Вавиловский журнал генетики и селекции. 2019;23(2):199-202
DOI 10.18699/VJ19.482

The study of the regulatory region of the *Drosophila melanogaster Notch* gene by new methods of directed genome editing

O.V. Andreyenkov ✉, E.I. Volkova, N.G. Andreyenkova, S.A. Demakov

Institute of Molecular and Cellular Biology, SB RAS, Novosibirsk, Russia
✉ e-mail: andreenkov@mcb.nsc.ru

The *Notch* gene plays a key role in the development of organs and tissues of neuroectodermic origin, including the nervous system. In eukaryotic organisms, the *Notch* pathway is involved in cell fate determination. The *Notch* gene was first discovered in *Drosophila melanogaster*. In mammals, the family of Notch receptors includes four homologues. In humans, mutations in the *Notch* gene cause several hereditary diseases and carcinogenesis. Studies of the regulatory zone of the *Notch* gene in *D. melanogaster* have been conducted for several decades. We review their results and methods. The regulatory zone of the *Notch* gene is in the region of open chromatin state that corresponds to the 3C6/3C7 interband on the cytological map of polytene chromosomes of *D. melanogaster* salivary glands. The development of new methods for directed genome editing made it possible to create a system for introducing directed changes into the regulatory zone of the gene. Using the CRISPR/Cas9 system, we obtained a directed 4-kilobase deletion including the 5′-regulatory zone, promoter, and the first exon of the *Notch* gene and introduced the *attP* site into the first intron of the *Notch* gene. This approach enabled targeted changes of the sequence of the regulatory and promoter regions of the gene. Thus, it provided a new powerful tool for studies of *Notch* gene regulation and the organization of the open chromatin state.

Key words: *Drosophila melanogaster*; *Notch* gene; CRISPR/Cas9 system; gene activity regulation; insulator protein; open chromatin state; interbands.

For citation: Andreyenkov O.V., Volkova E.I., Andreyenkova N.G., Demakov S.A. The study of the regulatory region of the *Drosophila melanogaster Notch* gene by new methods of directed genome editing. Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding. 2019;23(2):199-202. DOI 10.18699/VJ19.482

Исследования регуляторной зоны гена *Notch* у *Drosophila melanogaster* с использованием новых подходов направленного геномного редактирования

О.В. Андреенков ✉, Е.И. Волкова, Н.Г. Андреенкова, С.А. Демаков

Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ e-mail: andreenkov@mcb.nsc.ru

Ген *Notch* играет ключевую роль в развитии органов и тканей нейроэктодермального происхождения, в том числе нервной системы. У эукариотических организмов сигнальный каскад *Notch* вовлечен в процессы клеточной детерминации. Впервые ген *Notch* был открыт у *Drosophila melanogaster*. У млекопитающих семейство Notch рецепторов включает четыре гомолога. У человека мутации в гене *Notch* приводят к развитию ряда наследственных заболеваний, а также связаны с канцерогенезом. Исследования регуляторной зоны гена *Notch* на *D. melanogaster* насчитывают уже несколько десятилетий. В статье сделан обзор результатов и методов этих исследований. Регуляторная зона гена находится в районе открытого состояния хроматина, соответствующем междиску 3C6/3C7 на цитологической карте политенных хромосом слюнных желез *D. melanogaster*. Развитие новых методов направленного геномного редактирования позволило создать систему для введения направленных изменений в регуляторную зону гена. Используя систему CRISPR/Cas9, мы получили делецию в регуляторной и промоторной зоне гена *Notch*, включая его первый экзон, и ввели сайт *attP* в первый интрон гена *Notch*. Это сделало возможным направленное изменение последовательностей регуляторной и промоторной зон гена.

Ключевые слова: *Drosophila melanogaster*; ген *Notch*; система CRISPR/Cas9; регуляция активности генов; инсуляторные белки; открытое состояние хроматина; междиски.

Introduction

The *Notch* gene is the central element of the evolutionarily conservative gene network that controls embryonic development. It is involved in determining neuroectodermal cell fates through the mechanism of lateral inhibition. This mechanism is based on the ligand-receptor interaction between neighboring cells. Accordingly, the *Notch* gene product is a transmembrane receptor protein. Its ligands are Delta and Serrate. The extracellular domain of the protein, which includes 36 EGF repeats and 3 lin12/Notch repeats, is responsible for binding

to the ligand (Wharton et al., 1985). The protein also includes transmembrane and intracellular domains. Processing of this receptor protein involves several stages of cleavage by specific proteases. Interaction with the ligand also triggers the irreversible separation of the intracellular domain. The latter is transmitted to the nucleus, where it forms a part of a complex that activates the transcription of target genes. Thus, each stage of the signal transmission involves several proteins. One of the key features of the signal transduction system is that the Notch receptor is irreversibly split in the course of signal transduction. The signal transmission, therefore, is conducted without amplification at this stage.

