GENETICS

Organ-specific transcripts as a source of gene multifunctionality: lessons learned from the *Drosophila melanogaster sbr* (*Dm nxf1*) gene

Ludmila Mamon[1], Viktoria Ginanova[1], Sergey Kliver[2], Mariya Toropko[1], and Elena Golubkova[1]

[1]Department of Genetics and Biotechnology, Faculty of Biology, Saint Petersburg State University, Universitetskaya nab., 7–9, Saint Petersburg, 199034, Russian Federation
[2]Theodosius Dobzhansky Center for Genome Bioinformatics, Saint Petersburg State University, Sredniy Pr., 41, Saint Petersburg, 199004, Russian Federation

Address correspondence and requests for materials to Elena Golubkova, elena_golubkova@mail.ru, e.golubkova@spbu.ru

Abstract

Analysis of the transcriptomes of different organisms has demonstrated that a single gene can have multiple transcripts. The sources of transcriptional variability are the alternative promoters, polyadenylation sites, splicing, and RNA editing. A comparison of the organisms of different taxa has demonstrated that the complexity of organization during evolution arises not due to an increase in the number of protein-coding genes. The greatest variability of transcripts is specific to the nervous and germinal systems. A variety of mechanisms providing for the complexity of the transcriptome ensures a precise and coordinated regulation of organ-specific functions through a combination of cis-acting elements and trans-acting factors. The *D. melanogaster sbr* (*Dm nxf1*) gene has proven to be an excellent model for investigating mechanisms potentially leading to the emergence of multiple products with various functions.

Keywords: *nxf* (nuclear export factor), *D. melanogaster*, alternative splicing, intron retention, transcriptional variability, alternative polyadenylation.

Citation: Mamon, L., Ginanova, V., Kliver, S., Toropko, M., and Golubkova, E. 2019. Organ-specific transcripts as a source of gene multifunctionality: lessons learned from the *Drosophila melanogaster sbr* (*Dm nxf1*) gene. Bio. Comm. 64(2): 146–157. https://doi.org/10.21638/spbu03.2019.206

Author's information: Ludmila Mamon, Dr. of Sci. in Biology, Professor, orcid.org/0000-0001-5338-0703; Viktoria Ginanova, PhD, Junior Researcher, orcid.org/0000-0002-8404-9893; Sergey Kliver, MSc, Junior Researcher, orcid.org/0000-0002-2965-3617; Mariya Toropko, student; Elena Golubkova, PhD, Associate Professor, orcid.org/0000-0002-9528-5760

Manuscript Editor: Anton Nizhnikov, Department of Genetics and Biotechnology, Faculty of Biology, Saint Petersburg State University, Saint Petersburg, Russia

Received: May 6, 2019;

Revised: June 17, 2019;

Accepted: June 20, 2019;

Funding: This work was supported by the Russian Foundation for Basic Research (19-04-01255 a).

Competing interests: The authors have declared that no competing interests exist.

Contributors to Transcriptome Complexity

Analysis of the transcriptomes of different organisms has demonstrated that a single gene can have multiple transcripts. In *Drosophila melanogaster*, 47 genes can potentially code for more than 1000 transcript isoforms each (Brown et al., 2014). The sources of transcriptional variability are the alternative promoters, polyadenylation sites, splicing, and RNA editing. The same DNA sequence can encode both protein coding transcripts and non-coding RNAs, including regulatory antisense transcripts. The so-called nested genes provide an additional opportunity for an increased variability of transcripts that are products of the same genome locus. These can be separate genes located either on the same strand or on opposite strands within the host gene (Kumar, 2009). Such "nested" genes are located in intron 3 of the *D. melanogaster sbr* (small bristles) gene (Fig. 1). Ten exons and nine introns of the *sbr* gene with a length of 14341 bp are located in genomic region X:10,832,752..10,847,092 (FlyBase, 2019). Intron 3 (8892 nt) comprises 62 % of this gene sequence. Within this intron, there exist three genes: CG32669, CG15209, and CG15210. Initially they were predicted to be ORFs; their transcriptional activity was demonstrated later. The expression pattern and functions of these genes have been poorly investigated so far. One of them, CG32669, has the same orientation as the *sbr* gene, and the other two genes — CG15209 and CG15210 — are transcribed in the opposite direction.

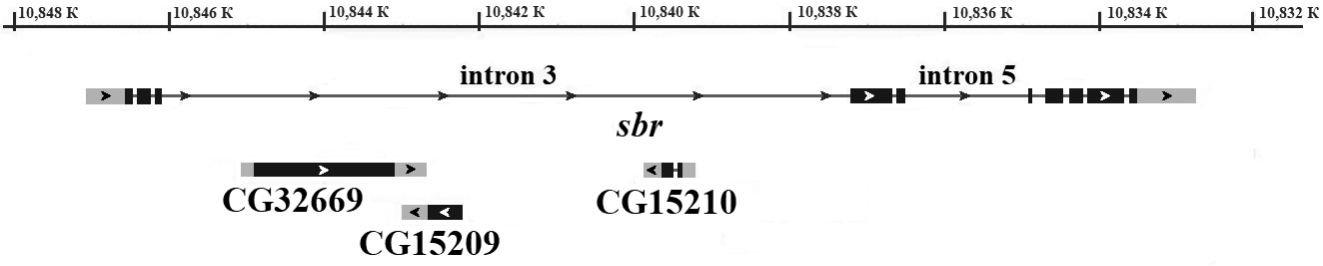

Fig. 1. The structure of the *sbr* (*Dm nxf1)* gene. The CG15210, CG15209 and CG32669 genes in intron 3 of the *sbr* gene.

Antisense transcripts that are complementary to the coding part of the gene are a common feature of transcriptomes. They include both long non-coding RNAs (lncRNAs) and short RNAs. Comprehensive analysis of the *Drosophila* transcriptome (Brown et al., 2014) has found both sense and antisense transcripts to be present in the same cell type at the same time. Sometimes antisense lncRNAs overlap with the 3′ and 5′ UTRs of adjacent genes, forming gene chains across contiguously transcribed regions. Antisense transcripts may overlap with both non-coding and coding sequences (CDSs). Moreover, orthologous genes have comparable antisense transcripts, which suggests a conserved regulation of gene expression with antisense transcripts (Brown et al., 2014).

The presence of paralogous genes with specific functions as a result of duplication followed by subsequent functional divergence may increase the flexibility of the gene expression network and, subsequently, adaptive reaction. The *Nxf* gene family is a good example of the research of the evolution of paralogous genes. Inside this family, only the *Nxf1* gene is homologous for all the species; other genes, even those sharing the same name, may be non-homologous in distant species. For example, the *nxf3* of *Drosophila* is not homologous to the mammalian *Nxf3* (Herold et al., 2001; Mamon et al, 2013), and the mouse *Nxf3* is not orthologous to the human *Nxf3* (Sasaki et al., 2005).

In mammals, the *Nxf* family includes genes with testis-specific and brain-specific expression (Jun et al., 2001; Tretyakova et al., 2005; Zhou et al., 2011). However, in *D. melanogaster*, the presence of organ-specific transcripts generated by different alternative events (such as the use of alternative promoters, alternative splicing, and polyadenylation) has been demonstrated only for the *sbr* (*Dm nxf1)* gene (Ivankova et al., 2010; Ginanova et al., 2016; Ginanova, 2017). An organ-specific pool of significantly different transcripts serves as the basis for the multifunctionality of the *sbr* gene. It is possible that certain tissue-specific functions are scattered throughout different paralogous genes in mammals.

Alternative sources of transcriptome complexity

DIFFERENTIAL PROMOTER USAGE

Core promoters consist of a variety of sequence elements, such as the TATA box, the Initiator (INR), and the downstream promoter element (DPE), recognized by the TATA-binding protein (TBP) and TBP-associated factors of the TFIID complex (Xu et al., 2016). No universal elements have been discovered in all core promoters, and the TATA box is found only in about 10 to 20 % of the metazoan core promoters (Kadonaga, 2012). The canonical basal transcription machinery, including RNA polymerase II (RNAPII), interacts with the core promoter, typically located within -40 to +40 of the transcription start site (Kadonaga, 2002). The activity of the basal transcription machinery itself is quite low. Additional factors, such as activators or repressors with a variety of co-regulators (co-activators or co-repressors), are necessary for an effective transcription initiation by the RNAPII basal transcription machinery at the core promoter (Xu et al., 2016). These sequence-specific regulators of transcription are bound to regulatory DNA sequences, located at promoter-proximal or more distal regions. Transcription initiation efficiency of the RNAPII transcription machinery depends on the core promoter elements, combinatorial assortments of a variety of transcriptional activators or repressors, and chromatin modifications that can be navigated by RNAs, including miRNAs (Kadonaga, 2002, 2012).

