Studying primary pair-rule gene eve in the early development of Drosophila by using Zelda as approach.

Muen Liu

Abstract

This paper mainly studies the primary pair-rule gene Even-skipped to reveal its position in the early transcription network established in the Drosophila embryo and gain a deeper view of the network. We used the genome activator Zelda as an approach. We found Zelda's role in regulating Eve by committing CHIP- seq towards RNA polymerase and Zelda protein in wt and Zld- embryo. With the PWM matrix, we know transcription factors that might bind to eve. By further researching its TFs, we find the relationship between timing and pattern of expression of these genes involved in the early transcriptional network, while Zld is still our approach. Also, by looking into the specific gene, we find the possible timing mechanism on Eve based on the accumulation of earlier transcribed transcription factors. The role of maternally loaded transcription factors in early development is once again shown by affecting the timing and patterning of primary patterning transcription factors.

1. Introdcution

As the genome activator Zld being discovered, bringing a new view to the transcriptional network of the early developing Drosophila, the role of Zld in regulating early expressed gene is revealed. Still its effect on particular aspects of the transcriptional network is yet to be further looking at. In this work we studied a specific Zld- targeted gene even-skipped as well as the transcriptional network involving early pair-rule genes and gap genes that appears in the early stage of the Drosophila embryo by using Zelda as a approach.

2. Result

2.1 Zld target eve and affect its expression

We compared the difference in expression of eve between wildtype embryos and Zld-knock down embryos by doing CHIP-seq on RNA polymerase II during NC13 and found that the RNA polymerase binding on eve gene is reduced in Zld-knockdown embryos (Fig1A). This means the expression of eve is to some extend affected by Zld

Our Zld-CHIP-seq data shows that there are Zld binding peaks right upstream the TSS of eve within 500 bp as well as regions that are downstream or upstream the gene. This shows eve is a Zld-target gene and these bindings indicate Zld's role in regulating the gene.

Also, since the Zld-bound region has a remarkable overlap with HOT region and Zld-bound peaks were seen over well-defined enhancers [1], we deduced that Zld bound region we found on the eve gene, is the potential enhancer position.

2.2 Zelda affects the timing and pattern of eve

From the CHIP-seq data we found that although the eve's expression is reduced in the Zld- embryos, it is still expressed to a certain extend

Eve exhibits a pair-rule pattern in the early stage of Drosophila embryo (Fig 1B). Comparing the expression pattern of eve in Zld- embryo and wt embryo, revealed that expression of eve is delayed for two nuclear cycle from NC10 to NC 12 and the stripe pattern has been disrupted during NC13 and NC14 in the Zld- embryo. We know that here eve isn't fully dependent on Zld for activation, but regulated by Zld in a different way, it dependent on Zld for proper timing of expression [1]. And the timing seems to affect the pattern too.

2.3 Other factors that affect eve

We next investigate the other transcriptional factors because Zld might not be the only factor that regulate eve. Knowing Eve's expression is modified by gap-gene mutations, [2].We use Position Weight Matrix of every known transcription factors of Drosophila melanogaster to scan the Zld-bound region we have on the eve gene which represent the potential enhancer region, we found remarkable number of high score binding sites of many other TFs besides Zld along the enhancer region, among which Cad, Kr, Hb, are most commonly found (Fig2), they are all expressed ubiquitously or gaped in the embryo and around the expression time of Eve.

