International Core Journal of Engineering Volume 7 Issue 5, 2021
ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0074

Conservation of Zelda's Binding Sites on SisterlessA and Wingless Genes

Hongrui Feng

Beijing National Day School, Beijing, 100039, China.

Abstract

Zelda, a transcriptional activator of Drosophila melanogaster, initiates the transformation from utilizing maternal deposit to individual's DNA sequence. Due to the importance of zygote's future development, the DNA binding site for Zelda protein is assumed as highly conserved in all subspecies of fruit flies. In this research, I looked into the details of two specific genes: sisA and wg, to validate the hypothesis and find the conservation pattern of Zelda binding site.

Keywords

Zelda; Transcriptional Activator; Conservation; Zelda Binding Site.

1. Introduction

Delayed transcriptional activation of the zygotic genome is the commonest phenomenon in the fruit fly. After fertilization, the development of zygote is controlled by maternal deposit, in this case, maternal protein and mRNA. However, activating the zygote's genome is vital and instrumental for an individual organism to develop. Otherwise, the consequence could be lethal for failure in such a conversion from relying on maternally deposited products to translating its DNA sequence. Even though the transition stage mechanism is still unclear, Zelda (ZLD) has been shown to play a significant role as a transcriptional activator in Drosophila melanogaster. [1] Research had shown that younger embryo had high levels of maternal zld transcripts, indicating that maternally loaded zld transcripts are degraded during cellularization and replaced with zygotic zld. [2] Distinct activation domains (involves Zelda sequence and other DNA sequences) have been identified and uploaded to a certain database.

2. Research procedure

Firstly, the basic dataset was downloaded in a gff and bdg file type, which the IGB browser could access. Before reading the specific information from the file, correspondent species and genome version should be selected. IGB browser only provided a general overview of the enzyme activity relating to the translation of the Zelda sequence among the different cycles of zygotic development. Utilizer could only infer whether the Zelda sequence exists on the gene and be translated during the process. Specific information includes the direction of the strand, length of the gene, starting bp and ending bp.

UCSC genome browser on D. melanogaster offers a clearer presentation on the ChIP-Seq of Zelda transcription factor and the Zelda peak. Zelda protein is supposed to bind at the enhancer site because Zelda protein is a transcriptional factor that boosts the zygotic DNA sequence transcription. UCSC genome browser allows the utilizer to have a closer view of the enhancer sites and check whether the IGB browser's dataset is up to date. Otherwise, unnecessary time is spent on insignificant exploration. Unlike the IGB browser, UCSC doesn't require to upload a file. Entering a specific position or gene term allows the utilizer to quickly view and check the Zelda peaks of the fly's gene.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0074

Thirdly, NCBI and Jaspar websites allow me to find out all the transcriptional motifs of two genes. The NCBI database enables the utilizers to access the gene's sequence, specific to a single nucleotide. FASTA form of the gene sequence is exported and pasted on the Jaspar website. Jaspar could create an excel document to show all the transcriptional motifs and relative scores. Relative scores infer to the activity or affinity of each transcriptional motifs. vfl motif is correspondent to the Zelda binding site. A document focusing on the Zelda binding site could be selected from all transcriptional motifs.

Fourthly, we utilized the UCSC genome browser again to check the Zelda sequence's conservation pattern on the genes. The dense black region means highly-conserved. As we can see from the figure, there are many dense black regions, but they are not necessarily all Zelda sequence. In this research, we were zooming in each piece of the gene's enhancer site: sisA and wg, marking out the Zelda sequence based on the document downloaded from the Jasper website. The different strands are marked out differently. A reverse (from right to left) and supplementary sequence should be found if the strand is positive. The negative-strand could be discovered the exact sequence from left to right.

Last but not least, we utilize Flyexpress website to check the skeletal development of the zygote of fruit flies before stage 8. The lean pattern presented on the website involves two categories: dorsal and ventral. Flyexpress help us to visualize the skeleton correspondent to the gene we studied on.

3. Conclusion

For sisA, a 779 bp long gene densely conserves in the coding region. However, for wg, a 9107 bp long gene is highly conserved instead of densely conserved. The conserved regions of wg are loosely packed compared to that of sisA. sisA has 7 vlf motifs that could be found on the right-hand side of the gene because sisA is a negative-strand gene. The right-hand side of sisA is precisely the place of enhancer and promoter, which are the place Zelda is supposed to bind.

On the contrary, wg only has 3 vfl motifs, and all three motifs are found inside the coding region. This phenomenon is in contrast with the theory that suggested above, in this case, wg is a positive-strand gene, so Zelda motifs should be found on the left-hand side of the TSS site of wg. Higher conserved regions of wg have one thing in common: TATA boxes. Most conserved sequences are TATA boxes for wg, which is another excellent point to notice.

4. Reflection

As I did all the research, I found out that some of the Zelda motifs are not highly conserved in all fruit fly species. This theory is in contrast with what I expected or the hypothesis that I give out at the beginning of the research. Zelda motifs are transcriptional activator that only regulates the activation of transcription. [3] The following coding region of Zelda motif is the region that is more valuable to research on. In this research, I only looked into the growth pattern (skeletal pattern) that two genes expressed, which is one type of representation of the function of the coding region. Highly conserved sequences are responsible for the appearance, or something that is in common in all the species of the fruit fly. Not-conserved region only accountable for specific development within one type of organism, for example, *Drosophila melanogaster*. Zelda motifs with a higher score and higher relative score are commonly highly conserved. Zelda motifs that have a lower score and lower relative score are usually not highly conserved. [4]

References

[1] Danielle C. Hamm, Eliana R. Bondra, and Melissa M. Harrison (2015) Transcriptional Activation Is a Conserved Feature of the Early Embryonic Factor Zelda That Requires a Cluster of Four Zinc Fingers for DNA Binding and a Low-complexity Activation Domain. The Journal of Biological Chemistry VOL.290, NO.6, pp.3508–3518.

[2] Hsiao-Lan Liang, Chung-Yi Nien, Hsiao-Yun Liu, Mark M. Metzstein, Nikolai Kirov & Christine Rushlow (2008) The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature 456, 400–403.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202105_7(5).0074

[3] Dufourt, J., Trullo, A., Hunter, J. et al. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. Nat Commun 9, 5194 (2018). https://doi.org/ 10.1038/s41467-018-07613-z.

[4] Liang, HL., Nien, CY., Liu, HY. et al. The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila Supplementary Information. Nature 456, 400–403 (2008). https:// doi. org/10.1038/ nature07388.