hypersensitive site near the 5'-end of the *slp1* transcription unit. The diagram on the *left* indicates the relationship of an ~5-kb interval of the *slp1* locus to the DNA fragments detected in the Southern blot shown on the *right*. The thin line on the *left* indicates the extent of a 5-kb fragment produced by BstXI digestion, and the adjacent thick vertical bar indicates the antisense riboprobe used to detect *slp1*-containing sequences. The horizontal line and downward arrow indicate the initiation site and direction of transcription, respectively. DNA samples isolated from wild-type and R+F embryos (3–4 h AED) are organized as indicated across the *top*, with the amounts of DNase I used on different samples indicated across the *bottom*. **(*B*)** The location of PCR products generated with different primer pairs used for ChIP studies is shown relative to a map of the *slp1* locus. The PCR products are numbered *left* to *right*, from 1 to 8. **(*C*)** Results of quantitative PCR on ChIP samples generated using an antibody against histone H3 with chromatin isolated from wild-type (blue bars) and R+F (red bars) embryos. The nonspecific background was determined using a control rabbit IgG antiserum with these same chromatin preparations (cyan and magenta bars, respectively). Results obtained for different intervals of the *slp1* locus are as labeled across the *bottom*. ChIP results using a primer pair that amplifies a segment of the *Bro* structural gene are given on the *right*. Error bars indicate the standard error in the percent precipitation values for each interval. **(*D*)** Results of quantitative PCR with ChIP samples generated using an antibody that detects acetylated H3, labeled as in *C*.

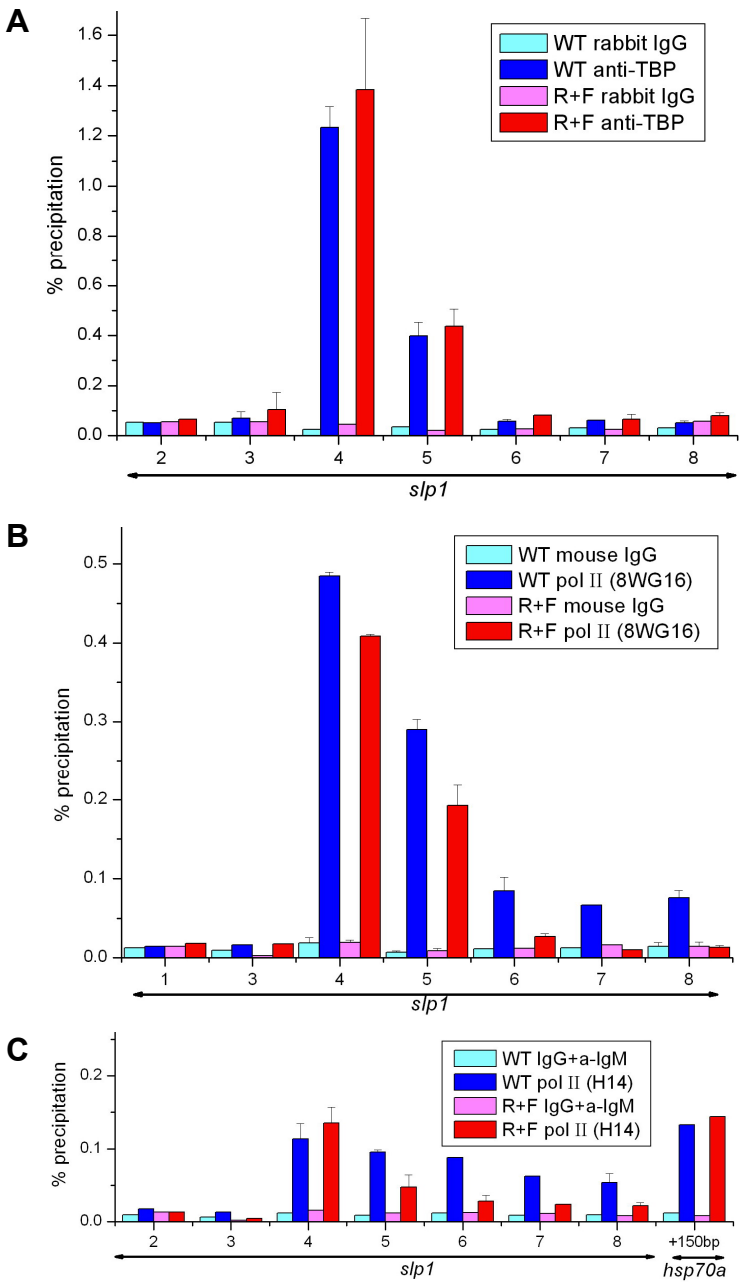


Figure 2-3 **Initiation of transcription in *slp1*-repressed embryos**. ChIP assays on chromatin from wild-type and R+F embryos using anti-TBP antiserum **(*A*)**, the monoclonal antibody 8WG16 **(*B*)**, which recognizes the heptad repeats of the CTD, irrespective of their phosphorylation status (Patturajan et al. 1999; Jones et al. 2004; Morris et al. 2005), and the monoclonal antibody H14 **(*C*)**, which specifically recognizes phosphorylation of serine residue 5 within the heptad repeats that comprise the CTD (Patturajan et al. 1999; Jones et al. 2004; Morris et al. 2005). Background controls for each antibody were determined using nonimmune serum from the appropriate species as indicated. The labeling scheme is the same as used in Figure. 2.

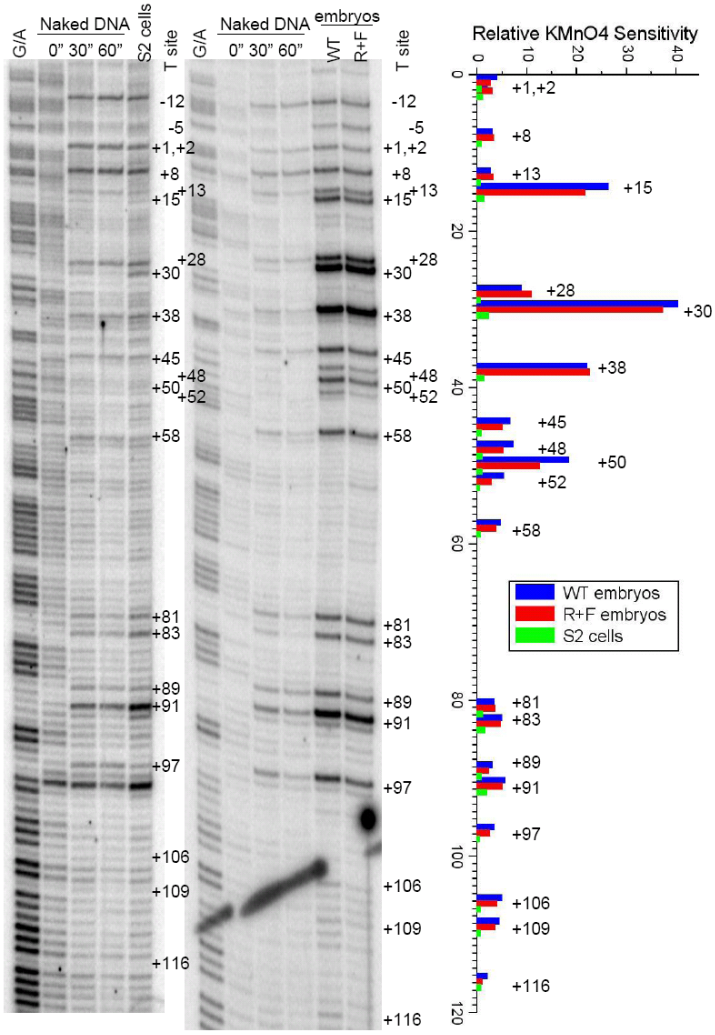
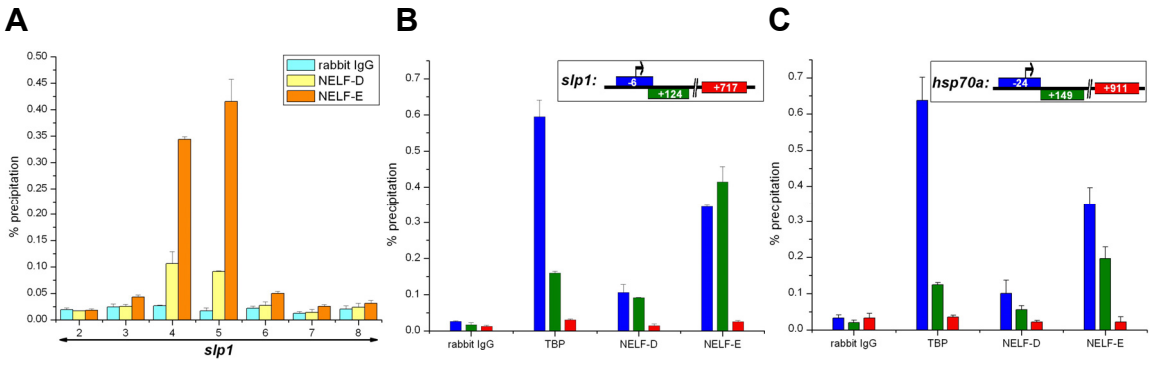


Figure 2-4 Developmental regulation of *slp1* promoter-proximal pausing. The two *left* panels are autoradiographs revealing the permanganate sensitivity of thymine residues in an interval extending from 12 nt upstream of to 116 nt downstream from the *slp1* transcription start site. Each of these panels includes a G/A ladder and three lanes showing results obtained by treating naked DNA with 10 mM $KMnO_4$ for 0, 30, and 60 sec. In addition, the autoradiograph on the *left* shows *in vivo* results obtained from treatment of *Drosophila* S2 cells with 10 mM $KMnO_4$ for 60 sec, whereas the autoradiograph on the *right* shows results obtained with similarly treated wild-type and R+F 3–4-h-AED *Drosophila* embryos. The graph on the far *right* presents a quantification of the raw PhosphorImager data for the S2, wild-type, and R+F samples using the Semi-Automated Footprinting Analysis software package SAFA(Das et al. 2005). The relative sensitivity of each thymine residue is the ratio of the signal for each experimental sample relative to the signal obtained from the 30-sec naked DNA control. A 10-fold or greater increase in $KMnO_4$ sensitivity is observed for thymine residues at +15, +28, +30, +38, and +50 in both embryo samples, whereas the sensitivity of these residues in S2 cells is nearly indistinguishable from that observed on naked DNA.



D

Factor: *gene*: ratio		TBP	NELF-D	NELF-E
slp1	-5/ +133bp	3.7±0.4	1.2±0.3	0.8±0.1
hsp	-24/+150bp	5.1±0.7	1.8±0.9	1.8±0.5
wg	-26/ +54bp	1.5±0.2	1.1±0.3	0.9±0.3
en	-47/ +69bp	1.4±0.1	1.0±0.3	0.8±0.3

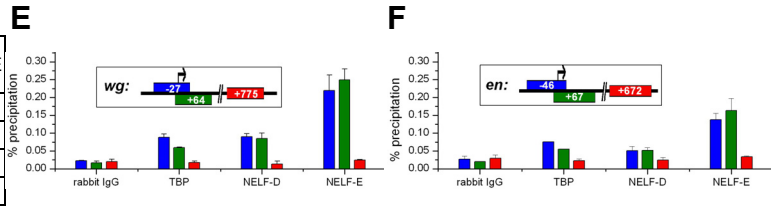


Figure 2-5 NELF associates with promoter-proximal 5'-UTR regions of segment-polarity genes in the blastoderm embryo. **(*A*)** ChIP assays on chromatin from wild-type 3–4-h-AED embryos using antisera against the NELF-D (yellow bars) and NELF-E (orange bars) subunits of NELF. The primer pairs used to detect NELF association with different intervals of the *slp1* locus are as described for Figure 2. Background signals obtained using control IgG are shown by cyan bars. **(*B*)** Comparison of the pattern of TBP, NELF-D, and NELF-E association with intervals of *slp1* detected by primer pairs 4, 5, and 7, centered 9 bp upstream of (blue bars), and 121 bp (green bars) and 717 bp (red bars) downstream from the transcription start site, respectively. **(*C*)** Pattern of TBP, NELF-D, and NELF-E association with *hsp70a* using a similar representation as in *B*. In this case, the ChIP assays were for intervals centered 24 bp upstream of (blue bars) and 149 bp (green bars) and 911 bp (red bars) downstream from the transcription start site. The preferential association of TBP that is detected with the promoter-proximal versus 5'-UTR *hsp70a* intervals is greater than that observed for *slp1*, presumably due to differences in the exact location of the primer pairs used for these two genes. The slightly stronger signal obtained with both NELF subunits and the *hsp70a* upstream promoter interval indicates NELF association maps closer to –24 bp than to +149 bp, relative to the transcription start site. This is consistent with *in vivo* data placing the paused Pol II complex 21–35 bp downstream from the *hsp70a* transcription start site (Rasmussen and Lis 1995). **(*D*)** A table providing the ratios of ChIP signals obtained with promoter-proximal, and 5'-UTR intervals of four different genes (*slp1*, *hsp70a*, *wg*, and *en*) with TBP, NELF-D, and NELF-E antisera. Raw data showing the association of TBP, NELF-D, and NELF-E with different regions of *wg* and *en* are shown in *E* and *F*, respectively. The intervals detected from the *wg* locus centered 27 bp upstream of (blue bars), and 64 bp (green bars) and 775 bp (red bars) downstream from the transcription start site. Similarly for *en*, the PCR-amplified intervals centered 46 bp upstream of (blue bars) and 67 bp (green bars) and 672 bp (red bars) downstream from the transcription start site.

Chapter 3 NELF function in *Drosophila* early embryonic development

Abstract

The negative elongation factor NELF is implicated to play a role in establishing a paused Pol II complex on the *hsp70*a promoter (Wu et al. 2003). The fact that NELF-D and NELF-E are present at promoter proximal regions of *sloppy paired 1* (*slp1*), *wingless* (*wg*) and *engrailed* (*en*) in 3-4 hour old *Drosophila* embryos suggests the NELF complex might also play an important role for cell specification during early embryonic development (Wang et al. 2007). In this chapter, I first determined by q-RT-PCR that all NELF subunits are maternally provided. Immuno-fluorescent detection showed that NELF-A and NELF-E are localized in nuclei at the cellularization stage. Chromatin Immuno-precipitation (ChIP) at different stages of early development suggested the NELF complex is pre-loaded on the promoter-proximal regions of many genes prior to their initial transcriptional activation. Maternally introducing a dominant negative form of NELF-E (NELF-EΔRRM) interferes with the activity of the pair-rule transcription factor Eve as a repressor in the embryo. Additionally, a 30 minute heat-shock leads to de-repression of certain pair-rule genes at the blastula stage in NELF-EΔRRM embryos but not in wild type. Moreover, by injecting NELF antisera, I was able to interfere with the transcription of certain genes and arrest embryonic development at cellularization. These results provide compelling evidence that NELF plays an important role in developmentally regulated transcription repression.

Introduction

The appropriate regulation of transcription is crucial to cell specification during development. Transcription starts with the recruitment of general and gene-specific factors, followed by assembly of the pre-initiation complex. After initiation, mRNA is synthesized and transcription is brought to an elongation step (Lis 1998). Although in many cases, the recruitment of RNA polymerase II (Pol II) to a promoter is the rate-limiting step in gene activation (Ptashne and Gann 1997), more and more studies have described genes with Pol II stalling at the promoter-proximal region (Saunders et al. 2006; Core and Lis 2008). We previously showed that in the *Drosophila* embryo, developmentally regulated transcription was controlled at the transcriptional elongation step (Wang et al. 2007). Later genome-wide Pol II ChIP-chip analysis on 2-4h old *Drosophila* embryos revealed that genes having paused Pol II are highly enriched for developmentally-controlled genes (Zeitlinger et al. 2007).

DSIF (DRB sensitivity-inducing factor) and NELF were initially discovered in the study of DRB inhibited transcription. They work together to inhibit transcription elongation *in vitro* (Wada et al. 1998; Yamaguchi et al. 1999). NELF and Pol II ChIP-chip assays identified 2111 genes having NELF at their 5' end in *Drosophila* S2 cells. Among these, 37% have Pol II stalled at promoter-proximal region. More interesting, 79% of genes having stalled Pol II also have NELF binding (Lee et al. 2008). The high correlation between NELF and stalled Pol II strongly suggests the participation of NELF in transcription pausing during elongation. In the early *Drosophila* embryo, the finding that NELF associates with the promoter region of segmentation genes *slp1*, *wg* and *en* indicates NELF is involved in developmentally regulated transcription elongation (Wang et al. 2007).

In this chapter, I describe the mRNA and protein distribution of NELF in early embryos. I found that NELF was maternally provided and recruited to the promoter-proximal region of several segmentation genes prior to their transcriptional activation. Genetic approaches showed that NELF is needed for effective transcriptional repression. Additional evidence suggests NELF functions in many biological processes through transcription regulation.

Results

Expression of NELF in early embryogenesis

Table 3-1 lists some basic information on all 4 NELF subunits provided by Flybase. All 4 genes' transcripts are detectable in embryos. To start investigating NELF function in embryonic development, I first did quantitative reverse transcribed PCR (q-RT-PCR) to examine the mRNA expression level in embryos. The location of each primer pair used for q-RT-PCR is marked in Figure 3-1A. Wild type embryos were collected for 2 hours and aged for 0, 2, 4, 6, 8, 10 hours. As shown in Figure 3-1B, all four transcripts are provided maternally. The continuous decline of NELF-A, B and E during embryo development indicates the contribution of zygotically provided transcripts is limited.

To examine mRNA expression in embryos, I made NELF-A and NELF-E RNA probes. Whole embryo *in situ* hybridization showed that they both have ubiquitous staining from stage 1 to stage 11 (data not shown). At BDGP (Berkeley *Drosophila* Genome Project) Gene Expression website, NELF-B and NELF-D also have ubiquitous staining from stage 1 to stage 12 (data not shown).

Next I looked at the NELF-A and NELF-E protein distribution in embryos by antibody staining. NELF-A has a ubiquitous distribution throughout early embryogenesis. NELF-E accumulation starts from the middle part of the embryos and quickly distributes to all cells after cellularization. Both proteins co-localize with nuclei at blastula stage, as shown in Figure 3-2.

Temporal recruitment of NELF

Embryos were collected every 45 minutes from 2 to 5 hours AED (after egg deposition). ChIP experiments using rabbit-IgG, NELF-A and NELF-E antisera were performed in each embryo collection. PCR reactions were done for nine promoter proximal regions. ChIP data for the 2h-2h45mins collection is presented in Figure 3-3A. From these data, it is clear that NELF complex has already associated to all promoter proximal regions at the 2h to 2h45min developmental stage, prior to the extensive expression of most pair-rule or segment polarity genes. This observation agrees with the theory that the initial cell fate specifications are controlled by transcriptional elongation, and NELF is one of the key factors in this process.

To compare signals among different time windows, I assumed that the association of NELF at *heat shock 70a* (*hsp70a*) promoter region was not

changed during the early embryonic development. Therefore, each ChIP signal (calculated in percent precipitation relative to input) was divided by the corresponding ChIP signal at the *hsp70*a promoter region. The ratio was adjusted to 1 for all primer pairs at 2h-2h45mins window and plotted versus time in Figure 3-3B. The interesting finding was that along with the activation of genes, NELF gradually disassociated from most promoter proximal regions except *slp1*. I think that the differences in descending slopes might reflect the different "plasticity" of each gene during these early stages.