Alternative promoter usage is yet another mechanism responsible for the complexity of the transcriptome and proteome (Vacik and Raska, 2017). If alternative promoters are located downstream of the canonical transcription start site, usually in one of the introns, they drive the expression of alternative RNA isoforms without upstream exons. As a result, some important functional domains coded for by the upstream exons will be lost in proteins coded for by alternative mRNA isoforms. Such shortened protein isoforms can be functionally distinct from the full-length protein, coded for by the canonical mRNA isoform.

ALTERNATIVE SPLICING

Alternative splicing (AS) is a common mechanism for increasing protein variety in eukaryotes (Black, 2000; Graveley, 2001; Nilsen and Graveley, 2010). Approximately 95 % of human genes are characterized by alternative pre-mRNA splicing (Wang et al., 2008). The major types of alternative splicing are exon "skipping", the use of exon cassettes that are sets of several adjacent exons with only a single exon from each cassette being chosen during the splicing of pre-mRNA, and the use of different 5′ and 3′ splice sites in exons or introns. A complex interplay of cis- and trans-acting factors promotes or represses the assembly of a splicing complex called a spliceosome at the splice site, and carries out the selection of canonic or an alternative 5′ or 3′ splice site. The spliceosome is a multi-protein-RNA complex, responsible for the precise excision of introns or intron/exon blocks from the pre-mRNA and the fusion of the remaining exons together.

If the pre-mRNA of a gene undergoes alternative splicing of different types, the number of protein isoforms may exceed the number of coding genes of the respective organism. The *Drosophila Dscam* (Down syndrome cell adhesion molecule) gene encoding axon guidance receptors are a great example thereof. This gene includes four exon cassettes of varying exon numbers, and can potentially generate over 38000 mRNA splice-isoforms. Different *Dscam* mRNAs probably label cell types or even single cells in the *Drosophila* brain (Schmucker, 2000; Graveley, 2005).

INTRON RETENTION

Intron retention (IR) is one of the variants of AS common in mammals. IR is considered to be a mechanism of gene expression regulation (Braunschweig et al., 2014; Jacob and Smith, 2017; Schmitz et al., 2017). IR enhances the complexity of the transcriptome (Schmitz et al., 2017). Within the main open reading frame, IR may result in the emergence of premature termination codons (PTCs), leading to the nonsense-mediated decay of mRNA with a retained intron, or, possibly, to the production of truncated proteins. mRNAs with retained introns give rise to alternative protein isoforms, and are therefore a source for the functional diversity of gene products. The existence of mRNA with an intron is a conservative feature of *nxf1* genes in different organisms (Mamon et al., 2013, 2014; Wang et al., 2015). The retained intron may be part of the protein coding sequence, as in the *Ce nxf1* mRNA, or it may only code for either the 17 C-terminus amino acids of the mammalian short NXF1 protein (Li et al., 2006) or the 5 C-terminus of the *Drosophila* short SBR (Mamon et al., 2013, 2014). The use of the PTC in the retained intron leads to the production of truncated proteins. In all cases, IR leads to the emergence of mRNAs with an extended 3′ UTR, including the canonical sequence of the *Nxf1* mRNAs downstream of the retained intron. This suggests the presence of special regulatory functions related to the control of the spatial-temporal properties of mRNAs with a retained intron (Schmitz et al., 2017).

ALTERNATIVE POLYADENYLATION

The 3′ end mRNA processing is a necessary step of mRNA maturation in eukaryotes, including endonucleolytic cleavage and untemplated polyadenylation. The cleavage and polyadenylation (CPA) machinery is composed of macromolecular complexes, such as the cleavage and polyadenylation stimulatory factor (CPSF), cleavage stimulatory factor (CStF), cleavage factor complexes — Im (CFIm) and IIm (CFIIm), and others (Proudfoot, 2011; Erson-Bensan, 2016). CPSF recognizes the polyadenylation signal (PAS), located ~10–30 nt upstream of the cleavage site. Most eukaryotic genes have multiple PASs used in alternative cleavage and polyadenylation (APA). There are two major PAS hexamers: AAUAAA and AUUAAA. Other weaker signal variants were later detected in different species: UAUAAA, AGUAAA, AAGAAA, AAUAUA, AAUACA, CAUAAA, GAUAAA, AAUGAA, UUUAAA, ACUAAA, AAUAGA, AAAUAA, AUAAAA, AUAAAU, AUAAAG, CAAUAA, UAAUAA, AUAAAC, AAAAUA, AAAAAA, AAAAAG, AACAAA, UUAUAU, AAAAAU, and UUUAUU (Beaudoing et al., 2000; Tian et al., 2005; Derti et al., 2012; Sanfilippo et al., 2017). Such noncanonical PASs may be recognized by specific trans-acting factors and serve for alternative polyadenylation. The choice of the cleavage and polyadenylation site is determined not only by PAS but also by the two motifs adjacent to it: U-rich/UGUA upstream elements (USEs) and the U-/GU-rich downstream element (DSE) (Colgan and Manley, 1997; Neve et al., 2017). The CstF complex interacts with DSE and mediates mRNA cleavage at the polyadenylation (pA) site (Neilson and Sandberg, 2010; Erson-Bensan, 2016). The CFIm complex binds to USE and mediates the cleavage reaction. The CFIIm complex promotes the termination of the RNA polymerase II-mediated transcription, and poly(A) polymerases (PAPs) catalyze the addition of untemplated adenosines downstream of the pA site (Proudfoot, 2011). The assembly process of the core factors of the cleavage and polyadenylation complex, including CPSF, CStF, CFIm, CFIIm, PAP, and others, is complicated by the interconnection between the 3′ end processing with 5′ capping, pre-mRNA splicing, and the transcription in the nucleus (Perales and Bentley, 2009; Szostak and Gebauer, 2012; Neve et al., 2017).

Depending on the location of the cleavage and polyadenylation (pA) site, APA events can be classified into two major groups: coding region APA (CR-APA) and UTR-

APA. CR-APA uses pA sites that are located within either CDS or the introns between CDS segments. As a result, alternative mRNA isoforms differ in their coding potential. UTR-APA uses the different pA sites that are located in the 3′ UTR, and the resulting coding potential of alternative mRNA isoforms does not change (Neve et al., 2017).

The majority of genes contain multiple PASs in their 3′ UTRs (Elkon et al., 2012); therefore, 3′ UTR-APA may be a very common event, increasing the variability of transcripts at a transcriptome level. mRNA is usually a part of the ribonucleoprotein (RNP) complex of RNAs and RNA-binding proteins attracted by a combination of cis-acting elements in the 3′ UTR or in another part of mRNA. The longer 3′ UTRs may contain additional regulatory cis-acting elements that can regulate mRNA localization and protein abundance, or create conditions and act a scaffolding for recruiting a protein complex containing RNA-binding proteins (Berkovits and Mayr, 2015; Mitra et al., 2015).

The use of the proximal PAS in the 3′ UTR may lead to a loss of cis-elements for binding with microRNAs and/or RNA-binding proteins (RBPs) (López de Salines et al., 2004; Flynt and Lai, 2008; Szostak and Gebauer, 2012; White et al., 2012). RBPs can recognize not only certain sequences, but also a specific secondary structure of the 3′ UTR. RBP target sites in 3′ UTR can overlap and form targets for miRNAs, and vice versa. In each case, mRNA's fate is determined by the combination of cis-elements in 3′ UTR and the combination of trans-acting factors (RBP and miRNA). A different length of 3′ UTR can serve as a mechanism for differential regulation of subcellular functions of the same proteins (An et al., 2008). There are two types of mRNA *BDNF* (brain-derived neurotrophic factor) with long 3′ UTR and short 3′ UTR in mice. It was shown that only the long 3′ UTR provides the targeting mRNA *BDNF* to dendrites, controlling the abundance of the dendritic BDNF protein and the pruning and enlargement of dendritic spines. miRNAs binding sites in the 3′ UTRs are cis-acting elements, which mediate mRNA decay and translational repression in animals (Iwakawa and Tomari, 2015). So, despite no changes in the protein-coding capacity difference in 3′ UTR, it may be functionally significant and can influence the transport, stability, translation efficiency, and subcellular localization of mRNAs (Colgan and Manley, 1997; Hilger et al., 2011).

