

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Drosophila Chromosomal Polymorphism: From Population Aspects to Origin Mechanisms of Inversions

Carolina Garcia and Vera L. S. Valente

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.73246>

Abstract

High rates of chromosomal rearrangements are remarkably abundant in *Drosophila* Fallén, 1832 (Insecta, Diptera) genus, highlighting the paracentric inversions. Since different species of this genus are paradigms for genetics, evolutionary, and population studies, polymorphism analyses for chromosomal inversions have provided basic knowledge for beautiful biological questions. Chromosomal inversions suppress meiotic recombination and thus, natural selection can act to preserve favorable gene complexes. Analyses of natural and laboratory populations show that these polymorphisms provide adaptive advantages to their carriers in relation to diverse factors, such as niche exploration and climatic factors. In addition, due to their monophyletic origin, they also serve as genetic markers for the construction of unrooted phylogenies. With the increasing domain of molecular techniques and genome sequencing, factors such as the reuse of breakpoints by different inversions and the mechanisms that give rise to these polymorphisms have been exploited with scientific refinement. These analyses show the presence of regions that are hot spots for breakpoints, fitting the fragile breakage chromosomal evolution model, as well as the involvement of transposition elements at the origin of chromosomal inversions.

Keywords: chromosomal evolution, chromosomal inversion, polytene chromosomes, staggered breaks, transposable elements

1. Introduction

Structural chromosome rearrangements originate from chromosomal breaks at different sites, followed by reconstitution of these breaks in a distinct combination. They involve large quantities of genetic material at the cytological level and can be visualized under light microscopy.

The analysis of different rearrangements in the karyotype of the species of this genus was favored due to the presence of the polytene chromosomes. These polytene chromosomes are formed in interphase nuclei and are the final product of successive replication cycles without the consequent separation of the daughter chromatids, resulting in a huge structure that presents natural banding, formed by the precise synapses of parallel chromomeres of the sisters' chromatids. It is estimated that the polytene chromosomes founded in the salivary glands undergo 2^{10} replication events, generating up to 1024 filaments in each chromosomal pair of a diploid cell [1], originating a unique visualization magnitude. Tissues and organs containing cells with polytene chromosomes are, in general, involved in intense short time secretory functions, in a fast-growing context. Another peculiarity of the interphase polytene chromosomes is the non-segregation after replication; the parental chromosomes remain united and paired in the same conformation only seen in meiosis I of most other organisms [2].

The physical structure of the polytene chromosomes enables the accurate analysis of the different chromosomal rearrangements in *Drosophila* focusing on inversions—the most frequent rearrangement of the genus. This rearrangement consists in the simultaneous break of two sites in a chromosome and the reorganization of this area with a 180° inverted order.

Inversions are classified in two types, in diploid organisms: paracentric (do not involve the centromere in its formation, occurring in the same chromosome arm) and pericentric (involve the centromere and more than one chromosome arm). This rearrangement can be visualized as heterozygous during the pairing of the homologous chromosomes in meiosis I when only one of the parental chromosomes carries the inversion, forming an inversion loop for the correct pairing of the homologous chromosomes; or as homozygous when both parental chromosomes carry the inversion. These chromosomal conformations can be visualized on the *Drosophila* interphase polytene chromosomes (**Figure 1**) [3].

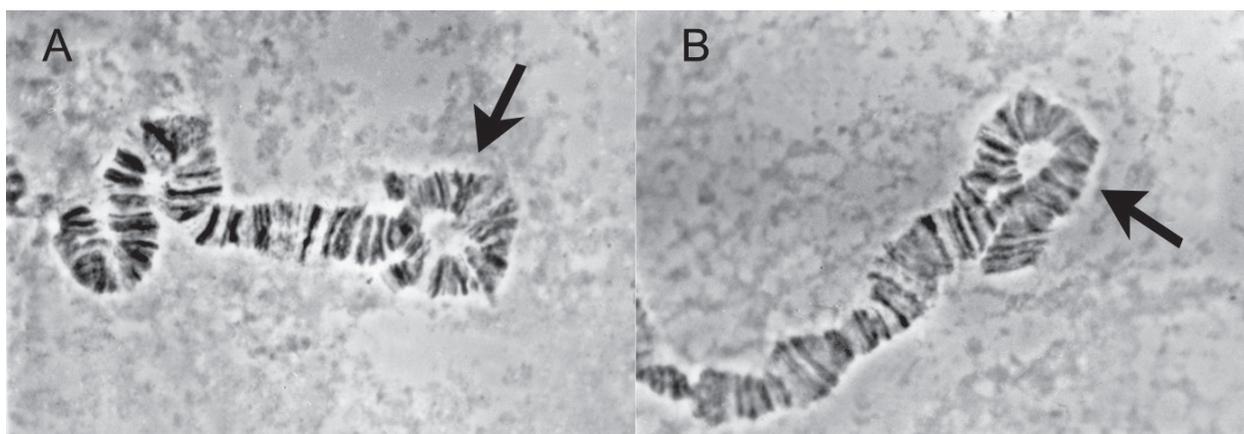


Figure 1. Chromosome inversions in heterozygosity in *Drosophila willistoni* Sturtevant, 1916 polytene chromosomes. (A) IIR-E inversion in the IIR chromosome arm. (B) III-J inversion in the chromosome III. Arrows point to the inversion loops formed by the correct pairing of the homologous chromosomes. Both inversions are physically at the distal end of the chromosome. Note that the precise synapse of the sister chromatids, according to the degree of compaction along the chromosome, forms a pattern of dark and light bands (bands and interbands, respectively). Regarding the chromosome polymorphism in *D. willistoni*, see the review in [56]. Source: Collection of images by Professor Vera L. S. Valente, deposited in the Laboratory of *Drosophila* at the Federal University of Rio Grande do Sul, Brazil.

Chromosomal inversions, compared to the other structural chromosomal rearrangements, use to be better tolerated by the organisms that carry them, since do not imply, theoretically, an increase or reduction of the genomic material. An inversion that occurs within a gene, however, can result in mutation, often lethal to the organism. The changing position of the genes, related to each other's and their controlling sequences, which is called the position effect, is another consequence of the inversion, resulting in alterations of gene expression and, consequently, alterations at the phenotypic level.

The behavior of a heterozygous inversion and the consequences it may entail differs during meiosis and mitosis. In meiosis I, the occurrence of crossing over inside of a paracentric inversion loop induces the formation of a dicentric chromosome (with two centromeres) and an acentric fragment (without centromere), resulting in gametes with deletions. In contrast, the occurrence of a meiotic recombination at the pericentric inversion loop results in the normal segregation of the chromosomes during meiosis I, since the centromeres are contained in the inversion, but originates gametes with deletion and duplications at meiosis II ending. During the mitosis, a heterozygous inversion does not imply major difficulties for the course of the cycle, since each chromosome duplicates and the sister chromatids are directed to the resulting daughter cells [4]. Illustrations of this are found in several genetics books, usually in Structural Chromosomal Alterations chapter.

Species of the *Drosophila* genus are model organisms for the study of chromosomal inversions, given the high resolution of the polytene chromosomes analysis, coupled with the fact that more than half of the studied species of *Drosophila* are naturally polymorphic for inversions [5]. However, based on the knowledge of the genomic destabilization and effects on the production of gametes that the inversions can originate, the high occurrence of chromosomal polymorphism is not expected *a priori* in the different living beings. The species of the genus *Drosophila* present a high rate of paracentric inversions, without a major deleterious effect on their reproductive success due the presence of defense mechanisms in males and females, preventing the production of gametes bearing unbalanced chromosomal rearrangements [6].

There is a mechanism in the meiosis of females of *Drosophila melanogaster* Meigen, 1830, carrying heterozygous inversion that selectively eliminates the recombinant gametes during the formation of the polar corpuscles. In this mechanism, the first polar corpuscle to be excluded is one of the balanced chromatids (standard order, or inverted order). The second polar corpuscle eliminated is the dicentric chromosome. The acentric fragment is not oriented in the meiotic spindle and is later degraded. The last polar corpuscle to be eliminated, which will be effectively fertilized, also presents the standard order, or inverted order [7].

The mechanism of protection against the production of inviable gametes in males of *D. melanogaster* seems to be the suppression of recombination in spermatogenesis [8]. Mutations in genes that affect the segregation of chromosomes that did not undergo meiotic exchange in *Drosophila* females do not have the same effect in males, suggesting that the exchange is not necessary for the correct segregation of homologous chromosomes in meiosis I in males of this genus [6].

Aside from the inferred suppression of recombination in males, reports of its occurrence at the meiotic level are present in the literature, evidencing some peculiarities. Among these,

the high occurrence in males showing the phenomenon of the hybrid dysgenesis of different species stands out. This phenomenon is also characterized by the presence of high frequencies of inviable offspring, mutations, structural chromosomal alterations, and distortion of the rate of transmission of alleles by one sex [9].

Another peculiarity is the spontaneous occurrence of recombination in males of species with a high degree of polymorphism for paracentric inversions, such as *D. melanogaster* [10], *Drosophila ananassae* Doleschall, 1858 [11], and *D. willistoni* [12].

Despite the exceptions, the presence of several cases of multiple heterozygosities occurring in many species of *Drosophila* support the great efficiency of these mechanisms and direct us to other biological aspects involving these chromosome rearrangements. The purpose of this chapter is to provide a basic overview of the knowledge of the evolutionary basis of its wide occurrence, and the adaptability conferred by the chromosomal polymorphism to the bearers of paracentric inversions found in this genus, converging in the present day in the analyses at the genomic level of the mechanisms that originate these inversions.

2. Population studies of chromosomal inversions in the genus *Drosophila*

The high polymorphism of chromosomal inversions has been used as a model for different adaptative processes, involved in the maintenance of the genetic variation. The concerns of Theodosius Dobzhansky and collaborators, more than 80 years ago, originated the early studies encompassing analyses of chromosomal inversions in natural populations of *Drosophila persimilis* Dobzhansky and Epling, 1944 and *Drosophila pseudoobscura* Frolova, 1929 [13]. Their findings were the stimuli for many of the discoveries that constituted the basis of modern evolutionary synthesis, which intricately combines Charles Darwin theory of evolution of species with Mendelian heritage patterns and population genetics.

The work of Dobzhansky “Genetics and the Origin of Species” [14] was a great incentive to the development of experimentation in evolutionary and population genetics.

Several experiments with *D. pseudoobscura* performed by Dobzhansky and colleagues were the basis to the postulation of the co-adaptation model of the genes contained in inversions [15]. Dobzhansky established that the reduced recombination in the inversions of this species is able to sustain positive combinations of genes in epistasis with other gene arrangement prevailing in the population. Therefore, gene complexes linked in an inversion in the different chromosome types are inherited as blocks and are rarely corrupted by meiotic recombination. Thus, the heterozygosity would be preferable to homozygosity, as predicted by the balancing selection [13, 14]. Thenceforth, the analysis and characterization of the chromosomal inversion polymorphism in natural populations of other species have become extensively explored. Also, indirect evidence of the association of chromosomal inversion with a better adaptation of the carrier individuals based on statistics was reported.

