An Alien Promoter Capture as a Primary Step of the Evolution of Testes-Expressed Repeats in the Drosophila melanogaster Genome

Lev A. Usakin, Galina L. Kogan, Alla I. Kalnykova, and Vladimir A. Gvozdev

Department of Animal Molecular Genetics, Institute of Molecular Genetics, Moscow, Russia

Fertility of Drosophila melanogaster males is impaired due to the disruption of the silencing of the X-linked, testis-expressed, repeated Stellate (Ste) genes. Ste silencing is mediated by symmetric transcription of the paralogous Y-linked repeats and exerted by an RNA interference (RNAi) mechanism. Here we present a scenario for the origin of the Ste genes and their suppressors. The primary intermediate of their evolution emerged as a result of the acquisition of a preformed alien, testis-specific promoter. This intermediate is identified as a chimeric gene containing coding region of an autosomal gene for testis-specific protein kinase CK2. The 5' region of the chimeric has been acquired from a member of a family of testis-expressed X-linked genes of unknown function. We propose that the evolution and amplification of the novel chimeric gene have led to the overproduction of the regulatory CK2 subunit in testes. The evolution of the Y-linked descendants of the primary intermediate resulted in the RNAi-mediated suppression of excessive expression of the X-linked paralogs. The newly detected "dead family" of cognate repeats on the Y chromosome has contributed to the evolution of Ste and its suppressors via gene conversion. Our results show that RNAi silencing, considered as a defense against viruses and transposable elements, may be involved in the evolution of eukaryotic genomes.

Introduction

The question of formation and evolution of new genes is the subject of a number of recent intriguing reports (Long et al. 2003). The study of the origin of the Drosophila melanogaster genome of the male fertility system of Stellate (Ste) repeats and their suppressors illustrates the new toolkits in eukaryotic genome evolution, including acquisition of a preformed alien promoter and insertion of a transposable element. These events initiate RNA interference (RNAi) silencing, beneficial for population maintenance and survival of newly evolved genes.

An unusual example of eukaryotic tandem repeats is represented by the paralogs Ste and Suppressor of Stellate (Su(Ste)) tandem repeats localized on the X and Y chromosomes, respectively, in the D. melanogaster genome (Hardy et al. 1984; Livak 1990). The variants of the Ste repeats are located in the euchromatin and constitutive heterochromatin of the X chromosome (Shevleyov 1992; Palumbo et al. 1994; Tulin et al. 1997). Ste repeats produce a protein with extensive homology to the regulatory β-subunit of protein kinase CK2 (Livak 1990; Bozzetti et al. 1995). Testis-specific overexpression of the Ste genes causes multiple meiotic abnormalities, and their silencing is required to maintain male fertility (Hardy et al. 1984; Bozzetti et al. 1995). The Su(Ste) repeats are thought to be suppressors of Ste expression because their deletion causes overexpression of the Ste genes (Hardy et al. 1984; Livak 1990). Su(Ste) repeats containing damaged open reading frame (ORF) (Kalnykova, Dobritsa, and Gvozdev 1997) maintain sense transcription and carry a hoppel transposon (Balakireva et al. 1992) that provides antisense transcription through these repeats (Aravin et al. 2001). The formation of double-stranded Su(Ste) RNA that is processed into small interfering RNA (siRNA) in testes by the RNAi machinery is the key point of Ste silencing (Aravin et al. 2001).

It was reported earlier that the autosomal testis-specific βCK2tes gene encoding a putative regulatory subunit of protein kinase CK2 may be considered a precursor of the Ste-Su(Ste) family (Kalnykova et al. 1997). The βCK2tes gene shows close homology and exon-intron structure similar to the Ste-Su(Ste) repeats but has a nonhomologous promoter region (Kalnykova et al. 1997). We have identified the primary intermediate in the course of the Ste-Su(Ste) repeat evolution, which has been spawned by the acquisition of a preformed testis-specific promoter of an alien gene by a copy of the autosomal βCK2tes gene. We propose that the illegitimate expression of the newly emerged gene was suppressed by the evolving of RNAi-silencing system as a result of transposon insertion and gene amplification on the sex chromosomes.

