# The Drosophila Zinc Finger Protein Aef1 Colocalizes with Enhancers and Is Involved in the Transcriptional Regulation of Numerous Genes

N. E. Vorobyeva[1], J. V. Nikolenko[2], A. N. Krasnov[1*]
[1]Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia
[2]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia
*E-mail: krasnov@genebiology.ru
Received November 07, 2024; in final form, March 06, 2025
DOI: 10.32607/actanaturae.27556


**ABSTRACT** **In our previous studies, we demonstrated that the Drosophila zinc finger protein Aef1 interacts with the SAGA DUB module. The Aef1 binding sites colocalize with the SAGA histone acetyltransferase complex and the dSWI/SNF chromatin remodeling complex, as well as the origin recognition complex (ORC). Aef1 predominantly localizes with the promoters of active genes (55% of all sites) and can be involved in transcriptional regulation. In this study, we showed that Aef1 binding sites in Drosophila S2 cells, located outside gene promoters, are nucleosome-depleted regions and colocalize with the SAGA, dSWI/SNF, and ORC complexes. Aef1 binding sites colocalize with the CBP protein and the H3K27Ac histone tag, which is considered to be an active enhancer mark. An RNA-Seq experiment was conducted in Drosophila S2 cells, both normal and with RNA interference targeting the Aef1 protein, to study the role played by the Aef1 protein in transcriptional regulation. The Aef1 protein was shown to affect the transcription of 342 genes, more than half of those (178 genes) containing Aef1 at their promoters or enhancers. Hence, we infer that the Aef1 protein is recruited to both promoters and enhancers and is involved, both directly and indirectly, in the regulation of the transcription of the respective genes.**
**KEYWORDS** **Aef1, SAGA, dSWI/SNF, ORC, CBP, H3K27Ac, enhancers.**
**ABBREVIATIONS** **SAGA – histone acetyltransferase complex; SWI/SNF – chromatin remodeling complex; ORC – origin recognition complex.**

## INTRODUCTION

Regulation of eukaryotic gene expression is a complex process involving several successive stages of transcription, mRNA processing, mRNA export from the nucleus, translation, and protein folding [1]. Local chromatin structure, gene location relative to functional nuclear compartments, and long-range interactions between cell regulatory elements constitute an additional level in the regulation of genetic processes in the context of the complex architecture of the eukaryotic genome in the 3D space of the cell nucleus [2–5].

The highly conserved SAGA coactivator complex, chromatin histone modification (acetylation and deubiquitination) being its main function, is composed of more than 20 protein subunits [6]. The SAGA complex subunits interact with various transcriptional activators, thereby recruiting the complex to the promoters of specific genes [7, 8]. There is a degree of synergism between the chromatin remodeling complex and chromatin modifying complex. The SAGA complex has been shown to acetylate the nucleosomes on gene promoters during transcription activation, leading to the recruitment of the dSWI/SNF chromatin remodeling complex and stimulation of its remodeling activity [9, 10]. The SWI/SNF chromatin remodeling complex and CBP/p300/Nejire acetyltransferase, which is responsible for tagging active enhancers with H3K27Ac, exhibit considerable functional action on the activation of the ecdysone-dependent genes *dhr3* and *hr4* in S2 cells [11–14]. Histone H3 acetylation (H3K27Ac) at enhancers was shown to be required for the activation of ecdysone-dependent genes [15].

The Drosophila SAGA and SWI/SNF complexes reside within various regulatory elements of the genome, including promoters, where they often colocalize with the origin recognition complex (ORC) [16]. Replication is initiated at numerous sites known as replication origin sites. The ORC composed of six subunits (ORC1–6) is recruited to the replication origins. The ORC binds specific genomic regions [17]; however, the subunits of this complex exhibit no explicit DNA sequence specificity. Therefore, a question arises: which factors are responsible for the positioning of ORCs in the genome?