The length of 3′ UTRs can change by increasing or decreasing the poly(A)-tail in the cytoplasm allowing for the regulation of the stability of mRNA and the efficiency of its translation (Richter, 1996). The cytoplasmic polyadenylation element (CPE) in 3′ UTR is the binding site for the CPE-binding protein (CPEB) that promotes polyadenylation-induced translation. CPE (UUUUUAU) is located upstream of the nuclear PAS (AAUAAA). In *Xenopus laevis*, CPEB associates with Maskin, binding the translation initiation factor 4E (Barnard et al., 2005). The interaction of Maskin with eIF4E excludes eIF4G and prevents the formation of the eIF4F initiation complex. Phosphorylation events within the CPE-binding protein complex disrupt its connection with the eIF4F initiation complex and allow the cytoplasmic poly(A) polymerase to elongate the poly(A) tail of mRNA. The elongated poly(A) tail is bound by the poly(A)-binding protein, which in turn binds eIF4G, disrupting the Maskin and eIF4E interaction, thereby initiating translation (Barnard et al., 2005; Richter and Sonenberg, 2005). Such strategy of translation regulation allows for the storage of mRNAs in a state of temporary unavailability for translation and for the activation of mRNA translation with the corresponding signal. This mechanism of translation regulation of specific mRNAs in cell cycle regulation (Barnard et al., 2005; Richter and Sonenberg, 2005) and the synapto-dendritic compartment of neurons (Wells et al., 2000; Du and Richter, 2005) has been found.

During the process of CR-APA, the use of an alternative PAS in an intron leads to the formation of a protein isoform without the C-terminus of a full-length protein. The U1snRNP actively suppresses PASs in introns. These PASs are called cryptic (Neve et al., 2017). However, these intronic PASs can become available during increased proliferation (Elkon et al., 2012). mRNAs with APA in the intron are a source of truncated proteins that are incapable of performing the functions of full-length proteins that are coded by the same gene. Identification of trans-acting APA regulators and cis-acting regulatory elements may promote understanding of the mechanisms of APA (White et al., 2012).

The majority of the sources of transcriptome complexity is the cornerstone of the origin of the *sbr* organ-specific transcripts.

ORGAN-SPECIFIC TRANSCRIPTS

The analysis of transcriptomes has made it possible to conclude that the greatest variability of transcripts is specific to the nervous and germinal systems. Genes expressed in the testes, brain, and ovaries have properties that promote the variability of transcriptomes.

BRAIN-SPECIFIC TRANSCRIPTS

The presence of mRNAs with an extended 3′ UTR is a conserved feature of the nervous system in contrast to the testis transcripts (Miura et al., 2013, 2014; Hilger et al., 2011; Smibert et al., 2012). In the nervous system, hundreds of genes have mRNAs being processed with the use of increasingly distant PASs, and their 3′ UTRs reach tens of kb in length (Hilgers, 2015). Specific RNA-binding proteins inhibit cleavage and polyadenylation (CPA) at proximal sites. In *Drosophila*, the ELAV (embryonic-lethal abnormal visual system) protein is known to play

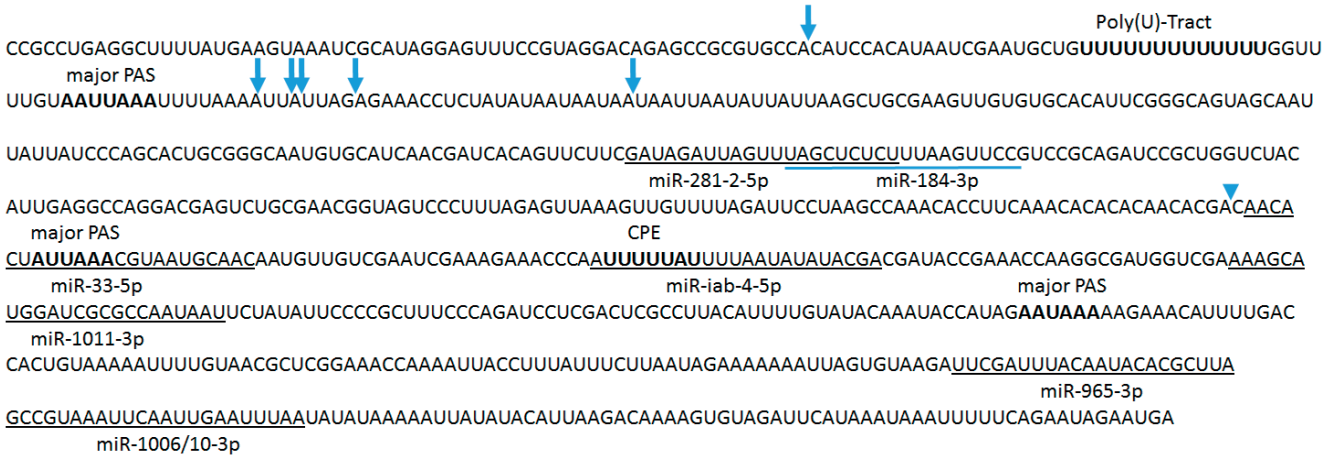

Fig. 2. Exon 10 sequence of the *sbr* gene. The end of the coding region is shown in bold. Arrows mark the ends of testis-specific transcripts detected by 3' RACE-PCR. Arrow head marks the 3' ends of the head and/or ovary transcripts. Important polyadenylation elements are depicted as poly(U)-tract, major PAS — major polyadenylation signal, CPE — cytoplasmic polyadenylation element. The sequences corresponding to predicted poorly conserved sites for conserved miRNA families are underlined.

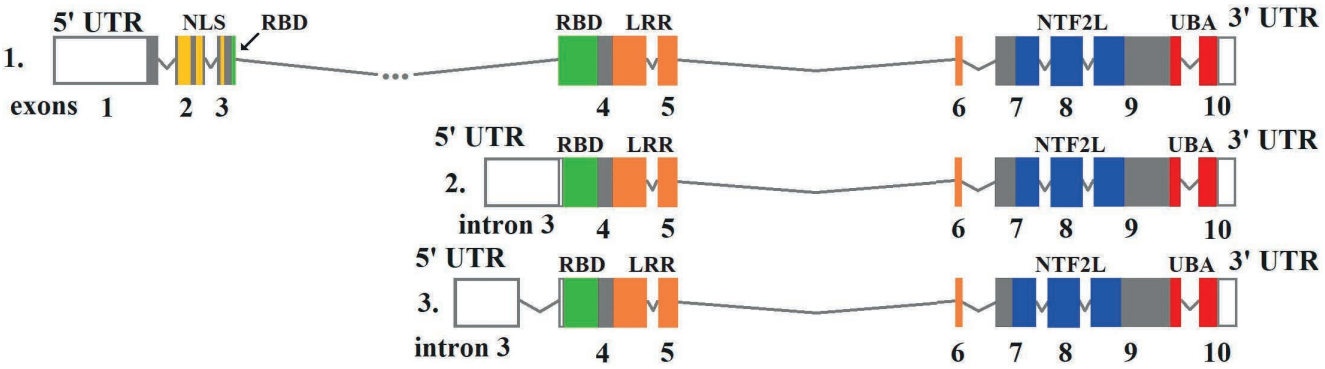

Fig. 3. The testis-specific *sbr* mRNAs are detected by 5' and 3' RACE-PCR. Boxes indicate exons included in the *sbr* mRNAs. Exon numbers are signed under each box. RBD — RNA-Binding Domain; LRR — Leucine-Rich Domain; NTF2L — Nuclear Transport Factor 2 Like; UBA — Ubiquitin-Associated domain; NLS — Nuclear Localization Signal. In testes, all *sbr* mRNAs have the shortened 3' UTR. In all transcripts, length of the 3′ UTR varies from 44 to 124 nucleotides. Two *sbr* mRNAs, using the alternative promoters in intron 3, give rise the shortened protein isoform without NLS and the part of RBD, which are present in the canonical full-length SBR.

the role of APA regulators. However, no binding sites for ELAV have been identified in the extended 3′ UTRs (Hilger et al., 2012; Smibert et al., 2012). It was suggested that the distant APA depends on specific sequences in the promoters of genes producing mRNAs with an extended 3′ UTR (Hilgers et al., 2012; Hilgers, 2015; Oktaba et al., 2015). A search of such sequences by means of computational analysis has revealed that the GAGA element is frequently found in the promoter regions of the genes producing mRNAs with extended 3′ UTR (Li and Gilmour, 2013; Oktaba et al., 2015). The GAGA element signals pausing to the RNA Polymerase II (Pol II). This confirms the hypothesis that there exists interrelation between the processes of transcription initiation and CPA.

CStF-64 is a brain-specific CStF. CStF-64 is a splice variant of CStF and is found in all regions of the brain and the peripheral nervous system (MacDonald and McMahon, 2010). CStF-64 may play a role in determining the polyadenylation site in brain-specific mRNAs.

Interestingly, the hippocampus exhibits the largest number of 3′ UTR extensions in comparison with all other tissues (Miura et al., 2013). 3′ UTR lengthening may increase the number of cis-acting elements used as target sites for neural-specific miRNAs.