Materials and Methods

Southern and northern blotting was performed as described earlier (Kalnykova et al. 1997; Tulin et al. 1997). The Batumi strain of D. melanogaster (Novosibirsk, Institute of Cytology and Genetics) as well as Drosophila species obtained from Tucson Center (Drosophila simulans, wild-type, Australia; Drosophila mauritiana, wild-type, North America; Drosophila yakuba, wild-type, Africa; Drosophila sechellia, wild-type, Seychelles; Drosophila teissieri, wild-type, Cameroon) were used for Southern blotting experiments. DNA samples from females were digested with EcoRI. For northern blotting, strain Df(1)w67c23(2)y was used. 32P-labeled probes to βNACtes genes for Southern and northern blotting were obtained using polymerase chain reaction (PCR) product generated with primers 5′-cagatctgtgagaatgg were used in PCR analysis of DNA isolated from males and females.

For detection of the AE003039 scaffold, the primers 5′-cagctgctgacattcacaacggaac and 5′-tagctgctgacattcacaacggaac were used in PCR analysis of DNA isolated from males and females.

The third release of D. melanogaster genome sequence was used for computational analysis (Celniker et al. 2002; Hoskins et al. 2002).

Sequences used correspond to the following genes: βNACtes1 (CG13402 in Flybase annotation), βNACtes2...
(CG18157), \(\beta\)NACtes3 (CG9404), \(\beta\)NACtes4 (CG18313), \(\beta\)NACtes5 (pseudogene, 171,915–173,074 bp of AE003499.3 scaffold), \(\beta\)NACtes6 (CG32598), AE003039 scaffold, \(\beta\)CK2tes (L49382 [Kalmykova et al. 1997]), euchromatic Ste (X15899 [Livak 1990]), heterochromatic Ste (X97134 [Tulin et al. 1997]) and Su(Ste) (Z11734 [Kalmykova, Dobritsa, and Gvozdev 1998]) repeats, and Dm665 clone (Danilevskaya et al. 1991). Pseudo-\(\beta\)CK2tes repeats (PCKR) were extracted from two \(D.\) melanogaster unmapped contigs, AABU01001632 and AABU01001889 (see Supplementary fig. 1, Supplementary Material online).

Expressed sequence tag (EST) data used for identification of putative transcriptional start sites for \(\beta\)NACtes and Ste-Su(Ste) are as follows: 5'EST for \(\beta\)NACtes—BF497602, BF502865 and BF502998; EST for Su(Ste)—L42286, L42287 and L42288 (Kalmykova, Dobritsa, and Gvozdev 1998); and EST for Ste—cDNA1-6 (not submitted to GenBank) (Livak 1990).

**Results**

A BLAST search of the \(D.\) melanogaster genome databases revealed homologous X-linked genes containing putative promoter regions similar to those of the Ste-Su(Ste) genes (fig. 1A). Five copies of these genes with intact ORFs are located in the 12DE region on the polytenic chromosome map, near the euchromatic Ste cluster (fig. 1B). The sixth copy is a damaged pseudogene located in the 13A region. The databases contain testis ESTs corresponding to these genes. The mentioned genes will be referred to hereafter as the \(\beta\)NACtes genes because their coding regions show strong similarity to the \(Drosophila\) bic gene, which is thought to encode the \(\beta\)-subunit of the nascent-associated complex capable of binding to a newly synthesized peptide on a ribosome (Wiedmann et al. 1994; Markesich et al. 2000).

The alignment of putative promoter sequences in the \(\beta\)NACtes genes shows a 190-bp stretch with 95% identity to the Ste-Su(Ste) promoter region (fig. 2, red letters). EST data (see Methods) and previous studies (Livak 1990; Kalmykova, Dobritsa, and Gvozdev 1998) allowed us to infer transcription starts for \(\beta\)NACtes and Ste-Su(Ste) that are indicated in figure 2 by white letters among violet and cyan. The 3' -end of the shared 190-bp region (called hereafter the conservative promoter, CP) lies adjacent to the transcriptional start sites of both Ste-Su(Ste) and \(\beta\)NACtes. No homology is detected between the transcribed sequence of the \(\beta\)NACtes and Ste-Su(Ste) genes. Northern analysis validated the preferential testis-specific expression of the \(\beta\)NACtes genes (fig. 3A) as well as the Ste-Su(Ste) repeats (Livak 1990). This suggests a key role for CP in providing the testis-specific expression of Ste-Su(Ste) and \(\beta\)NACtes genes.