Dominant Negative NELF-E suppresses transcription repression

Human NELF-E contains three motifs: an N-terminal leucine zipper motif (LZ), a central domain rich in Arg-Asp dipeptide repeats (RD), and a C-terminal RNA recognition motif (RRM). The RRM has been shown to be required for NELF-dependent repression. A NELF-E deletion derivative without the RRM domain is able to co-IP with other NELF subunits but loses its ability to repress transcription *in vitro*. (Yamaguchi et al. 2002)

In *Drosophila*, both the leucine zipper motif and the RNA recognition motif have 30% identity to the corresponding motif in human, while the RD motif is absent (Wu et al. 2005). A straightforward assumption is that the RRM has similar RNA binding ability as well as a transcription repression function in *Drosophila*. Ectopic expression of NELF-E without RRM (NELF-EΔRRM) could compete with endogenous NELF-E to form a defective NELF complex that is not able to repress transcription (Figure 3-4A). In other words, the NELF-EΔRRM might function as dominant negative NELF-E that can be used to investigate NELF's role during development.

To test this idea and examine NELF function in *Drosophila* embryonic development, I generated *UAS-NELF-EΔRRM* transgenic flies. The NELF-EΔRRM construct has the predicted RRM sequence deleted but still keeps the 3' UTR of NELF-E. 36 individual lines were obtained from transformation. In order to examine the expression of the NELF-EΔRRM protein, a cross between *NGT* females with males from two independent *UAS-NELF-EΔRRM* lines were set up and 2-4hour old embryos were collected for western blot. For both crosses, I was able to detect a band at the right size for NELF-EΔRRM that is not present in wild type embryos (Figure 3-4B).

In Figure 3-4B, the truncated NELF-E is expressed at about 50% the amount of endogenous NELF-E. In order to get more expression, I used meiotic recombination to generate recombinant chromosomes carrying multiple *UAS-NELF-EΔRRM* transgenes. Inverse PCR was used to map the insertion sites

of the different transgenes and also to confirm the presence of the multiple transgenes on the final recombinant chromosomes (See Appendix Table 5-1 for detailed information of inverse PCR results). Two 6-copy-NELF-EΔRRM-insertion fly lines were generated with only 1 insertion difference, as shown in Table 3-2.

Having these two inducible NELF-E dominant negative lines and the *NGT* system, I was able to study NELF function during early embryogenesis. First I found that maternally provided NELF-EΔRRM (about the same amount as endogenous NELF-E protein in 0-3h old embryos, Figure 3-4C) does not affect embryonic viability. Detailed microarray data confirmed that there is no global effect on gene transcription at the gastrula stage even though the NELF-EΔRRM mRNA is about 1.3-fold greater than the endogenous NELF-E mRNA.

In order to directly examine whether NELF participates in the regulation of gene expression during segmentation, I investigated the effects of NELF-EΔRRM expression on the potency of different pair-rule transcription factors. These experiments take advantage of the ability to reproducibly and quantitatively manipulate ectopic expression using the *NGT* system (Tracey et al. 2000). *Drosophila* Even-skipped (Eve) is generally thought to be a transcriptional repressor (Fujioka et al. 1995; Fujioka et al. 2002). Ectopic expression of Eve in all cells can repress many genes from stage 5 and finally cause the death of the embryos. I tested the effect of maternally introducing NELF-EΔRRM into embryos that also express Eve over a range of levels that have increasingly stronger effects on embryonic viability. The results clearly show that NELF-EΔRRM suppress the lethality associated with *NGT*-driven Eve expression (Figure 3-5).

Both NELF-EΔRRM lines partially rescue embryos from Eve-mediated lethality. Moreover, the *in situ* hybridization results reveal that Eve-dependent *slp1* and *odd skipped* (*odd*) repression are suppressed in embryos containing NELF-EΔRRM. As shown in Figure 3-6A,B, after 5 hours development, *slp1* and *odd* expression in *NGT*/NELF-EΔRRM → *UAS-eve* cross are restored to nearly wild type level, while in *NGT*/II;III →*UAS-eve* cross, *slp1* and *odd* remain repressed.

To prove the restoration is not due to the reduction of Gal4-driven Eve, these same females (NGT/NELF-EΔRRM; NGT/II;III) were crossed to *UAS-lacZ* males. Beta-Galactosidase activity was measured on single embryos at 3.5AED (onset of gastrulation) and 5.5hAED (2 hours after onset of gastrulation). For both time points, there is even more LacZ in both NELF-EΔRRM lines compared with control (NGT/II;III X *UAS-laz*) (Table 3-3). These results indicate that ectopic Eve expression is not reduced in *NGT*/ NELF-EΔRRM→*UAS*-eve cross.

I also did the similar viability test on the background of overexpression of Runt and Ftz, individually. Runt and *ftz* both can function as transcriptional activators and repressors. (Nasiadka and Krause 1999; Coffman 2003) The lethality caused by low level ubiquitous Runt or Ftz is very likely due to the repression on certain genes instead of activation. Results show that NELF-EΔRRM can also partially rescue embryos from the lethal effects of ectopic Runt or Ftz expression (Figure 3-7). *In situ* hybridization results confirm that Runt-dependent *engrailed* (*en*) repression and Ftz-dependent *wingless* (*wg*) and *slp1* repression are suppressed by NELF-EΔRRM (data not shown).

The above experiments all were based on providing NELF-EΔRRM maternally. I also investigated the effects of expressing this construct zygotically. Zygotic expression at the highest level possible produced much weaker effects on gene expression comparing with expressing maternally (data not shown). Combining this observation with NELF temporal recruitment data, I conclude that pre-loading of NELF to the promoter proximal region is important for transcription repression.

In order to see the effects of zygoticly provided NELF-EΔRRM, a 30 minutes heat-shock was given to 2-3hour old embryos from a cross of homozygous *NGT* females with homozygous *UAS-NELF-EΔRRM6A* males. The idea behind the heat-shock was to stress the system, hoping that heat-shock would pause most transcription and also potentially disassociate the original NELF complex from different promoter proximal regions. While recovering, NELF-EΔRRM will be recruited to NELF complexes and interfere with the restoration of normal transcription. After 30 minutes recovery, embryos were fixed and *in situ* hybridization was performed with several probes. I found that the expression of *runt* and *ftz* was dramatically expanded, as shown in Figure 3-8. This strongly suggested that NELF is important for establishing the initial transcription repression. I also noticed that embryos having NELF-EΔRRM can mostly recover from heat shock but take longer time compared with wild-type embryos. This was consistent with the idea that NELF is needed to restore Pol II pausing at the heat shock promoter (Personal communication with Dr. David Gilmour).

Injection of antisera against NELF alters transcription and arrests cell cycle

To further investigate the NELF function in early embryonic development, Rabbit IgG, NELF-A, NELF-D or NELF-E antisera were injected into 20min to 1 hour old embryos from either the dorsal or ventral side. After 2 or 3 hours aging, embryos were fixed, followed by *in situ* hybridization with *ftz* and *slp1* probes. Rabbit IgG is not able to affect early embryogenesis as well as any gene's expression. For all three NELF antisera, they were able to arrest cell cycle at certain point. The various effects of different antibodies were mostly caused by

the various amount of effective antibody in each serum.

In detail, embryos having NELF-A antiserum were immediately arrested at cycle 13 and no expression pattern was observed. Embryos having NELF-D or NELF-E antiserum were arrested at blastula stage. As shown in Figure 3-9, *ftz* expression was expanded in the middle of embryos in NELF-E antiserum injected embryos. This confirms the negative function of NELF in *ftz* transcription. *Slp1* expression in the middle part of embryos disappeared in both NELF-D and NELF-E antisera injected embryos. This maybe due to the cell cycle arrest in the segmented region of embryo at the time *slp1* should be expressed.

Discussion

NELF is maternally provided and recruited to gene promoters prior to their transcription

All NELF subunits are maternally provided, as shown in q-RT-PCR result. The ubiquitous staining of NELF-E RNA and the gradually spreading pattern of NELF-E protein suggest that NELF-E translation happens after the fertilization of the egg. I noticed that the intermediate NELF-E staining covers the entire segmented region of the body but not both ends. Although I did not investigate, it is potentially very interesting to know why NELF-E protein is expressed in this way, and what factors are responsible for the initial inhibition of NELF-E translation at both ends (or super-activation in the central part). The co-localization of both NELF-A and NELF-E in nuclei at celluarization is consistent with the idea that NELF is involved in initial cell specification.

NELF ChIP data indicated that NELF is recruited to individual gene promoters prior to their transcription. In 5 out 6 genes tested, NELF gradually disassociates from promoter regions during development. This suggests NELF is important for initial cell specification in development. Also I noticed that different genes have different NELF dissociation rates. This phenomenon potentially indicates the “plasticity” of each developmental gene.

The mRNA expression of both *ftz* and *serendipity α* (*sryα*) are absent after 4 hours development (Weiner et al. 1984; Ibnsouda et al. 1993). The finding that NELF association at these two genes promoter proximal region keeps dropping during the 4h15min to 5hours time window instead of increasing suggests that NELF is not needed to maintain repression at later stages.

NELF-E is needed for effective transcriptional repression

The fact that dominant negative NELF-E can alleviate Eve-induced repression on *slp1* and *odd*; Runt-induced repression on *en* and Ftz-induced repression on *slp1* and *wg* further indicates a role for NELF in repression. I showed that these effects are not due to a decrease in GAL4-driven gene expression of Eve, Runt and Ftz as ectopic gene expression is not reduced in both NELF-EΔRRM lines. In fact, for both lines, the ectopic gene expression is actually increased. Because yeast does not have NELF, I do not think NELF-EΔRRM will function at the step of potentiating Gal4 driven gene expression. Therefore, the increase of gene expression might come from an increase of nanos-Gal4 transgene expression during oogenesis.

Heat-shock provides a disturbance into a developing system. Embryos expressing NELF-EΔRRM recovered slower than wild-type embryos and showed expanded expression of certain genes. Thus, NELF-E is involved in heat shock recovery as well as the initial transcriptional repression on developmentally regulated genes.

Microarrays were repeated twice with biological replicates to investigate the effect of maternally provided NELF-EΔRRM on global gene transcription at precisely staged gastrulating embryos (in a 10min window). Compared with embryos having no NELF-EΔRRM, only 25 transcripts were convincingly increased or decreased by equal or more than 1.2-fold (See Appendix Table 5-2), which include NELF-E RNA (2.3-fold) and transcripts from 3 genes that have *UAS-NELF*-EΔRRM P-transposon insertion. Interestingly, 18S ribosome RNA was increased by at least 1.9-fold. This finding suggests that NELF might also play negative role in rRNA synthesis and/or processing. There was no evidence before showing that NELF might be involved in Pol I transcription. But in yeast, DSIF was found to play roles in transcription elongation by RNA polymerase I and rRNA processing (Schneider et al. 2006).

Microarray data also shows that at the gastrulation stage, the approximate ratio of NELF-A: NELF-B: NELF-D: NELF-E mRNA is 1.0 : 0.62 : 0.81 : 0.86. Thus, none of the four NELF subunits appears to be limiting in wild-type embryos.

NELF function in transcriptional control

Injection of antisera against NELF-A, NELF-D or NELF-E can lead to cell cycle arrest prior to 4 hours development. This result suggests that NELF also plays roles in cell cycle regulation. Therefore, NELF is involved in many biological processes in the regulation of gene expression, such as, stress responses, cell specification, growth control (down-regulation of rRNA) and cell cycle regulation. This agrees with the finding that about 2000 genes have NELF associated at their 5’ end in *Drosophila* S2 cells (Lee et al. 2008). The abundant NELF association at the 5’ end of highly or not expressed genes and broad biological function of NELF indicate that NELF is more like a general factor in the transcriptional elongation control. The dominant negative experiment and antibody injection experiment confirm the negative role of NELF in transcription. I propose that differing from the transcription repression maintained by PcG complex or specific histone methylation (Simon and Tamkun 2002; Bhaumik et al. 2007; Schwartz and Pirrotta 2007), NELF involved repression is temporary and reversible. At the time a developmentally-controlled gene is gradually shut down, very likely repression is spreading, chromatin is tightened up and NELF dissociates, as suggested by NELF ChIP experiments in Figure 3-3. For genes highly transcribed, NELF may

serve as a regulator of expression level. It buffers the transcription to fit the physiological requirement. At extreme situation, such as, heat shock, NELF fully dissociates from heat shock gene locus to allow the maximal transcription by Pol II machinery.

In summary, my work first demonstrates the NELF function in *Drosophila* early embryonic development. Notably, NELF is important for transcription repression during cell fate specification. Also, NELF is potentially involved in many other biological processes. Thus, regulating transcription at an elongation step through NELF might be a common mechanism to suit dynamic demands in the regulation of gene expression during development.

Materials and methods

Drosophila strains and transgenes

The *Drosophila* strain y w(67c23) was used as wild-type for all *in situ* hybridization and biochemical experiments. All different *NGT*, *UAS-runt*, *UAS-ftz* lines were generated by previous lab members. *P*-transposition mediated by the Δ2-3 third chromosome (Robertson et al. 1988) was used to generate hops from the original *UAS-eve*[12] line obtained from John Reinitz. The relative stre*ngt*h of different *UAS-eve* lines were arranged based on the result of viability test with different *NGT* lines (See Appendix Table 5-3).

The *UAS-NELF-EΔRRM* lines contains random insertions of the P-transposon P{*UAS-NELF-EΔRRM*}. We generated this construct by cloning *NELF-EΔRRM* from its cDNA plasmid RE14181 and inserting it into p*UAST*. Transformants were generated by standard P-element-mediated germ line transformation using the p:Δ2-3 helper plasmid.

Antibody staining, antibody injection and western blot

Anti-NELF-A (rabbit serum), anti-NELF-D (rabbit serum) and anti-NELF-E(rabbit serum) were all kindly provided by Dr. David Gilmour. (Wu et al. 2003) Anti-rabbit IgG was from Sigma. For antibody staining, embryos were fixed with 10% para-formaldehyde for 20min and incubated with the primary antibody (1:100 dilution) at 4^{o}C overnight. Embryos were then washed and incubated with pre-absorbed Donkey anti-Rabbit AF647 (1:1000 dilution) for 1hour in the dark. Embryos were also incubated with Picogreen (1:30,000 dilution) for 20min to stain nuclei. After washing and mounting, the image was taken on a Leica confocal microscope. For the antibody injection experiment, 0-30min old embryos were dechorionated, lined up and glued to a cover slip. After 10 to 15mins drying, embryos were covered with halocarbon oil and antiserum was injected from one side. After 2 to 3 hours aging, embryos were removed from the cover slip by heptane and fixed by 10% para-formaldehyde. Fixed embryos were glued back to a cover slip and the vitelline membrane was removed manually. Embryos were stored in methanol for later *in situ* hybridization. Western blot was done using 25 embryos with standard western protocol. Signal was detected by LI-COR Odessey system by using Goat anti-rabbit AF680 as secondary antibody.

ChIP

Embryos were collected every 45 mins and aged for 2h, 2h45min, 3h30min

and 4h15min individually from a cage of y w (67,23) flies. ChIP was performed essentially as described previously. ChIP was repeated twice using independent chromatin preparations. Primer sequences used for PCR are listed in Appendix Table 5-4.

Microarray

Control RNA was extracted from a cross between females heterozygous for the *NGT[40+A]* Gal4-driver and males homozygous for *UAS-lacZ*. Experimental RNA was extracted from a cross between females heterozygous for both the *NGT[40+A]* Gal4-driver and NELF-EΔRRM and males homozygous for *UAS-lacZ*. Embryos at the headfold stage were manually selected under microscope and RNA was extracted immediately using the High Pure RNA isolation Kit (Roche). Control and experimental RNAs were respectively hybridized to Affymetrix's *Drosophila* Tiling 2.0R Array and data was analyzed by Affymetrix provided software. Microarray was repeated twice with biological replicates.

Quantitative measurement of β-galactosidase activity

Quantitative measurement of β-galactosidase activity was performed essentially as described previously. (Tracey et al. 2000) All readings were done by $20/20^{n}$ luminometer (Turner Biosystem). The relative light units reported were light emitted in 1 second averaged from 10-sec reading.

***In situ* hybridization**

In situ hybridization was carried out as described previously (Swantek and Gergen 2004). RNA riboprobes to detect the *slp1, runt, ftz, odd, wg, en* mRNA transcripts were synthesized as described. (Wheeler et al. 2002; Swantek and Gergen 2004) The *NELF-A* probe was synthesized with SP6 polymerase using a BglII-linearized SD07139 template. The *NELF-E* probe was synthesized with T3 polymerase using a PstI-linearized RE14181 template. SD07139 and RE14181 cDNA constructs were obtained from *Drosophila* Gene Collection.

Inverse-PCR

Inverse-PCR to map the inserted position of *NELF-EΔRRM* P-transposons was performed as described on Berkeley *Drosophila* Genome Project (BDGP) website: http://www.fruitfly.org/about/methods/inverse.pcr.html HpaII and Sau3AI (NE Biolabs) were used for restriction enzyme digestion. Primer-pairs pry1/pry4, plac1/pwht1 and Phusion hot start high-fiderlity DNA polymerase (NE Biolabs, F-540) were used for PCR.

Viability test

In *eve*, *runt* and *ftz* viability test, females were generated from a cross between females homozygous for wild-type, or *UAS-NELF-EΔRRM6A*, or *UAS-NEL*F-EΔRRM6B and males homozygous for *NGT[40]*;*NGT[A]*. Males were generated from crossing homozygous *UAS-eve*, or *UAS-runt*, or *UAS-ftz* flies with flies containing correspondent chromosome balancer. We determined viability by dividing the number of *UAS*-eve/*runt*/*ftz* progeny by the number of chromosome balancer progeny.

Name	Annotation Symbol	Genomic Location		Transcript length
		Chromosome (arm)	Cytogenetic map	
NELF-A	CG5874	3R	93C7-93C7	4648nt
NELF-B	CG32721	X	7B7-7B7	2189nt
NELF-D (TH1)	CG9984	X	14C2-14C3	2269nt
NELF-E	CG5994	3L	66D8-66D8	1225nt

Table 3-1 Basic information on NELF subunits The above information is gathered from Flybase. The NELF complex contains 4 subunits, A, B, D and E.