There are transcripts that undergo post-transcriptional cleavage to release specific fragments, which then function independently (Tuck and Tollervey, 2011), and 3′ UTR extensions may be a source of ncRNAs after post-transcriptional cleavage (Miura et al., 2013). A significant increase in the 3′ UTR length is observed in the *sbr* mRNA with retained intron 5. This intron was dubbed a cassette intron because it is part of the evolutionarily conserved cassette exon — 110 bp — intron — exon — 37 bp, which is found in the *Nxf1* genes in different organisms. mRNAs with a cassette intron are also found in organisms of different taxa (Mamon et al., 2013, 2014; Wang et al., 2015). Cassette introns contain evolutionarily conserved motifs. In *Drosophilidae* they are evident in the form of

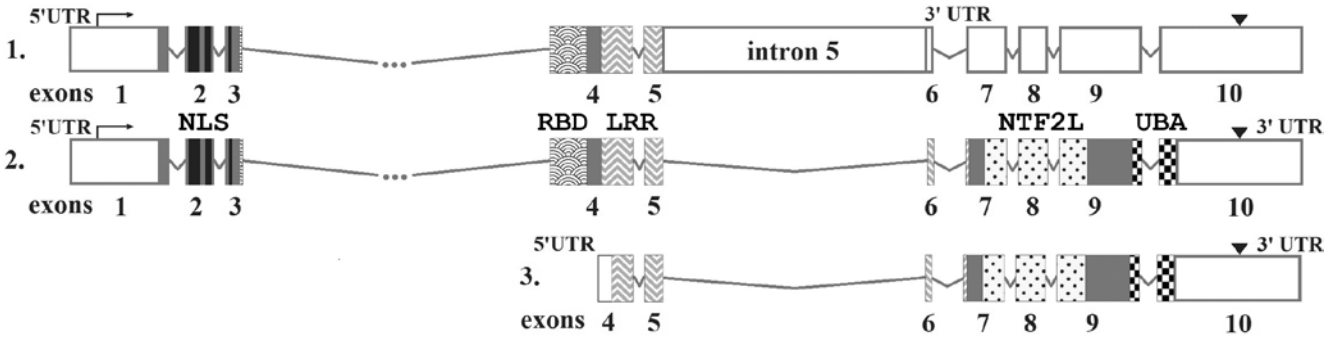

Fig. 4. The *sbr* mRNAs are detected by 5' and 3' RACE-PCR in the head RNA probe. The symbols are the same as in Fig. 3. Horizontal arrows show the alternative transcription start sites, arrow heads show the alternative site of cleavage and polyadenylation. The *sbr* mRNAs isoforms are listed in order of decrease of their quantity.

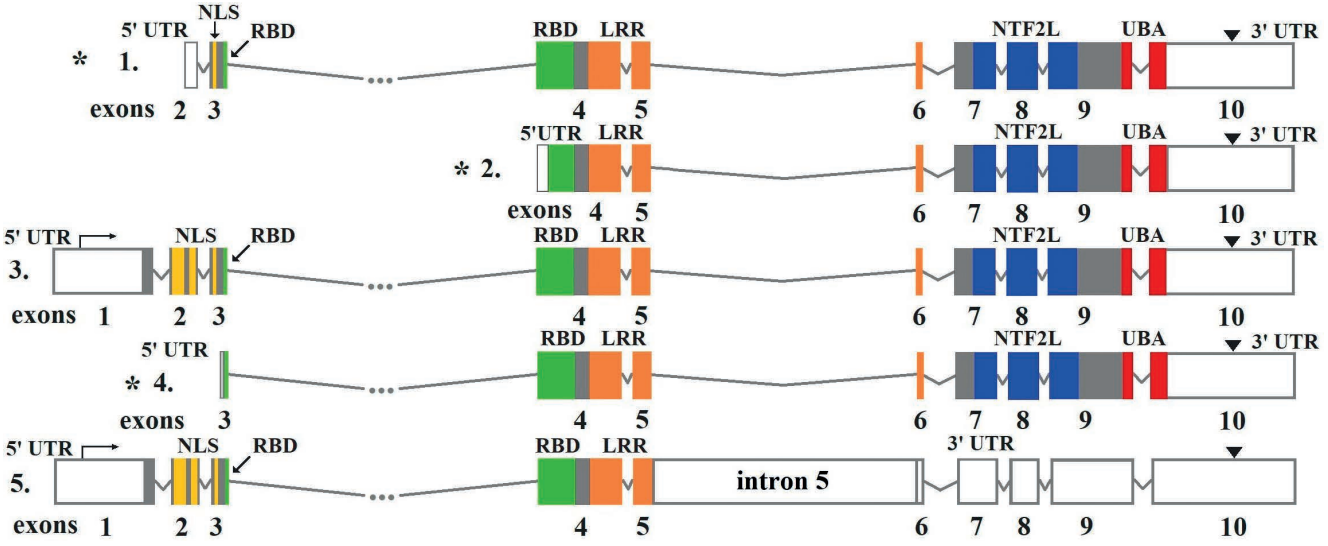

Fig. 5. The *sbr* mRNA isoforms are detected by 5' and 3' RACE-PCR in the ovary RNA probe. The symbols are the same as in Fig. 3. Horizontal arrows show the alternative transcription start sites, arrow heads show the alternative site of cleavage and polyadenylation. The *sbr* mRNAs isoforms are listed in order of decrease of their quantity. The ovary-specific *sbr* mRNA isoforms are marked by asterisk.

two extended poly(A) sequences. Such poly(A)-tracts are known to facilitate the inclusion of corresponding introns into the processed RNA (Jacob and Smith, 2017).

It should be noted that the *sbr* gene's intron containing mRNA is the most abundant among the *sbr* transcripts in adult fly heads (Fig. 4). This is also supported by the results of a northern blot analysis, and suggests neurospecificity of the mRNAs with the retained intron (Ivankova et al., 2010). The *sbr* mRNA with the retained intron gives rise to a truncated protein (unpublished). The investigation of the role of the truncated SBR protein in the formation and function of the nervous system is currently underway.

TESTIS-SPECIFIC TRANSCRIPTS

Translational regulation is a fundamental characteristic of gene expression in mammalian testes (Kleene, 2001, 2003). Long-living mRNAs that are synthesized long before translation have been found. They are stored in a translationally inactive state for several days, and are then translated in transcriptionally inactive elongated spermatids during the formation of spermatozoa from spermatids (Hawthorne et al., 2006). The widespread use of the alternative promoter, upstream PAS, and alternative splicing are features of gene expression in spermatogenic cells (Kleene, 2001, 2003). In spermatogenic cells, the transcripts of some housekeeping genes encode truncated proteins that cannot have the same functions as the full-length products of the same genes in somatic cells (Kleene, 2001). Alternative promoter usage forms an alternative 5′ UTR determining the efficiency of mRNA translation. Translation is repressed with the upstream reading frames (uORFs) in the 5′ UTR before the start codon (AUG) of the reading frame encoding the "functional" protein in the mRNA (Child et al., 1999). The presence of uORFs in mRNA's 5′ UTR is an additional mechanism of the translation regulation in spermatogenesis (Kleene,

2003). Paralogous genes within gene families with only testes-specific expression are yet another feature of spermatogenesis. In mice, round spermatids express a testis-specific isoform of poly(A) polymerase, which is localized in the cytoplasm in contrast to the somatic isoform with the nuclear localization (Kashiwabara et al., 2000, 2002). Thus, cytoplasmic polyadenylation is an additional way of translation regulation of mRNAs during spermatogenesis (Kashiwabara et al., 2002; Kleene, 2003). There are testis-specific translation repressors, such as Y-box proteins and the testis-brain RNA binding protein (TB-RBP) in mammals (Kleene, 2003). It is an additional source allowing each mRNA to use its own mechanism of translational regulation. The specifics of a testis-specific transcriptome consist of transcripts of many genes, which are expressed in both somatic and spermatogenic cells, and are modified by alternative transcription start sites, splicing, and upstream polyadenylation sites, that form the atypical patterns of gene expression in spermatogenical cells (Kleene, 2001, 2003).

Mammalian and invertebrate testes use many mRNAs with shorter 3′ UTRs compared to other tissues (Liu et al., 2007; Smibert et al., 2012; Neve et al., 2017). 3′ UTR shortening promotes mRNA stability and translational effectivity. It has been demonstrated that in mice, shortening the 3′ UTR eliminates destabilizing elements, such as AU-rich elements and transposable elements located downstream of testis-specific APA (tsAPA) in 3′ UTR (Li W. et al., 2016).