Our previous studies revealed that the insulator protein Su(Hw) carrying zinc finger domains interacts with the ENY2 protein (a subunit of the SAGA complex) and recruits the SAGA, SWI/SNF, and ORC complexes to Su(Hw)-dependent Drosophila insulators, thus being simultaneously involved in transcriptional regulation and the positioning of replication origins [18–22]. A hypothesis has been put forward that there also are other proteins carrying zinc finger domains which interact with the Drosophila SAGA complex and function in a similar manner at other regulatory elements of the genome, including promoters. Further experiments have identified four additional proteins having zinc finger domains: CG9890, CG9609, Aef1 (Adult enhancer factor 1), and CG10543 [23–27]. These proteins also colocalize with the SAGA, ORC, and dSWI/SNF complexes at their binding sites, preferentially at active gene promoters' sites, and can be involved in transcriptional regulation. As shown previously, the Aef1 protein is recruited to the enhancers of the *adh*, *yp1*, and *yp2* genes and affects their transcription [28–30]. Our earlier study [26] revealed that RNA interference targeting the Aef1 protein affects the transcription of several genes. In order to assess the impact of Aef1 on the transcription of all the genes in Drosophila S2 cells, we conducted an RNA-Seq experiment both under normal conditions and upon RNA interference targeting the Aef1 protein.

## MATERIALS AND METHODS

### Cultivation of Schneider 2 (S2) cells. RNA interference

S2 cells were cultured in Schneider's Insect Medium (Sigma, USA) supplemented with 10% fetal bovine serum (HyClone, USA) at 25°C. The cells were transfected using the Effectene Transfection Reagent (Qiagen, USA), according to the manufacturer's protocol. Knockdown of the *Aef1* gene was performed via RNA interference according to a published protocol [22]. The dsRNA corresponding to a fragment of plasmid pBluescript II SK(-) (Stratagene, USA) was used as nonspecific control for RNA interference. The dsRNA for knockdown of the *Aef1* gene and control was synthesized using the following primers: Aef1, GAATTAATACGACTCACTATAGGGAGAATGATGCATATCAAAAGCCT and GAATTAATACGACTCACTATAGGGAGATCCGGGATGCTCGCTATGT; pBluesciptIISK(-), GAATTAATACGACTCACTATAGGGAGAGTTACATGATCCCCCATG and GAATTAATACGACTCACTATAGGGAGATTTCGCCCCGAAGAACG.

For each RNA interference experiment, 30 μg dsRNA per $10^6$ cells was used. The experiment was conducted in three replicates. RNA was extracted after 5-day incubation.

### RNA-Seq and identification of differentially expressed genes

RNA-Seq libraries were constructed using a NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs). The library quality was verified using Bioanalyzer. The libraries were sequenced on an Illumina HiSeq 2000 genome sequencing system. Raw reads in the Fastq format were aligned to the Drosophila genome dmel_r6.54 using the Hisat2 software [31]; adapters had preliminarily been removed in the Atropos software [32]. The "-a" key enabling the search for multiple alignments to be excluded from the analysis was also employed. Only the unique mapped reads were used for further work by analyzing the NH:i tag among the output data in the Hisat2 software. Differentially expressed genes were identified in the CuffDiff2 software [33].

### Enrichment analysis of protein factors at Aef1 binding sites

The ChIP-Seq profiles of the Aef1, GCN5, OSA, ORC2, H3, CBP, and H3K27Ac proteins obtained earlier [14, 24, 26, 27, 34] were used to study the colocalization of Aef1 binding sites with various protein factors. This analysis and data visualization were performed using the deepTool2 suite [35].

## RESULTS AND DISCUSSION

### The Aef1 binding sites are nucleosome-depleted regions and colocalize with the SAGA, dSWI/SNF, and ORC complexes regardless of the genomic localization

We have previously demonstrated that the Aef1 protein predominantly resides at the promoters of active genes (55% of sites) and is involved in transcriptional regulation [26]. Aef1 binding sites colocalize with the chromatin modification and chromatin remodeling complexes SAGA and dSWI/SNF, as well as with the