Drosophila pseudoobscura presents a broad geographic distribution in North America, being founded since west Canada, USA, and part of Central America, with the presence of a subspecies

in Colombia (*D. pseudoobscura bogotana*), and individuals collected in New Zealand (Oceania) [16]. The *ST*, *AR*, *CH*, *PP*, *SC*, *OL*, *EP*, and *TL* arrangements founded on the chromosome 3 of this species are extensively monitored and traditionally present altitudinal cline on their frequencies. Among these, the *TL* inversion presents a frequency increase on the Pacific coast since the 70 decade, which seems to be related to environmental changes [17].

Drosophila subobscura Collin, 1936 is a species with high chromosomal inversion polymorphism. Their rearrangements have been traditionally associated with adaptation to environmental variables. This Palearctic species invaded the American continent in the 70/80 decades [18]. Studies encompassing the frequency of inversion in European, North-, and South American populations show an inversely proportional relation of the increase in the frequency of inversions occurring in low latitudes (hot climate areas) and a decrease of frequency of the inversions occurring in high latitudes (cold climate areas) [19]. The chromosomal polymorphism in this species has also been related to environmental heavy metal contamination [20].

Drosophila buzzatii Patterson and Wheeler, 1942 belongs to the cactophilic species of the *repleta* group. It is originally from Southern Latin America, and its occurrence has been reported in the 1970s in the Mediterranean region, the Canary Islands, equatorial Africa and Australia, associated with cactus species of the genus *Opuntia*, which have been disseminated by human interference [21]. In this species, latitudinal clines in the frequency of some inversions have been inferred for the populations of the original areas and the colonized areas. The polymorphism described for the second chromosome, for example, the *2j* arrangement has been related to the longer development time, and larvae viability [22].

The Neotropical species *Drosophila mediopunctata* Dobzhansky and Pavan, 1943, belongs to the *Drosophila* subgenus, *tripunctata* group (the second largest group of Neotropical species). The acrocentric chromosome II of this species is highly polymorphic, with 17 inversions described, which are distributed in the distal (inversions *DA*, *DI*, *DS*, *DP*, *DR*, *DL*, and *DJ*) and proximal (inversions *C0*, *PC1*, *PC2*, *PC3*, *PC4*, *PC5*, *PB0*, *PA0*, and *PA8*) regions. Based on the 72 haplotypes already described for this chromosome it is possible to infer that the inversions at the distal and proximal regions practically do not overlap, and there is strong linkage disequilibrium between them. Thus, *DA* inversion is mostly found in association with *PA0* inversion. In the same way, *DI* inversion is associated with *PB0* inversion, *DS* with *PC0*, *DP* with *PC0*, and *DS* with *PC0*. Thus, it is difficult to find one of these distal inversions not associated with the corresponding proximal inversions. These five haplotypes are the most frequent (>90%) in the natural populations of *D. mediopunctata* from Southeastern Brazil. Since 1980, the inversions of chromosome II of this species have been analyzed as potential bioindicator of genetic responses to environmental changes, under the action of natural selection. Collections conducted from 1986 to 1988 and from 1991 to 2002 in different places of Southeast and Southern Brazil showed that *DA*, *DP*, and *DS* inversions present seasonal variation of their frequencies, and the inversion *DA* increased in dry and cold periods, and *DP* and *DS* inversions during rainy and hot periods. In addition, this panorama is related to altitudinal clines. Later collections (2007–2010) in one of the sampled sites (Itatiaia National Park, Rio de Janeiro, Brazil) allowed the comparison of the mean frequencies of inversions at the distal region, with the previous frequencies for this site. It was observed that the mean frequencies of *DA* and *DI* inversions increased, while *DS*, *DP*, and *DV* (associated with higher temperatures) decreased

their frequencies; and the *DA* inversion no longer has a significant correlation with altitude. Considering the climatic changes that occurred during these two decades in the region of Itatiaia Park, this suggests that temperature change has little influence on the seasonal changes in the frequencies of inversions in this species. Climate changes may have affected other genetic or morphological features, which may be more directly related to the inversions in chromosome II of *D. mediopunctata* [23–26].

Although several characteristics are indirectly associated with the inversions, little progress has been made in defining the genetic-evolutionary basis of these associations [23]. Direct shreds of evidence associating chromosomal inversions and selective pressures have been presented with the advancement of molecular techniques and genome sequencing.

Increasing amounts of data tend to confirm the inhibition of the recombination within the inversion area and also in adjacent areas, which is fundamental to the maintenance of the adaptive role. The patterns of linkage disequilibrium (LD) located within these regions reflect the inversion history and the gene flow since its origin [27–31].

An example of this case comes from the study of genetic variation and the unbalance of cosmopolitan inversion *In(3R)P* in two *D. melanogaster* populations from Australia, one from a tropical region (subdivided between individuals with inversion and individuals with the standard arrangement) and another from a temperate region (whose individuals carried only the standard arrangement). Since their high frequencies are related to higher temperatures, this inversion is known to be associated with climatic adaptations and the success of an evolutionarily recent migratory event (100 years) of this species in Australia. The results of this analysis support the hypothesis that *In(3R)P* inversion is associated with capture of locally adapted alleles, which interact substantially with loci external to the inversion. However, it was not possible to clarify whether these alleles are either in an additive or epistatic mode. Interestingly, high rates of LD in the region within the inversion are also found in the corresponding genomic region of the individuals that carried the standard arrangement in the tropical population, evidencing selection of such loci. Another result showed a high differentiation of the genomic region that involves the *In(3R)P* inversion between the tropical and the temperate population [30].

Despite the confirmed association of chromosomal inversions with the maintenance of combinations of alleles that lie within this region, gene recombination in the inverted region of a chromosome is possible because viable recombinant gametes arise through double meiotic recombination within the inverted region and also in consequence of gene conversion [31].

The prediction of recombination rates analysis in chromosomes carrying a heterozygous inversion, based on two mathematical models (Poison and Counting), made by Navarro and collaborators [32] infer three main points about this: “(i) the lower the inversion, the greater the effect on the reduction of the double meiotic recombination rate; (ii) in short inversions and in regions around the breakpoint, inversion reduces the rate of recombination but does not have the same capacity to prevent gene conversion; (iii) reduction of the recombination rate is not uniform throughout the chromosome, generally reducing the gene flow between different arrangements to near zero close to the breakpoints, and higher recombination rates are found in the central regions of the inversion.” The inversion also influences the events of recombination of regions outside their limits. All these findings have implications for the analyses that use balancer chromosomes [32–34].

It should be noted that a fraction of these chromosomal polymorphisms occurring in the different species is adaptively neutral, and thus suffer less selective pressure (or none), and its fixation, or loss, depends on population size and migration. These inversions can also reach high frequencies through other mechanisms, such as the inversion *In(1)Be* of the X chromosome of *D. melanogaster*. This inversion, considered of recent origin, has its maintenance probably due to the distortion of the transmission ratio through males of the species [35].

Despite the high acceptance and diffusion of the co-adaptation model of the genes contained in the inversions in *Drosophila* [15], alternative hypotheses point to different scenarios for the propagation and distribution of chromosomal inversions in populations of living beings, as a result of the increasing acquisition of knowledge and domain of improved analysis techniques [33, 34, 36].

3. Inversions breakpoints in *Drosophila*: chromosomal distribution

Parallel to the evolutionary-population studies of chromosomal inversions in *Drosophila*, the concern about the cause and origin of these polymorphisms in populations was already present.

Krimbas and Powel [37] wrote the best definition of the traditional point of view for the genesis of inversions: "It is that they are the result of two independent breaks, occurring at the same time, followed by the reconnection of the broken parts of the chromosome in an inverted orientation with respect to neighboring regions. Thus, the multiple overlapping inversions found in many *Drosophila* species would have occurred sequentially, not due to the simultaneous occurrence of multiple breaks. Regarding *in tandem* inversions (side-by-side inversions), the coincidence of breakpoints is attributed to chance, in events that occurred at different times. The hypothesis of the unique origin of the inversion is reinforced by the rarity of a chromosomal inversion event. It is even rarer that two events originating the same inversion occur spontaneously at the same time in the same chromosome site [37, 38]."

The monophyletic origin of the inversions implies that different rearrangements in the same chromosome can clarify some aspect of the evolutionary history of the analyzed species (or distinct species, when inter-crossings are possible), establishing the inversions as genetic markers for the reconstruction of unrooted phylogenies [39, 40].

"The first genetic dataset used for phylogenetic construction were the inversions of the chromosome 3 of *D. pseudoobscura* [41]." For this, the karyotype of a given populations of this species was arbitrarily inferred as the standard arrangement, being named *ST*. The crossings of males collected in the wild (as well as male offspring of the collected females), with females of the *ST* lineage, showed the differences of the chromosomal arrangements between the populations due to the formation of inversion loops in the F1 offspring. This comparative methodology of chromosomal inversions allowed relating the different triads of overlapping heterozygous inversions found in an unrooted phylogenetic tree. Based on this, a hypothetical central arrangement in the phylogeny, which has never been found in nature in later works, has been inferred. However, the key point for this analysis was that all copies of a particular inversion would have a unique origin, the arrangement being seen in the individuals of a population as a replica of the single arrangement that arose in the past in a single

common ancestor; in other words, its monophyletic origin. Later, molecular phylogenies corroborated the unique origin of the inversions in the chromosome 3 in *D. pseudoobscura*, and the topology of molecular phylogeny is in accordance with the topology obtained from the cytogenetic data [41].

The analysis of the phylogenetic relationships of overlapping inversions [39] considers the most parsimonious route (those with the small amount of inversion) for the evolutionary inference. Phylogenies were constructed for various species groups, such as *melanogaster* [42], *cardini* [43], Hawaiian *Drosophila* [44], *virilis* [45], *fasciola* subgroup [46], *willistoni* subgroup [47], among others.

Considering the traditional point of view of an inversion genesis, the distribution of the inversions along the chromosomes occurs randomly [37]. Sometimes, this characteristic seems to be well suited to the chromosomal distribution of the arrangement of chromosome 3 of *D. pseudoobscura* in natural populations [48], sometimes does not seem [39]. The inversion breakpoints induced by X-ray in *Drosophila* (and many other organisms, in general), seem to cluster preferentially in regions closer to the centromere [37, 49]. Add to this postulate, the evolutionary random breakage model, which gained notoriety with analyses of genomic comparisons, mainly between humans and mouse, later extending to other mammals in the 1980s. This model, in a simplistic way, assumed that the chromosome rearrangements, responsible for the breakdown of the synteny between these organisms, had their breakpoints distributed randomly along the chromosomes [50–52].

However, increasingly consistent studies evidencing the occurrence of repeated breaks in the same site for different inversions in a considerable amount of species have raised doubts regarding the randomness of the breakpoints distribution. These sites were denominated “hot spots,” and may involve particular structural instabilities of these regions [37].

The availability of the complete human genome and other mammals showed the effects of the limitations of the random breakage model, since it did not consider countless regions of the genomes of these organisms, because they were not available. The analysis of 281 syntenic blocks up to 1 Mb shared between humans and mouse showed the presence of 190 additional blocks with less than 1 Mb in size, which was very difficult to identify by alignment, and were totally unknown until then. The comparison of the chromosomal rearrangements occurred during the divergence between the two species showed a large number of breakpoints close to each other. This characteristic did not fit the random breakage model theory, so the fragile breakage model was proposed [53, 54].