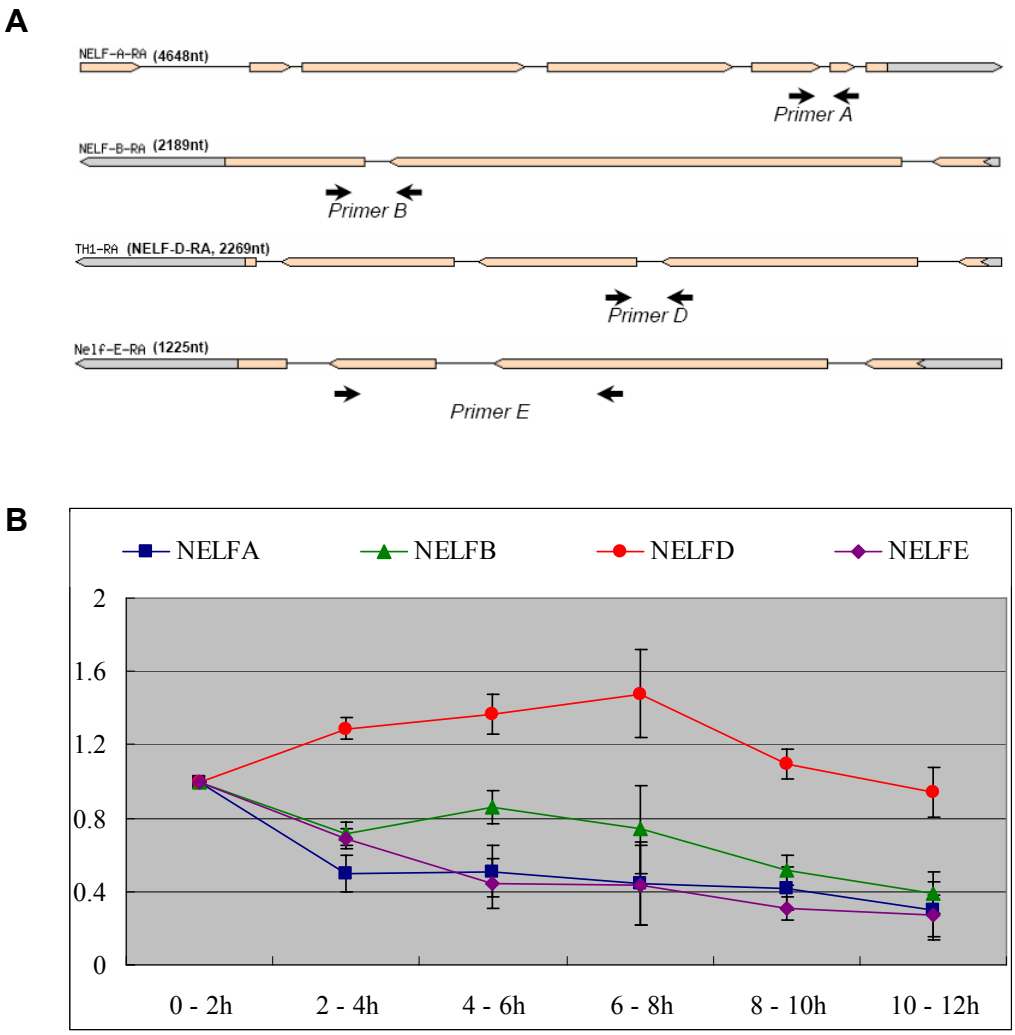


Figure 3-1 q-RT-PCR of NELF mRNA during *Drosophila* embryogenesis (A) The relative positions of each primer pair used for q-RT-PCR. The size of PCR products amplified by Primer A, B, D are between 150bp to 205bp. The size of PCR product amplified by Primer E is 371bp. **(B)** The relative amount of NELF RNA during 0-12h embryogenesis is determined by q-RT-PCR with 2 hour embryo collections aged for 0, 2, 4, 6, 8, 10 hours individually. The RNA amount in the 0-2h collection is adjusted to 1 for all NELF subunits.

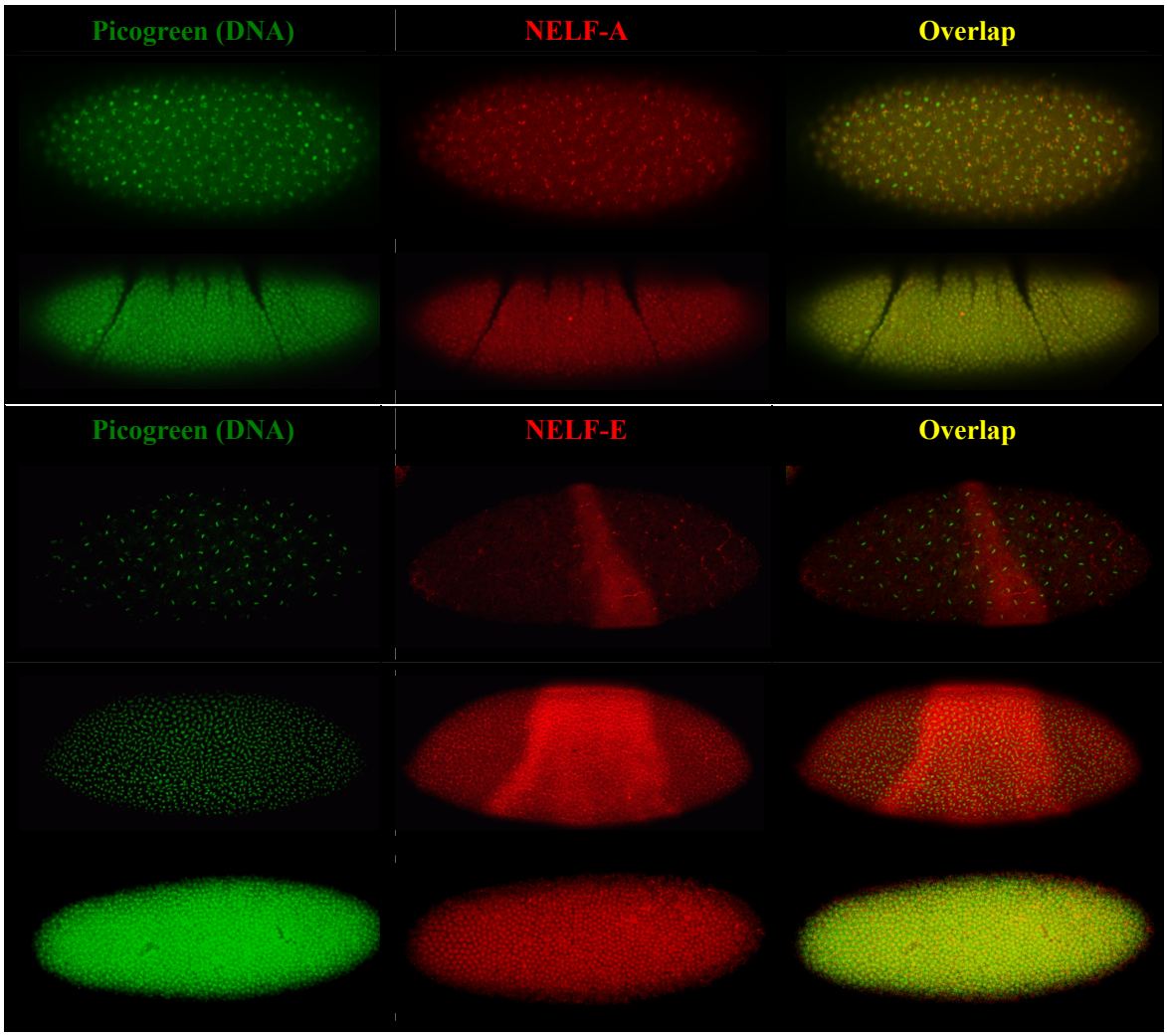


Figure 3-2 NELF-A and NELF-E protein distribution in early embryos Whole embryos are stained with Picogreen and NELF-A or NELF-E antiserum and detected by fluorescence. The upper two rows are embryos stained with Picogreen and NELF-A. The bottom three rows are embryos stained with Picogreen and NELF-E. NELF-A has a ubiquitous distribution throughout early embryogenesis. NELF-E accumulation starts from the middle part of the embryo and quickly distributes to all cells after cellularization. Both proteins co-localize with nuclei at the blastula stage.

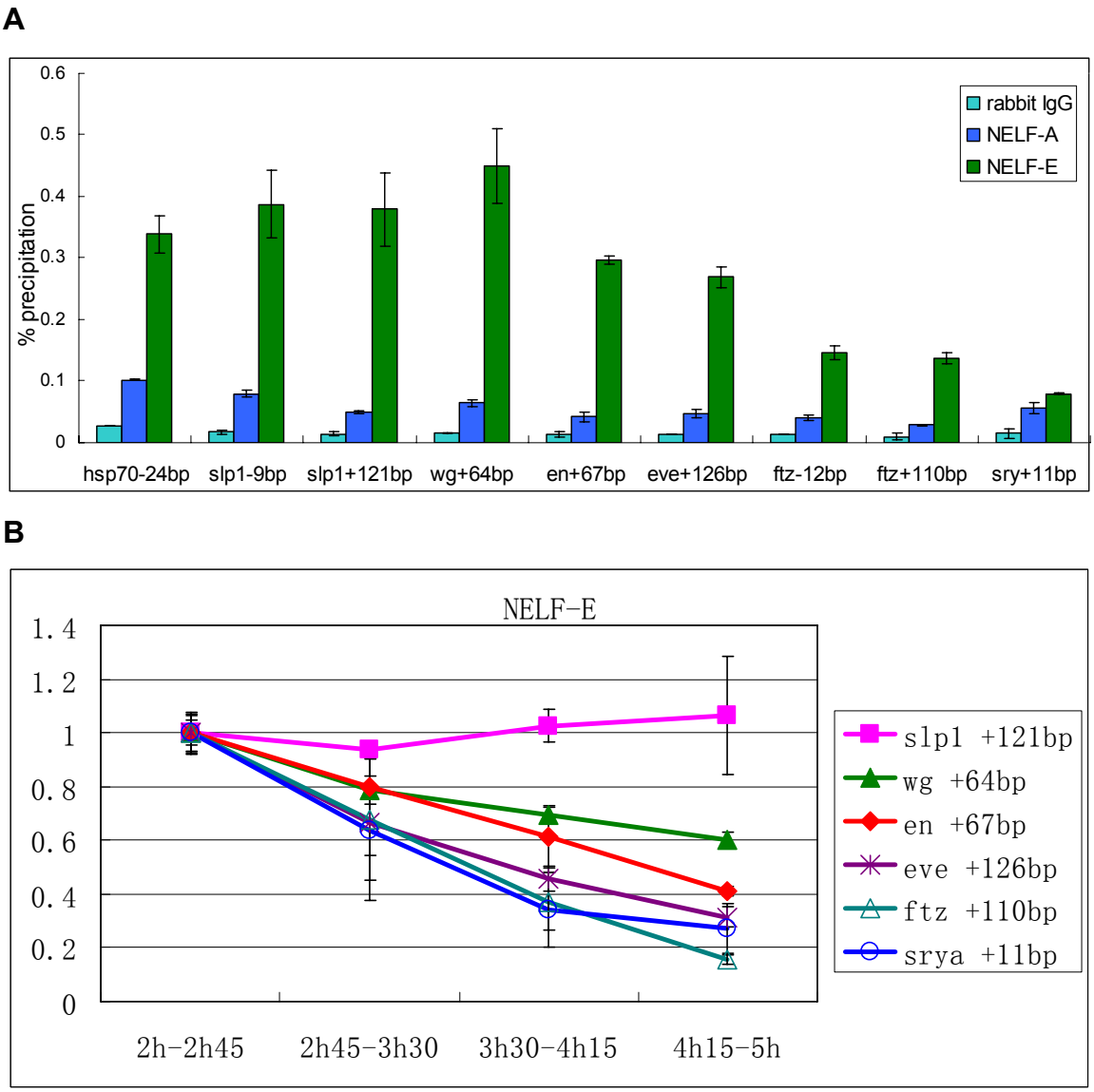


Figure 3-3 Temporal recruitment of NELF **(A)** Results of quantitative PCR on ChIP samples generated using antibodies against NELF-A and NELF-E with chromatin fragments isolated from 2h-2h45min wild-type embryos. The nonspecific background was determined using a control rabbit IgG antibody with the same chromatin preparation. Results obtained for 9 promoter or promoter-proximal regions are as labeled across the bottom. Error bars indicate the standard error in the percent precipitation values for each interval. **(B)** Results of quantitative PCR on NELF-E ChIP from 2h-2h45min, 2h45min-3h30min, 3h30min-4h15min and 4h15min-5h wild-type embryos. Results obtained for 6 promoter-proximal regions are presented with different colors and labels as shown on the right. Each signal (calculated in percent precipitation relative to input) is divided by the corresponding ChIP signal at the *hsp70*a promoter region. The ratio is adjusted to 1 for all primer pairs at 2h-2h45min window. The Y axis is the relative ChIP signal compared to the signal at 2h-2h45min window for each primer pair.

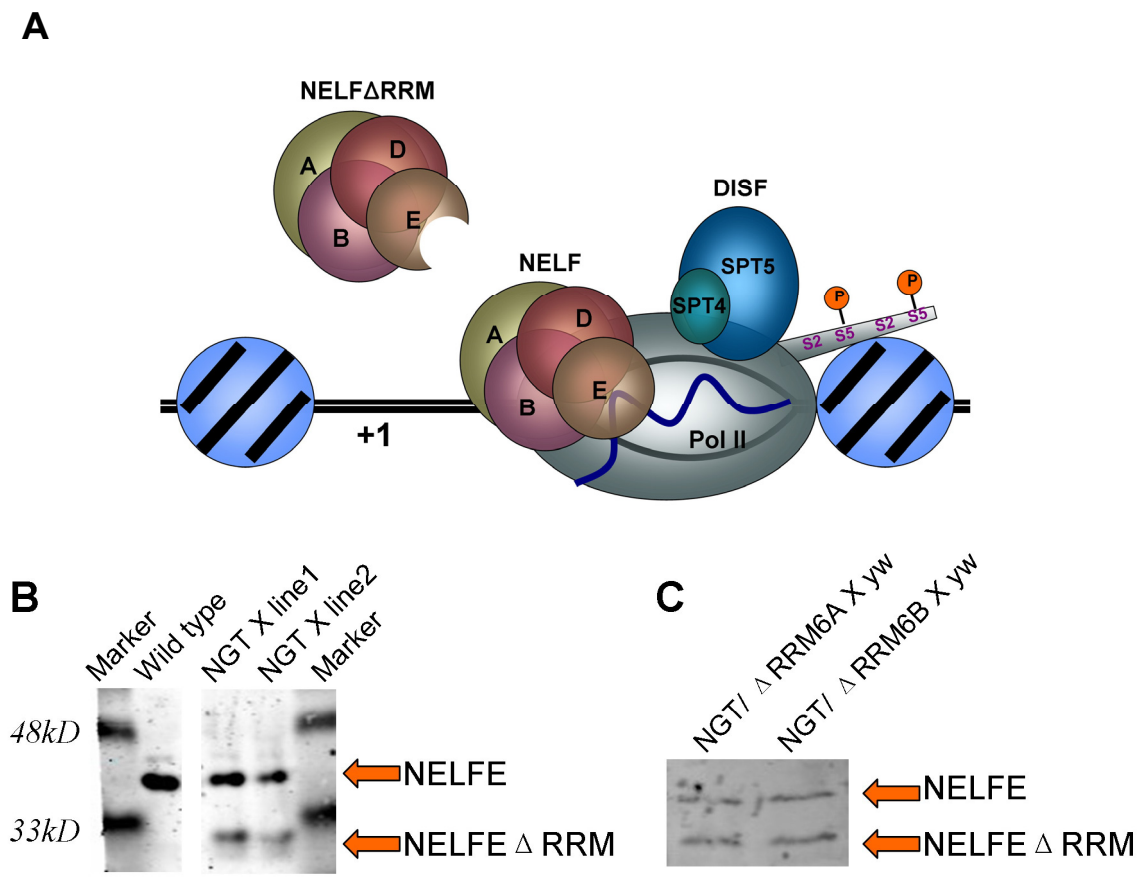


Figure 3-4 NELF-EΔRRM functions as a dominant negative form of NELF-E . (A) Predicted NELF-EΔRRM working model. NELF complex can bind to a nascent RNA and negatively affect Pol II transcription at the promoter proximal region. NELF-E without the RRM domain can compete with endogenous NELF-E to form a defective NELF complex with other NELF subunits. This defective NELF complex has no RNA binding ability. **(B)** Western blot result with NELF-E antibody shows that embryos generated by crossing two different *UAS-NELF-EΔRRM* males to homozygous *NGT[40]+NGT[A]* females express a novel cross-reacting protein at the right size for NELF-EΔRRM (about 8kD smaller than Wild-type NELF-E) that is not present in wild-type embryos as pointed out by arrows. **(C)** Western blot result with NELF-E antibody shows that maternally provide NELF-EΔRRM using the *NGT* system with two 6-copy *UAS-NELF*-EΔRRM lines respectively can generate NELF-EΔRRM in amounts that are comparable to wild-type NELF-E in 0-3hr old embryos.

Meiotic map	Insert position	Gene locus	*UAS*-NELF-EΔRRM 6A	*UAS*-NELF-EΔRRM 6B
2-32	2L:9,449,417	intron of *numb*	+	+
2-58	2R:4,121,229	upstream of *pnut*	+	
2-60	2R:6,363,010	intron of *CPTI*		+
2-90	2R:15,998,939	upstream of *18w*	+	+
3-0	3L:267,876	5'UTR of *miple*	+	+
3-48.3	3R:4,495,305	exon of *CG8036*	+	+
3-85	3R:20,369,628	upstream of *CG5789*	+	+

Table 3-2 Insert information of two *UAS-NELF*-EΔRRM lines *UAS-NELF*-EΔRRM6A and *UAS-NELF*-EΔRRM6B are two 6-copy NELF-EΔRRM-insertion fly lines with only 1 insertion difference. The presence of an insertion is indicated with a "**+**" mark. Inverse PCR was used to confirm the presence of the multiple transgenes on the final recombinant chromosomes.