The Ste-Su(Ste) family is known to be peculiar to the \(D.\) melanogaster genome, whereas Southern analysis revealed the presence of \(\beta\)NACtes genes in the other \(D.\) melanogaster subgroup species (fig. 3B). This observation indicates that the \(\beta\)NACtes genes are more ancient than the Ste-Su(Ste) genes. Thus, in the course of the \(D.\) melanogaster genome evolution, the acquisition of a preformed \(\beta\)NACtes promoter by a Ste-Su(Ste) precursor has occurred.

In the unmapped \(D.\) melanogaster genome data set, the AE003039 scaffold contains a pseudogene that may be considered to be a primary chimeric intermediate in the course of the Ste-Su(Ste) repeat evolution. PCR analysis of DNA isolated from males and females allowed us to attribute the pseudogene to the Y chromosome (see Methods). The pseudogene contains the \(\beta\)NACtes-like promoter fused to damaged exon-intron structure similar to those of the Ste-Su(Ste)-\(\beta\)CK2tes genes. The CP of this chimeric pseudogene conserves several diagnostic nucleotide positions of the \(\beta\)NACtes genes (fig. 2, yellow highlighting) that are substituted in the Ste-Su(Ste) family of repeats. Moreover, the 5' region of the AE003039 scaffold contains a stretch of sequence peculiar to all \(\beta\)NACtes genes (fig. 2, gray letters) and a fragment of about 200 bp that is homologous to the sequence juxtaposed to the promoter of the \(\beta\)NACtes3 gene (fig. 2, yellow letters). These data suggest that the chimeric gene in the AE003039 scaffold has emerged as a result of nonhomologous recombination of the \(\beta\)NACtes3 promoter region with the exon-intron structure of the \(\beta\)CK2tes gene.

The peculiarities of the structures of the 3' regions of the chimeric \(\beta\)CK2tes and Ste-Su(Ste) genes reveal the details of likely evolutionary relations between these genes (fig. 4, Supplementary fig. 2, Supplementary Material online). The chimeric gene contains the regions designated A and C, whereas the Su(Ste) repeats contain the A and a fragment of C, but euchromatic Ste genes carry C region and vestiges of the A (fig. 4). The parental \(\beta\)CK2tes gene includes only the region A. These data suggest that the observed chimeric gene is an intermediate ancestor of the Ste-Su(Ste) genes.
The Su(Ste) repeats contain the B fragment that was detected as a separate sequence on the X chromosome adjacent to Ste cluster (fig. 1, Supplementary fig. 2, Supplementary Material online). This observation indicates the putative involvement of the X chromosome in the maturation of the Y-linked Su(Ste) repeats.

The crucial event responsible for the creation of a suppressor function of Su(Ste) repeats is the insertion of hoppel transposon causing antisense transcription of the Su(Ste) repeats (Aravin et al. 2001). Transcription of both strands of the Su(Ste) repeats provides the formation of double-stranded Su(Ste) RNA that participates in RNAi-mediated silencing of Ste expression. The Y-linked Su(Ste) repeats have a hoppel transposon insertion into their CP regions (fig. 2, cyam highlighting of an insertion site). The euchromatic Ste repeats located in the 12D region on the X chromosome carry a short deletion that may be a vestige of hoppel insertion and its imprecise excision (fig. 2). The hoppel insertion was detected in the Su(Ste) orphon that is located outside of the Su(Ste) cluster on the Y chromosome, has conserved nucleotide substitutions peculiar to the Ste repeats, and is considered to be an intermediate in the Y-linked Su(Ste) repeat evolution (Kogan et al. 2000). Thus, the presence of hoppel insertion vestiges in the X-linked euchromatic Ste repeats argues in favor that the hoppel insertion event might occur on the X chromosome before the transmission of the Su(Ste) ancestor to the Y chromosome.
length, and besides, βCK2tes-like regions carry fragments of various transposable sequences (HeT-A, gtwin, copia). PCKR sequences have regions of close homology to the Dm665 clone earlier attributed to the Ste-like repeats located on the Y chromosome (Danilevskaya et al. 1991). Dm665 repeats were detected by Southern analysis in the D. yakuba genome (Danilevskaya et al. 1991), whereas Ste-Su(Ste) repeats are unique to D. melanogaster. This result indicates that PCKRs (including Dm665) have been acquired and amplified on the Y chromosome earlier than the Su(Ste) repeats.