This model was based on the inference that breakpoints of chromosome rearrangements occur mainly within fragile genome sites (hot spots), in other words, regions prone to breakage. These fragile sites may correspond to regions with lots of transposable elements (TEs), to segmental duplications, or to a palindromic sequence. “The reuse feature does not imply the use of the same genomic position (at the nucleotide level) repeatedly, but rather that, the breakpoint presents multiple genomic regions that originate chromosomal rearrangements [53, 54].”

Pioneering results at cytological level, on the reuse of breakpoints by different inversions, provided challenging data about the randomness of these breaks in *Drosophila*. Cáceres et al. [55] analyzed 86 paracentric (heterozygous and fixed) inversions described for species of the *D. buzzatii* complex and 18 inversions induced in *D. buzzatii* by introgression, through

crossings with *Drosophila koepferae* Fontdevila and Wasserman, 1988. The authors found that inversions of intermediate size are the most successful for the fixation in this species. They also observed that the breakpoints distribution of chromosome 2 inversions of these species, taking into account the location of the band involved in the break, is not random. The authors founded up to eight breakpoints at the same band in certain chromosomal segments. Similar results were observed in *D. subobscura* [56], Hawaiian *Drosophila* [57], and *D. willistoni* [58].

Although the reuse intra or interspecific of the inversions breakpoints, at cytological level, is common and well documented in the *Drosophila* genus, the characterization at DNA sequence level is still limited [59–61]. *In silico* comparisons of total genomes of different species available [62], estimate between 1.5 [63] and 2.27 [64] times the reuse of breakpoints throughout the evolutionary history of the species of this genus.

4. Characterization of inversion breakpoints in *Drosophila* and origin mechanisms

Delimitation and characterization of the inversion breakpoints are fundamental to determinate the mechanisms that originate them. In *Drosophila*, two main mechanisms have been highlighted in the origin of chromosomal inversions.

The first mechanism is the non-allelic homologous recombination (NAHR, also called ectopic recombination) between repetitive sequences, especially, the TEs [65–67]. The molecular machinery used by this mechanism is the same as allelic recombination, which has direct involvement with the genetic recombination in meiosis I. When ectopic recombination occurs between two copies of a repetitive sequence (very similar or identical), which are located physically at different chromosomal sites and in opposite orientations, the resulting inverted chromosome segment is flanked by two copies of these sequences, which are chimeric due to the exchange between them [66, 67]. The minimum identity between two sequences required for recombination is called minimal effective processing segment (MEPS). This parameter is not yet satisfactorily elucidated, *in vitro* analyses with prokaryotic organisms and mammalian cells infer that efficient MEPS for NAHR is 50 bp and between 270 and 280 bp, respectively [68]. However, the genomic approach of NAHR between copies of *Ty* retrotransposons in *Saccharomyces cerevisiae* Meyen ex E. C. Hansen, 1883 points out that more important than the identity between the copies of TEs is the genomic distance between them [69]. **Figure 2** illustrates a schematic for this mechanism. Based on this, it is important to note that when NAHR involves transposable elements, target site duplications (TSDs) of these can also be changed during recombination, a feature that has been very relevant for the recognition of this mechanism (see Section 4.1). For a long time, the TEs were considered junk DNA, and the involvement of these in the genesis of inversions of *Drosophila* genus provides solid knowledge to support the participation of these sequences in the molding of the genomes of living beings.

The second mechanism is via the erroneous repair of the free extremities, resulting from the chromosomal staggered breaks, by the non-homologous end joining (NHEJ). The physically close breaks in the chromosome cause failures in the correct pairing of the nitrogen bases, and

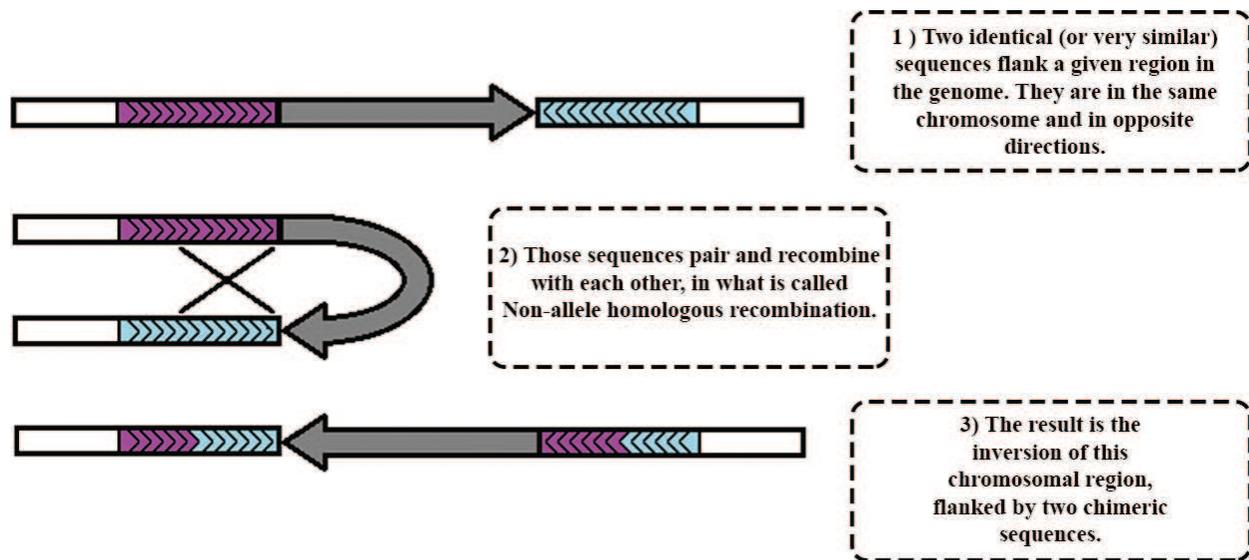


Figure 2. Standard schematic representation of the origin of chromosomal inversions via the non-allelic homologous recombination mechanism between identical or similar repetitive sequences. The arrow in dark gray represents the region that undergoes the inversion and its orientation. The repetitive sequences that flank this region in the left and right sides are represented with the tips indicating their orientations.

the chromosomal regions separate. The inversion is due to the junction of the 5' end with the 3' end of the other breakpoint [60, 70]. Duplicated DNA segments and in opposite orientations (delimiting the inverted chromosome segment) are the result of the repair and the main recognition mark of this mechanism [59]. In **Figure 3** it is possible to notice that staggered breaks occurred on both sides, duplicating two sequences that were originally single copies. However, based on the same figure, it is possible to extrapolate the occurrence of staggered breaks in only one side, and a simple break in the other side. The result is the duplication of just one originally single copy segment flanking the inversion. These duplicate sequences may involve genes. Gene duplication has been implicated as one of the main sources for the evolution of the genomes. The duplicate copy often does not undergo selective pressure, thus mutating more rapidly than the other essential regions of the genome. This may result in new gene functions, which is considered one of the most important results of these duplication events [71]. Thus, the repair of the free ends of staggered breaks by NHEJ gives rise to two different structural rearrangements: chromosomal inversion and duplication. Although in the case in question, duplications have small chromosomal magnitude compared to inversions, when they involve genes, they can also provide genomic variability in populations, and act on adaptive processes, speciation, and chromosome evolution.

The contribution of these two mechanisms is not completely clarified, and intriguing questions such as “whether these mechanisms are generalized among species of the genus and whether there are functional implications through the chromosomal evolution maintained by these inversions, remain open [59]”.

Table 1 presents a compilation of the different studies that characterized the inversion breakpoints at the molecular level in different species of the *Drosophila* genus. As can be

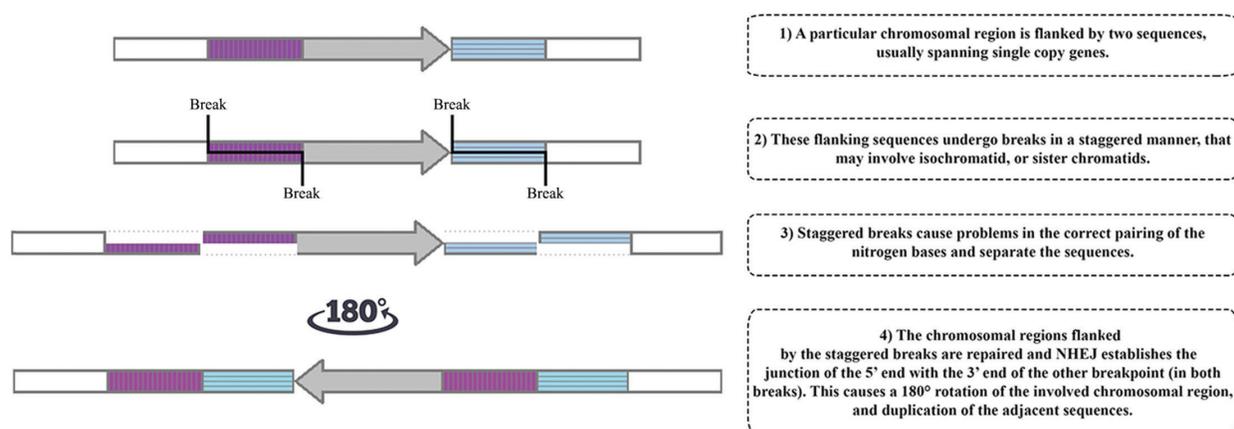


Figure 3. Standard schematic representation of the origin of chromosomal inversions via repair of the staggered breaks by non-homologous end joining. The gray arrow represents the region that suffers the inversion and its orientation. Adjacent sequences, originally single copies, are represented in the top of the image by vertical bars (left side) and horizontal bars (right side).

seen, the origin of different inversions, besides being via NAHR between TEs and other repetitive sequences, and staggered breaks followed by NHEJ, is also via simple breaks and repair. The breakpoint analysis does not always allow us to infer the probable origin of the inversions, a point that may be related to the antiquity of the inversion genesis, implying a greater amount of modifications in these regions, and loss of the signals that point to their origin mechanisms.

Species	Chromosomal inversion	Breakpoints description and mechanism of chromosomal inversion genesis
<i>D. melanogaster</i>	<i>In(3R)P</i>	Analysis by microdissection and sequencing of the inversion region in the chromosome. Absence of repetitive sequences at the breakpoints [72].
<i>D. melanogaster</i> x <i>D. subobscura</i>	Fixed inversion in the X chromosome of <i>D. subobscura</i>	Sequences of approximately 30–50 bp rich in thymines flanking the breakpoints [73].
<i>D. melanogaster</i>	<i>In(2L)t</i>	Analysis of the proximal breakpoint and presence of a TE – LINE [74].
<i>D. buzzatii</i>	<i>2j</i>	Presence of homologous copies of a TE denominated <i>Galileo</i> at the breakpoints, and origin of the inversion by NAHR between inverted copies of this TE [65].
<i>D. buzzatii</i>	<i>2q⁷</i>	Presence of homologous copies of a TE denominated <i>Galileo</i> at the breakpoints, and origin of the inversion by NAHR between inverted copies of this TE [66].
<i>D. pseudoobscura</i>	<i>Arrowhead</i>	Presence of 128 and 315 bp repetitive motifs in opposite orientation at the breakpoints of the inversion. Origin of the inversion by NAHR between the inverted copies of these repetitions [75].
<i>D. melanogaster</i>	<i>In(3R)Payne</i>	Small duplications in both breakpoints of the inversion [76].