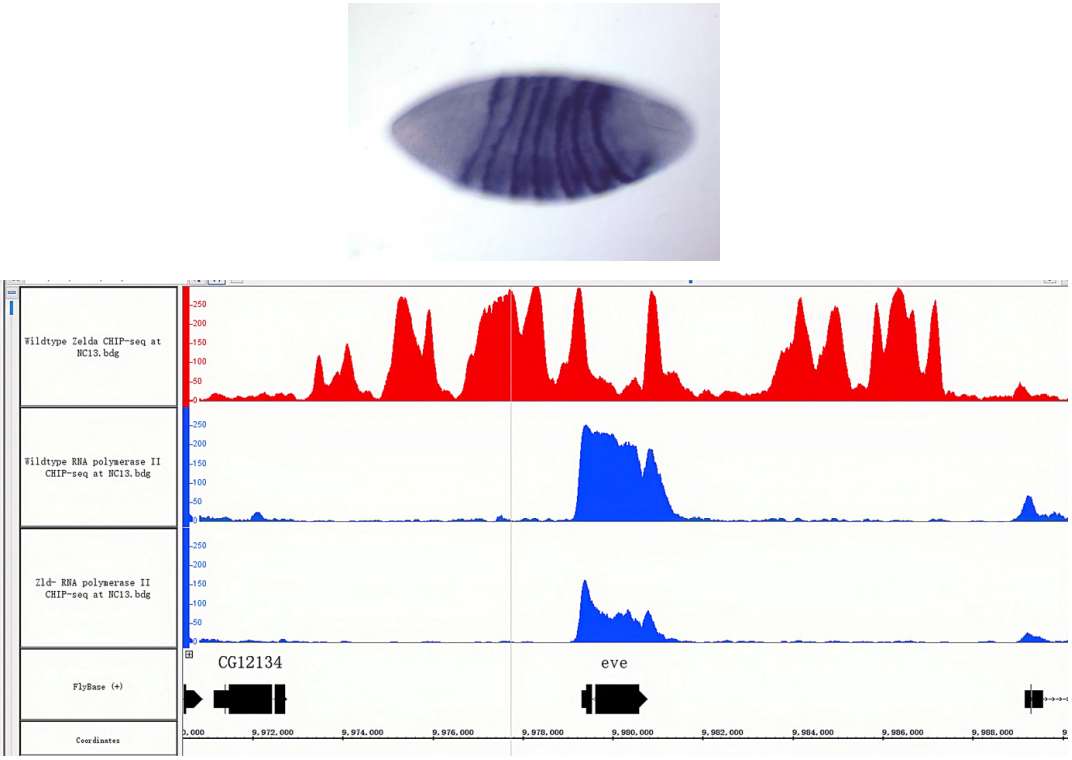
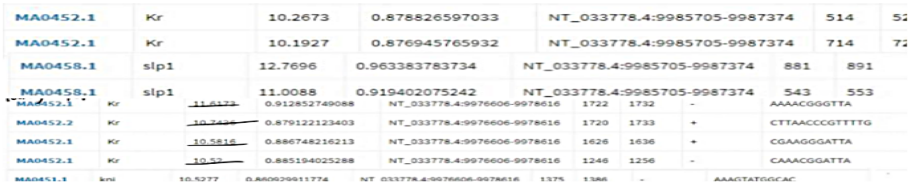


Fig 1. A: CHIP-seq data of RNA polymerase II(blue) and Zld bind in eve(red) B:pattern of eve shows 7 stripes across anterior-posterior axis.



MA0452.1	Kr	10.2673	0.878826597033	NT_033778.4:9985705-9987374	514	5
MA0452.1	Kr	10.1927	0.876945765932	NT_033778.4:9985705-9987374	714	7
MA0458.1	slp1	12.7696	0.963383783734	NT_033778.4:9985705-9987374	881	891
MA0458.1	slp1	11.0088	0.919402075242	NT_033778.4:9985705-9987374	543	553
MA0452.1	Kr	11.6173	0.912852749088	NT_033778.4:9976606-9978616	1722	1732	-	AAAACGGGTTA
MA0452.2	Kr	10.7426	0.879122123403	NT_033778.4:9976606-9978616	1720	1733	+	CTTAACCCGTTTTG
MA0452.1	Kr	10.5816	0.886748216213	NT_033778.4:9976606-9978616	1626	1636	+	CGAAGGGATTA
MA0452.1	Kr	10.52	0.885194025288	NT_033778.4:9976606-9978616	1246	1256	-	CAAACGGATTA
MA0451.1	kni	10.5277	0.860929911774	NT_033778.4:9976606-9978616	1375	1386	-	AAAGTATGGCAC

Fig.2: PWM scores obtain: Scanning on the eve's enhancer region. Higher score of binding shown higher possible of binding. We found multiple binding site Kr, cad and hb indicating their regulatory effect on Eve [3]

Kr is a zinc-finger repressor [4] first seen in fruit fly embryo in NC12 and perform a gap expression (Fig3A). Kr's expression is also seen delayed in the Zld- embryo [1].