The junction region of the parental βNACtes and βCK2tes sequences in the chimeric intermediate differs from the corresponding region of the Ste-Su(Ste) repeats by the insertion into Ste-Su(Ste) of 23-bp stretch nonhomologous to the parental genes (fig. 2, cyan letters). The insertion into Ste-Su(Ste) is positioned between the βNACtes- and βCK2tes-derived sequences. We detected this stretch within PCKRs (fig. 2, cyan letters). Besides the presence of this shared 23-bp stretch peculiar to the Ste-Su(Ste) and PCKR, there is a stretch of close homology between the PCKR region corresponding to βCK2tes promoter and Ste sequence upstream of CP (fig. 2, green letters). These data suggest the occurrence of gene conversion between PCKRs and the ancestors of the Ste-Su(Ste) genes.

Discussion

We present the detailed description of successive steps in the course of the evolution of heterochromatic repeat precursors as well as their further fate on the sex chromosomes in the D. melanogaster genome. The putative evolutionary pathway of the Y-linked Su(Ste) repeats and paralogous X-linked Ste repeats is presented in figure 4.

**Fig. 3.**—(A) Northern analysis of the βNACtes genes transcription. Lanes: 1, cell culture (Schneider 2); 2, testis; 3, carcasses lacking testis; 4, ovaries; 5, carcasses lacking ovaries; 6, embryos; 7, larvae; 8, pupae; and 9, heads. Hybridization with ribosomal protein 49 (rp49) RNA probe was used as a gel-loading control. (B) Presence of the βNACtes genes in the Drosophila melanogaster subgroup species. Lanes: 1, D. melanogaster; 2, Drosophila simulans; 3, Drosophila mauritiana; 4, Drosophila yakuba; 5, Drosophila sechellia; and 6, Drosophila teissieri.

**Fig. 4.**—The evolution of Ste-Su(Ste) repeats. The βNACtes- and the βCK2tes-Ste-Su(Ste)-coding regions are presented by open and fountain-filled boxes, respectively; introns are indicated by broken lines. The CP region (sequence shared by the βNACtes, chimeric, and Ste-Su(Ste)) genes is cross-hatched. The promoter region peculiar to the βNACtes genes and the chimeric gene is presented by gray box; latticed boxes depict the 5′ shared proximal regions of the chimeric and βNACtes3 genes. The regions A in the βCK2tes, Ste(Ste), and the chimeric gene and its remnants in the Ste repeats are shown by hatched boxes; regions C are shown by dotted boxes. The open box flanked by inverted arrows stands for the hoppel transposon in Ste(Ste) repeats. The site of putative hoppel excision in euchromatic Stellate (euSte) is indicated by an interruption in crosshatched box. The PCKRs contain βCK2tes homology region (indicated by fountain-filled box), βCK2tes promoter region (black box), and various transposon-derived flanking sequences (open boxes). The 23-bp region shared by PCKR and Ste-Su(Ste) is shown by small open box with a black cross. The regions of similarity between βCK2tes promoter, PCKR, and Ste sequence upstream of their CP are denoted by black boxes. Arrows denote the putative transcription start sites. The region B in Ste(Ste) repeats is presented by a light gray box.
The first step is depicted as a formation of the primary chimeric intermediate as a result of nonhomologous recombination between the X-linked βNActes3 promoter and the autosomal βCK2tes gene. The initial genomic location of this intermediate might be on the X-chromosome. Actually, the cluster of euchromatic Ste repeats is located adjacent to the parental βNActes genes, on the 12DE region of the X chromosome, and the B fragment observed in the Y-linked Su(Ste) copies is detected as a separate sequence in the 12D region, close to the euchromatic Ste cluster. The location of the chimeric intermediate on the Y chromosome may be considered as its subsequent translocation. At the same time, we cannot neglect the possibility of the initial chimera location on the Y chromosome or simultaneous translocation of its copy to the Y chromosome.

We propose a possibility of conversion events between the copies of the initially formed chimeric gene and cognate PCKRs. The hallmarks of these events are the presence of the shared 23-bp stretch peculiar to the Ste-Su copies and some PCKR as well as the homology between the PCKR region corresponding to βCK2tes promoter and Ste sequence upstream of CP. We consider that the PCKRs have emerged on the Y chromosome as a result of acquisition of a βCK2tes copy from the autosome that has been further subjected to amplification, divergence, and degeneration. PCKRs seem to be related to nonfunctional, so-called junk, elements of heterochromatin. However, these “nonfunctional” heterochromatic repeats played a role in the diversification and evolution of the Ste and Su(Ste) repeats. The importance of gene conversion in the diversification of the Su(Ste) repeats in the Y-linked heterochromatin has been shown (Balakireva et al. 1992), and it now can be extended to the evolution and emergence of new heterochromatic repeats.