Species	Chromosomal inversion	Breakpoints description and mechanism of chromosomal inversion genesis
<i>D. melanogaster</i> x <i>D. simulans</i> x <i>D. yakuba</i>	29 inversions	17 (59%) of the inversions presented inverted duplications at the breakpoints, including the <i>In(3R)84F1;93F6-7</i> inversion, which traditionally differentiates the karyotype of <i>D. melanogaster</i> and <i>D. simulans</i> . Origin of these inversions by staggered breaks mechanism [59].
<i>D. americana</i>	<i>In(4)a</i>	Repetitive sequences in opposite orientation of a <i>MITE</i> element in both breakpoints of the inversion [77].
<i>D. mojavensis</i> x <i>D. arizonae</i>	Inversion in the X chromosome	Absence of repetitive sequences at the breakpoints of the inversion [78].
<i>D. pseudoobscura</i> x <i>D. persimilis</i>	Inversion in the X and II chromosomes	<i>In tandem</i> repetitions of a 319 bp motif at the breakpoints of the inversion in the XR arm of <i>D. persimilis</i> [79].
<i>D. buzzatii</i>	2z ³	Presence of homologous copies of the TE <i>GalileoN</i> at the breakpoints and origin of the inversion by NAHR between the inverted copies of this TE [67].
<i>D. buzzatii</i>	5 g	Absence of significant repetitive sequences at the breakpoints [80].
<i>D. mojavensis</i>	Xe	Absence of significant repetitive sequences at the breakpoints. Probable origin by single breaks [70].
<i>D. americana</i> x <i>D. virillis</i>	Inversions (<i>In</i>) Xa and (<i>In</i>)5a	Presence of copies of the <i>MITE DAIBAM</i> at the breakpoints of the inversions in <i>D. americana</i> . Origin of the inversions by NAHR between the inverted copies of this TE [81].
<i>D. buzzatii</i>	Inversions 2m and 2n	2m inversion with 13 Kbp duplications in both breakpoints; origin of the inversion by staggered breaks [60]
<i>D. melanogaster</i>	Inversions <i>In(2L)t</i> , <i>In(2R)NS</i> , <i>In(3R)K</i> , <i>In(3R)Mo</i> , <i>In(3R)P</i> , <i>In(3L)P</i> , <i>In(1)A</i> , <i>In(1)Be</i>	Presence of inverted duplications at the breakpoints of the <i>In(2R)NS</i> , <i>In(3R)K</i> , <i>In(3R)P</i> , <i>In(1)A</i> , <i>In(1)Be</i> inversions [82].
<i>D. mojavensis</i>	Inversions 2c, 2f, 2g, 2h, 2q and 2r	Presence of copies of the TE <i>But-5</i> in both breakpoints of the 2s inversion by NAHR between the inverted copies of this TE. Presence of inverted duplications at the breakpoints of the 2h and 2q inversions; origin of the inversions by staggered breaks [83].
<i>D. subobscura</i>	O ₃	300 bp sequence in both breakpoints; origin of the inversion by staggered breaks [84].
<i>D. subobscura</i>	Inversions E ₁ and E ₂	Probable origin of the E ₁ inversion by staggered breaks and duplication of a region with approximately 400 bp, named β motif; origin of the E ₂ inversion by NAHR between α motifs (~ 700 bp) [61].

Species	Chromosomal inversion	Breakpoints description and mechanism of chromosomal inversion genesis
<i>D. subobscura</i>	Inversions E_9 and E_3	Presence of duplicated region (~8 Kbp) at the breakpoints of the E_9 inversion and of the duplicated region (~3.5 Kbp) at the breakpoints of E_3 inversion. Origin of the inversions by staggered breaks [85].
<i>D. subobscura</i>	E_{12}	Presence of the <i>Ugt58Fa</i> gene in both breakpoints. Origin of the inversions by staggered breaks [86].
<i>D. subobscura</i>	Inversions O_4 and O_8	Duplications in both breakpoints of the inversions; origin by staggered breaks [87].

Table 1. Molecular characterization studies of the inversion breakpoints in species of the *Drosophila* genus.

4.1. Involvement of the transposable elements at the origin of the inversions: non-allelic homologous recombination

Transposable elements are interesting and dominant components of the prokaryote and eukaryote genomes, meaning that the comprehension of their biology is a fundamental subject in genetics. Since their discovery by McClintock [88], much has been learned regarding the molecular properties of the TEs and their contribution to genome configuration of living beings.

These elements are classified according to their characteristics and transposition mode. Class I elements, also called retrotransposons, replicate through a “copy and paste” method and involve the production of an RNA intermediate, processed by reverse transcription to DNA and re-inserted in the genome. The retrotransposons subdivide into elements with Long Terminal Repeats (LTRs), for example, *copia* and *Gypsy* elements in *Drosophila*, that are similar to retroviruses; and the retrotransposons without LTRs, as Long Interspersed Elements (LINEs) and Short Interspersed Nuclear Elements (SINEs), which do not encode their reverse transcriptase and are also called retrotransposons [89, 90].

Class II elements, or DNA transposons, replicate, generically by a “cut and paste” mechanism, where the elements are physically excised from the genome and inserted into another site. In this case, there is an increase in the number of copies during the repairing of the excision sites of the DNA transposon by the host during DNA synthesis, or by the insertion of the TE in a genome site which has not been replicated [90, 91]. Still, among Class II elements, there is a non-autonomous element group denominated MITEs (Miniature Inverted-repeat Transposable Elements). These elements are short sequences with several copies in the genome and without coding capacity, as suggested by *Mar* element, which seems to be restricted to the *D. willistoni* subgroup [92].

The TEs of both classes are also classified in Subclass, Order, Superfamily, Family, and Subfamily based on their sharing of certain structures and sequence similarities [91].

The studies associating TEs with chromosomal rearrangements breakpoints in *Drosophila* genus begin mostly with the analysis of lineages presenting hybrid dysgenesis syndrome. This syndrome is caused by crossing certain lineages of *Drosophila* and is characterized by high mutation rates in germinative cells, causing a high frequency of inviable offspring,

recombination in males, mutation and structural chromosomal abnormalities [9]. The cause of hybrid dysgenesis has been reported to the activation of several TEs families, including *P*, *I*, and *hobo* elements in *D. melanogaster* [93] and *Penelope*, *Ulysses*, *Helena*, and *Telemac* in *Drosophila virilis* Sturtevant, 1916 [94]. Subsequently, studies involving programmed crossings and also cytogenetic and molecular analyses of the offspring followed the movement of the involved TEs and the appearance of chromosomal rearrangements associated with this movement [95].

The association of TEs insertions at cytological level with inversions breakpoints in natural populations of *Drosophila* has also been reported. Among them, stands out the analysis of the transposon *hobo* in *D. melanogaster* [96], the *P* element in *D. willistoni* [97], and the retroelements *Penelope* and *Ulysses* in *D. virilis* species group [98, 99].

The first analysis that directly evidenced the involvement of a TE at the origin of an inversion in a natural *Drosophila* population was made by Cáceres et al. [65]. This study analyzed the breakpoints of the polymorphic inversion $2j$ (of the second chromosome) of the species *D. buzzatii* (subgenus *Drosophila*, *repleta* group), which originated from the $2st$ (standard) chromosome arrangement. For the analysis of the breakpoints of the $2j$ inversion, these were delimited by chromosome walk, cloned, and sequenced in two lineages of *D. buzzatii*, which presented the $2st$ (lineage $st-1$) and $2j$ (lineage $(j-1)$) arrangements in homozygosis. For organization purposes, the breakpoints were designated AB and AC (distal breakpoint), CD and BD (proximal breakpoint) in the $2st$ and $2j$ lines, respectively. Sequencing and alignment of these regions in both lineages showed large insertions at the two inversion breakpoints, which were not present in the $2st$ standard arrangement. The insertion between A and C had 392 bp with long inverted repeats terminals (ITRs) of 106 bp. The insertion between B and D had 4319 bp, with ITRs as those of the 106 and 47 bp AC inserts. The central 180 bp of the AC insert and the BD sequence had 95% homology but was in opposite orientations. Sequences of 7 bp separated and inverted flanked each insert and resembled TSDs, which are the result of the TE insertion event. These characteristics pointed out that inversion $2j$ was generated by intrachromosomal pairing and recombination between the two homologous sequences inserted at distant sites and opposite orientations. The original structure of these inserts was homologous at approximately 274 bp and sustained a NAHR in *Drosophila*. These same insertions of the inversion breakpoints $2j$ were characterized as copies of a Class II TE, which was named *Galileo* [65].

Subsequently, the *Galileo* element was classified as a member of the *P* Superfamily of Class II Transposons [100] and subdivided into three subfamilies: *GalileoG* (*Galileo*), *GalileoN* (*Newton*), and *GalileoK* (*Kepler*) [67]. The involvement of this family was also pointed on the origin of two more polymorphic inversions of the chromosome 2 of *D. buzzatii*: $2q^7$ [66] e $2z^3$ [67]. These analyses showed, through cytological, molecular, and *in silico* analyses, that the origin of these inversions was due to the occurrence of NAHR between two copies of the TE *Galileo*, present at the breakpoints of these inversions.

Still, with respect to inversion $2j$ of *D. buzzatii* chromosome 2, its effect on the *CG13617* gene was analyzed. This gene was chosen because it is very close to the proximal breakpoint of this inversion (12 bp), and the embryos of homozygous lineage for the $2j$ arrangement have the expression five times lower compared to the standard lineages, without the presence of the

inversion. Based on the characterization of this region in the *D. buzzatti* genome and analysis of the mRNA levels, the authors pointed that the TE denominated *Kepler* is responsible for originating an antisense RNA, which forms a complex with the mRNA of the *CG13617* gene, performing a post-transcriptional regulation, making it inactive. *Kepler* TE is inserted adjacent to the proximal breakpoint in the lineages that carry the inversion *2j* and is not found in this same region of the breakpoint in the lineages without the inversion. The results of this study show a scenario of the interaction of antisense RNA with the *CG13617* gene via position effect. "Thus, the silencing of the *CG13617* gene is not due to the influence of the inversion *2j* itself, but rather due to the performance of sequences associated with them [101]."

There are also analyses of fixed inversions *2m* and *2n* in *D. buzzatti*, which are distributed *in tandem* and share the central breakpoint at the cytological level [60]. The delimitation and molecular characterization of the breakpoints were based on the genomic library of bacterial artificial chromosomes (BACs), and physical map of this species [102], and in the genome of the related species *Drosophila mojavensis* Patterson, 1940 [62], which did not exhibit such inversions. It was possible to establish which clones contained the regions of the three breakpoints in *D. buzzatii* (breakpoints denominated AC, BE, DF, whose direction from the left to the right is from the telomere to the centromere), by means of chromosomal walk by *in situ* hybridization, using BACs as probes. These positive BACs had their terminal portions sequenced, and these sequences served as a basis for delimiting the breakpoints (denominated AB, CD, EF, although not fully representative of the ancestral karyotype) in the genome of *D. mojavensis*. Subsequently, probes based on this genome were physically mapped on the polytene chromosomes of *D. buzzatii*, thus allowing the gene delimitation of the three breakpoints of *2m* and *2n* inversions. The comparison of these regions at the molecular level presented a very complex scenario. Small fragments of the *BuT-5* TE were found at both breakpoints of the *2n* inversion (breakpoints BE and DF), which may indicate their probable origin by ectopic recombination between these copies. However, due to the age of inversion, this assumption cannot be strongly based since these regions have already undergone many modifications and the TSDs have not been found. On the other hand, the *2m* inversion (AC and BE breakpoints) is flanked by ~13 Kbp duplications, which contain the *CG4673* gene. Thus, its most probable origin is via staggered breaks followed by NHEJ (See Section 4.2).