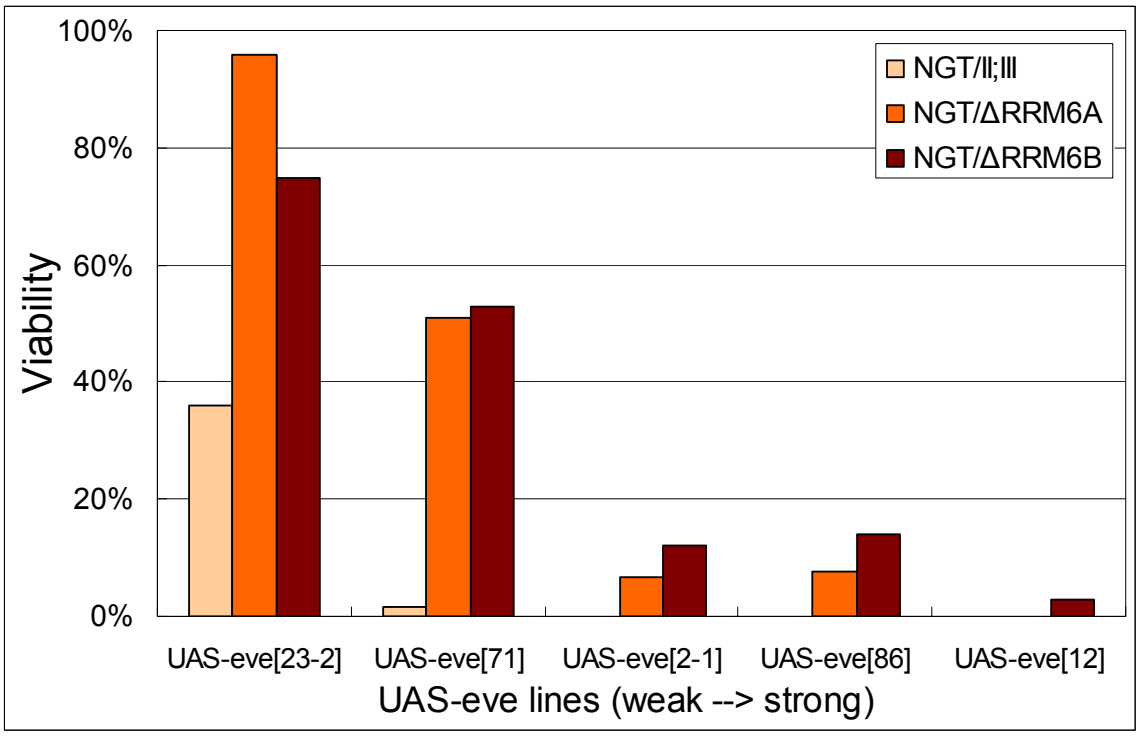


Figure 3-5 Dominant Negative NELF-E suppresses Eve-dependent lethality. The effects of maternally provided NELF-EΔRRM on the lethality associated with *NGT*-driven Eve expression were determined by mating males heterozygous for different *UAS*-eve transgenes and an appropriate balancer chromosome to three different genotypes of females. The five different transgenic *UAS*-eve lines are arranged in increasing order of their potency from left to right as indicated. The tan bars show the relative viability of these different transgenes in crosses with control females that carry single copies of both the *NGT[40]* and *NGT[A]* drivers. The relative viability increases for all of these *UAS*-eve lines in crosses with *NGT[40] NGT[A]* females that are also heterozygous for the six-copy NELF-EΔRRM6A (orange bars) and NELF-EΔRRM6B (brown bars) transgenes.

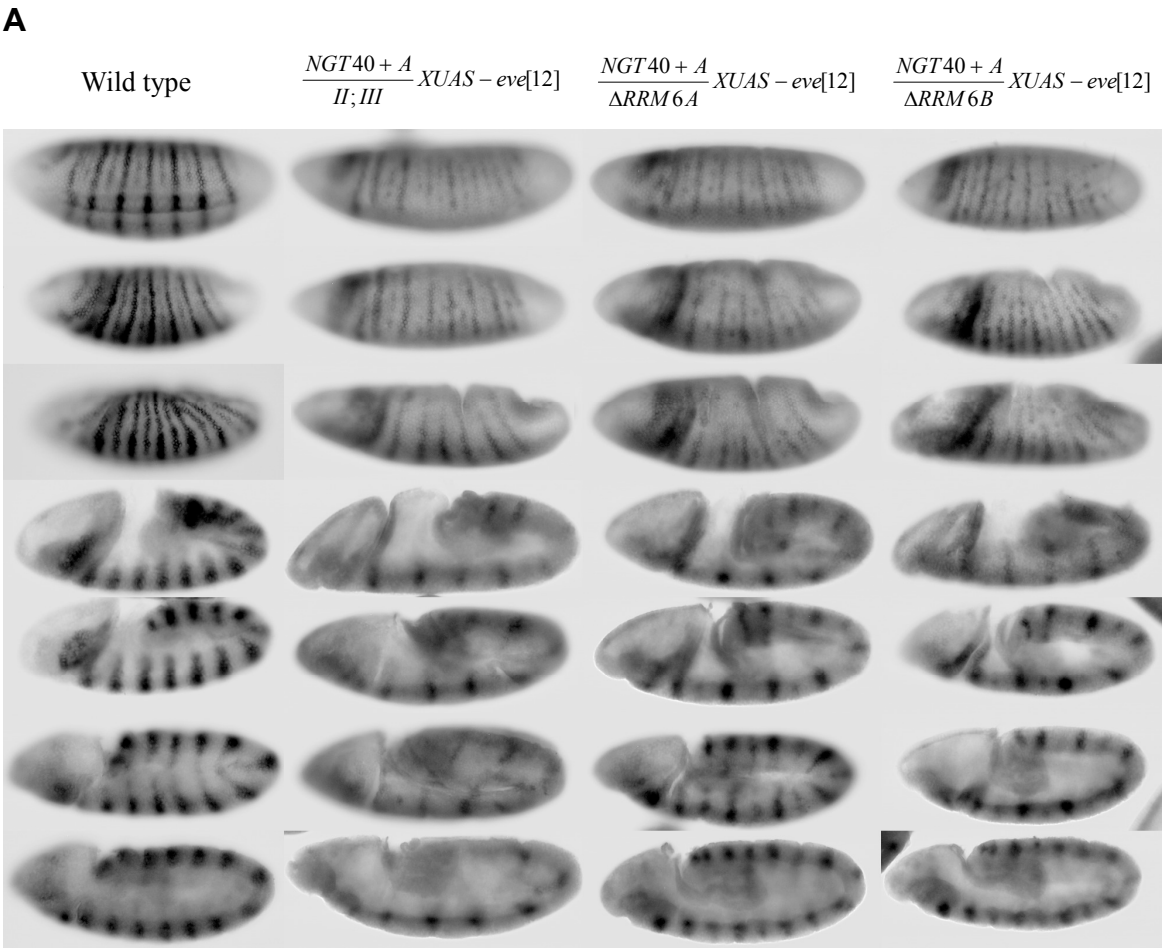
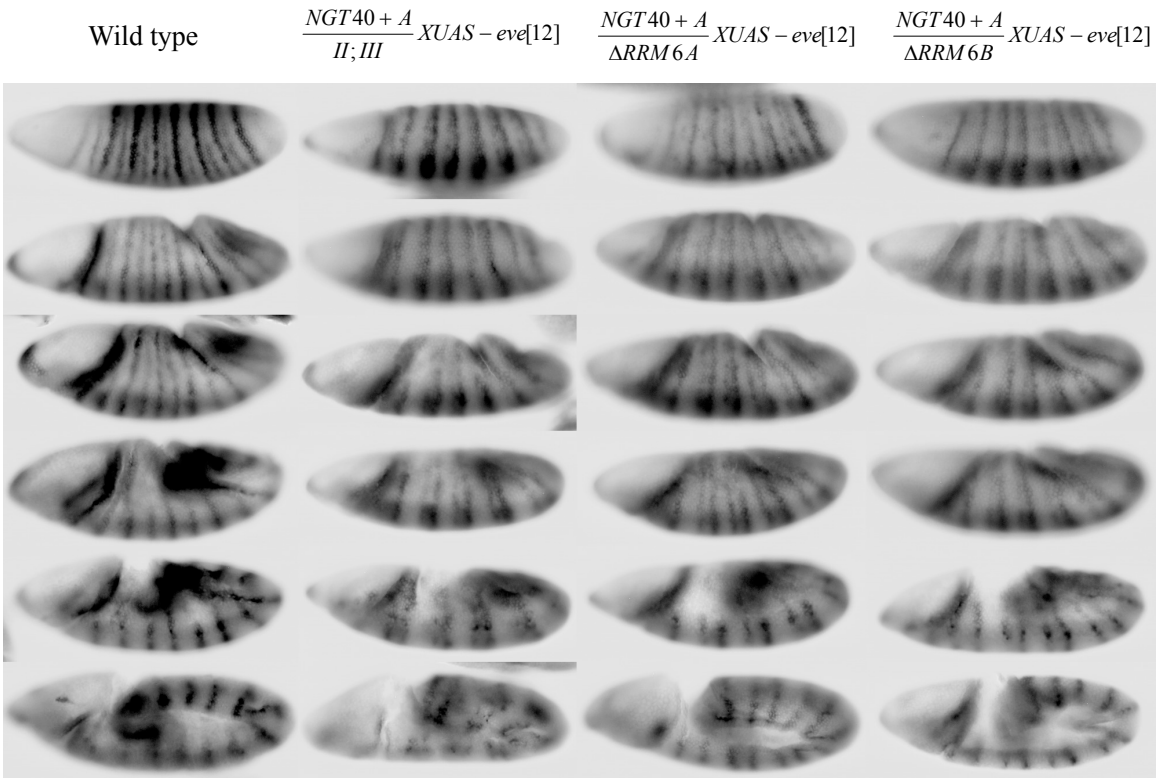


Figure 3-6 Eve-dependent *slp1* and *odd* repression are suppressed in embryos containing NELF-EΔRRM **(A)** Expression of *slp1* mRNA in blastula to germ band extension stage (from the top down) of wild-type (first column), eve-overexpressed only (second column), eve-overexpressed with maternal NELF-EΔRRM6A (third column) or 6B (fourth column) embryos as revealed by *in situ* hybridization. **(B)** (Next Page) Expression of *odd* mRNA in blastula to germ band extension stage (from the top down) of wild-type (first column), eve-overexpressed only (second column), eve-overexpressed with maternal NELF-EΔRRM6A (third column) or 6B (fourth column)

B


Wild type
$\frac{NGT40+A}{II;III}XUAS-eve[12]$
$\frac{NGT40+A}{\Delta RRM6A}XUAS-eve[12]$
$\frac{NGT40+A}{\Delta RRM6B}XUAS-eve[12]$

Paternal Genotype		*UAS*-lacZ/*UAS*-lacZ			
Maternal Genotype		Wild Type	$\frac{NGT40+A}{II;III}$	$\frac{NGT40+A}{\Delta RRM6A}$	$\frac{NGT40+A}{\Delta RRM6B}$
β-Galactosidase activity	**3.5hr**	463±28%	18,084±27%	29,261±28%	28,667±24%
	5.5hr	---	310,808±25%	354,452±18%	393,946±12%

Table3-3 Effects of NELF-EΔRRM on Gal4-*UAS* system as measured by β-Galactosidase activity. Embryos were selected at the onset of gastrulation based on the appearance of the cephalic and ventral furrows. Embryos at this stage were defined as 3.5 hr old. 5.5hr old embryos were obtained by aging them 2hr from the time of gastrulation. β-Galactosidase activity was measured on aliquots prepared from single staged embryos as described in materials and methods. The numbers shown indicate the relative light units±percent SD for measurements made on at least 30 embryos from each stage and genotype.

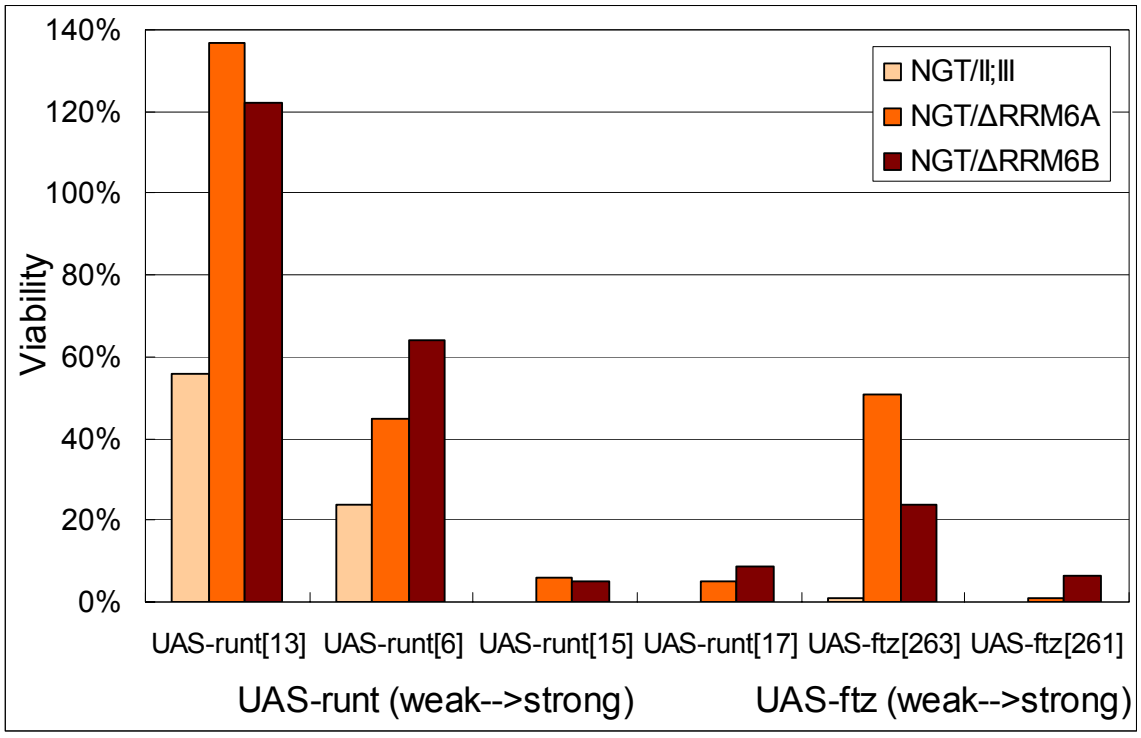


Figure 3-7 Dominant Negative NELF-E suppresses Runt or Ftz-dependent lethality.

The effects of maternally provided NELF-EΔRRM on the lethality associated with NGT-driven Runt or Ftz expression were determined by mating males heterozygous for different *UAS-Runt* or *UAS-ftz* transgenes and an appropriate balancer chromosome to three different genotypes of females. The labeling scheme is the same as used in Figure 3-5.

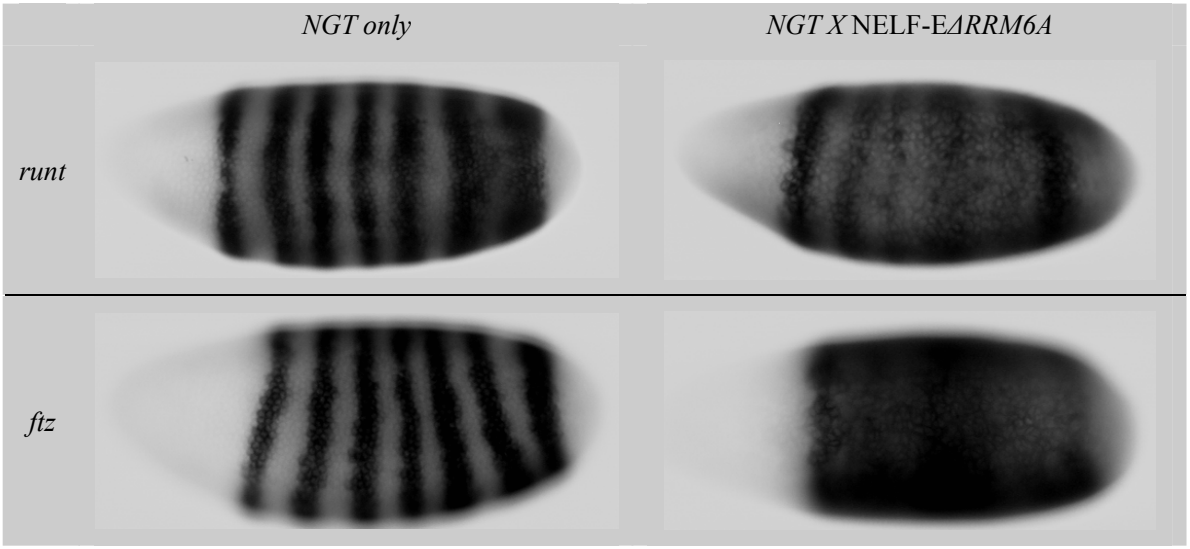


Figure 3-8 Heat shocking embryos leads to de-repression of *runt* and *ftz*. Expression of *runt* (first row) and *ftz* (second row) mRNA in blastula stage embryos that received a 30min heat shock followed by a 30min recovery. Embryos in the left column are from *NGT[40]* and *NGT[A]* homozygous females mated to wild-type males. Embryos in the right column are from a cross between same females and homozygous *UAS-NELF-EΔRRM6A*.

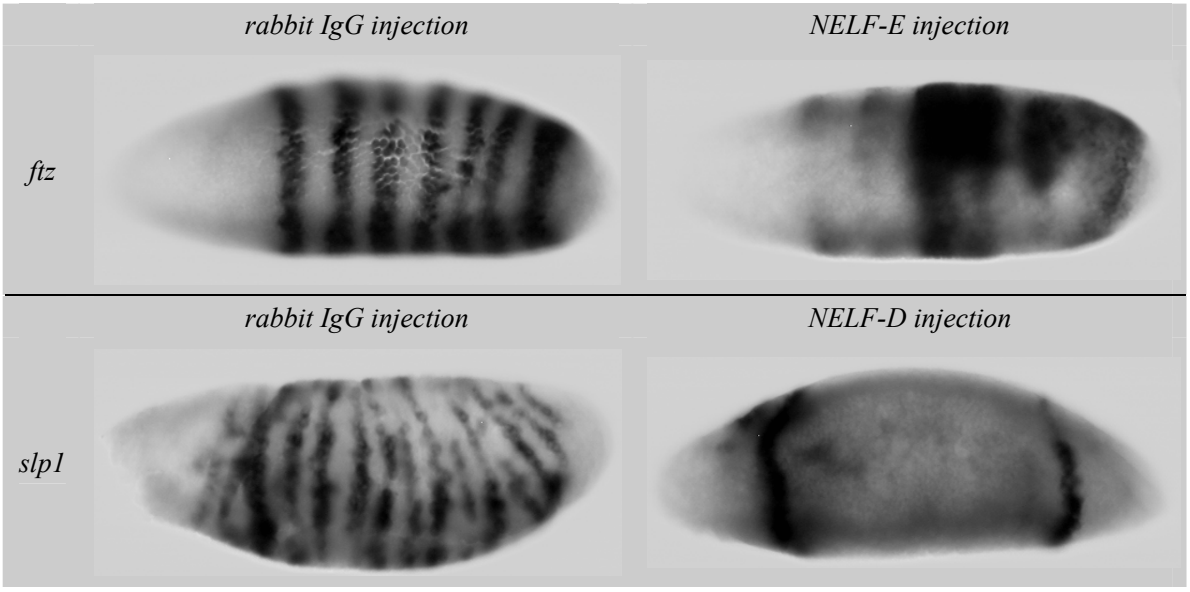


Figure 3-9 **Injection of antisera against NELF alters transcription and arrests cell cycle** Expression of *ftz* (first row) and *slp1* (second row) in wild-type embryos injected with rabbit IgG (left column) or with antisera against NELF-E (top right) or NELF-D (bottom right) injected as revealed by *in situ* hybridization

Chapter 4 General Discussion, Future Prospects, Ongoing Work and Miscellaneous Work

General Discussion and Future Prospects

Regulating at the elongation step in development

In this thesis, I used *Drosophila melanogaster* as the model system to study cell specification during early embryonic development. The *Drosophila* segmentation pattern is established by maternal effect genes that form gradients and regions of morphogenetic proteins along the anterior-posterior axis. These morphogenetic determinants create a gradient that differentially activates the gap genes, then pair rule genes, then segment polarity genes. The entire segmentation gene hierarchy relies on extremely precise transcription regulation. Each transcription factor is turned on at the right time and right place, and immediately cooperates with other factors to regulate downstream genes. Unlike pure cell lines, a developing embryo is a complex pool of all kinds of cells, which makes biochemical studies on embryos more challenging than on cell lines. So the majority of knowledge about segmentation is built by genetic experiments and *in vitro* protein-protein interactions. Therefore, although we can describe the regulation network of each factor during development, we barely know the molecular mechanisms. In the first part of my study, I employed a genetic manipulation to generate embryos that have segment polarity gene *sloppy paired* (*slp1*) repressed in all cells at the blastula stage. This provided a unique opportunity (or a unique *in vivo* test tube) to look at pair-rule gene dependent repression in the process of cell specification. By comparing these embryos with wild type embryos in which about 30% of the cells express *slp1*, I found that transcription was regulated at the elongation step. It has been widely assumed that the key rate-limiting step in gene activation is the recruitment of RNA polymerase II (Pol II) to the core promoter (Ptashne and Gann 1997), while my finding revealed a general role for regulation the elongation step in *Drosophila* development.