The *Notch* signaling cascade plays a key role in lateral inhibition. This term refers to intercellular interactions in the differentiation of peripheral nervous system cells, leading to the separation of proneuronal cluster cells through the neuronal development pathway and to the suppression of proneuronal gene expression in neighboring cells. This process suppresses the neuronal differentiation potential of neighboring cells and switches them to the epidermal pathway of development (Artavanis-Tsakonas et al., 1999). However, there is evidence that the gene functions are not limited to the processes of lateral inhibition; in particular, the gene also takes part in the formation of organs of neuronal origin (Brennan et al., 1997).

The molecular approaches to study of the Notch regulatory region

A very large number of works have been devoted to the functioning of the gene network as a whole and various stages of signal transmission (Greenwald, 2012). However, it is obvious that the regulation of expression of the gene itself can also influence the functioning of the *Notch* gene network profoundly.

This statement is supported by the existence of numerous mutations caused by various disorders in the *Notch* locus. In *Drosophila melanogaster*, the *Notch* gene occupies a locus of more than 37 kb in size located on the X chromosome. Its position on the cytological map of salivary gland polytene chromosomes corresponds to the 3C6/3C7 region. The entire structural region of the gene, including exons and introns, is in the 3C7 band, and the 5′ regulatory region of the gene upstream of the transcription initiation sites is in the open chromatin of the 3C6/3C7 interband (Rykowski et al., 1988). The long length of the locus is due to several large introns in the gene structure. In general, the *Notch* locus is one of the most fully characterized *D. melanogaster* genes. It has long been the subject of molecular and cytogenetic studies.

In accordance with the function of the gene, the mutations associated with this locus are manifested in disorders emerging during neuroectoderm differentiation in organs of neuroectodermal origin. Phenotypically, they include the eye, wing, gonad, and bristle aberrations. However, the presence of several independent functions of a gene at different developmental stages makes it difficult to analyze its mutations by studying their phenotypic consequences (Brennan et al., 1997).

Along with mutations affecting the structural part of the *Notch* gene, a lot of them being associated with EGF repeats, there are mutations in introns and the regulatory region of the gene that affect the fly development. The described mutations in the gene introns are associated with mobile element insertions. They include, for example, mutations of the *facet* group (*fa*, fa^3, fa^g, $fa^{g\text{-}2}$, fa^{fx} and fa^{sw}) (Markopoulou, 1989), which are associated with insertions of *copia*-like mobile elements clustered on a 4-kb segment in the second intron of the gene. All mutations in this group hinder the correct formation of the facet eye. In the severest cases, the correct hexagonal pattern of facets is lost, the order and number of bristles are disturbed, and additional interstitial tissue separating the facets appears. Some mutations of the *facet* group also cause notches on the wing.

Along with insertions into the introns of the gene, a deletion in the regulatory region of the *Notch* gene is well known. The *strawberry* (fa^{swb}) phenotype of rough eyes in males is associated with a deletion of 880 bp in the 5′ regulatory region of the *Notch* gene in the immediate vicinity of the transcription initiation sites. Together with the deletion, the 3C6/3C7 interband disappears from the salivary gland polytene chromosome cytological preparations, and the disappearance is accompanied by the replacement of the 3C6 band with a longer band, probably resulting from the fusion of the 3C6 and 3C7 bands.

The authors of works concerning the cytology of this region interpret the interband disappearance either as a consequence of deletion of all sequences that normally comprise the interband, or as the result of the disappearance of a cis-activating sequence inside the deleted sequence. This sequence is meant to be necessary for the formation of the decondensed chromatin state, and its disappearance leads to closure of a large interband sequence into a more condensed band. “It would be interesting to know whether a band and an interband accept their state actively, or one of these states is initial and is realized in the absence of sequences necessary for the other state” (Rykowski et al., 1988).

The subdivision into several structurally and functionally independent units is characteristic of eukaryotic chromosomes. The 5′ regulatory region of the *Notch* gene performs barrier functions. According to the high-resolution Hi-C data, the boundary of topological domains is mapped to this locus (Stadler et al., 2017). It also comprises the binding peaks of insulator proteins CTCF, CP190, and GAF (the *trl* gene product), which may be responsible for the insulator activity of promoters (Ohtsuki, Levine, 1998).