There is an extensive analysis of the mechanisms of origin of fixed inversions in *Drosophila mojavensis*, another representative of the *repleta* group. This species is the only representative of the *mulleri* complex that inhabits the Sonora desert, one of the aridest known environments, with fauna and flora quite peculiar [83]. The analysis of the chromosome evolution of *D. mojavensis* shows 10 fixed inversions in relation to the primitive arrangement I of the *repleta* group, along the evolution: one on the chromosome X (*Xe*), seven on chromosome 2 (*2c*, *2f*, *2g*, *2h*, *2q*, *2r*, and *2s*) and two on chromosome 3 (*3a* and *3d*) [83, 103]. The molecular characterization of the breakpoints of the seven inversions of chromosome 2 of this species occurred by means of end sequencing of clones of chromosome 2 of the genomic library of BACs of *D. buzzatii* [102]. Subsequently, these sequences were mapped in the genome of *D. mojavensis* and compared with the genome of *D. virilis* (external species with the karyotype without inversions). The breakpoints of *2c*, *2r*, and *2s* inversions showed copies of TEs flanking both sides of the inversion. However, the *2s* inversion stood out, due to the

presence of the *BuT-5* transposon at its breakpoints. The distal copy had 981 bp delimited by 9 bp AAGGCAAGT and CTGTATAAT sequences. At the proximal breakpoint, the copy of *BuT-5* TE was a 27 bp fragment, with 12 bp identical to one end and the remaining 15 bp were identical to the other end of this TE, delimited by sequences of 9 bp ACTTGCCTT and ATTATACAG. The sequences ACTTGCCTT and CTGTATAAT are the inverted complementary sequences of AAGGCAAGT and ATTATACAG, respectively, and constitute the TSDs derived from the insertion of the element. These characteristics indicate that the origin of the *2s* inversion of *D. mojavensis* is due to ectopic recombination between the two copies of the *BuT-5* TE. Functional inference of this inversion in the *D. mojavensis* genome indicates that the proximal copy of *BuT-5* TE acts on the *Dmoj\CG10375* gene promoter (which probably relates to the *Hsp40* gene family). *In silico* analyses show that *2s* inversion and the proximal copy of *BuT-5* TE increase the expression of this gene and may have direct implication with the thermotolerance regulation in this species [83].

Another species that clearly presents the involvement of TEs in the genesis of their inversions is the *Drosophila americana* Spencer 1938 (subgenus *Drosophila, virilis* group). The neo-X chromosome of this species is derived from a centromeric fusion segregating between the X-chromosome (Muller element A) and chromosome 4 (Muller Element B) in this species. This chromosomal fusion is positively correlated with latitude and has a polymorphic *In(4)a* inversion [77]. In addition, the arrangement of *D. americana* chromosome 4 is homosequential to the arrangement of the same chromosome in *D. virilis*, a related species that has its genome sequenced. Thus, its genome served as the basis for the design of the analysis, associated with the construction of a genomic library of BACs of *D. americana*. The analysis of *In(4)a* inversion of neo-X indicated its probable origin by means of ectopic recombination between two copies of a repetitive MITE element, which was widely dispersed in the genome of *D. virilis*. These same sequences were not present in the corresponding region in strains of the species without inversion (analysis made by PCR) and in *D. virilis*. The characteristics of this repeating sequence that support its identity as TE is the presence of 240 bp TIRs flanking an internal region of 869 bp. Comparisons of the multiple copies present in the genome of *D. virilis* with the sequences found at breakpoints in *D. americana* indicate that the copy present at the proximal breakpoint is a canonical element, whereas the copy present at the distal breakpoint is a rearranged element. From the functional point of view, the proximal breakpoint of this inversion presents allelic associations consistent with co-adaptation [77].

Subsequently, sequencing with low genome coverage of two strains of *D. americana* allowed the analysis of the *Xa* inversion fixed in *D. americana* and absent in *D. virilis* and the polymorphic *5a* inversion in *D. americana* [81]. The alignment of the breakpoints of both inversions between the two species indicated that in the regions where the alignment was corrupted, there was always a sequence varying between 500 and 1130 bp, present only in the lineages carrying the inversions. These sequences showed by BLASTN high similarity to an incomplete MITE sequence, with TIRs of 240 bp. In this study, the authors named it *DAIBAM* (*Drosophila americana* Inversion Breakpoints Associated MITE). In *Xa* inversion, it was possible to find clear TSDs and defective copies of TE *DAIBAM* flanking the inversion. In *5a* inversion, copies of the *DAIBAM* element flanking the inversion had more than 70% nucleotide similarity. Considering that TE *DAIBAM* copies are defective and that the analyzed inversions are old,

the authors infer that the data found supported the origin of inversions *Xa* and *5a* by ectopic recombination between the *DAIBAM* elements present at the breakpoints of these inversions. It was also found that this element was the same as that present at the breakpoints of inversion *In(4)a* [77]. Thus, the *DAIBAM* element is involved in the origin of at least 20% of the inversions occurring in the *virilis* group [77, 81].

4.2. Inversion origin via staggered breaks and repair by non-homologous end joining.

It has now been characterized that the origin of the inversions via staggered breaks followed by repair by NHEJ, is prevalent in two chromosomal systems: between the fixed chromosomal inversions that differentiate the *D. melanogaster* karyotype from those of *D. simulans* Sturtevant, 1919 and *D. yakuba* Burla, 1954 [59]; and the chromosomal polymorphism of the E and O chromosomes of *D. subobscura* [61, 84–87].

Drosophila melanogaster, *D. simulans*, and *D. yakuba* are members of the *D. melanogaster* subgroup (*Sophophora* subgenus). The main karyotypic difference between *D. melanogaster* and its cryptic *D. simulans* is the occurrence of inversion in the right arm of the 3 chromosome, denominated *In(3R)84F1;93F6-7* [104]. *Drosophila yakuba*, on the other hand, has at least 28 paracentric inversions differentiating its chromosomes from those of *D. melanogaster*.

The study of Ranz et al. [59] analyzed the breakpoints of 29 interspecific inversions in these species through experimental and computational methods.

The analysis of the breakpoints of the *In(3R)84F1;93F6-7* inversion highlighted that the breakpoints were proximally flanked by the *CG2708* and *CG7918* genes, and distally by *CG31176* and *CG34034* in *D. melanogaster*. Among these regions, there are occurrences of expressed sequences, and three of these sequences (*HDC14862*, *pdf800* e *HDC12400*) are duplicated and in opposite directions, in both breakpoints of the inversion in *D. melanogaster*, with 95% of identity between them. These sequences are single copies in *D. simulans* and *D. yakuba*, indicating that these duplications are a derived state with respect to the chromosomal arrangement of these species. Comparisons of the 3R chromosomal arm of *D. melanogaster*, *D. simulans*, and *D. yakuba* at the molecular level, highlighted a fixed inversion in the latter species (*In3R(7)*), that reuses the breakpoints of the *CG7918-CG34034* region, also used by the *In(3R)84F1;93F6-7* inversion. In both breakpoints of the *In3R(7)* inversion, there were two duplicated sequences (*CG34034* e *CG31286*) and in opposite orientation [59].

Due to the presence of inverted duplications associated with the *In(3R)84F1;93F6-7* and *In3R(7)* inversion breakpoints, the most parsimonious mechanism involved on its origins is through staggered breaks, proposed and schematized for the first time in this analysis. These staggered breaks can be isochromatid, occurring during the premeiotic mitosis and involving staggered single-strand breaks; or chromatid, occurring during the meiotic prophase involving staggered double-strand breaks [59].

The same *in silico* study analyzed the breakpoints of 28 paracentric inversions that differentiate the *D. melanogaster* chromosomes from those of *D. yakuba*, as well as a pericentric inversion in the chromosome 2. The genomic and phylogenetic evidences suggest that among these 29 inversions, 28 originated in the *D. yakuba* lineage. The analysis of the inversions

breakpoints showed that in approximately 62% of the cases (18 of 29 inversions), occurred the presence of duplicate sequences, which were presents with just a single copy in the *D. melanogaster* genome. Sequences of both breakpoints were inverted and duplicated in six of these inversions (as in **Figure 3**), and the sequence of just one of the breakpoints was duplicated in 12 inversions, which can be explained by several factors, for example, modifications occurred along the time. Most of these duplications (except three) did not prove to be functional. The comparative analysis of these breakpoints among *D. yakuba*, *D. melanogaster*, and other species, regarding the occurrence of TEs and its involvement in the origin of these inversions via NAHR, showed little support for this mechanism. It is clear in this analysis that most of the inversions that differentiate the *D. melanogaster* chromosomes from those of *D. yakuba* originated by staggered breaks in the latter species (17 of 29 analyzed inversions) and point to a rapid chromosomal evolution in the lineage that leads to *D. yakuba* [59].

The polymorphism of the Palearctic species *D. subobscura* (*Sophophora* subgenus, *obscura* group) has been extensively characterized and monitored for more than seven decades, which allows associating its variation with climate changes [18–20, 56]. Its karyotype is composed of six pairs of chromosomes, with the highest level of polymorphisms for inversions in all of them (except in the dot chromosome). This polymorphism is well characterized for the presence of complex chromosomal arrangements, formed by the occurrence of overlapping inversions, being the E and O chromosomes the ones with the highest occurrence of these arrangements in natural populations.

One of the pioneering analyzes in this species involved the characterization of the breakpoints of the O_3 inversion, which can be found in the O_{st} lineages (corresponding to the current standard arrangement of the species). This inversion originated from the extinct ancestral O_3 arrangement, that also gave rise to the O_{3+4} arrangement, which segregates the O_4 inversions in the different populations. For this analysis, breakpoints of the O_3 inversion in the extinct arrangement and without the O_3 inversion were denominated AB (proximal breakpoint) and CD (distal breakpoint). The O_{3+4} chromosome arrangement differs from the O_3 arrangement due to a small inversion of its distal breakpoint (called DC), presenting the same order of the proximal breakpoint (AB). In turn, the O_{st} chromosomal arrangement differs from O_3 by inversion O_3 (note that the extinct O_3 arrangement does not involve the O_3 inversion, which occurs in the O_{st} chromosome), involving B and C regions (their breakpoints being then called AC and BD). The analysis was a strategy that mixed *in situ* hybridization and *in silico* tests, together with the knowledge of the location of previously established probes [84].

New probes were established via comparisons with the available genomes of *D. melanogaster* and *D. pseudoobscura*, which made it possible to delimit the genomic region containing the breakpoints of the O_3 inversion in the chromosomes of the O_{3+4} arrangement. The posterior sequencing of this region in the O_{st} e O_{3+4} lineages allowed the comparisons between the breakpoints of the O_3 e O_{3+4} inversions, respectively. As a result, it was found that the breakpoints AB and DC in the O_{3+4} inversion comprised two small regions of 309 and 63 bp, respectively. The 63 bp sequence was the same 309 bp sequence, which was deleted at the origin of the O_{3+4} inversion. In turn, the same 309 bp sequence was present at both O_3 inversion breakpoints, indicating that at the origin of this inversion such region was duplicated, being present in

regions B and C of the breakpoints. The AB and DC regions in the lineage that carries the O_{3+4} inversion showed no similarity to any known TE. This meticulous analysis, based on the absence of TEs and the duplication of the 309 bp fragment, infers that the origin of the inversion O_3 , present in the chromosomal arrangement O_{st} , was by means of staggered double-strand breaks [84].