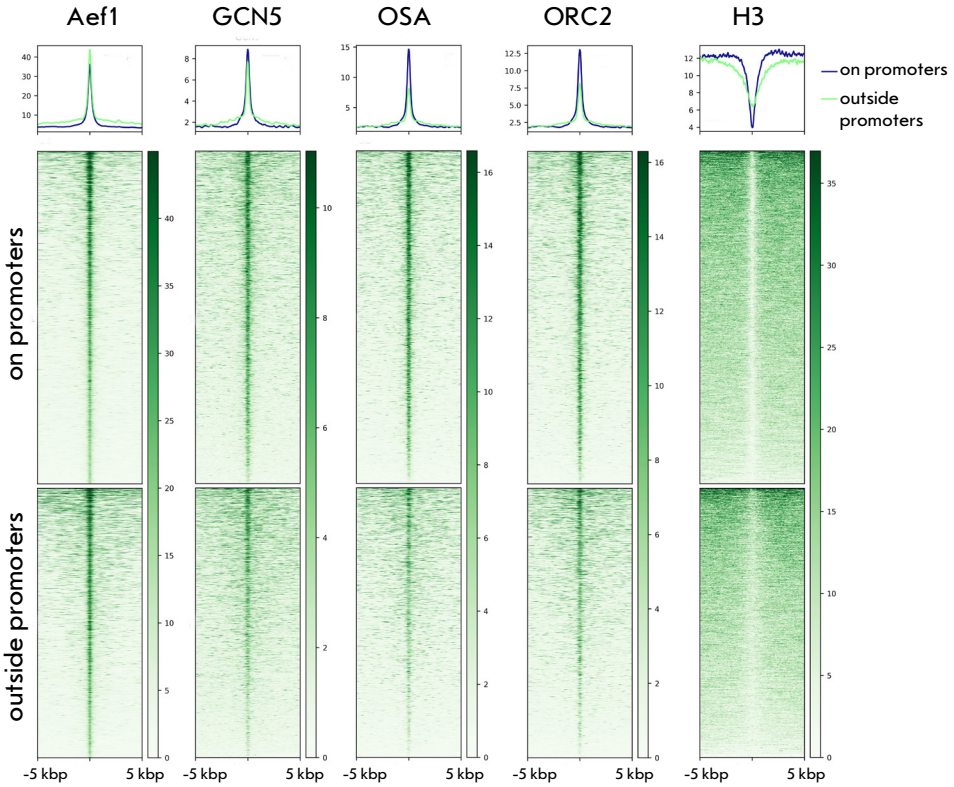


Fig. 1. Genomic colocalizations of Aef1 binding sites with GCN5 (the SAGA complex), OSA (the dSWI/SNF complex), ORC2 (the ORC complex), and histone H3. The names of the respective proteins are displayed at the top of the panels. Data are shown for Aef1 sites located in promoters (the middle panel) and outside promoters (the lower panel). The upper panel displays averaged profiles. The blue line represents the profile of proteins located at Aef1 promoter sites, while the green line represents the profile of the proteins on Aef1 sites outside promoters

ORC replication complex. A considerable portion of Aef1 sites (35%) are located within gene bodies (excluding promoters) and intergenic regions (10%). This study focuses on the properties of the binding sites residing outside promoters. The ChIP-Seq profiles of the proteins Aef1, GCN5 (the SAGA complex), OSA (the SWI/SNF complex), ORC2 (the ORC complex), and histone H3 were utilized for the analysis [14, 24, 26, 27]. We analyzed the enrichment in each of these proteins at two categories of sites: the Aef1 binding sites at promoters and outside them (*Fig. 1*). The results show that the protein complexes under study are recruited to both groups of sites with approximately equal efficiency, although the sites outside the promoters are characterized by lower levels of the OSA and ORC2 proteins. An analysis of the histone H3 distribution revealed that all the Aef1 binding sites were nucleosome-depleted regions, which is typical of active regulatory elements involved in transcriptional regulation [36].