The most detailed work on the insulator function of the sequence directly adjacent to the promoter region of the *Notch* gene was carried out by J. Vazquez and P. Schedl (Vazquez, Schedl, 2000). The authors focused on the study of the fa^{swb} deletion effects. The conclusion that the sequence deleted by fa^{swb} normally protects the promoter of the *Notch* gene from the position effect is based on the observation of the impact of chromosome rearrangements that save the mutant phenotype. In this work, the insulator activity of the sequence deleted by fa^{swb} was checked in transgenic constructs.

First, the distribution of DNase-I hypersensitive sites (DHSs) in the promoter region of the *Notch* gene was investigated. In the regulatory region of the gene outside the promoter region, three hypersensitive sites were found. They occupied a total of about 200 nucleotide pairs. One of the sites overlapped the proximal break point of the fa^{swb} deletion, and two others

were mapped to the region removed by this deletion. It is known that the positions of DHSs correlate with the locations of regulatory DNA sequences in eukaryotes, and this fact was used for their mapping (Thomas et al., 2011).

The insulator activity of the sequence removed by the fa^{swb} deletion was tested in two regards: the ability to block enhancer-promoter interactions and the ability to protect reporter genes from the position effect (Vazquez, Schedl, 2000). The cloned sequence of 2.3 kb in length included all the transcription initiation sites of the gene, its promoter region, the sequence removed by the fa^{swb} deletion, and about 1 kb of more distal DNA sequence. As part of the *P*-transposon, it disturbed the reporter gene expression when the cloned sequence was placed between the enhancer and promoter, and it did not affect the expression when the cloned sequence was situated upstream of the enhancer. With the second approach applied in the work, the same sequence as part of a transgenic construct protected the reporter gene from the position effect. Expression of the reporter gene with the presence of the sequence from the 5′ region of the *Notch* gene was constant regardless of the genetic environment at the site of the transgenic construct insertion in the genome. With the absence of the sequence in question, the rate of the expression of the reporter gene depended on the position effect.

In all the tests described above, the removal of the region corresponding to the fa^{swb} deletion from the transgenic sequences resulted in disappearance or substantial weakening of the insulator properties of the sequence studied. Further analysis of the insulator activity of some restriction fragments of the sequence under study within transposons brought the authors to the conclusion that the insulator activity was associated with the sequence removed by the fa^{swb} deletion; specifically, a sequence including about 60 bp localized proximal to the gene might also be responsible for this activity.

Thus, in the work by J. Vazquez and P. Schedl (2000), the approach proposed thirty years ago was practically used for the first time: "Sequencing the material deleted by fa^{swb} have revealed a 47 bp repeat that has no supposed regulatory motifs. Experiments on transformation with the removal of short sequences could help determine whether a 47 bp repeat or some other part of the 880 bp sequence is necessary to determine the interband state" (Ramos et al., 1989). Here it is necessary to clarify that the tandem repeat having a complex organization of two sections, 47-bp long each, occupying a total of 97 bp is one of the characteristic features of the sequence deleted by fa^{swb}. The authors of early works devoted to the *Notch* gene regulatory region had certain expectations of it.

The obvious continuation of using this approach was the search of the *Notch* gene regulatory region for the shortest nucleotide sequence that would be necessary and sufficient for the formation of the interband chromatin state. The work began with the design of the pICon genetic construct, which allowed the studied sequences of the *Notch* gene regulatory region to be embedded into the same given position in the genome (Zimin et al., 2004). To do this, an FRT site was introduced into the construct. With the presence of FLP recombinase, every sequence studied could be integrated as part of a special donor construct into the FRT site introduced into the genome within pICon. The sequences studied in the donor construct were flanked by FRT sites and, as a result of an FRT/FLP recombination event, integrated into the genome.

However, at the second stage of the work, a 1504-bp long sequence including the region removed by the fa^{swb} deletion and more distal regions was integrated into the genome as part of the pICon construct by means of *P*-element-mediated transformation (Semeshin et al., 2008). The transforming construct was called p3C-Icon after the cytological region of the origin of the sequence under study. Thus, the potential of the site-specific transformation system was not used at this stage. Seven fly stocks were obtained, and four of them, carrying the insertion of the sequence from the regulatory region of the *Notch* gene, were examined. In all cases, the insertions of p3C-Icon were characterized by emergence of a new interband. The FRT/FLP system was then used to cut the material of the sequence under study, and in all cases the interband disappeared from the sites of p3C-Icon insertion. It was also shown that when transferred to a new genetic environment as part of the construct, the DHSs from the *Notch* 5′ regulatory region retained their properties.