Still, in the O chromosome of *D. subobscura*, the breakpoints of the O_4 and O_8 inversions were delimited, sequenced and analyzed. Just as the inversion O_4 segregates only with the O_3 arrangement (giving rise to the complex chromosome arrangement O_{3+4}), the inversion O_8 segregates only with the arrangement O_{3+4} (giving rise to the chromosomal arrangement O_{3+4+8}). Comparisons of the O_4 inversion breakpoints with the respective regions in the O_{st} arrangement (without the inversion) pointed the occurrence of *Pxd*, *CG5225*, *Acf*, and *Set8* gene fragments at the proximal breakpoint, and the *CG5225*, *Pxd*, and *Acf* gene fragments at the distal breakpoint [87].

In the regions corresponding to the breakpoints of the O_{st} arrangements, fragments of the *Pxd*, *CG5225*, and *CG4009* genes were found at the proximal breakpoint. The distal breakpoint of the O_{st} arrangement encompasses fragments of the *Set8* and *Acf* genes. It is evident that at the origin of inversion O_4 fragments of the *Set8* and *Acf* genes were duplicated at the proximal breakpoint, and fragments of the *CG5225* and *Pxd* genes were duplicated at the distal breakpoint. This scenario fits the origin of inversion O_4 by the staggered double-strand break mechanism. The O_8 inversion breakpoints in the O_{3+4} arrangement (without the inversion) and O_{3+4+8} arrangements presented a similar picture to that of the O_4 inversion. The presence of the *Prosβ2R2* gene at both O_8 inversion breakpoints shows that this was doubled and fits the origin of this inversion O_8 by the staggered double-strand break mechanism. This analysis also found that genes *CG5225* and *Prosβ2R2* are involved in multiple rearrangements (duplications and transpositions, in addition to inversions) occurring along the chromosomal evolution of the species of the genus *Drosophila* [87].

The *D. subobscura* species also had the breakpoints of the E_1 , E_2 , E_9 , $E_{3'}$, and E_{12} inversions of the acrocentric chromosome E delimited, sequenced and analyzed. These inversions give rise to the complex arrangements $E_{1+2'}$, $E_{1+2+9'}$, $E_{1+2+9+3}$, and $E_{1+2+9+12}$. These chromosome constitutions, besides providing a great system for the analysis of the mechanisms of origin of inversions, also provide a basis for studying the reuse of the inversion breakpoints at the molecular level [61, 85–87].

The E_1 and E_2 inversions share, cytologically, one of the breakpoints. The comparison of the breakpoints of the standard lineage E_{st} (AB, EF, GH breakpoints) with the E_{1+2} lineage (AG, FB, EH breakpoints) showed two motifs, denominated α and β , which share the terminal portion named δ , in opposite orientations. The α motif was present at the AB and AG breakpoints with the same orientation, but with inverted orientation in the GH breakpoint (two copies with inverted orientation in the E_{st} chromosome and a single copy in the E_{1+2} chromosome). The β motif was present with the same orientation at the EF and EH breakpoints, and with inverted orientation at the FB breakpoint (a single copy in the *Est* chromosome and two copies in the E_{1+2} chromosome). The α motif exhibits small fragments similar to the SGM element, whereas the β motif is not similar to any described TE. Based on this scenario, the probable origin of the E_1 inversion was inferred by staggered breakpoints, that lead to the duplication of the β motif

present at the FB and EH breakpoints. The origin of the E_2 inversion, on the other hand, was inferred due to the ectopic recombination between two α motifs, present in both AB and GH breakpoints. The reuse was inferred by the presence of 400–700 bp repetitions at the breakpoints; however, it was impossible to elucidate which of the two inversions originated first [61].

The extensive analysis done in the classical rearrangements of the E and O chromosomes, mentioned above, showed that, with the exception of the E_2 inversion, the other chromosomal arrangements originated via staggered double-strand break mechanism. Thus, *D. subobscura* resembles *D. melanogaster*, and both emphasize a possible predominance of this mechanism in the origin of the inversions of the species belonging to the subgenus *Sophophora*. In addition, duplicate regions in these events range from a few hundred base pairs to about 8 Kbp (see **Table 1**), encompassing whole and partial genes in some of these duplications. However, no dose effect or generation of new transcripts was detected in the analyses [61, 84–87].

Still considering staggered break mechanism followed by erroneous repair by NHEJ, the molecular characterization of the inversion breakpoints in *D. mojavensis* indicates that its inversions $2h$ and $2q$ originated by this route. The $2h$ inversion would have originated by staggered single-break at the distal breakpoint in the parental chromosome, resulting in a duplicated region of approximately 7 Kb, encompassing *CG1792*, *Dmoj\GI23402*, and *pasha* genes. This event resulted in the origin of the gene *Dmoj\GI23123*, located at the proximal breakpoint of inversion $2h$. This gene, by similarity, showed a relationship with the *pasha* gene, and according to the prediction of the modENCODE software, it is also functional. Thus the *Dmoj\GI23123* gene originated from the duplication of the *pasha* gene (the extra copy of the gene giving rise to a new gene) in the event that resulted in the $2h$ inversion [83].

Staggered single-break occurred in the two breakpoints of the parental chromosome in the $2q$ inversion. This event resulted in a duplication of an approximately 4 Kb region containing a partial fragment of the *CG1208* gene. The duplication of this gene resulted in the origin of a new gene, called *Dmoj\GI22075*, at the distal breakpoint of the $2q$ inversion. The new gene maintained the MFS domain (Major Facilitator Superfamily) as an important feature of the *CG1208* gene [83].

The $2h$ and $2q$ inversions of *D. mojavensis* are pioneer examples of the origin of new genes with possible new functions, via duplication, based on the origin of an inversion by the staggered break mechanism followed by NHEJ [83].

5. Concluding remarks

Inversions are structural chromosomal alterations that, most of the time, neither imply genetic unbalance, nor phenotypic modifications in its carriers. However, one of its characteristics is to be a source of genetic variability, in which natural selection acts. Thus, the inversions participate in the chromosomal evolution of numerous species, including *Homo sapiens*. The basic knowledge about the biological influence of inversions is largely based on the analysis of the polytene chromosomes of the *Drosophila* model organism, which extends to other living beings.

The first works, with descriptive approaches to the frequency of chromosome polymorphism in different natural populations, while indirectly pointing out that the inversions provided advantages to its users, raised questions that until now guide the analysis on this theme: How does natural selection work in inversions? How do inversions offer greater adaptability to living beings? What is the role of the inversions in the speciation processes? What are the functional consequences of inversions in living beings? Are inversions randomly distributed on chromosomes? How do inversions originate?

The *Drosophila* model organism provides knowledge and answers to these questions nowadays, with the availability of complete genomes of different species, improved cytomolecular techniques, as well as a solid knowledge about the cytogenetics of polytene chromosomes.

The molecular characterization of the inversion breakpoints tells us about the mechanisms that originate these rearrangements, the genomic composition of the region involved in the inversion—which allows to analyze the nucleotide variation and to show which genes are under selection—the reuse of certain regions for the breakage of different inversions that occurred at different times, the age of the inversion, its monophyletic origin, possible positional effect and its influence on the genes that are inside and outside the inversion, among others. Valuable understandings emerge, but are still incipient.

These analyses go far from being simplistic, but, with the current resources, we have never had so much opportunity to acquire knowledge. Let us live the new time in science, and avail the most of the knowledge already established, with the certainty that many other questions will arise.

As the eminent geneticist Michael Ashburner of the University of Cambridge, United Kingdom, compiles: “What a wonderful time to be a biologist [105].”

Acknowledgements

We are grateful to Mr. Leonardo Lindenmeyer for all the assistance provided with the images of this manuscript. Grants and fellowships of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, process 455101/2014-0) and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) from Brazil.

Author details

Carolina Garcia* and Vera L. S. Valente

*Address all correspondence to: cafagibio@hotmail.com

Department of Genetics, Institute of Biosciences, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

References

- [1] Swift H. Nucleic acids and cell morphology in dipteran salivary glands. In: Allen JM, editor. *The Molecular Control of Cellular Activity*. New York: McGraw Hill; 1962. pp. 73-125
- [2] Zhimulev IF, Koryakov DE. Polytene chromosomes. *Encyclopedia of Life Sciences (ELS)*. Chichester: John Wiley & Sons, Ltd; 2009. DOI: 10.1002/9780470015902.a0001183.pub2
- [3] Hartl DL, Jones EW. Human karyotypes and chromosome behavior. In: *Genetics: Analysis of Genes and Genomes*. 7th ed. London: Jones and Bartlett; 2009. pp. 250-294
- [4] Griffiths AJF, Wessler SR, Carroll SB, John Doebley J. Large-scale chromosomal changes. In: *Introduction to Genetic Analysis*. 9th ed. New York: WH Freeman & Company; 2015. pp. 481-520
- [5] Powell JR. Population genetics—Inversions. In: *Progress and Prospects in Evolutionary Biology—The *Drosophila* Model*. New York: Oxford University; 1997. pp. 48-80
- [6] Orr-Weaver TL. Meiosis in *Drosophila*: seeing is believing. *Proceedings of the National Academy of Sciences of the United States of America*. 1995;**92**(23):10443-10449. PMID: PMC40628
- [7] Hinton CW, Lucchesi JC. A Cytogenetic study of crossing over in inversion heterozygotes of *Drosophila melanogaster*. *Genetics*. 1960;**45**(1):87-94. PMID: PMC1210005
- [8] Morgan TH. Complete linkage in the second chromosome of the male of *Drosophila*. *Science*. 1912;**36**(934):719-720. DOI: 10.1126/science.36.934.719
- [9] Kidwell MG, Kidwell JF, Sved JA. Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility, and male recombination. *Genetics*. 1977;**86**(4):813-833. PMID: PMC1213713
- [10] Henderson SA, Woodruff RC, Thompson JN. Spontaneous chromosome breakage at male meiosis associated with male recombination in *Drosophila melanogaster*. *Genetics*. 1978;**88**(1):93-107. PMID: PMC1213793
- [11] Goñi B, Matsuda M, Yamamoto MT, Vilela CR, Tobar YN. Crossing over does occur in males of *Drosophila ananassae* from natural populations. *Genome*. 2012;**55**(7):505-511. DOI: 10.1139/g2012-037
- [12] Santos-Colares MC, Degrand TH, Valente VLS. Cytological detection of male recombination in *Drosophila willistoni*. *Cytologia (Tokyo)*. 2004;**69**(4):359-365. DOI: 10.1508/cytologia.69.359
- [13] Dobzhansky T. *Genetics of the Evolutionary Process*. New York: Columbia University Press; 1970
- [14] Dobzhansky T. *Genetics and the Origin of Species*. 1st, 2nd, 3rd ed. New York: Columbia University Press, 1937, 1941, 1951