Cad expresses ubiquitously in the early stage of embryo (Fig1D) Alongside Bicoid (bcd), caudal forms concentration gradients down the anterior-posterior (A-P) axis providing positional information and subsequent induction of the gap genes. Acts as a key regulator of the Hox gene network and activates transcription via the downstream core promoter element (DPE) relative to the TATA box [5]. The comparison of its RNA polymerase CHIP-seq in wt and Zld- embryo also shows that without Zld's activation Cad's expression is reduced in an significant amount (Fig3B).

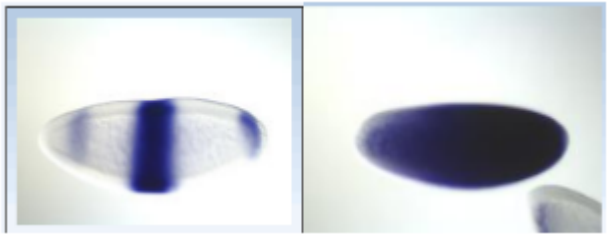


Fig.3 Expression pattern of Kr show gapped pattern.(left) Expression of cad shown ubiquitously with a gradient. (right) *[6]*

Ubiquitous gene Cad and gap gene Kr are both Zld-target genes and their binding to primary pair-rule gene Eve may give a hint to how Zld affect the timing and patterning of the early expressed pair-rule genes via the transcriptional network.

We also use PWM matrix to scan two other early-expressed pair-rule genes odd and ftz, and also sited multiple Cad site and Kr site meaning they might be regulated by Cad and Kr as well as Zelda. Note that these three early pair-rule genes ftz odd and eve kinds of interact in a way as another part of transcriptional network.

3. Discussion:

3.1 Relationship between the timing and patterning.

Eve's product is first seen in NC 10 and its repressor Kr is not observed until NC 11-12 [6]. And in wild-type embryos, eve, ftz, hairy, and runt were initially expressed in broad domains as early as nc 10, which refine into the respective seven-stripe patterns by the end of nc 14 [7]. We know pair-rule stripes are formed by localized gap repressors acting on stripe enhancers [8]. Stripes of eve may be an result of repression from Kr and other interacting gap gene on specific location of eve's expression in the embryo and setting up the boundaries of eve's stripe according to their pattern.

In the Zld- embryo, the expression of Kr is delayed [1] while the Eve's expression is also delayed. Both of their pattern are disrupted with the delay, so as other gap genes. The delayed Kr and Eve may be a cause of the failure in the formation of boundaries of eve stripes in the NC14 that gap repressors haven't formed their pattern for eve

And the disrupted timing of expression may also be the reason why the pattern of gap genes is disrupted too in the Zld-. Gap genes are expressed at different stage of embryo gt genes is first seen in nc 10 [1] while Kr is first observed at nc 11-12, evidence show that gap genes interact with each and form complementary or overlapping pattern. Hb and Gt set the anterior Kr border, while Hb and Kr establish the anterior border of the posterior gt domain. Gt and Kr is known to establish their complementary domains [9].

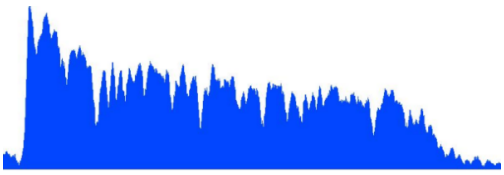
The idea of maternally loaded transcription factor determines the position, spread and pattern of expression of gap genes and primary pair-rule genes by interaction can't work cause almost every maternal transcription factors are activators and express ubiquitously which make them not capable to guide the forming of irregular shaped expression pattern. Also, the interaction between gap genes themselves must be considered that their interaction within the network is involved in the patterning.