At the third stage of the work, we used the FRT site integrated in the pICon(dv) genetic construct situated in the 84F region of chromosome 3 to insert the sequences from the *Notch* regulatory region. This time, an integrated FRT site provided an opportunity to test different sequences in the same genetic environment and then to cut out the integrated sequences. For this purpose, the tested sequences as part of the pFRTV vector were introduced into various sites in the *Drosophila* genome and then transferred site-specifically to the 84F region. The previously tested 1504-bp sequence was used as a basic one to design a shorter sequence of 914 bp including the whole fragment deleted by the fa^{swb} deletion. Also, a second test sequence of 1258 bp in length was obtained by removing a part of the 1504 bp sequence proximal to the *Notch* promoter. It was assumed that the DHSs found in (Vazquez, Schedl, 2000) were located in the removed part. We showed that the deletion of 246 bp led to the complete disappearance of the interband from the 84F region, whereas the presence of the 914-bp sequence completely including the fragment deleted by the fa^{swb} deletion caused the formation of an interband (Andreenkov et al., 2010). Thus, the shortest sequence required for the formation of decondensed chromatin state of the 3C6/3C7 interband was identified.

After obtaining these data, we cloned this minimal sequence into pFRTV, inserted it into the genome, and transferred to the acceptor construct pICon(dv) in the 84F region of chromosome 3 (a slightly larger 276-bp fragment was used). The presence of this fragment in the 84F region led to the appearance of an interband. It also turned out that the appearance of the interband was accompanied by the emergence of a DHS in the insertion region. Specifically, the insertion of sequences of 914 bp and 276 bp in size caused the formation of a DHS and an interband and the insertion of the 1258-bp sequence did not lead to the formation of these elements. Similarly, we cloned three tandem repeated sequences of 914 bp in length (fragment deleted by fa^{swb}) into the 84F region. As a result, an interband and three DHSs formed.

To summarize, studies of the *Notch* gene regulatory region were carried out using different approaches united by the idea of transferring individual sequences from the regula-

tory region to ectopic genetic environment. The autonomous ability of individual short fragments to cause the formation of decondensed chromatin state was shown.

However, the possibility of introducing changes directly to the sequence of the *Notch* gene regulatory region offers much greater opportunities. One of them arises with the development of directed genome editing methods. Previously, we used homologous recombination for the directional deletion of the *Drosophila RNaseZ* gene. The gene sequence was replaced by the attP site (Andreenkov et al., 2016). Such sites are used to integrate sequences into the genome by using the *ϕC31*-mediated recombination method. The attP/attB method of directed recombination permits one to embed then a sequence that will completely rescue the deletion, as well as any other artificially constructed sequence, at the attP site introduced at the deletion site.

Recently, this method has been modified. Now the approach employs the CRISPR/Cas9 system to obtain a directed deletion and insertion of the attP site, which allows more efficient construction of systems for making directed changes in a genome (Gratz et al., 2014).

We used this approach to obtain a directed 4-kilobase deletion in the *Notch* gene. With the CRISPR/Cas9 system, the regulatory and promoter zones, the first exon, and a part of the first intron of the *Notch* gene were removed and replaced by the attP site. For this purpose, we created a genetic construct that contained sequences homologous to the 3′ and 5′ sites flanking the deletion, the attP site, the *miniwhite* reporter gene, and the GMR enhancer. The *miniwhite* reporter gene of the construction was flanked by lox sites, which allowed us to remove them after inserting the attP site. In this way, a fly stock carrying the attP site and a 4-kilobase deletion in the *Notch* gene was obtained. To check the functionality of the created system, we inserted the 4-kb sequence deleted when creating an attP site into this attP site and thereby restored the functions of the gene.

Conclusion

Thus, a system that allows introducing any directed changes into the regulatory region of the *Notch* gene was first created. It will be used to obtain directed deletions, to replace functionally significant sites, and to introduce single nucleotide substitutions. Now we have an opportunity to accurately reproduce the natural fa^{swb} deletion and to remove the minimal sequence found in the regulatory region of the *Notch* gene, which has the properties of interband formation when transferred as part of a transposon to an ectopic environment. The system allows removing various *Notch* gene transcription initiation sites and analyzing their influence on gene activity. It can also be used to replace the binding site of the CTCF and CP190 insulator proteins with an alternative BEAF-32 binding site. It will allow us to make inferences as to the functionality of various insulator proteins. We conclude that a fundamentally new and powerful tool has been developed and it can be applied to studies of the *Notch* gene regulation and the organization of the open chromatin state.