- [15] Dobzhansky T. On some of the problems of population genetics and evolution. Symposium sui Fattori Ecologici e Genetici della Speciazione Degli Animali. La Reserca Scientifica. 1949;**19**(Suppl):11
- [16] Lambert DM, McLea MC. *Drosophila pseudoobscura* from New Zealand. *Drosophila Information Service*. 1983;**59**:72-73
- [17] Anderson WW, Arnold J, Baldwin DG, Beckenbach AT, Brown CJ, Bryant SH, Coyne JA, Harshman LG, Heed WB, Jeffery DE, et al. Four decades of inversion polymorphism in *Drosophila pseudoobscura*. *Proceedings of the National Academy of Sciences of the United States of America*. 1991;**88**(22):10367-10371. PMID: PMC52929
- [18] Brncic D, Prevosti A, Budnik M, Monclús M, Ocaña J. Colonization of *Drosophila subobscura* in Chile I. First population and cytogenetic studies. *Genetica*. 1981;**56**(1):3-9. DOI: 10.1007/BF00126923
- [19] Balanyà J, Huey RB, Gilchrist GW, Serra L. The chromosomal polymorphism of *Drosophila subobscura*: A microevolutionary weapon to monitor global change. *Heredity (Edinburgh)*. 2009;**103**(5):364-367. DOI: 10.1038/hdy.2009.86
- [20] Kenig B, Kurbalija Novičić Z, Patenković A, Stamenković-Radak M, Anđelković M. Adaptive role of inversion polymorphism of *Drosophila subobscura* in lead stressed environment. *PLoS One*. 2015;**10**(6):e0131270. DOI: 10.1371/journal.pone.0131270
- [21] Fontdevila A, Ruiz A, Ocaña J, Alonso G. Evolutionary history of *Drosophila buzzatii*. II. How much has chromosomal polymorphism changed in colonization? *Evolution*. 1982;**36**(4):843-851. DOI: 10.1111/j.1558-5646.1982.tb05450.x
- [22] Hoffmann AA, Sgrò CM, Weeks AR. Chromosomal inversion polymorphisms and adaptation. *Trends in Ecology & Evolution*. 2004;**19**(9):482-488. DOI: 10.1016/j.tree.2004.06.013
- [23] Schaeffer SW, Goetting-Minesky MP, Kovacevic M, Peoples JR, Graybill JL, et al. Evolutionary genomics of inversions in *Drosophila pseudoobscura*: Evidence for epistasis. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;**100**(14):8319-8324. DOI: 10.1073/pnas.1432900100
- [24] Galina A, Peixoto AA, Bitner-Mathé CB, Souza WN, Silva LB, Valente VLS, Klaczko LB. Chromosomal inversion polymorphism in *Drosophila mediopunctata*: Seasonal, altitudinal, and latitudinal variation. *Genetics and Molecular Biology*. 2004;**27**(1):61-69. DOI: 10.1590/S1415-47572004000100011
- [25] Klaczko LB. Evolutionary genetics of *Drosophila mediopunctata*. *Genetica*. 2006;**126**(1):43-55. DOI: 10.1007/s10709-005-1431-6
- [26] Batista MR, Ananina G, Klaczko LB. Unexpected long-term changes in chromosome inversion frequencies in a Neotropical *Drosophila* species. *Climate Research*. 2012;**53**(2):131-140. DOI: 10.3354/cr01088
- [27] Pegueroles C, Ordóñez V, Mestres F, Pascual M. Recombination and selection in the maintenance of the adaptive value of inversions. *Journal of Evolutionary Biology*. 2010;**23**(12):2709-2717. DOI: 10.1111/j.1420-9101.2010.02136.x

- [28] McGaugh SE, Noor MA. Genomic impacts of chromosomal inversions in parapatric *Drosophila* species. *Philosophical Transactions of the Royal Society B*. 2012;**367**(1587):422-429. DOI: 10.1098/rstb.2011.0250
- [29] Fuller ZL, Haynes GD, Zhu D, Batterton M, Chao H, et al. Evidence for stabilizing selection on codon usage in chromosomal rearrangements of *Drosophila pseudoobscura*. *G3 (Bethesda)*. 2014;**4**(12):2433-2449. DOI: 10.1534/g3.114.014860
- [30] Rane RV, Rako L, Kapun M, Lee SF, Hoffmann AA. Genomic evidence for role of inversion *3RP* of *Drosophila melanogaster* in facilitating climate change adaptation. *Molecular Ecology*. 2015;**24**(10):2423-2432. DOI: 10.1111/mec.13161
- [31] Chovnick A. Gene conversion and transfer of genetic information within the inverted region of inversion heterozygotes. *Genetics*. 1973;**75**(1):123-131. PMID: PMC1212990
- [32] Navarro A, Betran E, Barbadilla A, Ruiz A. Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics*. 1997;**146**:695-709. PMID: PMC1208008
- [33] Andolfatto P, Depaulis F, Navarro A. Inversion polymorphisms and nucleotide variability in *Drosophila*. *Genetical Research*. 2001;**77**(1):1-8. DOI: 10.1017/S0016672301004955
- [34] Miller DE, Cook KR, Yeganeh Kazemi N, Smith CB, Cockrell AJ, Hawley RS, Bergman CM. Rare recombination events generate sequence diversity among balancer chromosomes in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*. 2016;**113**(10):E1352-E1361. DOI: 10.1073/pnas.1601232113
- [35] Corbett-Detig RB, Hartl DL. Population genomics of inversion polymorphisms in *Drosophila melanogaster*. *PLoS Genetics*. 2012;**8**(12):e1003056. DOI: 10.1371/journal.pgen.1003056
- [36] Kirkpatrick M, Barton N. Chromosome inversions, local adaptation and speciation. *Genetics*. 2006;**173**(1):419-434. DOI: 10.1534/genetics.105.047985
- [37] Krimbas CB, Powell JR. Chromosomal polymorphism in natural and experimental populations. In: Krimbas CB, Powell JR, editors. *Drosophila Inversion Polymorphism*. Florida, USA: CRC; 1982. pp. 2-52
- [38] Sperlich D, Pfriem P. Chromosomal polymorphism in natural and experimental population. In: Ashburner M, Carson HL, Thompson JN Jr, editors. *The Genetics and Biology of Drosophila*. Vol. 3e. London: Academic Inc.; 1986. pp. 257-309
- [39] Sturtevant AH, Dobzhansky T. Inversions in the third chromosome of wild races of *Drosophila pseudoobscura*, and their use in the study of the history of the species. *Proceedings of the National Academy of Sciences of the United States of America*. 1936;**22**(7):448-450. PMID: PMC1076803
- [40] Dobzhansky T, Sturtevant AH. Inversions in the chromosomes of *Drosophila pseudoobscura*. *Genetics*. 1938;**23**:28-64. PMID: PMC1209001

- [41] Wallace AG, Detweiler D, Schaeffer SW. Evolutionary history of the third chromosome gene arrangements of *Drosophila pseudoobscura* inferred from inversion breakpoints. *Molecular Biology and Evolution*. 2011;**28**(8):2219-2229. DOI: 10.1093/molbev/msr039
- [42] Lemeunier F, Aulard S. Inversion polymorphism in *Drosophila melanogaster*. In: Krimbas CB, Powell JR, editors. *Drosophila Inversion Polymorphism*. Florida, USA: CRC; 1982. pp. 339-405
- [43] Heed WB, Russell JS. Phylogeny and population structure in island and continental species of the *cardini* group of *Drosophila* studied by inversion analysis. University of Texas Publishing. 1971;**7****103**:91-130
- [44] Carson HL, Kaneshiro KY. *Drosophila* of Hawaii: Systematics and ecological genetics. *Annual Review of Ecological Systems*. 1976;**7**:311-345. DOI: 10.1146/annurev.es.07.110176.001523
- [45] Throckmorton LH. The *virilis* species group. In: Ashburner M, Carson HL, Thompson JN, editors. *The Genetics and Biology of Drosophila*. Vol. 3b. London: Academic Press; 1982. pp. 227-296
- [46] Diniz NM, Sene FM. Chromosomal phylogeny of the *Drosophila fasciola* species subgroup revisited (Diptera, Drosophilidae). *Genetics and Molecular Biology*. 2004;**27**(4):561-556. DOI: 10.1590/S1415-47572004000400016
- [47] Rohde C, Garcia AC, Valiati VH, Valente VL. Chromosomal evolution of sibling species of the *Drosophila willistoni* group. I. Chromosomal arm IIR (Muller's element B). *Genetica*. 2006;**126**(1-2):77-88. DOI: 10.1007/s10709-005-1433-4
- [48] Olvera O, Powell JR, de la Rosa ME, Salceda VM, et al. Populations genetics of Mexican *Drosophila*. VI. Cytogenetics aspects of the inversion polymorphism in *Drosophila pseudoobscura*. *Evolution*. 1979;**33**:381-395. DOI: 10.2307/2407628
- [49] Jackson WD, Barber HN. Patterns of chromosome breakage after irradiation and ageing. *Heredity*. 1958;**12**:1-25. DOI: 10.1038/hdy.1958.1
- [50] Ohno S. Ancient linkage groups and frozen accidents. *Nature*. 1973;**244**(5414):259-262. DOI: 10.1038/244259a0
- [51] Sachs RK, Levy D, Chen AM, Simpson PJ, Cornforth MN, et al. Random breakage and reunion chromosome aberration formation model; an interaction-distance version based on chromatin geometry. *International Journal of Radiation Biology*. 2000;**76**(12):1579-1588. DOI: 10.1080/09553000050201064
- [52] Nadeau JH, Taylor BA. Lengths of chromosomal segments conserved since divergence of man and mouse. *Proceedings of the National Academy of Sciences of the United States of America*. 1984;**81**:814-818. PMID: PMC344928
- [53] Pevzner P, Tesler G. Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;**100**(13):7672-7677. DOI: 10.1073/pnas.1330369100