Since tsAPA is not used efficiently in somatic cells, it has been suggested that there are testis-specific factors involved in the cleavage and polyadenylation process in non-canonical proximal APAs, such as the testis-specific form of CStF-64 (Wallace et al., 1999). The 64-kDa subunit of the CStF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences the cleavage site choice (MacDonald et al., 1994; MacDonald and McMahon, 2010). Somatic and testis-specific forms of CStF-64 have been found (Wallace et al., 1999). The CStF-64 gene for the somatic form is encoded on the X chromosome in both mice and humans. This suggests that the somatic form of CStF-64 is absent in meiotic cells because of the X chromosome inactivation, whereas the autosomal *CStf2t* gene encodes the testis-specific form τCStF-64 (Das et al., 2007). The τCStF-64 is found in spermatocytes and early spermatids (Wallace et al., 1999).

Alternative promoters in the intron 3 of the *Drosophila melanogaster sbr (Dm nxf1)* gene drive the expression of two testis-specific mRNA isoforms without exons 1, 2, and 3. 5′ UTRs of these mRNAs isoforms contain either uORFs, or a small intron (Ginanova et al., 2016). Both features of the 5′ UTR are known to control translation efficiency of the corresponding mRNAs in mammals (Mededbach et al., 2011; Bicknell et al., 2012); their functionality in *Drosophila* is unknown. The testis-specific *sbr* mRNAs initiate the synthesis of the testis-specific short protein (tSBR), which excludes nuclear localization signals (NLS) present in the canonical SBR protein (Fig. 3). This suggests that tSBR is unable to get into the nucleus by itself and carries out some functions in the cytoplasm, likely participating in the biogenesis of long-living mRNAs during meiosis and spermiogenesis (Ginanova et al., 2016).

All *sbr* (*Dm nxf1)* mRNAs have a shortened 3′ UTR in testes (Ginanova et al., 2016) (Fig. 2). The 3′ UTR of the full-length *sbr* mRNAs has several cis-acting elements, such as AU-rich, alternative PASs, CPE, and the predicted target-sites for miRNAs. All of them are located downstream of the cleavage and polyadenylation sites, which are used during the processing of the testis-specific *sbr* mRNAs. The poly(U)-tract is among the cis-elements in the 3′ UTR of *sbr* (*Dm nxf1)* mRNAs, and remains in the shortened 3′ UTR of *sbr* (*Dm nxf1)* mRNAs. The polypyrimidine-tract-binding (PTB) protein 2 regulates meiotic male germ cell mRNAs in mice (Iguchi et al., 2006). The dmPTB plays an important role during spermatogenesis in *Drosophila* (Sridharan et al., 2016).

Testis-specific transcripts are unknown for the mammalian *Nxf1* gene. There are testis-specific paralogous genes among *Nxf* gene family in mice (Sasaki et al., 2005) and humans (Yang et al., 2001).

TRANSCRIPTS IN OOGENESIS

A subset of maternal mRNAs and proteins are synthesized during oogenesis and are retained in the oocyte to direct the first mitotic divisions and specify the patterning of the embryo. Early embryogenesis passes through a stage when developmental control is handed from maternally provided gene products to those synthesized from (by) the zygotic genome (Tadros and Lipshitz, 2009). The oocyte-to-embryo transition (OET) can be subdivided into two interrelated processes: first, a subset of maternal mRNAs and proteins is eliminated; then the transcription of the zygotic genome begins.

There are several pathways of maternal mRNA degradation. One of them is the ARE (AU-rich element)-mediated pathway known in *Xenopus laevis* (Voeltz and Steitz, 1998). The Embryonic Deadenylation Element Binding Protein (EDEN-BP) triggers the deadenylation of maternal transcripts upon fertilization via the recognition of AU-rich cis-elements. The activity of EDEN-BP is regulated by phosphorylation (Detivaud et al., 2003). This may explain the different activity of EDEN-BP in oocytes and early embryos. Another way to mediate the maternal mRNA degradation is accomplished through the binding of a miRNA (Guo et al., 2008; Lund et al., 2009).

Yet another way of post-transcriptional regulation of the maternal mRNAs in early embryogenesis is the

CPE-dependent mRNA polyadenylation and its consequent translation. CPEB-mediated mRNA silencing/re-activation is essential to oogenesis (Tay and Richter, 2001) and in developing hippocampal neurons (Kundel et al. 2009) in mice. Cytoplasmic polyadenylation is a major mRNA regulator during oogenesis and embryogenesis in *Drosophila* (Coll et al., 2010; Cui et al., 2013).

The sequence motifs in the 3′ UTRs are the "combinatorial code" allowing for a precise spatial-temporal regulation of mRNA translation during OET (Evsikov et al. 2006), and the keys to a specific code combination are provided by a combination and/or modifications of the corresponding trans-acting factors with protective or destructive functions (Svoboda et al., 2015; Schultz et al., 2018).

The greatest variety of the *sbr* mRNAs is found in the ovaries (Fig. 5). Moreover, the canonical isoform is not the most numerous. Several different promoters are used as transcription starts for different mRNAs *sbr*. Alternative promoters are located in exons 2, 3, and 4 in addition to the canonical one. The *sbr* expression patterns in early embryos (0–2 h) and in the ovaries are similar, according to the northern blot analysis (Ivankova et al., 2010). One may hypothesize that the majority of *sbr* mRNA isoforms are transcribed during oogenesis, and the maternal products of the *sbr* gene are necessary during early development of the embryo. The mutations of the *sbr* gene have a maternal effect on early development. The SBR protein is abundant in early embryos and marks the spindles of nuclear divisions (Golubkova et al., 2015). The fraction of the shorter *sbr* mRNAs that is present in embryos 1–2 h of age disappears when zygotic transcription begins (Ivankova et al., 2010). These results imply an essential role of the *sbr* gene in the early embryogenesis of *D. melanogaster*. Since the early embryogenesis of *D. melanogaster* occurs in the absence of a transcription, the functions of SBR (Dm NXF1) as a factor of nuclear export of mRNAs are unnecessary. Nuclear division in syncytial embryos of *D. melanogaster* is characterized by semiclosed mitosis within the nuclear membranes, which are only disrupted at the poles adjacent to the centrioles (Stafstrom and Staehelin, 1984). The NPC (nuclear pore complex) assembly and disassembly at the site of fusion between the mitotic nuclear envelope and the overlying spindle membrane are detected during early synchronous mitosis in *D. melanogaster* embryos (Kiseleva et al., 2001). The predicted truncated products of the *sbr* gene during oogenesis can have alternative functions. The *sbr* gene products participate in complicated processes that can be relatively independent from transcription. These processes are provided by the cyclic translation of mRNAs of factors that are involved in DNA replication, nuclear divisions, and cytoskeleton dynamics.

A variety of SBR isoforms allow each of them to be part of a particular RNP complex that includes mRNAs for nuclear divisions in early embryos. mRNAs with different functional and regulatory capabilities require specialized RBPs. SBR (Dm NXF1) interacts with the subunits of the origin recognition complex, which associates with replication origins and initiates the pre-replication complex assembly (Kopitova et al., 2016). The *sbr* gene is essential for nuclear divisions (Golubkova and Mamon, 2012) and the cytoskeleton (Mamon et al., 2017). Since the RNA-binding protein SBR is found not only in the nucleus or nuclear envelope but also in the cytoplasm, it may be part of RNP complexes storing organ-specific long-living mRNAs. The nucleoporin-binding domains in the truncated SBR isoforms allow for the localization of the corresponding RNP complexes in the nuclear envelope (Mamon et al., 2017).

The *sbr* mRNA diversity in ovaries reflects the specificity of the early embryonic development of *D. melanogaster*. Various maternal mRNAs are involved in control of the synchronous nuclear divisions. The specific RNA-binding proteins are needed for functioning of different mRNA. This also applies to the SBR proteins.

Conclusion

A comparison of the organisms of different taxa has demonstrated that the complexity of organization during evolution arises not due to an increase in the number of protein-coding genes. One of the reasons is alternative transcript processing, which generates functional diversity from the same or similar number of genes. A variety of mechanisms providing for the complexity of the transcriptome ensures a precise and coordinated regulation of organ-specific functions through a combination of cis-acting elements and trans-acting factors. The *D. melanogaster sbr* (*Dm nxf1)* gene has proven to be an excellent model for investigating these mechanisms, potentially leading to the emergence of multiple products with various functions.