Examples of regulation during elongation have been noted for over 25 years, but until lately, they were generally considered exceptions. Several recent genome-wide ChIP-chip experiments from *Drosophila* and human cell lines found

that about 20% to 30% of genes have enriched Pol II density at the 5' end of the genes (Kim et al. 2005; Lee et al. 2006; Guenther et al. 2007; Muse et al. 2007). This class included genes with high to undetectable expression. Another genome-wide ChIP-chip assay with *Drosophila* embryos found that of the genes having a paused polymerase, genes that respond rapidly to developmental and cell signaling were highly enriched (Zeitlinger et al, 2007). This strongly supports my hypothesis that developmentally regulated transcriptional elongation is central to the process of cell fate specification.

Regulating at a transcriptional elongation step provides developmental genes with the ability to rapidly respond to developmental signals during embryogenesis. This may be critical for *Drosophila* embryogenesis because of its rapid development. The initial establishment of segment-polarity gene-expression patterns in response to the pair-rule transcription factors occurs within a relatively brief developmental window of ~30min, spanning the completion of cellularization and the beginning of germ band extension. The entire process of *Drosophila* embryogenesis is completed in 24 hours. Most mammalian embryos develop more slowly than *Drosophila*. One question raised here is whether it is a common mechanism of regulating developmental genes at the elongation step, or is it just a unique feature for *Drosophila* embryogenesis. Although there is no direct evidence shown in mammalian embryos, recent evidence shows that in mammalian embryonic stem (ES) cells the chromatin of many genes is configured in developmentally-potentiated states that index them for later transcriptional outcomes (Rasmussen 2008). In particular, genome-wide ChIP-chip on human ES cells found genes encoding most developmental regulators have transcription initiated but not elongated (Guenther et al. 2007). Based on the data of ES cells, very likely this regulation manner is common in metazoan development. The possible explanation is that even though the whole developmental process is slow, the ability to react rapidly to developmental signals can maximally insure the uniformity of the downstream response.

Factors contribute to Pol II pausing and elongating

Several mechanisms may work together to block Pol II progression at the promoter-proximal region. First, the first nucleosome downstream of the transcriptional start site (TSS) could be a physical barrier for Pol II progression. Second, positive factors might be required for dissociation of Pol II from the pre-initiation complex (PIC) as well as for full activation. Third, negative factors might inactivate Pol II or block Pol II from further elongation. Table 4-1 lists a summary of elongation factors that might be involved in this process.

Because only a subset of genes have Pol II paused downstream of the TSS, it

is very interesting to know if the phenomenon of Pol II pausing is sequence-dependent or is the result of combinatorial signals. One way to answer this question is to analyze the basal promoters and 5' UTR of all genes having paused Pol II by computational methods to see if any common character emerges, such as, GC rich sequence-clusters or sequences that may form hairpin structures. Secondly, *in vitro* transcription with active nuclear extracts could also be used to test this idea.

Recent work points out that not only the 5' end but also the 3' end of temporarily quiescent genes contribute to Pol II pausing (Glover-Cutter et al. 2008). Human *p21* was known to be regulated at the elongation step. Before and after its activation, Pol II, capping factors, Spt5 (subunit of DSIF) and 3' processing factors were found at both 5' end and 3' end of *p21*. The dual localization of these factors at both ends is not likely due to the loop formed between both ends. Therefore, Pol II undergoes a dual-pausing wherein elongation arrests near the transcription start site and in the 3' flank to allow co-transcriptional processing by factors recruited to the Pol II complex. We have not carefully investigated the function of 3' ends in gene regulation. But my finding that Groucho (a transcription co-repressor) binds to the 3' end of *ftz* in early embryos (data not shown) suggests that the 3' end is potentially important.

Function of P-TEFb in developmentally regulated transcription elongation

P-TEFb (positive transcription elongation factor b) is a cyclin-dependent kinase composed of CDK9 and CycT. During *HIV-1* (Human Immunodeficiency Virus type 1) transcription, P-TEFb can specifically stimulate the processivity of Pol II elongation and antagonize the actions of negative elongation factors. Besides being a cellular cofactor for *HIV-1*, P-TEFb is also a general transcription factor required for efficient expression of the vast majority of cellular genes (Zhou and Yik 2006). Moreover, in the developing embryo, the onset of new transcription is delayed in the germ cells relative to somatic cells. Evidence suggests that in *C. elegans* and *Drosophila* embryos, transcription is blocked after the initiation step, and P-TEFb recruitment is inhibited in the germ cells (Seydoux and Dunn 1997; Zhang et al. 2003; Hanyu-Nakamura et al. 2008). Therefore, P-TEFb becomes one of the most obvious candidates to overcome the influences of the negative factors and positively control the elongation phase of transcription during development.

We did a few experiments related to P-TEFb, but none of them were convincing. Because it is such an obvious candidate, I think there are several more experiments we should do. First, in a recent paper, antibodies against *Drosophila* CDK9 and CycT were generated and successfully used for a ChIP

experiment in *Drosophila* S2 cell line (Hanyu-Nakamura et al. 2008). We plan to request antibodies from them and try some ChIP experiments in early embryos. If P-TEFb is essential for reversing the negative effects of NELF and stimulating the transcription of developmental regulated genes, we should see a correlation between an association of P-TEFb at promoter-proximal regions and the expression of many developmental genes. Also, in *slp1*-repressed embryos, the association of P-TEFb at the *slp1* locus should be greatly reduced. Second, it was reported that a single residue mutation of CDK9(D→N) can kill the CDK9 kinase activity *in vitro* (Garriga et al. 1996; Lis et al. 2000). Therefore, we can generate dominant-negative CDK9 lines by transforming *UAS-CDK9*(D→N) into flies. Third, there are fly lines having P-element inserted at *CDK9* or *CycT* locus. Potentially they could be CDK9 and CycT mutants. Our preliminary dose-dependent experiment gives us indirect results. We can further study P-TEFb function by generating germ line clones to eliminate maternally provided P-TEFb. Altogether, these experiments should give us more information on the role of P-TEFb in development.

Chromatin structure in development

In this thesis, I only examined chromatin structure by looking at H3 and acetylated H3 distribution at the *slp1* and *brother* (*bro*) loci. The conclusion I made from these experiments was that the H3 and Ac-H3 patterns are not changed in *slp1*-repressed embryos compared with wild-type embryos. To extend this conclusion, I think not only H3 or Ac-H3, but also the entire chromatin structure of the *slp1* locus is well prepared for activation even that the gene may never be turned on in a large subset of cells. This idea can be tested by doing more ChIP experiments. The histone modification we are interested in includes both transcription activation marks (H3.3, H4Ac, H3K4me3, H3K4me2, H3K4me, H3K36me3, H3K79me3) and repression marks (H3K9me3, H3K27me3). Moreover, it is potentially interesting to look at chromatin structure at different stages of development. We can easily generate *slp1*-repressed embryos. We know this repression can be maintained well for at least 3 hours. In my experiment, I only focused on the blastula stage, in which the transcription repression is just being established. Chromatin structure may undergo some changes to maintain the repression status. It is easy for us to do different time collections from 2 hour to 5 hours AED and trace the change in histone modification after transcription repression is established.

Chromatin looping structure may also form between upstream regulatory elements and promoter region. We can use 3C technique (Dekker et al. 2002) to identify the looping structure and observe potential changes during the

development.

NELF function in *Drosophila* early embryonic development

Chapter 3 suggests a widespread function for NELF in development. I have proposed that NELF negatively regulates transcription elongation at different levels. The repression established by NELF is temporary and reversible, which provides genes with the ability to quickly and accurately respond to developmental signals. It is very interesting that yeast do not have NELF. This ability is probably gained later in metazoan evolution.

We also know NELF is critical for embryonic development based on the fact that our NELF-A mutant flies are recessive embryonic lethal and fail to complement a small deletion of the corresponding region. More experiments on these NELF-A mutants are ongoing now. By characterizing the NELF-A mutant, we hope to obtain further insight on the *in vivo* role of NELF.

Lee et al. proposed that NELF's interaction with the nascent transcript associated with paused Pol II is independent of sequence and RNA secondary structure(Lee et al. 2008). This explains the global function of NELF. We do not know whether nascent transcripts are sufficient to recruit NELF to the Pol II machinery or there are some other factors that recruit NELF. But the quick localization of NELF into nuclei and the broad association of NELF at gene 5' end hint that an opened chromatin and a nascent transcript with Pol II probably is enough to attract the NELF complex.

Model of developmentally regulated transcription

Figure 4-1 shows the model of developmentally regulated transcription based on my data and current publications. In early *Drosophila* embryonic development, the loci of genes involved in segmentation will be open and modified regardless of the cell fate. Pol II then binds to the promoter regions of these genes to initiate transcription. NELF is recruited by the nascent RNA simultaneously to repress the transcription temporarily. DISF is another factor which is also thought to be recruited to the Pol II complex and plays negative role in transcription. This temporary repression stage will soon be affected due to developmental signals and the current cell environment. In response to the negative signals, some loci will enter the persistent repression stage, in which the originally opened loci will have histone 3 tri-methylated at lysine 9 and lysine 27. Pol II transcriptional machinery will leave the loci and the chromatin is condensed to resist activation. In response to the limited positive signals, some loci will be activated restrictedly by controlling the phosphorylation of NELF-E. The partially phosphorylated NELF

complex will frequently contact the Pol II machinery downstream of the promoter to regulate the efficiency of transcription. At some certain conditions in which positive signals are abundant, NELF will become fully phosphorylated and leave the gene loci. Under this circumstance, transcription will be fully activated and synthesize plentiful transcripts. In summary, transcription during early development is dynamically regulated at the elongation step through the association of NELF to nascent RNA and Pol II machinery. The scale of activation depends on the classes and concentration of positive signals.

For human NELF-E protein, serines at positions 181, 185, 187, 191 in its RD domain are potentially phosphorylated by p-TEFb (Fujinaga et al. 2004). But *Drosophila* NELF-E doesn't have an RD domain. Sequence alignment between *Drosophila* and human NELF-E shows that none of the serine residue above is present in *Drosophila* NELF-E. So either *Drosophila* uses other serines as phosphorylation sites, or the intermediate status of NELF-E is not regulated by phosphorylation.

Several experiments potentially can be done to test this model. First, several serine residues in NELF-E can be gradually replaced by alanines. *In vitro* transcription experiment can quantitatively determine the efficiency of transcription. Also, flies having this mutation should have similar phenotype to the decreased-dosage or null phenotype of P-TEFb (if P-TEFb is responsible for NELF-E phosphorylation). Second, if cells having a gene activated or repressed can be directly separated at an early developmental stage, it will provide great material to look at the association and phosphorylation of NELF-E in opposite cell environments. In the miscellaneous work part of this thesis, I will discuss the possibility of this approach.

Our data suggested the broad and important functions of NELF during *Drosophila* development. Therefore, it is not a surprise to see the association of human NELF with diseases. NELF-A was found to be encoded by *WHSC2*, a candidate gene for a multiple malformation syndrome called the Wolf-Hirschhorn syndrome (Wright et al. 1999; Yamaguchi et al. 2001); Low NELF-B (also named as COBRA1) expression is strongly associated with metastatic breast cancer (Sun et al. 2008). Higher expression of NELF-B and NELF-E is associated with tumorigenesis in the upper gastrointestinal tract (Midorikawa et al. 2002; McChesney et al. 2006). I believe further NELF studies in development will help reveal the molecular basis of these NELF-related diseases and hint the way of effective treatment.

Ongoing work

NELF-A mutant

A homozygous-lethal stock NELF-A[KG09483] was obtained from the *Drosophila* Stock Center. The chromosome was first cleaned up by genetic methods so that the lethality was only caused by P-transposon insertion. PCR confirmed that KG P-transposon was inserted in the middle of the NELF-A first intron. In order to determine the effect of this insertion on NELF-A transcription, we made NELF-A Germ Line Clone (GLC) females and crossed to wild-type males. Therefore, all embryos from this cross should lack maternal NELF-A. q-RT-PCR on 0-2hour old embryos indicated less than 10% of NELF-A transcript was present in NELF-A GLC embryos comparing with wild-type embryos (data not shown) We also balanced NELF-A[KG09843] with TM3 containing twi-Gal4 and UAS-GFP. Embryos having two copies of NELF-A[KG09843] can be selected by the absence of fluorescence. Combining GLC and fluorescence selecting, we will be able to generate and select embryos missing both maternal and zygotic NELF-A. Currently we are doing a large-scaled GLC experiment. Hopefully we can get result to support the finding from NELF-E dominant negative lines and obtain further insight on the *in vivo* role of NELF.

NELF-D RNA Interference

A genome-wide library of *Drosophila melanogaster* RNAi transgenes is publicly available since 2007(Dietzl et al. 2007), These RNAi transgenes consist of short gene fragments cloned as inverted repeats and expressed using the GAL4/UAS system. We “imported” several UAS-IR lines last year and started testing them for activity. Recently, we found that one *UAS-IR-NELFD* line can slightly reduce the lethality caused by over-expression of Eve and Runt. By crossing *UAS-IR-NELFD* line to *Sev-Gal4, nina.GMR-Gal4* or ey-Gal4 line (they all express Gal4 in eye), I obtained rough eye or eyeless phenotypes, which indicates a function for NELF-D in eye development. More viability tests and functional studies are ongoing now. Also, we plan to generate stronger *UAS-IR-NELFD* lines by hopping P-element insertion into other positions and combining several lines together (like the way I did to generate stronger *UAS-eve* and *UAS-NELFEΔRRM* lines). When we get these lines, we can measure the NELF-D RNA amount by q-RT-PCR and rank a series of *UAS-IR-NELFD* lines in strength. Potentially this is a useful tool for us to study NELF function quantitatively.

Miscellaneous work

Detection of short transcripts

Both ChIP data and permanganate footprinting data showed that *slp1* transcription is blocked at the early elongation step. I also tried several approaches to determine if *Drosophila* embryos contain short RNAs from the 5' end of the *slp1* transcription unit that would be produced by blocking transcriptional elongation. Northern blot and RNase protection assay were not able to give me convincing results due to their limited sensitivity. I therefore applied a technique called RNA ligation mediated RT-PCR (RLM-RT-PCR), which includes a signal amplification step. The products of RLM-RT-PCR are defined at one end by the DNA primer used in the RNA Ligase reaction and at the other end by a primer specific for the gene of interest. (Figure 4-2A)

To test this method, I synthesized a 270-nucleotide RNA containing *slp1* sequence from -81 to +125 relative to transcription start site (TSS). *slp1* primers with a 5'-end 13 and 39 bp downstream of TSS (for short, primer+13 and primer+39) were used individually for final PCR step. For both primers, I was able to get a single clear band at expected size. Purification and sequencing of these products demonstrated they were the predicted products. (Figure 4-2B)

RLM-RT-PCR on RNA isolated from nuclei of 3-4h AED wild-type embryos using a *slp1* primer with a 5'-end 8 nucleotides downstream of TSS (primer+8) generates multiple bands. When corrected for contributions from the primer and linker, the actual short RNA transcripts size ranged from 30nt to 80nt, with the peak from 30nt to 50nt. Therefore, Pol II paused at multiple places. Synthesis of this product is dependent on both the RNA Ligase and Reverse Transcriptase steps of the RLM-RT-PCR protocol (Figure 4-2C).

The above result is consistent with the *in vivo* permanganate footprinting data, in which strong Pol II pausing sites were observed from +15 to +50 downstream from TSS and followed by weak pausing sites till +90. This provides additionally evidence that the patterned expression of *slp1* in the early embryo depends on regulating transcriptional elongation.

There are several advantages of this technique. It provides an approach for detecting prematurely truncated RNA transcripts in embryos where a subset of the cells express the full-length mRNAs. It is very sensitive, 2ug of nuclear RNA is enough to get a positive signal even through theoretically only 20% of the cells have truncated RNA transcripts. Moreover, it is a relatively easy experiment to do.

Once the cDNA is made, it can be used to check different genes by simply doing several more PCR reactions. The disadvantage of this technique is that sometimes you will get PCR artifacts. Therefore each gene specific primer needs to be optimized. Also, it is better to perform southern blot afterwards to confirm that PCR products are gene specific, Overall, I believe RLM-RT-PCR, together with *in vivo* footprinting are potentially a good combination to study transcription elongation of specific genes in developmental organisms.

Investigate the role of the GAGA factor in cell specification

Background The *Drosophila* GAGA factor is encoded by the Trithorax-like locus. It has been shown to bind the GA/CT rich sequence in many *Drosophila* promoters (*en, eve, E74, ftz, kruppel, ultrabithorax, hsp26, hsp70, H3, H4*) *in vitro* by DNase I protection assay (Soeller et al. 1993; Granok et al. 1995). *In vivo* ChIP-chip assay identified 2907 regions have GAGA association in *Drosophila* S2 cell lines (Lee et al. 2008). GAGA RNA was detectable throughout embryogenesis, with maximum relative levels at 9 to 12h AED (after egg deposition). Multiple RNA species were detected, two of which were most prominent. From 0-12h AEL, an RNA of approximately 2.4kb predominated. After 9h AEL, a 3.0kb RNA was present in equivalent or greater levels. GAGA protein is ubiquitously expressed in embryos and localized in nuclei after the blastula stage (Soeller et al. 1993).

GAGA is generally known as a transcription activator. It can locally alter chromatin structure and allow other transcription factors access to the DNA. It directs histone H3.3 replacement that prevents heterochromatin spreading (Nakayama et al. 2007). As a global factor, it is also involved in cell cycle and nuclear division (Bhat et al. 1996).

Results and Discussion GAGA antibody was obtained from Susumu Hirose. I did GAGA ChIP on 3-5h old wild-type embryos. Results of quantitative PCR on ChIP samples were shown in Figure 4-3A. It was known that GAGA is need for hsp70 gene activation. It binds to hsp70 promoter region *in vivo* before and after heat shock. As expected, we got high GAGA binding signal at hsp70 promoter region but not at structural gene region. The observation that GAGA also binds to *slp1*, *en* promoter region strongly suggests that GAGA is involved in cell specification during early embryogenesis.