## Aef1 binding sites colocalize with active enhancers

The Aef1 binding sites located outside promoters were analyzed to better understand the nature of the binding sites. Several studies have shown that Aef1 is recruited to enhancers of the *adh*, *yp1*, and *yp2* genes, affecting their transcription [28–30]. Histone tag H3K27Ac mediated by acetyltransferase CBP/p300/Nejire is an active enhancer mark. The chromatin remodeling complex SWI/SNF and acetyltransferase CBP/p300/Nejire are recruited to ecdysone-dependent enhancers, which is needed for transcription activation [14]. We aimed to examine the genomic colocalization of the CBP protein and histone tag H3K27Ac at Aef1 binding sites. The previously obtained ChIP-Seq profiles of CBP and H3K27Ac [14, 34] were used. *Figure 2* demonstrates that Aef1 binding sites are characterized by CBP and H3K27Ac recruitment. Therefore, it can be inferred that a significant portion of Aef1 binding sites colocalize with active enhancers.

We noticed that CBP and H3K27Ac also tag the Aef1 binding sites residing in promoters and analyzed another group of sites; namely, the promoters lacking the Aef1 protein (*Fig. 2*, lower panels). An analysis of these sites revealed low CBP and H3K27Ac levels. Hence, recruitment of CBP and the high-intensity H3K27Ac signal at Aef1-carrying promoters correlate with the presence of an Aef1 binding site rather than with the promoter in general.

We decided to identify the potential consensus motif responsible for Aef1 binding at promoters and outside them using the MEME-ChIP software. An identical consensus motif (CAA)n was identified in

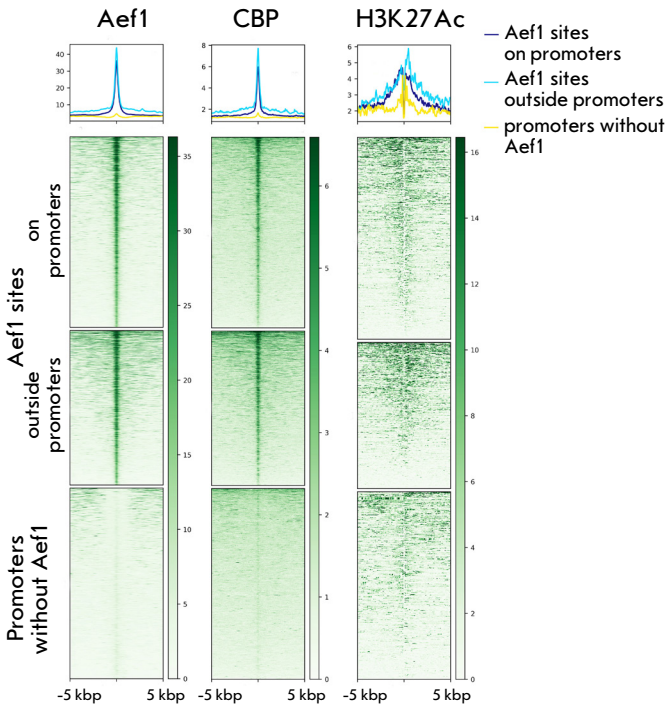


Fig. 2. Genomic colocalizations of the Aef1, CBP, and H3K27Ac proteins at three groups of sites: at Aef1 sites on promoters (upper panels); at Aef1 sites outside promoters (middle panels); and at promoters without Aef1 (lower panels). The names of the respective proteins are displayed at the top of the panels

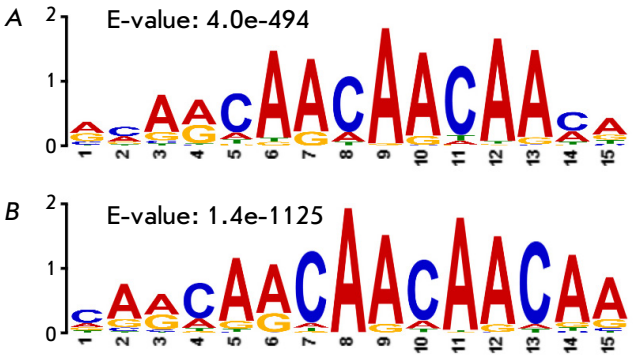


Fig. 3. The potential consensus binding motif of the Aef1 protein identified at promoter sites (*A*) and outside promoters (*B*). The E-value shows the statistical significance of the result and represents the probability of a random match

both groups of sites (*Fig. 3*), which had been identified previously across the entire set of sites [26], suggesting that both groups of sites appear owing to the DNA-binding properties of Aef1 *per se* rather than via looping of other regulatory elements. These findings are consistent with the data showing that an experimentally confirmed Aef1 binding site within the *adh* gene enhancer contains the CAACAA sequence.