References

An, J. J., Gharami, K., Liao, G.-Y., Woo, N. H., Lau, A. G., Vanevski, F., Torre, E. R., Jones, K. R., Feng, Y., Lu, B., and Xu, B. 2008. Distinct role of long 3'UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell* 134(1):175–187. https://doi.org/10.1016/j.cell.2008.05.045

Barnard, D. C., Cao, Q., and Richter, J. D. 2005. Differential phosphorylation controls Maskin association with eukaryotic translation initiation factor 4E and localization on the mitotic apparatus. *Molecular and Cellular Biology* 25(17):7605–7615. https://doi.org/10.1128/MCB.25.17.7605-7615.2005

Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J.-M., and Gautheret, D. 2000. Patterns of variant polyadenylation signal usage in human genes. *Genome Research* 10:1001–1010. https://doi.org/10.1101/gr.10.7.1001

Berkovits, B. D. and Mayr, C. 2015. Alternative 3´ UTRs act as scaffolds to regulate membrane protein localization. *Nature* 522:363–367. https://doi.org/10.1038/nature14321

Bicknell, A. A., Cenik, C., Chua, H. N., Roth, F. P., and Moore, M. J. 2012. Intron in UTRs: why we should stop ignoring them. *Bioessays* 34:1025–1034. https://doi.org/10.1002/bies.201200073

Black, D. L. 2003. Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry* 72:291–336. https://doi.org/10.1146/annurev.biochem.72.121801.161720

Braunschweig, U., Barbosa-Morais, N. L., Pan, Q., Nachman, E. N., Alipanahi, B., Gonatopoulos-Pournatzis, T., Frey, B., Irimia, M., and Blencowe, B. J. 2014. Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Research* 24:1774–1786. https://doi.org/10.1101/gr.177790.114

Brown, J. B., Boley, N., Eisman, R., May, G. E., Stoiber, M. H., Duff, M. O., Booth, B. W., Wen, J., Park, S., Suzuki, A. M., Wan, K. H., Yu, C., Zhang, D., Carlson, J. W., Cherbas, L., Eads, B. D., Miller, D., Mockaitis, K., Roberts, J., Davis, C. A., Frise, E., Hammonds, A. S., Olson, S., Shenker, S., Sturgill, D., Samsonova, A. A., Weiszmann, R., Robinson, G., Hernandez, J., Andrews, J., Bickel, P. J., Carninci, P., Cherbas, P., Gingeras, T. R., Hoskins, R. A., Kaufman, T. C., Lai, E. C., Oliver, B., Perrimon, N., Graveley, B. R., and Celniker, S. E. 2014. Diversity and dynamics of the *Drosophila* transcriptome. *Nature* 512:393–399. https://doi.org/10.1038/nature12962

Child, S. J., Miller, M. K., and Geballe, A. P. 1999. Translational control by an upstream open reading frame in the HER-2/*neu* transcript. *Journal of Biological Chemistry* 274:24335–24341. https://doi.org/10.1074/jbc.274.34.24335

Colgan, D. F. and Manley, J. L. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes and Development* 11:2755–2766. https://doi.org/10.1101/gad.11.21.2755

Coll, O., Villalba, A., Bussotti, G., Notredame, C., and Gebauer, F. 2010. A novel, canonical mechanism of cytoplasmic polyadenylation operates in *Drosophila* embryogenesis. *Genes and Development* 24:129–134. https://doi.org/10.1101/gad.568610

Cui, J., Sartain, C. V., Pleiss, J., and Wolfner, M. F. 2013. Cytoplasmic polyadenylation is a major mRNA regulator during oogenesis and egg activation in *Drosophila*. *Developmental Biology* 383:121–131. https://doi.org/10.1016/j.ydbio.2013.08.013

Dass, B., Tardif, S., Park, J. Y., Tian, B., Weitlauf, H. M., Hess, R. A., Carnes, K., Griswold, M. D., Small, C. L., and MacDonald, C. C. 2007. Loss of polyadenylation protein τCstF-64 causes spermatogenic defects and male infertility. *Proceeding of the National Academy of Science of USA* 104:20374–20379. https://doi.org/10.1073.pnas.0707589104

Derti, A., Garrett-Engele, P., MacIsaac, K. D., Stevens, R. C., Sriram, S., Chen, R., Rohl, C. A., Johnson, J. M., and Babak, T. 2012. A quantitative atlas of polyadenylation in five mammals. *Genome Research* 22(6):1173–1183. https://doi.org/10.1101/gr.132563.111

Detivaud, L., Pascreau, G., Karaiskou, A., Osborne, H. B., and Kubiak, J. Z. 2003. Regulation of EDEN-dependent deadenylation of Aurora A/Eg2-derived mRNA via phosphorylation and dephosphorylation in *Xenopus laevis* egg extracts. *Journal of Cell Science* 116:2697–2705. https://doi.org/10.1242/jcs.00477

Du, L. and Richter, J. D. 2005. Activity-dependent polyadenylation in neurons. *RNA* 11:340–1347. https://doi.org/10.1261/rna.2870505

Elkon, R., Drost, J., van Haaften, G., Jenal, M., Schrier, M., Oude Vrielink, J. A., and Agami, R. 2012. E2F mediates enhanced alternative polyadenylation in proliferation. *Genome Biology* 13:R59. https://doi.org/10.1186/gb-2012-13-7-r59

Erson-Bensan, A. E. 2016. Alternative polyadenylation and RNA-binding proteins. *Journal of Molecular Endocrinology* 57:F29–F34. https://doi.org/10.1530/JME-16-0070

Evsikov, A. V. and de Evsikova, C. M. 2009. Gene expression during the oocyte-to-embryo transition in mammals. *Molecular Reproduction and Development* 76(9):805–818. https://doi.org/10.1002/mrd.21038

Evsikov, A. V., Graber, J. H., Brockman, J. M., Hampl, A., Holbrook, A. E., Singh, P., Eppig, J. J., Solter, D., and Knowles, B. B. 2006. Cracking the egg: molecular dynamics and evolutionary aspects of the transition from the fully grown oocyte to embryo. *Genes and Development* 20:2713–2727. https://doi.org/10.1101/gad.1471006

Flynt, A. S. and Lai, E. C. 2008. Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nature Reviews Genetics* 9:831–842. https://doi.org/10.1038/nrg2455

Ginanova, V., Golubkova, E., Kliver, S., Bychkova, E., Markoska, K., Ivankova, N., Tretyakova, I., Evgen'ev, M., and Mamon, L. 2016. Testis-specific products of the *Drosophila melanogaster sbr* gene, encoding nuclear export factor 1, are necessary for male fertility. *Gene* 577:153–160. https://doi.org/10.1016/j.gene.2015.11.030

Ginanova V. 2017. Raznoobraznyye i tkanespetsifichnyye izoformy Nxf1 (yadernogo faktora eksporta 1), osnovnogo faktora yadernogo eksporta mRNK u *Drosophila melanogaster* [Variety and tissue-specific isoforms of the Nxf1 (nuclear export factor 1), a main factor of nuclear export of mRNA in *Drosophila melanogaster*]. PhD thesis. St.-Petersburg.

Golubkova, E. V., Atsapkina, A. A., and Mamon, L. A. 2015. The role *of sbr / Dm nxf1* gene in syncytial development in *Drosophila melanogaster*. *Cell and Tissue Biology* 9(4):271–283. https://doi.org/10.1134/S1990519X15040057

Golubkova, E. V. and Mamon, L. A. 2012. The role of Dm NXF1 in controlling early embryonic mitoses in *Drosophila melanogaster*. *Encyclopedia of Cell Biology Research* 543–549. Nova Science Publishers. Inc.

Graveley, B. R. 2001. Alternative splicing: increasing diversity in the proteomic world. *Trends in Genetics* 17:100–107. https://doi.org/10.1016/S0168-9525(00)02176-4

Graveley, B. R. 2005. Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. *Cell* 123:65–73. https://doi.org/10.1016/j.cell.2005.07.028

Guo, X., Gourronc, F., Audic, Y., Lyons-Levy, G., Mitchell, T., and Hartley, R. S. 2008. ElrA and AUF1 differentially bind cyclin B2 mRNA. *Biochemical and Biophysical Research Communications* 377:653–657. https://doi.org/10.1016/j.bbrc.2008.10.029

Hawthorne, S. K., Busanelli, R. R., and Kleen, K. C. 2006. The 5´ UTR and 3´ UTR of the sperm mitochondria-associated cysteine-rich protein mRNA regulate translation in spermatids by multiple mechanisms in transgenic mice. *Developmental Biology* 297:118–126. https://doi.org/10.1016/j.ydbio.2006.04.468

Herold, A., Klymenko, T., and Izaurralde, E. 2001. NXF1/ p15 heterodimers are essential for mRNA nuclear export in *Drosophila*. *RNA* 7:1768–1780. https://doi.org/10.1017.S1355838201013565

Hilgers, V. 2015. Alternative polyadenylation coupled to transcription initiation: Insights from ELAV-mediated 3′ UTR extension. *RNA Biology* 12(9):918–921. https://doi.org/10.1080/15476286.2015.1060393

Hilgers, V., Lemke, S. B., and Levine, M. 2012. ELAV mediates 3′ UTR extension in the *Drosophila* nervous system. *Genes and Development* 26:2259–2264. https://doi.org/10.1101/gad.199653.112