Timing maybe a significant factor here. Timing maybe controlled by maternal loaded transcription factors which we will talk about later. The relationship between the timing and pattern is so delicate that a delay can cause the disrupted pattern rather than just delayed pattern. We consider this as a result of the interaction of gap genes within their class. The changing timetable causes the changes in patterning via a "butterfly effect" based on

the interaction of these transcription factors. Due to the complicate regulatory effect of these transcription factors have on each other in the network, it is clear to say that a single change might cause a big difference. Timing here controls the extend of these interaction effect by determine the amount of a transcription factors via its accumulation. and also at some point the time order of genes being transcribed might also contribute to this specificity of interaction such as determine which genes will be dominant in certain region while others will be controlled. At first the time order decides the basic pattern or potential pattern in this way, and in further on nuclear cycle the interaction take over its role and forces the pattern develop from its basic form to the final complementary and overlapping pattern we see in the cellularization stage while determine the primary cell fate through control over the further on network with its unique expression pattern.
The relationship between timing and patterning is shown here.

3.2 Timing Eve via Cad's accumulation

Our result of PWM matrix shown that eve has many high affinity Cad binding sites which indicate Cad's role in activate eve.
There is a possible mechanism of timing, which is provided by the intrinsic properties of the regulatory motifs established by Zld [1], a delay in the activation of the third gene in the loop occurs because of its dependence on accumulation of the second gene product [10-12].
In the case of eve, Cad is expressed ubiquitously in the 1-hour embryo. Despite its maternal loaded parts, Cad is also transcribed in the 0-2 hr embryo, due to its rapidly disappearance in the stage 4-6 embryo during which most of genes we are studying appeared ,we think the expression at this time maybe is for the supplement of Cad proteins and maintaining a concentration gradient for a short time.
The concentration gradient maybe affecting the activation of eve and other genes, it might also be the reason of timing. Cad is shown down-regulated in Zld- embryo significantly during stage 4-6 (Fig 3B).
The expression of eve maybe dependent on the accumulation of Cad transcription factors. The decrease in amount of Cad during the stage it express might related to the delay it experience.



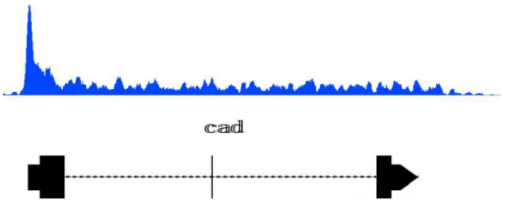


Fig.4: CHIP-seq data of RNA polymerase II in Cad (Zld- lower, wt upper). This indicates that Cad is down-regulated in Zld- embryo

Further on research need to be done on finding if Cad accumulation affect the timing of Eve directly or it still depend on other factors in the network. But, surely ubiquitous gene which Cad is, is most close to the primary pair-rule gene in segmentation hierarchy time order.
Another hypothesis is that Zld, Cad and Bcd activate Eve together while the absent of Zld slower the process of binding of RNA polymerase. But it can't explain what role the reduction of CAD expression takes in the delaying Eve. Again further research need to be done focusing on the timing mechanism maternally loaded factors have on zygotic genes.

4. Methods

4.1 CHIP-seq

We crossed link Zelda protein to its target DNA sequence by adding formaldehyde. Sheared by sonication to break chromatin into small pieces and add antibody specific to Zelda protein to bind the protein forming complex that heavier than other substances and then precipitate and purify. Then separate the DNA from the protein and sequence the DNA where Zld binds and exhibit it on the browser. Same procedure was used on RNA polymerase II.