### The Aef1 protein is involved in transcriptional regulation

As mentioned earlier, Aef1 binding sites are localized both within and outside gene promoter regions. Both groups of sites colocalize with active enhancer marks, indicating that the Aef1 protein could play a role in transcriptional regulation. For this purpose, we conducted a RNA-Seq analysis on Drosophila S2 cells (under normal conditions and with RNA interference targeting the Aef1 protein). The analysis identified 342 genes whose expression was significantly altered (q-value < 0.05) upon RNA interference targeting Aef1. All the genes affected by RNA interference targeting the Aef1 protein were categorized into several groups, depending on whether there were Aef1 binding sites within the gene. It turned out that 57 (17%) genes contained Aef1 exclusively in their promoters, 52 (15%) genes contained Aef1 only in potential enhancers, and 69 (20%) genes had Aef1 in both promoters and enhancers. A total of 164 (48%) genes lacked Aef1 binding sites. These findings suggest that the Aef1 protein localized in both promoters and enhancers is involved in transcriptional regulation. The results also imply that Aef1 may act either directly or indirectly, since there are no binding sites in half of the genes. It is fair to assume that there may exist looping between Aef1 binding sites (potential enhancers) and distal promoters.

## CONCLUSIONS

This study has demonstrated that the Aef1 zinc finger protein is involved in transcriptional regulation. RNA interference targeting the Aef1 protein affects the transcription of 342 genes in Drosophila S2 cells. Approximately half of these genes carry Aef1 binding sites in neither the promoter region nor the gene's body, which may be indicative of the indirect mechanisms of transcriptional regulation (e.g., via looping between enhancers and promoters). An analysis of Aef1 binding sites demonstrated that they colocalize with active enhancer marks: CBP protein and histone tag H3K27Ac. It is the general property of Aef1 binding sites that is independent of whether they reside in promoters or in the intergenic regions. Aef1-carrying promoters are much more enriched in the CBP protein and H3K27Ac compared to promoters lacking Aef1. That suggests that this property is specific to Aef1 binding sites rather than to promoters in general. It is known that several Drosophila enhancers containing Aef1 reside near transcription start sites (e.g.,

the *adh* gene enhancer) [30]. It can be hypothesized that Aef1 is a purely enhancer-associated protein, and that its localization at promoters may result from its recruitment to adjacent enhancers.

Our earlier studies [18, 19] have demonstrated that the Su(Hw) protein recruits the SAGA and dSWI/SNF complexes to its binding sites, resulting in the formation of nucleosome-depleted regions and recruitment of the ORC replication complex. Further experiments identified four additional zinc finger proteins (CG9890, CG9609, Aef1, and CG10543) colocalized with the SAGA, SWI/SNF, and ORC complexes [23–27]. The Su(Hw) protein is predominantly located in intergenic regions at insulators, while the CG9890, CG9609, and CG10543 proteins mainly localize with promoters. Our study has demonstrated that Aef1 binding sites colocalize with active enhancer marks. Despite the differences in genomic localization, all these proteins share properties with respect to the SAGA, SWI/SNF, and ORC complexes. We hypothesize that positioning of the ORC complexes in the genome is regulated by the DNA-binding proteins responsible for the formation of various regulatory elements, including insulators, promoters, and enhancers. We have demonstrated that Aef1 can be an example of such a protein. ●

*This work was supported by the Russian Science Foundation (grant No. 20-14-00269).*