Hilgers, V., Perry, M. W. Hednrix, D., Stark, A., Levine, M., and Haley, B. 2011. Neural-specific elongation of 3' UTRs during *Drosophila* development. *Proceeding of the National Academy of Science of USA* 108:15864–15869. https://doi.org/10.1073/pnas.1112672108

Iguchi, N., Tobias, J. W., and Hecht, N. B. 2006. Expression profiling reveals meiotic male germ cell mRNAs that are translationally up- and down-regulated. *Proceeding of the National Academy of Science of USA* 103:7712–7717. https://doi.org/10.1073/pnas.0510999103

Ivankova, N., Tretyakova, I., Lyozin, G., Avanesyan, E., Zolotukhin, A., Zatsepina, O. G., Evgen'ev, M. B., and Mamon, L. A. 2010. Alternative transcripts expressed by *small bristles,* the *Drosophila melanogaster nxf1* gene. *Gene* 458:11–19. https://doi.org/10.1016/j.gene.2010.02.013

Iwakawa, H. O. and Tomari, Y. 2015. The functions of microRNAs: mRNA decay and translational repression. *Trends in Cell Biology* 25:651–665. https://doi.org/10.1016/j.tcb.2015.07.011

Jacob, A. G. and Smith, C. W. 2017. Intron retention as a component of regulated gene expression programs. *Human Genetics* 136:1043–1057. https://doi.org/10.1007/s00439-017-1791-x

Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour, C. D., Santos, R., Schadt, E. E., Stoughton, R., and Shoemaker, D. D. 2003. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302:2141–2144. https://doi.org/10.1126/science.1090100

Jun, L., Frints, S., Duhamel, H., Herold, A., Abad-Rodrigues, J., Dotti, C., Izaurralde, E., Marynen, P., and Froyen, G. 2001. NXF5, a novel member of the nuclear RNA export factor family, is lost in a male patient with a syndromic form of mental retardation. *Current Biology* 11(18):1381–1391. https://doi.org/10.1016/S0960-9822(01)00419-5

Kadonaga, J. T. 2012. Perspectives on the RNA polymerase II core promoter. *Wiley Interdisciplinary Reviews-Developmental Biology* 1(1):40–51. https://doi.org/10.1002/wdev.21

Kadonaga, J. T. 2002. The DPE, a core promoter element for transcription by RNA polymerase II. *Experimental and Molecular Medicine* 34:259–264. https://doi.org/10.1038/emm.2002.36

Kashiwabara, S., Zhuang, T., Yamagata, K., Noguchi, J., Fukamizu, A., and Baba, T. 2000. Identification of a novel isoform of poly(A) polymerase, TPAP, specifically present in the cytoplasm of spermatogenic cells. *Developmental Biology* 228:106–115. https://doi.org/10.1006/dbio.2000.9894

Kashiwabara, S, Noguchi, J, Zhuang, T, Ohmura, K, Honda, A, Sugiura, S, Miyamoto, K, Takahashi, S, Inoue, K, Ogura, A, and Baba, T. 2002. Regulation of spermatogenesis by testis-specific, cytoplasmic poly(A) polymerase TPAP. *Science* 298:1999–2002. https://doi.org/10.1126/science.1074632

Kiseleva, E., Rutherford, S., Cotter, L. M., Allen, T. D., and Goldberg, M. W. 2001. Step of nuclear pore complex disassembly and reassembly during mitosis in early *Drosophila* embryos. *Journal of Cell Science* 114:3607–3618.

Kleene, K. C. 2001. A possible meiotic function of the peculiar patterns of gene expression in mammalian spermatogenic cells. *Mechanisms of Development* 106:3–23. https://doi.org/10.1016/S0925-4773(01)00413-0

Kleene, K. C. 2003. Patterns, mechanisms, and functions of translation regulation in mammalian spermatogenic cells. *Cytogenetic and Genome Research* 103:217–224. https://doi.org/10.1159/000076807

Kopytova, D., Popova, V., Kurshakova, M., Shidlovskii, Y., Nabirochkina, E., Brechalov, A., Georgiev, G., and Georgieva, S. 2016. ORC interacts with THSC/TREX-2 and its subunits promote Nxf1 association with mRNP and mRNA export in *Drosophila*. *Nucleic Acids Research* 44:4920–4933. https://doi.org/10.1093/nar/gkw192

Kumar, A. 2009. An overview of nested genes in eukaryotic genomes. *Eukaryotic Cell* 8:1321–1329. https://doi.org/10.1128/EC.00143-09

Kundel, M., Jones, K. J., Shin, C. Y., and Wells, D. G. 2009. Cytoplasmic polyadenylation element binding protein regulates neurotrophin 3-dependent β-catenin mRNA translation in developing hippocampal neurons. *Journal of Neuroscience* 29:13630–13639. https://doi.org/10.1523/JNEUROSCI.2910-08.2009

Li, J. and Gilmour, D. S. 2013. Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *The EMBO Journal* 32:1829–1841. https://doi.org/10.1038/emboj.2013.111

Li, W., Park, J. Y. Zheng, D., Hoque, M. Yehia, G., and Tian, B. 2016. Alternative cleavage and polyadenylation in spermatogenesis connects chromatin regulation with post-transcriptional control. *BMC Biology* 14:6. https://doi.org/10.1186/s12915-016-0229-6

Li, Y., Bor, Y. C., Fitzgerald, M. P., Lee, K. S., Rekosh, D., and Hammarskjold, M. L. 2016. An NXF1 mRNA with a retained intron is expressed in hippocampal and neocortical neurons and is translated into a protein that functions as an Nxf1 cofactor. *Molecular Biology of the Cell* 27:3903–3912. https://doi.org/10.1091/mbc.E16-07-0515

Li, Y., Bor, Y. C., Misawa, Y., Xue, Y., Rekosh, D., and Hammarskjöld, M. L. 2006. An intron with a constitutive transport element is retained in a Tap messenger RNA. *Nature* 443:234–237. https://doi.org/10.1038/nature05107

Liu, D., Brockman, J. M., Dass, B., Hutchins, L. N., Singh, P., McCarrey, J. R., MacDonald, C. C., and Graber, J. H. 2007. Systematic variation in mRNA 3'-processing signals during mouse spermatogenesis. *Nucleic Acids Research* 35:234–246. https://doi.org/10.1093/nar/gkl919

López de Salines, I., Zhan, M., Lal, A., Yang, X., and Gorospe, M. 2004. Identification of a target RNA motif for RNA-binding protein HuR. *Proceeding of the National Academy of Science of USA* 101:2987–2992. https://doi.org/10.1073.pnas.0306453101

Lund, E., Liu, M., Hartley, R. S., Sheets, M. D., and Dahlberg, J. S. 2009. Deadenylation of maternal mRNAs mediated by miR-427 in *Xenopus laevis* embryos. *RNA* 15:2351–2363. https://doi.org/10.1261/rna.1882009

MacDonald, C. C. and McMahon, K. W. 2010. Tissue-specific mechanisms of alternative polyadenylation: testis, brain, and beyond. *Wiley Interdisciplinary Reviews-RNA* 1:494–501. https://doi.org/10.1002/wrna.29

MacDonald, C. C., Wilusz, J., and Shenk, T. 1994. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Molecular and Cellular Biology* 14:6647–6654. https://doi.org/10.1128/MCB.14.10.6647

Mamon, L. A., Kliver, S. F., and Golubkova, E. V. 2013. Evolutionarily conserved features of the retained intron in alternative transcripts of the *nxf1* (Nuclear eXport Factor) genes in different organisms. *Open Journal of Genetics* 3:159–170. https://doi.org/10.4236/ojgen.2013.33018

Mamon, L. A., Kliver, S. F., Prosovskaya, A. O., Ginanova, V. R., and Golubkova, Ye. V. 2014. The intron-containing transcript: an evolutionarily conserved characteristic of the genes orthologous to *nxf1* (nuclear export factor 1). ISSN2079-0597. *Russian Journal of Genetics and Applied Research* 4(5):434–443. https://doi.org/10.1134/S2079059714050104

Mamon, L. A., Ginanova, V. R., Kliver, S. F., Yakimova, A. O., Atsapkina, A. A., and Golubkova, E. V. 2017. RNA-binding proteins of the NXF (nuclear export factor) family and their connection with the cytoskeleton. *Cytoskeleton* 74:161–169. https://doi.org/10.1002/cm.21362

Medenbach, J., Seiler, M., and Hentze, M. W. 2011. Translational control via protein-regulated upstream open reading frames. *Cell*. 145:902–913. https://doi.org/10.1016/j.cell.2011.05.005

Mitra, M., Johnson, E. L., and Coller, H. A. 2015. Alternative polyadenylation can regulate post-translational membrane localization. *Trends in Cell and Molecular Biology* 10:37–47.