The ancestral copies on the sex chromosomes have been amplified. Amplification of chimeric precursor copies on the X chromosome spawned the formation of the Ste repeats lacking the transposon insertion and containing an intact ORF for the regulatory subunit of protein kinase CK2. Amplification of Ste repeats may cause harmful overproduction of a regulatory subunit of protein kinase. However, transposon insertion into the intermediate copies, modifications of the 3′ end region, and amplification on the Y chromosome resulted in the evolving of Su(Ste) repeats transcribed in opposite directions. Production of double-stranded RNA by Su(Ste) repeats forced siRNA generation and silencing of homologous Ste repeats.

The outlined scenario of the evolution of Ste-Su(Ste) genes using a preformed alien promoter is similar to the earlier reported way for the generation in teleosts of an additional copy of receptor tyrosine kinase (Adam, Dimitrijevic, and Schartl 1993). The accidentally acquired promoter regulated by a repressor has caused overexpression of a new tyrosine kinase gene and tumor formation occurring in the absence of the repressor locus. The preexisting promoter repression mechanism ensured the evolutionary survival of the new oncogene (Adam, Dimitrijevic, and Schartl 1993). We propose that the expression of the emerged new βCK2tes-like repeated genes (Ste precursors or mature Ste) driven by an alien βNActes promoter might contribute to the elevated testis-specific protein kinase CK2 expression and change the spatiotemporal activity of this protein in the germinal tissue. The protein kinase CK2 has many functionally important targets (Meggio and Pinna 2003). In the D. melonagaster testes, at least three tissue-specific regulatory CK2 subunits encoded by βCK2tes, β′CK2tes (Kalmykova et al. 2002), and Ste genes are produced. One of the possible scenarios of the evolutionary survival of the evolved Ste gene can implicate a suggestion that novel regulatory CK2 subunit has been utilized for the regulation of the new testis-specific targets. The problem of the novel CK2 overexpression was challenged when the X-linked Ste had started to amplify. To knock down Ste expression to an “optimal” level, the number of Su(Ste) has emerged and has been maintained by selective pressure. The coupled coevolution of the tandem repeat clusters of the Ste and Su(Ste) on the sex chromosomes resulted in Ste silencing and ensured male fertility.

The described mode of a new testis-expressed gene emergence in the D. melanogaster genome using a preformed alien promoter capture differs in essence from the earlier reported evolutionary ways for novel genes. The testis-specific promoter of a dynein subunit was shown to be spawned by the reshuffling of the non-promoter coding and intron regions of the parental genes (Nurinski et al. 1998). In another case, the promoter region of a testis-expressed retroposed gene has been proposed to evolve as a result of gradual nucleotide substitutions under selective pressure (Betran and Long 2003).

The proposed pathway of the origin of the Ste and Su(Ste) gene family demonstrates a crucial role of transposable elements in the creation of the natural RNAi-based silencing system. The presence of transposon insertions on the Drosophila Y chromosome is considered a hallmark of the degeneration of genes that is thought to occur on the Y chromosomes. However, a hoppel insertion in the precursors of the Su(Ste) repeats contributes to the evolution of a new function directed to the Ste genes silencing. A similar evidence of transposon taming to cause double-stranded RNA production and gene silencing has been recently put forward (Volpe et al. 2002; Puig, Caceres, and Ruiz 2004). Thus, RNAi silencing, considered as a defense against viruses and transposable elements, may also be involved in the evolution of eukaryotic genomes.

Supplementary Material

Supplementary figs. 1 and 2 (alignments) are available at Molecular Biology and Evolution online (www.mbe.oupjournals.org).

Acknowledgments

We are indebted to Manyuan Long, James Mason, and two anonymous reviewers for valuable comments and suggestions. We thank M. B. Evgen’ev for gift of Drosophila species and colleagues of Gvozdev laboratory for discussion and text improvement. This work is supported by Molecular and Cellular Biology Program of Russian Academy of Sciences, by grant from Russian Foundation for Basic Researches (05-04-48034), and a Scientific School (SS-2074.2003.4) grant.
Literature Cited


Marta Wayne, Associate Editor

Accepted March 29, 2005