Next, embryos were collected every 45 minutes from 2-5 hours AED. ChIP experiment using rabbit-IgG and GAGA antisera were performed in each embryo collection. Besides *hsp70*, en and *slp1* promoter, I was also able to detect positive GAGA signal at *eve*, *wg*, *ftz* and *srya* promoter. This indicates the general role of GAGA in early embryogenesis. However, GAGA signal at the *wg*, *ftz* and *srya*

promoter regions are much weaker comparing with the signal at the *hsp70*, *slp1*, *en* and *eve* promoters (data not shown). Therefore, only the signal at the *slp1, en* and *eve* locus along the time scale was analyzed. To compare the signal among different time windows, each signal (calculated in percent precipitation relative to input) is divided by the corresponding ChIP signal at hsp70 promoter region. The ratio is adjusted to 1 for all primer pairs at 2h-2h45mins window and plotted versus time in Figure 4-3B.

Different from the descending trend of NELFE binding during development, GAGA binding is almost constant from 2 hours to 5 hours. This agrees with the idea that GAGA alters chromatin structure to keep the transcription locus active. The ability that we can detect differences in GAGA and NELF association trends provides strong evidence that NELF dissociates along with the activation of genes.

All genes we tested here are transcriptionally active at this stage. It is interesting to know on a genome-wide scale, how many and what kind of genes have GAGA involved in early *Drosophila* embryogenesis. If GAGA is involved, is its binding necessary and sufficient for chromatin opening and Pol II recruitment at this stage. Also, when does GAGA leave? Can we consider GAGA as a transcription activation marker, as the way we think about histone acetylation? It is potentially useful for us to pursue a little bit further studies on GAGA function in cell specification.

The study of transcription activation

Most of my work has focused on transcription repression. I did not do much work on transcription activation, not because I did not want to, but because I can not. Ectopic expression of Runt and Ftz provides a clean system to study repression, whereas I did not have a good genetic tool to look at activation. To address this problem, I have already tried several approaches, as list below.

1. Mutate Ftz in all cells

   As shown in Figure 2-1A, Ectopic expression of Runt and Opa in embryos that are mutant for Ftz leads to *slp1* activation in all somatic blastoderm cells. This type of embryo comes from the following cross: *NGT;Ftz/TM3*→ *UAS[runt,opa];Ftz/TM3*. Therefore, only 25% of the embryos are the embryos we want. To sort these embryos out, we tried to put GFP on the *TM3* chromatin balancer. So embryos with no fluorescence should be Ftz double mutant. Unfortunately, the fluorescence emitted from zygotic GFP or EGFP comes out relatively late. We can not find obviously fluorescent embryos at gastrulation

stage even though embryos having GFP or EGFP. A third chromatin marker Sry-α-deficient X3F also isn't able to give a detectable phenotype at this stage. Next I generated a RNAi knockdown to eliminate ftz in all cells. But by crossing *UAS-IR-Ftz* line to *NGT* females, I can't get a clear Ftz-null phenotype. I also tried to knock out *ftz* from both male and female germ lines (Detailed germ line clone cross can be viewed in Appendix Figure 5-2). Unfortunately, these flies were sterile for unknown reasons, presumably due to other mutations.

2. Slp1-positive cell sorting

   Because of the difficulties encountered with all of our genetic attempts, I decided to try and use cell flow sorting to isolate Slp1-positve cells. Cells were first isolated from staged embryos and fixed by formaldehyde. Then, all fixed cells were stained with Slp1 antibody + fluorescence conjugated secondary antibody, followed by cell sorting. The flaw of this method is that the quick nuclear localization of *slp1* makes the staining and sorting very inefficient. Therefore, the yield was low after sorting. To solve this problem, we are thinking to introduce mammalian CD2 into *Drosophila*. Mammalian CD2 has been shown to be an effective heterologous marker of the cell surface in *Drosophila* embryos (Dunin-Borkowski and Brown 1995). Theoretically, flies containing a transgenic CD2 gene driven by *slp1* upstream element should express CD2 in the same pattern as Slp1 and locate in the cell membrane right after translation. This makes it possible to sort Slp1-positive cells with commercially available CD2 antibodies. *UAS-CD2* flies are available at *Drosophila* stock center. Therefore before making constructs, we can test this idea by sorting CD2 positive cells from a cross between *UAS-CD2* flies and any flies containing a gene-specific Gal4 driver. An alternative way is to generate *slp1*-driven-GFP flies and sort Slp1-positive cells at relatively late stages. Both ideas are potentially feasible if we made the transgenic flies.

3. ChIP Slp1-ChIP

   Based on Mark Biggin's ChIP-chip data, Slp1 physically binds to its own promoter upstream region. So, instead of sorting Slp1-positive cells, potentially, we can get Slp1-assocated chromatin by doing Slp1-ChIP. The precipitated chromatin can be sonicated into small fragments and immuno-precipitated with another antibody. By doing PCR at *slp1* gene locus, we can answer some specific questions, such as, does NELF leave the *slp1* promoter-proximal region when *slp1* is actively transcribed. We have obtained Slp1 antibody from Mark Biggin. The regular *slp1* ChIP showed that Slp1 protein binds to the *slp1* gene locus at the -400bp region. We will test our ChIP Slp1-ChIP idea soon.

Material and Methods

Fly strains

NELFA[KG09843] was obtained from *Drosophila* Stock Center; UAS-IR-NELFD (42217) was obtained from Vienna *Drosophila* RNAi Center.

UAS-IR-ftz construct was cloned with the strategy described previously (Lee and Carthew 2003) and transformed by standard P-element-mediated germ line transformation using the p:Δ2-3 helper plasmid. Over 10 lines were generated but none of them can give convincing *ftz* null phenotype.

RLM-RT-PCR

Control sense RNA was synthesized with T3 polymerase using EcoRI-linearized plasmid containing a PCR fragment of slp1 -81 to +125 (relative to transcriptional start site) subcloned into the EcoRV and EcoRI site of pBS(KS+).

WT total RNA was extracted from ~200 WT 2.5h-4.5h AED embryos using the High Pure RNA isolation Kit (Roche). Nuclear RNA was purified from nuclei isolated from WT 2.5~4.5h AED embryos. Nuclei from ~200mg of embryos were treated with 40units of DNase (RNase free, Roche) for 30mins at 37^{o}C. Following treatment with SDS and proteinase K, RNA was extracted by phenol/chloroform and ethanol precipitation and then re-suspended in nuclease-free water.

RNA ligation was carried out at 37^{o}C for 2hours with 2μg WT total RNA+1pg of control RNA or 2μg nuclei RNA, 400pmol 5'-phosphorylated linkerA, 20% PEG8000, 1ul RNaseOut, 1Xligation buffer and 40units of T4 RNA ligase (NE Biolabs #M0204) in a 20ul reaction. After ligation, 400pmol Reversed-linkerA was added, the reaction was heated at 95 ^{o}C for 5min and cooled down slowly. 2ul of ligation sample was used for cDNA synthesis with First Strand cDNA synthesis kit (Roche). After cDNA synthesis, 1/20 of each sample was used for 30 cycles of PCR with 10pmol gene specific primer and Reversed-linkerA. Products of the RLM-RT-PCR assay were separated on polyacrylamide gels and visualized by staining with ethidium bromide. Primer information is listed in Appendix Table 5-4.

ChIP

GAGA antiserum was kindly provided by Susumu Hirose. ChIP experiment was performed as described previously.

Isolation of *Drosophila* embryonic cells and cell flow sorting

Stage collected embryos were dechorionated and transferred to a Dounce homogenizer in 1ml Schneider's *Drosophila* medium and homogenized with 4 slow strokes of the pestleB. The homogenate was transferred into a 1.5ml eppendorf and centrifuged at 860g for 2mins at 4 ^{o}C. The pellet was suspended in 1ml cold S2 medium and centrifuged again at same speed. The pellet was resuspended in 500ul of Ca^{2+} free S2 medium containing 0.25% trypsin and incubated at 25 ^{o}C for 10mins for 2.5h-4.5h old embryos. The concentration of trypsin used depended on the developmental stage of embryos. The reaction was stopped by adding 500ul of S2 medium+20% fetal bovine serum. The cells were filtered through 70μm nylon mesh and centrifuged at 860g for 2min. The pellet was suspended in 1ml PBS. Repeat centrifuge and suspend pellet in 200μl PBS+4% paraformadehyde to fix for 15 minutes at RT. After fixation, cells were washed twice with PBS/BSA/Azide buffer (PBS+0.5%BSA+0.02%NaN3) and counted by hemocytometer. Cells were then stored in PBS/BSA/Azide buffer at 10^{7} to 10^{8}/ml at 4 ^{o}C. If not used immediately or in few days, cells can be snap-frozen by liquid nitrogen and stored at -80^{o}C.

For cell immunostaining, 10^{7}-10^{8} cells were penetrated by permeabilization buffer (0.5%Saponin in PBS/BSA/Azide) for 10min and pelleted at 2000g for 3min. Cell pellet was suspended in permeabilization buffer containing primary antibody and incubated for 30min. The cells were then washed twice with permeabilization buffer and incubated with secondary antibody for 30min if necessary. Then cells were washed twice with permeabilization buffer and twice with PBS/BSA/Azide buffer. Finally, cells were suspended in PBS/BSA/Azide buffer and transferred to FACS tubes for FACS analysis.

Table 4-1 Transcription elongation factors

Category	Factor	Potential functions	*Drosophila* subunit
Establishment of paused Polymerase complex	DSIF	DRB sensitivity-inducing factor, travels along with Pol II; positively and negatively regulates Pol II elongation.	Spt4
			Spt5
	NELF	Negative elongation factor, binds to nascent RNA to pause transcription.	NELF-A
			NELF-B
			NELF-D
			NELF-E
Modulation of promoter-associated polymerase	Mediator	Has more than 25 subunits, activates transcription through direct interactions with Pol II; also cooperatively interacts with other coactivators.	Med17 (Srb4)
			Med20 (Srb2)
			Med23 (TRAP150b)
			Med24 (TRAP100)
			Med31; et al
	TFIIF	Stimulates the rate of transcription; required for efficient promoter escape.	TFIIFα (RAP74)
			TFIIFβ (RAP74)
	P-TEFb	Positive transcription elongation factor b, phosphorates NELF, DSIF, and ser2 of Pol II CTD.	Cdk9
			CycT
	TFIIS	Important for the efficient escape of Pol II from the pause. Stimulates the intrinsic RNA-cleavage activity of Pol II to create a new RNA 3′-OH in the active site and once again enables transcription elongation.	TFIIS
	Kismet	Associates with transcriptionally active regions and facilitates an early step in transcription elongation:	Kis-L
			KIs-S
Elongation complex components	FACT	Directs H2A/H2B binding, dissociates histone octamers.	Spt16
			SSRP1
	Spt6	Nucleosome disassembly/reassembly factor, facilitates Pol II elongation.	Spt6
	PAF1	Regulates histone monoubiquitylation and H3-K4 trimethylation. Coordinates histone modifications and nucleosome structure with transcription activation	Paf1
			Rtf1
			Cdc73

Elongation complex components (continued)		and Pol II elongation.	
	TAC1	Histone modification complex, facilitates Pol II elongation by methylation and acetylation of nucleosomal histones in the 5'-coding region.	Trx
			Nej (CBP)
			Sbf1
	ELL	Associates with Pol II, enhances Pol II elongation.	ELL
	Elongin	Stimulates Pol II elongation; enhances Pol II processivity.	EloA
			EloB
			EloC
	Pcf11	3'-end processing factor; promotes the dissociation of Pol II elongation complexes from DNA.	Pcf11

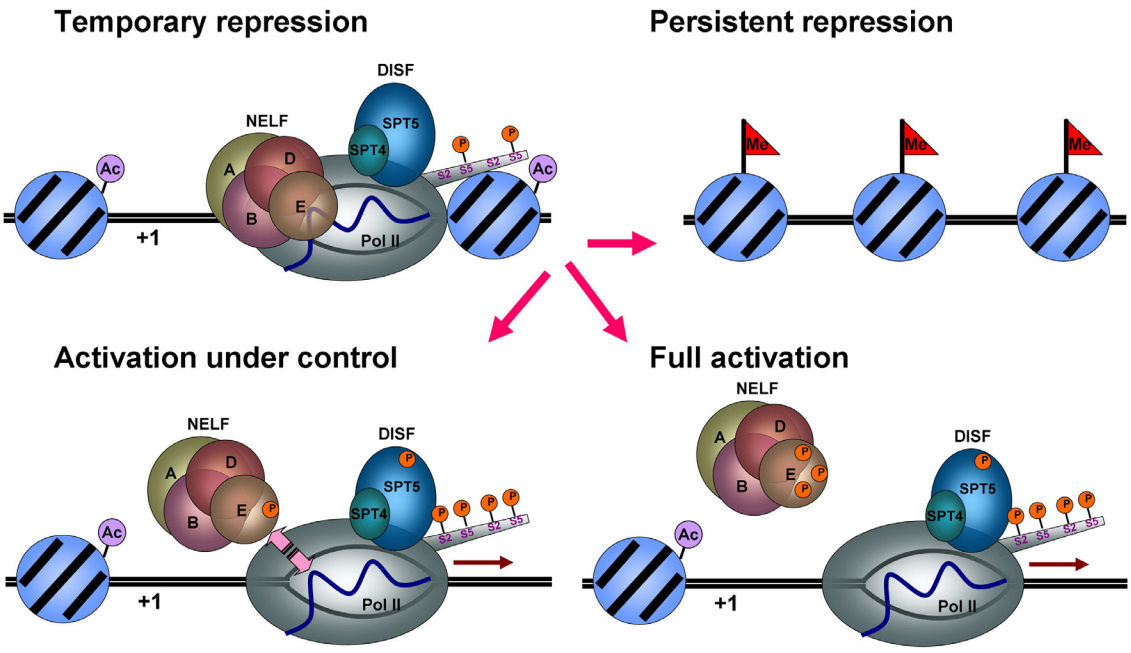


Figure 4-1 Model of developmentally regulated transcription Transcription during early development is dynamically regulated at the elongation step through the association of NELF to nascent RNA and Pol II machinery. The scale of activation depends on the classes and concentration of positive signals. At early stage of development, the loci of genes involved in segmentation will be open and modified regardless of the cell fate. Pol II then binds to the promoter regions of these genes to initiate transcription. NELF is recruited by the nascent RNA simultaneously to repress the transcription temporarily. DISF is another factor which is also thought to be recruited to the Pol II complex and plays negative role in transcription. In response to the negative signals, some loci will enter the persistent repression stage, in which the originally opened loci will have histone 3 tri-methylated at lysine 9 and lysine 27. Pol II transcriptional machinery will leave the loci and chromatin is condensed to resist activation. In response to the limited positive signals, some loci will be activated restrictedly by controlling the phosphorylation of NELF-E. The partially phosphorylated NELF complex will frequently contact the Pol II machinery at promoter downstream to regulate the efficiency of transcription. At some certain conditions in which positive signals are abundant, NELF will become fully phosphorylated and leave the gene loci. Under this circumstance, transcription will be fully activated and synthesize plentiful transcripts.

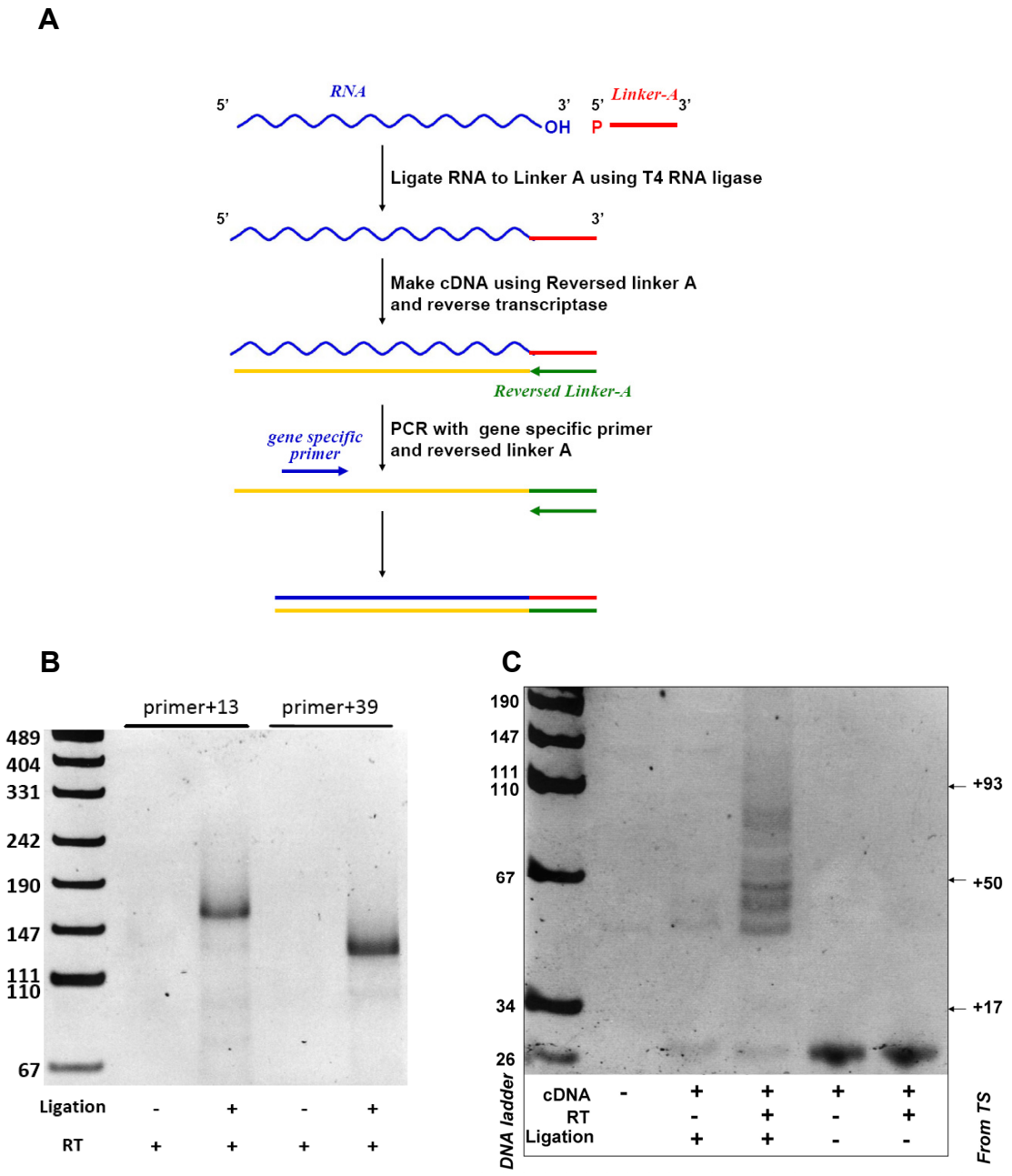
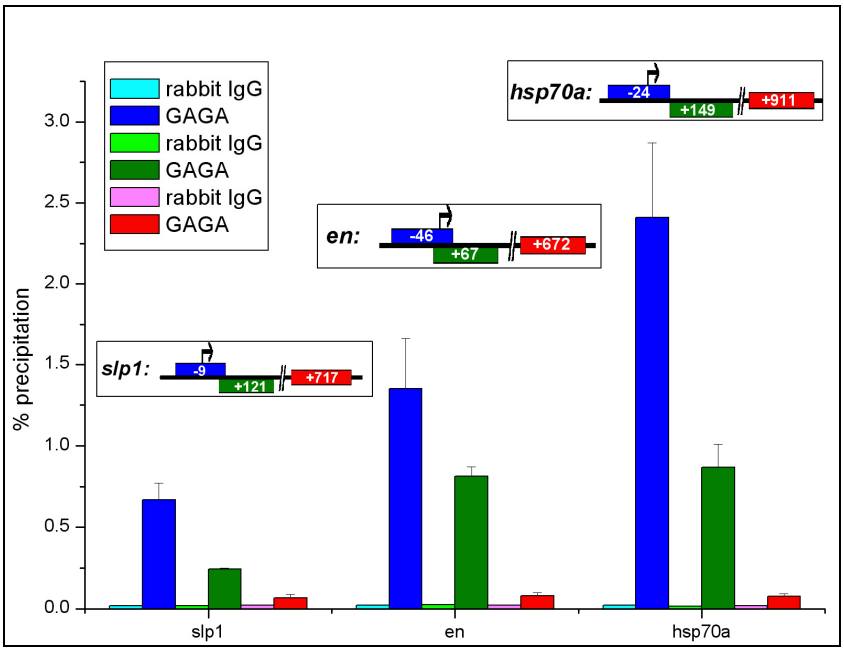


Figure 4-2 Expression of short RNAs from the 5' UTR regions of *Drosophila slp1* genes **A,** Overview of the RLM-RT-PCR protocol used to detect RNA transcripts with promoter proximal region. **B,** Results of RLM-RT-PCR on synthetic RNA using primer starting 13 and 39 nucleotides downstream of the *slp1* promoter. **C,** Results of RLM-RT-PCR on nuclear RNA isolated from 3-4h AED wild type embryos using primer starting 8 nucleotides downstream of the *slp1* promoter. Size marked on the right side of the gel has been adjusted for the linker and primer used. Therefore it represents the length of short RNA transcripts. Multiple bands amplified in RT and ligation positive sample indicate that *slp1* transcription has been stalled at more than one site.