REFERENCES

1. Orphanides G., Reinberg D. // Cell. 2002. V. 108. № 4. P. 439–451. doi: 10.1016/s0092-8674(02)00655-4.
2. Maksimenko O., Georgiev P. // Front Genet. 2014. V. 5. P. 28. doi: 10.3389/fgene.2014.00028.
3. van Bemmel J.G., Pagie L., Braunschweig U., Brugman W., Meuleman W., Kerkhoven R.M., van Steensel B. // PLoS One. 2010. V. 5. № 11. P. e15013. doi: 10.1371/journal.pone.0015013.
4. Rando O.J., Chang H.Y. // Annu Rev. Biochem. 2009. V. 78. P. 245–271. doi: 10.1146/annurev.biochem.78.071107.134639.
5. Tchurikov N.A., Krasnov A.N., Ponomarenko N.A., Golova Y.B., Chernov B.K. // Nucleic Acids Res. 1998. V. 26. № 13. P. 3221–3227. doi: 10.1093/nar/26.13.3221.
6. Koutelou E., Hirsch C.L., Dent S.Y. // Curr. Opin Cell Biol. 2010. V. 22. № 3. P. 374–382. doi: 10.1016/j.ceb.2010.03.005.
7. Baker S.P., Grant P.A. // Oncogene. 2007. V. 26. № 37. P. 5329–5340. doi: 10.1038/sj.onc.1210603.
8. Brown C.E., Howe L., Sousa K., Alley S.C., Carrozza M.J., Tan S., Workman J.L. // Science. 2001. V. 292. № 5525. P. 2333–2337. doi: 10.1126/science.1060214.
9. Chatterjee N., Sinha D., Lemma-Dechassa M., Tan S., Shogren-Knaak M.A., Bartholomew B. // Nucleic Acids Res. 2011. V. 39. № 19. P. 8378–8391. doi: 10.1093/nar/gkr535.
10. Li B., Carey M., Workman J.L. // Cell. 2007. V. 128. № 4. P. 707–719. doi: 10.1016/j.cell.2007.01.015.
11. Mazina M.Y., Kovalenko E.V., Derevyanko P.K., Nikolenko J.V., Krasnov A.N., Vorobyeva N.E. // Biochim. Biophys. Acta. 2018. V. 1861. № 2. P. 178–189. doi: 10.1016/j.bbagrm.2018.01.016.
12. Mazina M.Y., Nikolenko J.V., Fursova N.A., Nedil'ko P.N., Krasnov A.N., Vorobyeva N.E. // Cell Cycle. 2015. V. 14. № 22. P. 3593–3601. doi: 10.1080/15384101.2015.1100772.
13. Mazina M.Y., Kocheryzhkina E.V., Nikolenko J.V., Krasnov A.N., Georgieva S.G., Vorobyeva N.E. // Dokl. Biochem. Biophys. 2017. V. 473. № 1. P. 145–147. doi: 10.1134/S1607672917020144.
14. Krasnov A.N., Evdokimova A.A., Mazina M.Y., Erokhin M., Chetverina D., Vorobyeva N.E. // Int. J. Mol. Sci. 2023. V. 24. № 14. P. 11844. doi: 10.3390/ijms241411844.
15. Cheng D., Dong Z., Lin P., Shen G., Xia Q. // Int. J. Mol. Sci. 2022. V. 23. № 18. P. 10791. doi: 10.3390/ijms231810791.
16. MacAlpine D.M., Rodriguez H.K., Bell S.P. // Genes Dev. 2004. V. 18. № 24. P. 3094–3105. doi: 10.1101/gad.1246404.
17. Eaton M.L., Prinz J.A., MacAlpine H.K., Tretyakov G., Kharchenko P.V., MacAlpine D.M. // Genome Res. 2011. V. 21. № 2. P. 164–174. doi: 10.1101/gr.116038.110.
18. Vorobyeva N.E., Mazina M.U., Golovnin A.K., Kopytova D.V., Gurskiy D.Y., Nabirochkina E.N., Georgieva S.G., Georgiev P.G., Krasnov A.N. // Nucleic Acids Res. 2013. V. 41. № 11. P. 5717–5730. doi: 10.1093/nar/gkt297.
19. Mazina M., Vorob'eva N.E., Krasnov A.N. // Tsitologiia. 2013. V. 55. № 4. P. 218–224. doi:
20. Kurshakova M., Maksimenko O., Golovnin A., Pulina M., Georgieva S., Georgiev P., Krasnov A. // Mol. Cell. 2007. V. 27. № 2. P. 332–338. doi: 10.1016/j.molcel.2007.05.035.
21. Vorobyeva N.E., Erokhin M., Chetverina D., Krasnov A.N., Mazina M.Y. // Sci. Rep. 2021. V. 11. № 1. P. 16963. doi: 10.1038/s41598-021-96488-0.
22. Vorobyeva N.E., Krasnov A.N., Erokhin M., Chetverina D., Mazina M. // Epigenetics Chromatin. 2024. V. 17. № 1. P. 17. doi: 10.1186/s13072-024-00541-x.
23. Nikolenko J.V., Fursova N.A., Mazina M.Y., Vorobyeva N.E., Krasnov A.N. // Mol. Biol. (Moskow). 2022. V. 56. № 4. P. 557–563. doi: 10.31857/S0026898422040085.
24. Fursova N.A., Mazina M.Y., Nikolenko J.V., Vorobyova N.E., Krasnov A.N. // Acta Naturae. 2020. V. 12. № 4. P. 114–119. doi: 10.32607/actanaturae.11056.
25. Fursova N.A., Nikolenko J.V., Soshnikova N.V., Mazina M.Y., Vorobyova N.E., Krasnov A.N. // Acta Naturae. 2018. V. 10. № 4. P. 110–114. doi: 10.32607/20758251-2018-10-4-110-114.
26. Nikolenko J.V., Kurshakova M.M., Kopytova D.V., Vdovina Y.A., Vorobyova N.E., Krasnov A.N. // Mol. Biol. (Moskow). 2024. V. 58. № 4. P. 700–707. doi: 10.1134/S0026893324700286.
27. Nikolenko J.V., Kurshakova M.M., Kopytova D.V., Vdovina Y.A., Vorobyova N.E., Krasnov A.N. // Mol. Biol. (Moskow). 2024. V. 58. № 4. P. 693–699. doi: 10.1134/S0026893324700274.
28. An W., Wensink P.C. // Genes Dev. 1995. V. 9. № 2. P. 256–266. doi: 10.1101/gad.9.2.256.