References

Andreenkov O.V., Volkova E.I., Demakov S.A., Semeshin V.F., Zhimulev I.F. The decompact state of interchromomeric chromatin from the 3C6/C7 region of *Drosophila melanogaster* is determined by short DNA sequence. Doklady Biochemistry and Biophysics. 2010; 431:57-59. DOI 10.1134/S1607672910020018.

Andreenkov O.V., Volkova E.I., Demakov S.A., Xie X., Dubrovsky E.B., Zhimulev I.F. Targeted mutagenesis of *Drosophila RNaseZ* gene by homologous recombination. Doklady Biochemistry and Biophysics. 2016;471(1):399-402. DOI 10.1134/S1607672916060065.

Artavanis-Tsakonas S., Rand M.D., Lake R.J. Notch signaling: cell fate control and signal integration in development. Science. 1999;284: 770-776. DOI 10.1126/science.284.5415.770.

Brennan K., Tateson R., Lewis K., Martinez Arias A. A functional analysis of *Notch* mutations in Drosophila. Genetics. 1997;147:177-188.

Gratz S.J., Ukken F.P., Rubinstein C.D., Thiede G., Donohue L.K., Cummings A.M., O'Connor-Giles K.M. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. Genetics. 2014;196:961-971. DOI 10.1534/genetics.113.160713.

Greenwald I. *Notch* and the awesome power of genetics. Genetics. 2012;191:655-669. DOI 10.1534/genetics.112.141812.

Markopoulou K. Phenotypic and molecular analysis of the facet, a group of intronic mutations at the Notch locus of *Drosophila melanogaster* which affect postembryonic development. Genetics. 1989; 122:417-428.

Ohtsuki S., Levine M. GAGA mediates the enhancer blocking activity of the *eve* promoter in the *Drosophila* embryo. Genes Dev. 1998; 12:3325-3330.

Ramos R.G.P., Grimwade B.G., Wharton K., Scottgale T.N., Artavanis-Tsakonas S. Physical and functional definition of the *Drosophila Notch* locus by *P* element transformation. Genetics. 1989;123:337-348.

Rykowski M.C., Parmelee S.J., Agard D.A., Sedat J.W. Precise determination of the molecular limits of a polytene chromosome band: regulatory sequences for the *Notch* gene are in the interband. Cell. 1988;54:461-472.

Semeshin V.F., Demakov S.A., Shloma V.V., Vatolina T.Y., Gorchakov A.A., Zhimulev I.F. Interbands behave as decompacted autonomous units in *Drosophila melanogaster* polytene chromosomes. Genetica. 2008;132:267-279. DOI 10.1007/s10709-007-9170-5.

Stadler M.R., Haines J.E., Eisen M.B. Convergence of topological domain boundaries, insulators, and polytene interbands revealed by high-resolution mapping of chromatin contacts in the early *Drosophila melanogaster* embryo. ELife. 2017;6:e29550. DOI 10.7554/eLife.29550.

Thomas S., Li X.-Y., Sabo P.J., Sandstrom R., Thurman R.E., Canfield T.K., Giste E., Fisher W., Hammonds A., Celniker S.E., Biggin M.D., Stamatoyannopoulos J.A. Dynamic reprogramming of chromatin accessibility during *Drosophila* embryo development. Genome Biology. 2011;12:R43. DOI 10.1186/gb-2011-12-5-r43.

Vazquez J., Schedl P. Deletion of an insulator element by the mutation *facet-strawberry* in *Drosophila melanogaster*. Genetics. 2000;155: 1297-1311.

Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S. Nucleotide sequence from the neurogenic locus notch implies a gene product that shares homology with proteins containing EGF-like repeats. Cell. 1985;43:567-581.

Zimin P.I., Gortchakov A.A., Demakov S.A., Zhimulev I.F. A new construct for cloning DNA and modeling the structure of *Drosophila melanogaster* polytene chromosomes. Mol. Biology. 2004;38(2): 205-209.

Acknowledgements. This work is supported by the Russian Science Foundation, project 14-14-00934.

Conflict of interest. The authors declare no conflict of interest.

Received November 24, 2018. Revised December 17, 2018. Accepted December 21, 2018.