A



B

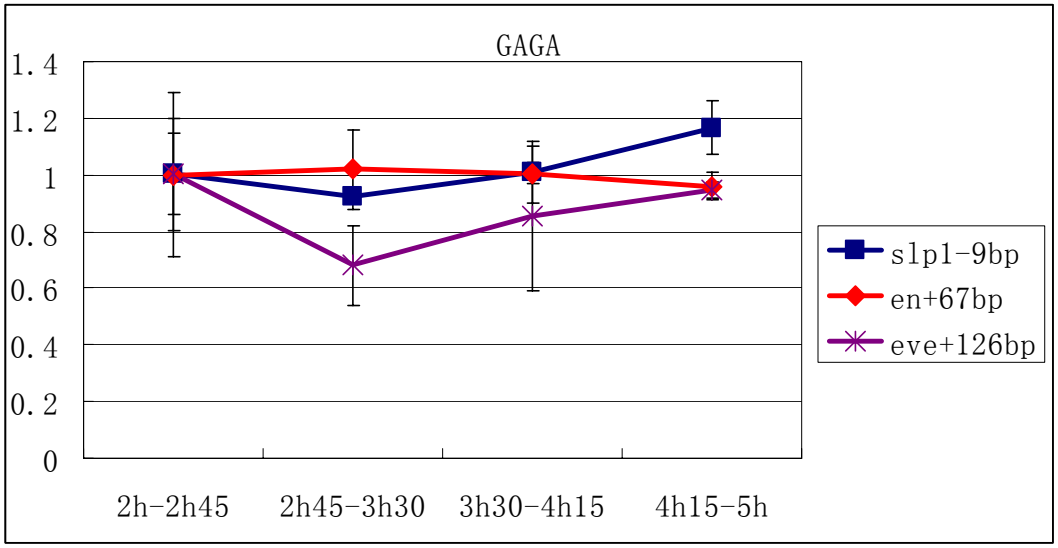


Figure 4-3 Temporal recruitment of GAGA **(A)** Results of quantitative PCR on ChIP samples generated using antibodies against GAGA with chromatin fragments isolated from 3-5h wild-type embryos. The nonspecific background was determined using a rabbit IgG antibody with the same chromatin preparation. Results for *slp1*, *en* and *wg* loci are showed as labeled. **(B)** Results of GAGA ChIP from 2h-2h45min, 2h45min-3h30min, 3h30min-4h15min and 4h15min-5h wild-type embryos. The labeling scheme and the way to normalize and present data is the same as used in Figure 3-3.

Bibliography

Aida, M., Y. Chen, et al. (2006). "Transcriptional pausing caused by NELF plays a dual role in regulating immediate-early expression of the junB gene." Mol Cell Biol **26**(16): 6094-104.

Aiyar, S. E., J. L. Sun, et al. (2004). "Attenuation of estrogen receptor alpha-mediated transcription through estrogen-stimulated recruitment of a negative elongation factor." Genes Dev **18**(17): 2134-46.

Akam, M. (1987). "The molecular basis for metameric pattern in the Drosophila embryo." Development **101**(1): 1-22.

Albert, I., T. N. Mavrich, et al. (2007). "Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome." Nature **446**(7135): 572-6.

Bender, T. P., C. B. Thompson, et al. (1987). "Differential expression of c-myb mRNA in murine B lymphomas by a block to transcription elongation." Science **237**(4821): 1473-6.

Berger, S. L. (2002). "Histone modifications in transcriptional regulation." Curr Opin Genet Dev **12**(2): 142-8.

Bhat, K. M., G. Farkas, et al. (1996). "The GAGA factor is required in the early Drosophila embryo not only for transcriptional regulation but also for nuclear division." Development **122**(4): 1113-24.

Bhaumik, S. R., E. Smith, et al. (2007). "Covalent modifications of histones during development and disease pathogenesis." Nat Struct Mol Biol **14**(11): 1008-16.

Boehm, A. K., A. Saunders, et al. (2003). "Transcription factor and polymerase recruitment, modification, and movement on dhsp70 in vivo in the minutes following heat shock." Mol Cell Biol **23**(21): 7628-37.

Brock, H. W. and M. van Lohuizen (2001). "The Polycomb group--no longer an exclusive club?" Curr Opin Genet Dev **11**(2): 175-81.

Buratowski, S., S. Hahn, et al. (1989). "Five intermediate complexes in transcription initiation by RNA polymerase II." Cell **56**(4): 549-61.

Cartwright, I. L., D. E. Cryderman, et al. (1999). "Analysis of Drosophila chromatin structure in vivo." Methods Enzymol **304**: 462-96.

Chen, C. and A. J. Sytkowski (2001). "Erythropoietin activates two distinct signaling pathways required for the initiation and the elongation of c-myc." J Biol Chem **276**(42): 38518-26.

Coffman, J. A. (2003). "Runx transcription factors and the developmental balance between cell proliferation and differentiation." Cell Biol Int **27**(4): 315-24.

Core, L. J. and J. T. Lis (2008). "Transcription regulation through promoter-proximal pausing of RNA polymerase II." Science **319**(5871): 1791-2.

Courey, A. J. and S. Jia (2001). "Transcriptional repression: the long and the short of it." Genes Dev **15**(21): 2786-96.

Das, R., A. Laederach, et al. (2005). "SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments." RNA **11**(3): 344-54.

Dekker, J., K. Rippe, et al. (2002). "Capturing chromosome conformation." Science **295**(5558): 1306-11.

Dietzl, G., D. Chen, et al. (2007). "A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila." Nature **448**(7150): 151-6.

Dunin-Borkowski, O. M. and N. H. Brown (1995). "Mammalian CD2 is an effective heterologous marker of the cell surface in Drosophila." Dev Biol **168**(2): 689-93.

Fujinaga, K., T. P. Cujec, et al. (1998). "The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and Tat." J Virol **72**(9): 7154-9.

Fujinaga, K., D. Irwin, et al. (2004). "Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element." Mol Cell Biol **24**(2): 787-95.

Fujioka, M., J. B. Jaynes, et al. (1995). "Early even-skipped stripes act as morphogenetic gradients at the single cell level to establish engrailed expression." Development **121**(12): 4371-82.

Fujioka, M., G. L. Yusibova, et al. (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**(19): 4411-21.

Gariglio, P., M. Bellard, et al. (1981). "Clustering of RNA polymerase B molecules in the 5' moiety of the adult beta-globin gene of hen erythrocytes." Nucleic Acids Res **9**(11): 2589-98.

Garriga, J., X. Mayol, et al. (1996). "The CDC2-related kinase PITALRE is the catalytic subunit of active multimeric protein complexes." Biochem J **319 (Pt 1)**: 293-8.

Glover-Cutter, K., S. Kim, et al. (2008). "RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes." Nat Struct Mol Biol **15**(1): 71-8.

Gomes, N. P., G. Bjerke, et al. (2006). "Gene-specific requirement for P-TEFb activity and RNA polymerase II phosphorylation within the p53 transcriptional program." Genes Dev **20**(5): 601-12.

Granok, H., B. A. Leibovitch, et al. (1995). "Chromatin. Ga-ga over GAGA factor." Curr Biol **5**(3): 238-41.

Green, M. R. (2000). "TBP-associated factors (TAFIIs): multiple, selective transcriptional mediators in common complexes." Trends Biochem Sci **25**(2): 59-63.

Guenther, M. G., S. S. Levine, et al. (2007). "A chromatin landmark and transcription initiation at most promoters in human cells." Cell **130**(1): 77-88.

Guo, S., Y. Yamaguchi, et al. (2000). "A regulator of transcriptional elongation controls vertebrate neuronal development." Nature **408**(6810): 366-9.

Hanyu-Nakamura, K., H. Sonobe-Nojima, et al. (2008). "Drosophila Pgc protein inhibits P-TEFb recruitment to chromatin in primordial germ cells." Nature **451**(7179): 730-3.

Howard, K. (1990). "The blastoderm prepattern." Semin Cell Biol **1**(3): 161-72.

Ibnsouda, S., F. Schweisguth, et al. (1993). "Relationship between expression of serendipity alpha and cellularisation of the Drosophila embryo as revealed by interspecific transformation." Development **119**(2): 471-83.

Ingham, P. W. (1988). "The molecular genetics of embryonic pattern formation in Drosophila." Nature **335**(6185): 25-34.

Jennings, B. H., S. Shah, et al. (2004). "Locus-specific requirements for Spt5 in transcriptional activation and repression in Drosophila." Curr Biol **14**(18): 1680-4.

Jones, J. C., H. P. Phatnani, et al. (2004). "C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats." J Biol Chem **279**(24): 24957-64.

Kagoshima, H., K. Shigesada, et al. (1993). "The Runt domain identifies a new family of heteromeric transcriptional regulators." Trends Genet **9**(10): 338-41.

Kim, J. B. and P. A. Sharp (2001). "Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase." J Biol Chem **276**(15): 12317-23.

Kim, T. H., L. O. Barrera, et al. (2005). "A high-resolution map of active promoters in the human genome." Nature **436**(7052): 876-80.

Kim, Y., J. H. Geiger, et al. (1993). "Crystal structure of a yeast TBP/TATA-box complex." Nature **365**(6446): 512-20.

Kim, Y. J., S. Bjorklund, et al. (1994). "A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II." Cell **77**(4): 599-608.

Komarnitsky, P., E. J. Cho, et al. (2000). "Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription." Genes Dev **14**(19): 2452-60.

Kornberg, R. D. and Y. Lorch (1999). "Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome." Cell **98**(3): 285-94.

Kouzarides, T. (2002). "Histone methylation in transcriptional control." Curr Opin Genet Dev **12**(2): 198-209.

Lanzuolo, C., V. Roure, et al. (2007). "Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex." Nat Cell Biol **9**(10): 1167-74.

Lebedeva, L. A., E. N. Nabirochkina, et al. (2005). "Occupancy of the Drosophila hsp70 promoter by a subset of basal transcription factors diminishes upon transcriptional

activation." Proc Natl Acad Sci U S A **102**(50): 18087-92.
Lee, C., X. Li, et al. (2008). "NELF and GAGA factor are linked to promoter-proximal pausing at many genes in Drosophila." Mol Cell Biol **28**(10): 3290-300.
Lee, T. I., R. G. Jenner, et al. (2006). "Control of developmental regulators by Polycomb in human embryonic stem cells." Cell **125**(2): 301-13.
Lee, Y. S. and R. W. Carthew (2003). "Making a better RNAi vector for Drosophila: use of intron spacers." Methods **30**(4): 322-9.
Levanon, D. and Y. Groner (2004). "Structure and regulated expression of mammalian RUNX genes." Oncogene **23**(24): 4211-9.
Li, X. Y., S. MacArthur, et al. (2008). "Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm." PLoS Biol **6**(2): e27.
Lis, J. (1998). "Promoter-associated pausing in promoter architecture and postinitiation transcriptional regulation." Cold Spring Harb Symp Quant Biol **63**: 347-56.
Lis, J. T., P. Mason, et al. (2000). "P-TEFb kinase recruitment and function at heat shock loci." Genes Dev **14**(7): 792-803.
Maldonado, E., M. Hampsey, et al. (1999). "Repression: targeting the heart of the matter." Cell **99**(5): 455-8.
Marshall, N. F., J. Peng, et al. (1996). "Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase." J Biol Chem **271**(43): 27176-83.
McChesney, P. A., S. E. Aiyar, et al. (2006). "Cofactor of BRCA1: a novel transcription factor regulator in upper gastrointestinal adenocarcinomas." Cancer Res **66**(3): 1346-53.
Midorikawa, Y., S. Tsutsumi, et al. (2002). "Identification of genes associated with dedifferentiation of hepatocellular carcinoma with expression profiling analysis." Jpn J Cancer Res **93**(6): 636-43.
Morris, D. P., G. A. Michelotti, et al. (2005). "Evidence that phosphorylation of the RNA polymerase II carboxyl-terminal repeats is similar in yeast and humans." J Biol Chem **280**(36): 31368-77.
Muse, G. W., D. A. Gilchrist, et al. (2007). "RNA polymerase is poised for activation across the genome." Nat Genet **39**(12): 1507-11.
Nakayama, T., K. Nishioka, et al. (2007). "Drosophila GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading." Genes Dev **21**(5): 552-61.
Narita, T., Y. Yamaguchi, et al. (2003). "Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex." Mol Cell Biol **23**(6): 1863-73.
Nasiadka, A. and H. M. Krause (1999). "Kinetic analysis of segmentation gene interactions in Drosophila embryos." Development **126**(7): 1515-26.
O'Brien, T. and J. T. Lis (1993). "Rapid changes in Drosophila transcription after an instantaneous heat shock." Mol Cell Biol **13**(6): 3456-63.

Orlando, V., E. P. Jane, et al. (1998). "Binding of trithorax and Polycomb proteins to the bithorax complex: dynamic changes during early Drosophila embryogenesis." EMBO J **17**(17): 5141-50.

Orphanides, G., T. Lagrange, et al. (1996). "The general transcription factors of RNA polymerase II." Genes Dev **10**(21): 2657-83.

Patturajan, M., N. K. Conrad, et al. (1999). "Yeast carboxyl-terminal domain kinase I positively and negatively regulates RNA polymerase II carboxyl-terminal domain phosphorylation." J Biol Chem **274**(39): 27823-8.

Peterson, C. L. and M. A. Laniel (2004). "Histones and histone modifications." Curr Biol **14**(14): R546-51.

Pinaud, S. and J. Mirkovitch (1998). "Regulation of c-fos expression by RNA polymerase elongation competence." J Mol Biol **280**(5): 785-98.

Ptashne, M. and A. Gann (1997). "Transcriptional activation by recruitment." Nature **386**(6625): 569-77.

Raschke, E. E., T. Albert, et al. (1999). "Transcriptional regulation of the Ig kappa gene by promoter-proximal pausing of RNA polymerase II." J Immunol **163**(8): 4375-82.

Rasmussen, E. B. and J. T. Lis (1995). "Short transcripts of the ternary complex provide insight into RNA polymerase II elongational pausing." J Mol Biol **252**(5): 522-35.

Rasmussen, T. P. (2008). "Developmentally-poised chromatin of embryonic stem cells." Front Biosci **13**: 1568-77.

Robertson, H. M., C. R. Preston, et al. (1988). "A stable genomic source of P element transposase in Drosophila melanogaster." Genetics **118**(3): 461-70.

Roth, S. Y., J. M. Denu, et al. (2001). "Histone acetyltransferases." Annu Rev Biochem **70**: 81-120.

Rougvie, A. E. and J. T. Lis (1988). "The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged." Cell **54**(6): 795-804.

Saunders, A., L. J. Core, et al. (2006). "Breaking barriers to transcription elongation." Nat Rev Mol Cell Biol **7**(8): 557-67.

Schneider, D. A., S. L. French, et al. (2006). "RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing." Proc Natl Acad Sci U S A **103**(34): 12707-12.

Schubeler, D., D. M. MacAlpine, et al. (2004). "The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote." Genes Dev **18**(11): 1263-71.

Schwartz, Y. B. and V. Pirrotta (2007). "Polycomb silencing mechanisms and the management of genomic programmes." Nat Rev Genet **8**(1): 9-22.

Seydoux, G. and M. A. Dunn (1997). "Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of Caenorhabditis elegans and Drosophila melanogaster." Development **124**(11):

2191-201.
Simon, J. A. and J. W. Tamkun (2002). "Programming off and on states in chromatin: mechanisms of Polycomb and trithorax group complexes." Curr Opin Genet Dev **12**(2): 210-8.
Sims, R. J., 3rd, R. Belotserkovskaya, et al. (2004). "Elongation by RNA polymerase II: the short and long of it." Genes Dev **18**(20): 2437-68.
Soeller, W. C., C. E. Oh, et al. (1993). "Isolation of cDNAs encoding the Drosophila GAGA transcription factor." Mol Cell Biol **13**(12): 7961-70.
Sun, J., G. Watkins, et al. (2008). "Deregulation of cofactor of BRCA1 expression in breast cancer cells." J Cell Biochem **103**(6): 1798-807.
Swantek, D. and J. P. Gergen (2004). "Ftz modulates Runt-dependent activation and repression of segment-polarity gene transcription." Development **131**(10): 2281-90.
Thompson, C. M., A. J. Koleske, et al. (1993). "A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast." Cell **73**(7): 1361-75.
Tracey, W. D., Jr., X. Ning, et al. (2000). "Quantitative analysis of gene function in the Drosophila embryo." Genetics **154**(1): 273-84.
Wada, T., T. Takagi, et al. (1998). "DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs." Genes Dev **12**(3): 343-56.
Wada, T., T. Takagi, et al. (1998). "Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro." EMBO J **17**(24): 7395-403.
Wang, X., C. Lee, et al. (2007). "Transcription elongation controls cell fate specification in the Drosophila embryo." Genes Dev **21**(9): 1031-6.
Weber, J. A., D. J. Taxman, et al. (1997). "Molecular architecture of the hsp70 promoter after deletion of the TATA box or the upstream regulation region." Mol Cell Biol **17**(7): 3799-808.
Weiner, A. J., M. P. Scott, et al. (1984). "A molecular analysis of fushi tarazu, a gene in Drosophila melanogaster that encodes a product affecting embryonic segment number and cell fate." Cell **37**(3): 843-51.
Westwood, J. T., J. Clos, et al. (1991). "Stress-induced oligomerization and chromosomal relocalization of heat-shock factor." Nature **353**(6347): 822-7.
Wheeler, J. C., C. VanderZwan, et al. (2002). "Distinct in vivo requirements for establishment versus maintenance of transcriptional repression." Nat Genet **32**(1): 206-10.
Woodcock, C. L. (2006). "Chromatin architecture." Curr Opin Struct Biol **16**(2): 213-20.
Woodcock, C. L., A. I. Skoultchi, et al. (2006). "Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length." Chromosome Res **14**(1): 17-25.