- [54] González J, Casals F, Ruiz A. Testing chromosomal phylogenies and inversion breakpoint reuse in *Drosophila*. *Genetics*. 2007;**175**(1):167-177. DOI: 10.1534/genetics.106.062612
- [55] Cáceres M, Barbadilla A, Ruiz A. Inversion length and breakpoint distribution in the *Drosophila buzzatii* species complex: Is inversion length a selected trait? *Evolution*. 1997;**51**(4):1149-1155. DOI: 10.2307/2411044
- [56] Krimbas CB, Loukas M. The inversion polymorphism of *Drosophila subobscura*. In: Hecht MK, Steere WC, Wallace B, editors. *Evo biol*. Vol. 12. New York: Plenum Press; 1980. pp. 163-234
- [57] Tonzetich J, Lyttle TW, Carson HL. Induced and natural break sites in the chromosomes of Hawaiian *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*. 1988;**85**(5):1717-1721. PMID: PMC279846
- [58] Rohde C, Valente VLS. Three decades of studies on chromosomal polymorphism of *Drosophila willistoni* and description of fifty different rearrangements. *Genetics and Molecular Biology*. 2012;**35**(4):966-979. DOI: 10.1590/S1415-47572012000600012
- [59] Ranz JM, Maurin D, Chan YS, von Grotthuss M, Hillier LW, Roote J, et al. Principles of genome evolution in the *Drosophila melanogaster* species group. *PLoS Biology*. 2007;**5**:1366-1381. DOI: 10.1371/journal.pbio.0050152
- [60] Calvete O, González J, Betrán E, Ruiz A. Segmental duplication, microinversion, and gene loss associated with a complex inversion breakpoint region in *Drosophila*. *Molecular Biology and Evolution*. 2012;**29**(7):1875-1889. DOI: 10.1093/molbev/mss067
- [61] Puerma E, Orengo DJ, Salguero D, Papaceit M, Segarra C, Aguadé M. Characterization of the breakpoints of a polymorphic inversion complex detects strict and broad breakpoint reuse at the molecular level. *Molecular Biology and Evolution*. 2014;**31**(9):2331-2341. DOI: 10.1093/molbev/msu177
- [62] *Drosophila* 12 Genomes Consortium. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*. 2007;**480**:203-218. DOI: 10.1038/nature06341
- [63] Bhutkar A, Schaeffer SW, Russo SM, et al. Chromosomal rearrangement inferred from comparisons of 12 *Drosophila* genomes. *Genetics*. 2008;**179**(3):1657-1680. DOI: 10.1534/genetics.107.086108
- [64] von Grotthuss M, Ashburner M, Ranz JM. Fragile regions and not functional constraints predominate in shaping gene organization in the genus *Drosophila*. *Genome Research*. 2010;**20**(8):1084-1096. DOI: 10.1101/gr.103713.109
- [65] Cáceres M, Ranz JM, Barbadilla A, Long M, Ruiz A. Generation of a widespread *Drosophila* inversion by a transposable element. *Science*. 1999;**285**:415-418. DOI: 10.1126/science.285.5426.415
- [66] Casals F, Cáceres M, Ruiz A. The foldback-like transposon *Galileo* is involved in the generation of two different natural chromosomal inversions of *Drosophila buzzatii*. *Molecular Biology and Evolution*. 2003;**20**:674-685. DOI: 10.1093/molbev/msg070
- [67] Delprat A, Negre B, Puig M, Ruiz A. The transposon *Galileo* generates natural chromosomal inversions in *Drosophila* by ectopic recombination. *PLoS One*. 2009;**4**:e7883. DOI: 10.1093/molbev/msg070

- [68] Jinks-Robertson S, Michelitch M, Ramcharan S. Substrate length requirements for efficient mitotic recombination in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*. 1993;**13**:3937-3950. DOI: 10.1128/MCB
- [69] Hoang ML, Tan FJ, Lai DC, Celniker SE, Hoskins RA, et al. Competitive repair by naturally dispersed repetitive DNA during non-allelic homologous recombination. *PLoS Genetics*. 2010;**6**(12):e1001228. DOI: 10.1371/journal.pgen.1001228
- [70] Runcie DE, Noor MA. Sequence signatures of a recent chromosomal rearrangement in *Drosophila mojavensis*. *Genetica*. 2009;**136**(1):5-11. DOI: 10.1007/s10709-008-9296-0
- [71] Ohno S. *Evolution by Gene Duplication*. New York Inc.: Springer-Verlag; 1970. 160p
- [72] Wesley CS, Eanes WF. Isolation and analysis of the breakpoint sequences of chromosome inversion *In(3L)Payne* in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America*. 1994;**91**(8):3132-3136. PMID: PMC43529
- [73] Cirera S, Martin-Campos JM, Segarra C, Aguade M. Molecular characterization of the breakpoints of an inversion fixed between *Drosophila melanogaster* and *D. subobscura*. *Genetics*. 1995;**139**:321-326. PMID: PM7705632
- [74] Andolfatto P, Wall JD, Kreitman M. Unusual haplotype structure at the proximal breakpoint of *In(2L)t* in a natural population of *Drosophila melanogaster*. *Genetics*. 1999;**153**:1297-1311. PMID: PM10545460
- [75] Richards S, Liu Y, Bettencourt BR, Hradecky P, Letovsky PS, et al. Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome Research*. 2005;**15**:1-18. DOI: 10.1101/gr.3059305
- [76] Matzkin LM, Merritt TJ, Zhu CT, Eanes WF. The structure and population genetics of the breakpoints associated with the cosmopolitan chromosomal inversion *In(3R)Payne* in *Drosophila melanogaster*. *Genetics*. 2005;**170**:1143-1152. DOI: 10.1534/genetics.104.038810
- [77] Evans AL, Mena PA, McAllister BF. Positive selection near an inversion breakpoint on the neo-X chromosome of *Drosophila americana*. *Genetics*. 2007;**177**(3):1303-1319. DOI: 10.1534/genetics.107.073932
- [78] Cirulli ET, Noor MA. Localization and characterization of X chromosome inversion breakpoints separating *Drosophila mojavensis* and *Drosophila arizonae*. *The Journal of Heredity*. 2007;**98**(2):111-114. DOI: 10.1093/jhered/esl065
- [79] Machado CA, Haselkorn TS, Noor MA. Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*. 2007;**175**(3):1289-1306. DOI: 10.1534/genetics.106.064758
- [80] Prazeres da Costa O, González J, Ruiz A. Cloning and sequencing of the breakpoint regions of inversion 5g fixed in *Drosophila buzzatii*. *Chromosoma*. 2009;**118**(3):349-360. DOI: 10.1007/s00412-008-0201-5
- [81] Fonseca NA, Vieira CP, Schlötterer C, Vieira J. The *DAIBAM MITE* element is involved in the origin of one fixed and two polymorphic *Drosophila virilis* phylad inversions. *Fly (Austin)*. 2012;**6**(2):71-74. DOI: 10.4161/fly.19423

- [82] Corbett-Detig RB, Cardeno C, Langley CH. Sequence-based detection and breakpoint assembly of polymorphic inversions. *Genetics*. 2012;**192**:131-137. DOI: 10.1534/genetics.112.141622
- [83] Guillén Y, Ruiz A. Gene alterations at *Drosophila* inversion breakpoints provide prima facie evidence for natural selection as an explanation for rapid chromosomal evolution. *BMC Genomics*. 2012;**13**(1):53. DOI: 10.1186/1471-2164-13-53
- [84] Papaceit M, Segarra C, Aguadé M. Structure and population genetics of the breakpoints of a polymorphic inversion in *Drosophila subobscura*. *Evolution*. 2012;**67**:66-79. DOI: 10.1111/j.1558-5646.2012.01731.x
- [85] Orengo DJ, Puerma E, Papaceit M, Segarra C, Aguadé MA. Molecular perspective on a complex polymorphic inversion system with cytological evidence of multiply reused breakpoints. *Heredity (Edinburgh)*. 2015;**114**(6):610-618. DOI: 10.1038/hdy.2015
- [86] Puerma E, Orengo DJ, Aguadé M. The origin of chromosomal inversions as a source of segmental duplications in the *Sophophora* subgenus of *Drosophila*. *Scientific Reports*. 2016;**6**:30715. DOI: 10.1038/srep30715
- [87] Puerma E, Orengo DJ, Aguadé M. Multiple and diverse structural changes affect the breakpoint regions of polymorphic inversions across the *Drosophila* genus. *Scientific Reports*. 2016;**6**:36248. DOI: 10.1038/srep36248
- [88] McClintock B. Maize genetics. *Carnegie Institution of Washington Year Book*. 1946;**45**:176-186
- [89] Capy P. Classification and nomenclature of retrotransposable elements. *Cytogenetic and Genome Research*. 2005;**110**(1-4):457-461. DOI: 10.1159/000084978
- [90] Finnegan DJ. Eukaryotic transposable elements and genome evolution. *Trends in Genetics*. 1989;**5**:103-107. DOI: 10.1016/0168-9525(89)90039-5
- [91] Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, et al. A unified classification system for eukaryotic transposable elements. *Nature Reviews. Genetics*. 2007;**8**(12):973-982. DOI: 10.1038/nrg2165
- [92] Deprá M, Ludwig A, Valente VLS, Loreto ELS. Mar, a MITE family of hAT transposons in *Drosophila*. *Mobile DNA*. 2012;**3**:13. DOI: 10.1186/1759-8753-3-13
- [93] Kidwell MG. Transposable elements and the evolution of dipteran genomes. In: Yeates DK, Wiegmann BM, editors. *The Evolutionary Biology of Flies*. New York: Columbia University; 2005. pp. 145-173
- [94] Zelentsova H, Shostak N, Kozitsina M, Barskyi V, et al. *Penelope*, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. *Proceedings of the National Academy of Sciences of the United States of America*. 1997;**94**(1):196-201. PMID: PMC19282
- [95] Engels WR, Preston CR. Formation of chromosome rearrangements by *P* factors in *Drosophila*. *Genetics*. 1984;**107**:657-678. PMID: PMC1202383

- [96] Lyttle TW, Haymer DS. The role of the transposable element *hobo* in the origin of endemic inversions in wild populations of *Drosophila melanogaster*. *Genetica*. 1992;**86**:113-126. DOI: 10.1007/BF00133715
- [97] Regner LP, Pereira MSO, Alonso CEV, Abdelhay E, Valente VLS. Genomic distribution of *P* elements in *Drosophila willistoni* and a search for their relationship with chromosomal inversions. *The Journal of Heredity*. 1996;**87**:191-198. DOI: 10.1093/oxfordjournals.jhered.a022984
- [98] Zelentsova H, Poluectova H, Mnjoian L, Lyozin G, Veleikodvorskaja V, et al. Distribution and evolution of mobile elements in the *virilis* species group of *Drosophila*. *Chromosoma*. 1999;**108**(7):443-456. DOI: 10.1007/s004120050396
- [99] Evgen'ev M, Zelentsova H, Mnjoian L, Poluectova H, Kidwell MG. Invasion of *Drosophila virilis* by the Penelope transposable element. *Chromosoma*. 2000;**109**(5):350-357. DOI: 10.1007/s004120000086
- [100] Marzo M, Puig M, Ruiz A. The *Foldback-like* element *Galileo* belongs to the *P* superfamily of DNA transposons and is widespread within the *Drosophila* genus. *Proceedings of the National Academy of Sciences of the United States of America*. 2008;**105**(8):2957-2962. DOI: 10.1073/pnas.0712110105
- [101] Puig M, Cáceres M, Ruiz A. Silencing of a gene adjacent to the breakpoint of a widespread *Drosophila* inversion by a transposon-induced antisense RNA. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;**101**(24):9013-9018. DOI: 10.1073/pnas.0403090101
- [102] González J, Nefedov M, Bosdet I, Casals F, Calvete O, et al. A BAC-based physical map of the *Drosophila buzzatii* genome. *Genome Research*. 2005;**15**:885-892. DOI: 10.1101/gr.3263105
- [103] Ruiz A, Heed WB, Wasserman M. Evolution of the *mojavensis* cluster of cactophilic *Drosophila* with descriptions of two new species. *The Journal of Heredity*. 1990;**81**(1):30-42. DOI: 10.1093/oxfordjournals.jhered.a110922
- [104] Sturtevant AH, Plunkett CR. Sequence of corresponding third chromosome genes in *Drosophila melanogaster* and *Drosophila simulans*. *The Biological Bulletin*. 1926;**50**:56-60. DOI: 10.2307/1536631
- [105] Ashburner M. *Drosophila* Genomes by the Baker's Dozen. *Genetics*. 2007;**177**(3):1263-1268. PMID: PMC2147995