Miura, P., Sanfilippo, P., Shenker, S., and Lai, E. C. 2014. Alternative polyadenylation in the nervous system: to what lengths with 3´ UTR extensions take us? *BioEssays* 36:766–777. https://doi.org/10.1002/bies.201300174

Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J. O., and Lai, E. C. 2013. Widespread and extensive lengthening of 3´ UTRs in the mammalian brain. *Genome Research* 23:812–825. https://doi.org/10.1101/gr.146886.112

Neilson, J. R. and Sandberg, R. 2010. Heterogeneity in mammalian RNA 3' end formation. *Experimental Cell Research* 316:1357–1364. https://doi.org/10.1016/j.yexcr.2010.02.040

Neve, J., Patel, R., Wang, Z., Louey, A., and Furger, A. M. 2017. Cleavage and polyadenylation: Ending the message expands gene regulation: Ending the message expands gene regulation. *RNA Biology* 14:865–890. https://doi.org/10.1080/15476286.2017.1306171

Nilsen, T. W. and Graveley, B. R. 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463:457–463. https://doi.org/10.1038/nature08909

Oktaba, K., Zhang, W., Lotz, T. S., Jun, D. J., Lemke, S. B., Ng, S. P., Esposito, E., Levine, M., and Hilgers, V. 2015. ELAV links paused Pol II to alternative polyadenylation in the *Drosophila* nervous system. *Molecular Cell* 57(2):341–348. https://doi.org/10.1016/j.molcel.2014.11.024

Perales, R. and Bentley, D. 2009. "Co-transcriptionality" — the transcription elongation complex as a nexus for nuclear transactions. *Molecular Cell* 36:178–191. https://doi.org/10.1016/j.molcel.2009.09.018

Proudfoot, N. J. 2011. Ending the message: poly(A) signals then and now. *Genes and Development* 25:1770–1782. https://doi.org/10.1101/gad.17268411

Richter, J. D. 1996. Dynamics of poly(A) addition and removal during development. pp 481-503 in J. W. B. Hershey, M. B. Mathews, and N. Sonenberg (eds) "Translational control". Cold Spring Harbor Laboratory Press. Plainview. NY.

Richter, J. D. and Sonenberg, N. 2005.Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433:477–480. https://doi.org/10.1038/nature03205

Sanfilippo, P., Wen, J., and Lai, E. C. 2017. Landscape and evolution of tissue-specific alternative polyadenylation across Drosophila species. *Genome Biology* 18:229. https://doi.org/10.1186/s13059-017-1358-0

Sasaki, M., Takeda, E., and Takano, K. 2005. Molecular cloning and functional characterization of mouse *Nxf* family gene products. *Genomics* 85:641–653. https://doi.org/10.1016/j.ygeno.2005.01.003

Schmitz, U., Pinello, N., Jia, F., Alasmari, S., Ritchie, W., Keightley, M.-C., Shini, S., Lieschke, G. J., Wong, J. J.-L., and Rasko, J. E. J. 2017. Intron retention enhances gene regulatory complexity in vertebrates. *Genome Biology* 18:216. https://doi.org/10.1186/s13059-017-1339-3

Schmucker, D., Clemens, J. C., Shu, H., Worby, C. A., Xiao, J., Muda, M., Dixon, J. E., and Zipursky, S. L. 2000. *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* 101:671–684. https://doi.org/10.1016/S0092-8674(00)80878-8

Schultz, R. M., Stein, P., and Svoboda, P. 2018. The oocyte to embryo transition in mouse: past, present, and future. *Biology of Reproduction* 99:160–174. https://doi.org/10.1093/biolre/ioy013

Smibert, P., Miura, P., Westholm, J. O., Shenker, S., May, G., Duff, M. O., Zhang, D., Eads, B. D., Carlson, J., Brown, J. B., Eisman, R. C., Andrews, J., Kaufman, T., Cherbas, P., Celniker, S. E., Graveley, B. R., and Lai, E. C. 2012. Global patterns of tissue-specific alternative polyadenylation in *Drosophila*. *Cell Reports* 1:277–289. https://doi.org/10.1016/j.celrep.2012.01.001

Sridharan V., Heimiller J., Robida M. D., and Singh, R. 2016. High throughput sequencing identifies misregulated genes in the *Drosophila* polypyrimidine tract-binding protein (*hephaestus*) mutant defective in spermatogenesis. *PLoS One* 11(3):e0150768. https://doi.org/10.1371/journal.pone.0150768

Stafstrom, J. P. and Staehelin, L. A. 1984. Dynamics of the nuclear envelope and of nuclear pore complexes during mitosis in the *Drosophila* embryo. *European Journal of Cell Biology* 34:179–189.

Svoboda, P., Franke, V., and Schultz, R. M. 2015. Chapter Nine — Sculpting the Transcriptome During the Oocyte-to-Embryo Transition in Mouse. *Current Topics in Developmental Biology* 113:305–349. https://doi.org/10.1016/bs.ctdb.2015.06.004

Szostak, E. and Gebauer, F. 2012. Translational control by 3´ -UTR-binding proteins. *Briefings Functional Genomics* 12:58–65. https://doi.org/10.1093/bfgp/els056

Tadros, W. and Lipshitz, H. D. 2009. The maternal-to-zygotic transition: a play in two acts. *Development* 136:3033–3042. https://doi.org/10.1242/dev.033183

Tay, J. and Richter, J. D. 2001. Germ cell differentiation and synaptonemal complex formation are disrupted in CPEB knockout mice. *Developmental Cell* 1:201–213. https://doi.org/10.1016/S1534-5807(01)00025-9

Tian, B., Hu, J., Zhang, H., and Lutz, C. S. 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Research* 33:201–212. https://doi.org/10.1093/nar/gki158

Tretyakova, I., Zolotukhin, A. S., Tan, W., Bear, J., Propst, F., Rothler, G., and Fleber, B. K. 2005. Nuclear export factor family protein participates in cytoplasmic mRNA trafficking. *Journal of Biological Chemistry* 280(36):31981–31990. https://doi.org/10.1074/jbc.M502736200

Tuck, A. C. and Tollervey, D. 2011. RNA in pieces. *Trends in Genetics* 27:422–432. https://doi.org/10.1016/j.tig.2011.06.001

Vacik, T. and Raska, I. 2017. Alternative intronic promoters in development and disease. *Protoplasma* 254:1201–1206. https://doi.org/10.1007/s00709-016-1071-y

Voeltz, G. K. and Steitz, J. A. 1998. AUUUA sequences direct mRNA deadenylation uncoupled from decay during *Xenopus* early development. *Molecular and Cellular Biology* 18:7537–7545. https://doi.org/10.1128/MCB.18.12.7537

Wallace, A. M., Dass, B., Ravnik, S. E., Tonk, V., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., and MacDonald, C. C. 1999. Two distinct forms of the 64,000 M_r protein of the cleavage stimulation factor are expressed in mouse male germ cells. *Proceeding of the National Academy of Science of USA* 96:6763–6768. https://doi.org/10.1073/pnas.96.12.6763

Wang, B., Rekosh, D., and Hammarskjold, M.-L. 2015. Evolutionary conservation of machinery for export and expression of mRNAs with retained introns. *RNA* 21:1–12. https://doi.org/10.1261/rna.048520.114

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456:470–476. https://doi.org/10.1038/nature07509

Wells, S. E., Hillner, P. E., Vale, R. D., and Sachs, A. B. 1998. Circularization of mRNA by eukaryotic translation initiation factors. *Molecular Cell* 2:135–140. https://doi.org/10.1016/S1097-2765(00)80122-7

White, E. J. F., Brewer, G., and Wilson, G. M. 2013. Post-transcriptional control of gene expression by AUF1: Mechanisms, physiological targets, and regulation. *Biochimica et Biophysica Acta* 1829:680–688. https://doi.org/10.1016/j.bbagrm.2012.12.002

Xu, M., Gonzalez-Hurtado, E., and Martinez, E. 2016. Core promoter-specific gene regulation: TATA box selectivity and Initiator-dependent bi-directionality of serum response factor activated transcription. *Biochimica et Biophysica Acta* 1859:553–563. https://doi.org/10.1016/j.bbagrm.2016.01.005

Yang, J., Bogert, H. P., Wang, P. J., Page, D. C., and Cullen B. R. 2001. Two Closely Related Human Nuclear Export Factors Utilize Entirely Distinct Export Pathways. *Molecular Cell* 8:397–406. https://doi.org/10.1016/S1097-2765(01)00303-3

Zhou, J., Pan, J., Eckardt, S., Leu, N. A., McLaughlin, K. J., and Wang, P. J. 2011. Nxf3 is expressed in Sertoli cells, but is dispensable for spermatogenesis. *Molecular Reproduction and Development* 78(4):241–249. https://doi.org/10.1002/mrd.21291