Wright, T. J., J. L. Costa, et al. (1999). "Comparative analysis of a novel gene from the Wolf-Hirschhorn/Pitt-Rogers-Danks syndrome critical region." Genomics **59**(2): 203-12.

Wu, C. H., C. Lee, et al. (2005). "Molecular characterization of Drosophila NELF." Nucleic Acids Res **33**(4): 1269-79.

Wu, C. H., Y. Yamaguchi, et al. (2003). "NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila." Genes Dev **17**(11): 1402-14.

Yamada, T., Y. Yamaguchi, et al. (2006). "P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation." Mol Cell **21**(2): 227-37.

Yamaguchi, Y., J. Filipovska, et al. (2001). "Stimulation of RNA polymerase II elongation by hepatitis delta antigen." Science **293**(5527): 124-7.

Yamaguchi, Y., N. Inukai, et al. (2002). "Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA." Mol Cell Biol **22**(9): 2918-27.

Yamaguchi, Y., T. Takagi, et al. (1999). "NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation." Cell **97**(1): 41-51.

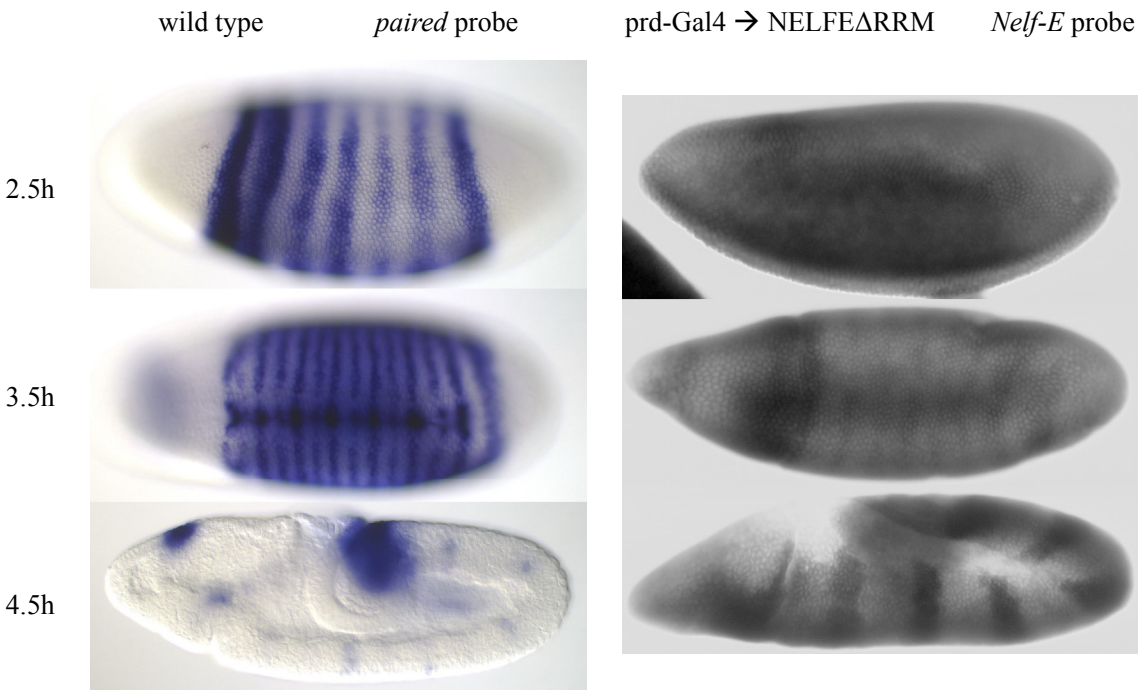
Zeitlinger, J., A. Stark, et al. (2007). "RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo." Nat Genet **39**(12): 1512-6.

Zhang, F., M. Barboric, et al. (2003). "A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb." Genes Dev **17**(6): 748-58.

Zhou, Q. and J. H. Yik (2006). "The Yin and Yang of P-TEFb regulation: implications for human immunodeficiency virus gene expression and global control of cell growth and differentiation." Microbiol Mol Biol Rev **70**(3): 646-59.

Appendix

Figure 5-1 NELF-EΔRRM is induced by Gal4/UAS system with 20min to 1h delay



To confirm we were able to induce the expression of NELF-EΔRRM in embryo, a cross between paired-Gal4 with one NELF-EΔRRM line was set up and embryos were collected for *in situ* hybridization. NELF-EΔRRM RNA is expressed in paired pattern as expected (right column). There is 20 min to 1h expression delay compared with wild type paired pattern (left column). This is due to the lag of Gal4 expression, translation and transcriptional regulation.

Table 5-1 Inverse PCR results on random UAS-NELF-EΔRRM insertion

LINES	lethality	insertion	position	polytene	meiotic	gene	orientation	iPCR RE
60-1		single	X:1,103,892	1E4	1-0.3	upstream of *CG11382*	right	HpaII
2-1		double	X:2,536,740	3A8	1-1	intron of *sgg*	right	HpaII
			X:2,579,461	3B1	1-1.2	upstream of *per*	right	HpaII
64-3		single	X:6,737,282	6D5	1-18	exon of *CG14435*	inverse	HpaII
65-1		single	X:6,594,535	6C7	1-18	downsteam of *CG32741*	inverse	HpaII
28-3		single	X:8,136,542	7D22	1-22	upsteam of *Trxr-1*	inverse	HpaII
55-2	lethal	single	X:15,527,660	13E3	1-51		inverse	HpaII
2-3		single	X:19,399,955	18C8	1-62.2	upstream of Pfrx	inverse	HpaII
93-1		single	2L:9,449,417	30B4	2-32	intron of *numb*	inverse	HpaII
40-4		single	2L:15,264,832	35C5	2-51	upstream of *yuri*	inverse	HpaII
28-5		single	2R:4,121,229	44C1	2-58	upstream of *pnut*	right	HpaII
38-2		single	2R:4,543,134	44D5	2-58	1st intron of *ptc*	right	HpaII
55-1		single?	2R:4,214,727	44C4	2-58		right	HpaII
71-3C		single	2R:6,363,010	47A10	2-60	intron of *CPTI*	right	HpaII
38-5	sterile	single	2R:6,462,610	47A13	2-60	2nd intron of *psq*	inverse	HpaII
40-2		single	2R:7,284,962	47F8	2-62	upstream of *sprt*	right	HpaII
36-1		single	2R:15,998,939	56F8	2-90	upstream of *18w*	inverse	HpaII
36-2		>=2	3L:267,876	61B3	3-0	5'UTR of *miple*	inverse	HpaII
			3R:4,495,305	85A5	3-48.3	exon of *CG8036*	right	
40-1		single	3L:699,171	61C8	3-0	upstream of *CG13894*	right	
16-3		** ^	3L:14,798,236	70F7	3-42	upstream of *CG17839*	right	HpaII
36-5		single	3L:16,893,877	73D5	3-44	intron of *Lmpt*	inverse	HpaII

14-1		single	3L:20,191,343	77A1	3-46	upstream of *CG32223*	inverse	HpaII
50-1	lethal	** ^	3L:21,627,517	78E1	3-47	upsteam of *MED1*	right	HpaII
72-1		single	3R:4,983,822	85C11	3-48.5	intron of *pum*	inverse	HpaII
56-1		single	3R:6,090,013	86A3	3-49	upstream of *CG11870*	inverse	HpaII
64-4			3R:10,924,467	88E1	3-60	exon of CG31304	inverse	HpaII
71-1	lethal	single	3R:19,607,537	95B5	3-81	5'UTR of *mbc*	inverse	HpaII
71-3B		single	3R:20,369,628	96A7	3-85	upstream of *CG5789*	inverse	HpaII
16-2	semi-lethal		3R: 27,513,541	100D1	3-105		right	Sau3AI

Table 5-2 Genes affected by maternal NELF-EΔRRM6B are revealed by microarray.

gene	Experiment1 fold change	Experiment2 fold change	Description
CG5994 (NELF-E)	2.3	2.3	
CG8036	2.3	1.9	INSERTION
CG7920	2	2.6	acetyl-CoA metabolic process
CG40274	1.9	1.4	
18S rRNA(710-1268)	1.9	5.7	
CG13512	1.7	1.4	
CG7037 (cbl)	1.6	2.3	
CG6171	1.5	2.6	
CG4949	1.4	1.2	
CG17531 (GstE7)	1.3	1.2	
CG31607	1.3	1.4	having GAP-like pattern
CG9953	1.3	1.2	proteolysis
CG6137(aub)	1.2	2.3	
CG5880	1.2	1.2	
CG11777	1.2	1.3	
CG12891 (CPT1)	-1.2	-1.2	INSERTION
CG8896 (18w)	-1.2	-1.4	INSERTION
CG3838	-1.2	-1.3	
CG18076	-1.2	-1.5	
CG17962	-1.2	-1.2	
CG18543	-1.3	-1.5	
CG32491 (mod)	-1.3	-1.3	
CG7855 (timeout)	-1.4	-1.5	
CG9027	-1.6	-1.4	
CG8588 (pst)	-4.6	-4.9	learning and memory

Table 5-3 Viability tests used to determine the strength of 5 UAS-eve lines.

♀ NGT line (LacZ activity) \ ♂ uas-eve	$\frac{uas-eve^{23-2}}{TM3}$	$\frac{uas-eve^{71}}{TM3}$	$\frac{uas-eve^{2-1}}{Cyo}$	$\frac{uas-eve^{86}}{TM3}$	$\frac{uas-eve^{12}}{Cyo}$
	Sb+/Sb	Sb+/Sb	Cy+/Cy	Sb+/Sb	Cy+/Cy
NGT31/II ($307\pm91\times10^3$)	154%(72)	136%(70)	96%(87)	72%(72)	71%(113)
NGT40/II ($603\pm78\times10^3$)	136%(69)	34%(97)	18%(126)	11%(119)	0.63%(159)
NGT31/NGT31 ($780\pm104\times10^3$)	51%(51)	10%(98)	4.4%(137)	2.4%(85)	0%(82)
NGT40/NGT40 ($1776\pm212\times10^3$)	1.1%(88)	0%(31)	0%(93)	0%(107)	0%(122)

The numbers report the percentage viability of *UAS–eve* progeny relative to their stubbled, or curly winged siblings. The number of siblings recovered is given in parentheses. The relative activity of NGT lines determined by mating to UAS-lacZ is taken from (Tracey et al. 2000)

Figure 5-2 Strategy to generate clone of *ftz* mutant cells in both the male and female germ line

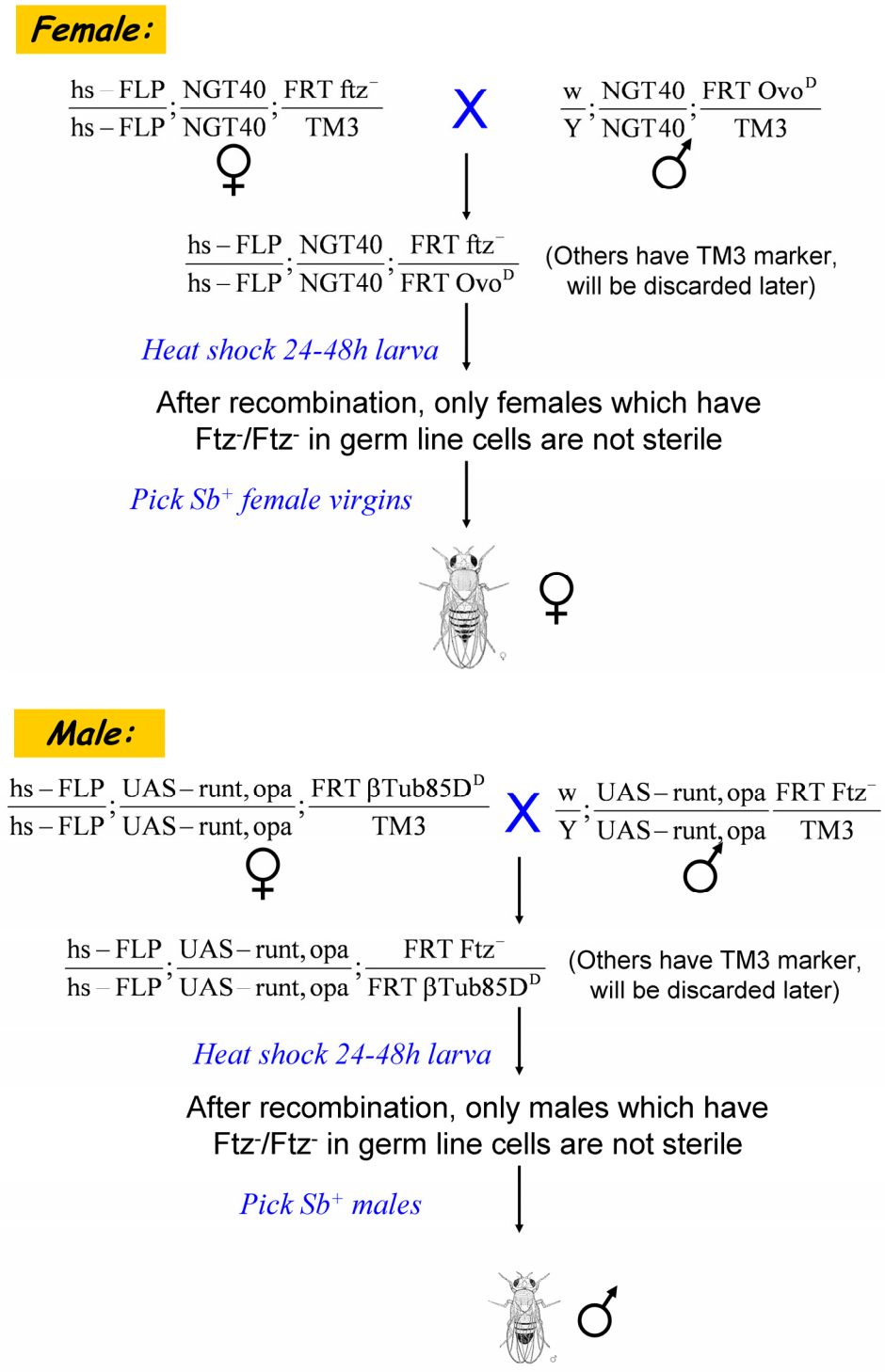


Table 5-4 Primers information

A. Primers for ChIP

Primer		Middle point *	Left sequence (5'→3')	Right sequence (5'→3')	Product size(bp)
slp1	1	-10,213	TCCTACAACTTTGCCTTCAGTTCG	TACTTGCCGCTGCTTTTGTTGTAT	156
	2	-852	GAAACTAGAACAAATCCCGGTGCT	CCAAACTAGTTCAAGGGTCGCTTT	169
	3	-402	TGTAAATTACTCTCGGATCCCAAT	CGTCAAAATATCCATTACATCCAA	211
	4	-6	TGTGGGCTCTCTTCGTGTAGACTT	TGGTTTCTGATTCTCACACACGAC	151
	5	+124	ATCCCAAGTCGTGTGTGAGAATCA	GTTCGGTTTTGATGGGTTGAGTTG	167
	6	+504	TCCCAGCAAGAAACAGAAAATGAC	TTGGCCTTGAAGTAGGGGAATCTA	165
	7	+717	CCTGAACAAGTGCTTCACGAAGAT	GCGAGAATATGGCCTGACGATAG	182
	8	+1,236	CTTGGGGAAGAGCTAGGACTGGTA	ATTCATTGCCGTTAACCATCGACT	175
hsp70a	1	-24	GTCACACAGTAAACGGCGCACT	CGCTTGTTTGTTTGCTTAGCTTT	172
	2	+149	CGCAGCTGAACAAGCTAAACAATC	TTCCTCGGTAACGACTTGTTGAAA	176
	3	+911	CGCAATGTGCTTATCTTCGACTT	CTTCTTGTACTTGCGCTTGAACTC	180
wg	1	-27	GTATGCGATCCGAAGACGAAGAC	GAGCGGAGGAGTGAAGATTGTT	150
	2	+64	ACCGTGTGTTTCAGTTAAGCGTTG	TGTATCGTAAATTTGGCTGTTTCG	151
	3	+775	CAGCAACAAACCTACAAAGCAAAC	CGAGAAGATGAACGAAGTTGAGAA	152
en	1	-46	CTCTCTTTCCACCGTGACAGTTC	GGCTTATGCTTTCGTGTTTGATTT	160
	2	+67	CTGCGATTCCGAAGTAGTCAAC	CTGAGCCACTGATTCTTCTGATTG	169
	3	+672	AGTTTCCACAATCAGACGCACAC	AGGATGTTGGAGATGGAAAAGG	170
bro		+490	GGCAAAGTTCATTTGCGTTCTAGT	GCTGGTTGTAGCTTTGAATCTGCT	169
rp49		N/A	TACAGGCCCAAGATCGTGAA	ATCTCCTTGCGCTTCTTGGA	311
sry-α		+11	CCCCAAAATAGGTCCTTTCTACCT	CAGCTCACTGCAAGTGTGTAATTG	157
eve		+126	GAGCGCAGCGGTATAAAAGGG	CAACATGGAGAGCCACCATG	179
ftz	1	-12	AAGTGTCATGCGCAGGGATATTTA	GGCCATATCGGATGTGTATTGCTA	177
	2	+110	GCACATCGCAGAGTTAGAGAAGAA	TAGTAGGCATTGCTGCCTGAATTA	167

B. Primers for NELF RT-PCR

Primer	Start*	Left Sequence (5'→3')	Start*	Right Sequence (5'→3')	Prod (bp)	w/o Intron
NELF-A	+4601	AGCAAGCTACGCAACAGCAATC	+4858	GACCAGCACGTTTTCGTTATTTG	257	194
NELF-B	+1519	GCATATGAGGATCCCTTCCTG	+1730	CAGCATGTGACGAGAGGTCA	212	150
NELF-D	+909	AGCGGCTCTCTCAGGAAATTATTA	+1132	ATATCGGCGGGATTCAGTGTATT	224	154
NELF-E	+577	CGACAGAATCTCTACCAACACTTC	+1036	GACCAGACTGCCGACGAG	460	371

* The transcription start site for each gene is set at +1.

C. Primers for R-LMPCR

slp1+39: ATCCCAAGTCGTGTGTGAGAATCA
slp1+13: GCATCTCAAGAGCAAACATCTCAA
slp1+8: AGCTAGCATCTCAAGAGCAAACAT
5phos_linkerA: [Phos]GCGGTGATTTAAAAGATCTGAATTC
Reversed_linkerA: GAATTCAGATCTTTTAAATCACCGC