29. Brodu V., Mugat B., Fichelson P., Lepesant J.A., Antoniewski C. // Development. 2001. V. 128. № 13. P. 2593–2602. doi: 10.1242/dev.128.13.2593.
30. Falb D., Maniatis T. // Genes Dev. 1992. V. 6. № 3. P. 454–465. doi: 10.1101/gad.6.3.454.
31. Kim D., Paggi J.M., Park C., Bennett C., Salzberg S.L. // Nat. Biotechnol. 2019. V. 37. № 8. P. 907–915. doi: 10.1038/s41587-019-0201-4.
32. Didion J.P., Martin M., Collins F.S. // Peer. J. 2017. V. 5. P. e3720. doi: 10.7717/peerj.3720.
33. Trapnell C., Hendrickson D.G., Sauvageau M., Goff L., Rinn J.L., Pachter L. // Nat. Biotechnol. 2013. V. 31. № 1. P. 46–53. doi: 10.1038/nbt.2450.
34. Mazina M.Y., Kovalenko E.V., Vorobyeva N.E. // Sci. Rep. 2021. V. 11. № 1. P. 172. doi: 10.1038/s41598-020-80650-1.
35. Ramirez F., Ryan D.P., Gruning B., Bhardwaj V., Kilpert F., Richter A.S., Heyne S., Dundar F., Manke T. // Nucleic Acids Res. 2016. V. 44. № W1. P. W160–165. doi: 10.1093/nar/gkw257.
36. McKay D.J., Lieb J.D. // Dev. Cell. 2013. V. 27. № 3. P. 306–318. doi: 10.1016/j.devcel.2013.10.009.