4.2 PWM matrix

Base on the database from http://jaspar.genereg.net/faq/. The matrix is based on the data of CHIP-seq, and represent the frequency of each base appeared on an certain position of the protein bound sequence. During scanning, the Jasper website points out the potential binding site of TFs on a given DNA sequence and calculate the binding score of the sequence base on the frequency matrix

5. Conclusion

Eve is a primary pair-rule gene that involved in the segmentation of Drosophila embryo. The study of its position in the early transcriptional network
Maternal loaded genes such as Zelda acting as key activators on not only genes but also to the whole network once again show its role in the early development by

effecting timing and pattern of expression of early-transcribed transcription factors.

From the result we know that maternally morphogen such as Zelda, Bicoidand Cad interact during the earliest stage of the embryo development, which determine the basic timing of expression of the zygotic transcription factors. The delayed in the transcription of gap gene and primary pair rule gene along with the disruption in their pattern may also show that the timing and patterning of the expression of these interlinked gene has a kind of relationship that based on the repressing or promoting effect their coded transcription factors have on each other. These maybe a explanation to how primary patterning of transcription factors is established from the earliest embryo. Further research and experiment must be taken for verification.

Reference

[1] Surkova, Y. , E V, Golubkova, L A, Mamon, & M G, et al. (2018). [morphogenetic networks which determine the spatial expression of zygotic genes in early drosophila embryo]. Ontogenez.

[2] Abbie, S. , Core, L. J. , Catherine, S. , Lis, J. T. , & Ashe, H. L. . (2019). Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription.

[3] Aure, M. R. , Fleischer, T. , Bjorklund, S. , Ankill, J. , Castro-Mondragon, J. A. , & Borresen-Dale, A. L. , et al. (2020). Crosstalk between microRNA expression and DNA methylation drive the hormone-dependent phenotype of breast cancer. Cold Spring Harbor Laboratory.

[4] Nirusimhan, V. , Gideon, D. A. , Parashar, A. , Jeyachandran, S. , Jeyaraman, J. , & Subbaraj, G. , et al. (2022). Structural modeling of drosophila melanogaster gut cytochrome p450s and docking comparison of fruit fly gut and human cytochrome p450s. Current drug metabolism (4), 23.

[5] Caccuri, F. , Messali, S. , Zani, A. , Campisi, G. , Giovanetti, M. , & Zanussi, S. , et al. (2022). Hiv-1 mutants expressing b cell clonogenic matrix protein p17 variants are increasing their prevalence worldwide. Proceedings of the National Academy of Sciences of the United States of America. (27), 119.

[6] Tomancak P, Beaton A, Weiszmann R, Kwan E, Shu S, Lewis SE, et al. (2002). Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol.;3(12):RESEARCH0088. PMCID: 151190Retrieved fromhttps://insitu.fruitfly.org/

[7] Sureka, R. , Avvaru, A. K. , Sowpati, D. T. , Pathak, R. U. , & Mishra, R. K. . (2022). Structural and developmental dynamics of matrix associated regions indrosophila melanogastergenome. BMC Genomics, 23(1). [8] Stanojevic D, Small S, Levine M (1991) Regulation of a segmentation stripe by overlapping activators and repressors in the Drosophila embryo. Science 254: 1385–1387. (cited in Nien, Liang Plos Gen 2011Temporal Coordination of Gene Networks by Zelda in the Early Drosophila Embryo)

[9] Kraut R, Levine M (1991) Mutually repressive interactions between the gap genes giant and Kruppel define middle body regions of the Drosophila embryo. Development 111: 611–621. (cited in Nien,Liang Plos Gen 2011Temporal Coordination of Gene Networks by Zelda in the Early Drosophila embryo

[10] Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet 31: 64–68. (cited inNien,Liang Plos Gen 2011Temporal Coordination of Gene Networks by Zelda in the Early Drosophila Embryo)

[11] Mangan S, Alon U (2003) Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci U S A 100: 11980–11985.. (cited inNien,Liang Plos Gen 2011Temporal Coordination of Gene Networks by Zelda in the Early Drosophila Embryo)

[12] Davydov, A. , Jafarpour, S. , Abate, M. , Bullo, F. , & Coogan, S. . (2022). Comparative analysis of interval reachability for robust implicit and feedforward neural networks.