Peer-reviewed version available at *Genes* **2018**, *9*, 515; doi:10.3390/genes9100515

Evolution of Plant B Chromosome Enriched Sequences

André Marques[1,*], Sonja Klemme[2] & Andreas Houben[3]

[1]Laboratory of Genetic Resources, Federal University of Alagoas, Av. Manoel Severino Barbosa, Bom Sucesso, Arapiraca – AL, CEP:57309-005, Brazil

[2]Biology Centre, Czech Academy of Sciences, Institute of Plant Molecular Biology, Branišovská 31, CZ-37005 České Budějovice, Czech Republic

[3]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, 06466 Stadt Seeland, Corrensstrasse 3, Germany

*Corresponding author: andre.marques@arapiraca.ufal.br

Abstract: B chromosomes are supernumerary chromosomes which are found in addition to the normal standard chromosomes (A chromosomes). B chromosomes are well known to accumulate several types of repeats. Although the evolution of B chromosomes has been subject of numerous studies, the mechanisms of accumulation and evolution of repetitive sequences is not fully understood. Recently, new genomic approaches have shed light on the origin and accumulation of different classes of repetitive sequences in the process of B chromosome formation and evolution. Here we discuss the impact of repetitive sequences accumulation on the evolution of plant B chromosomes.

Key words: B chromosome; satellite DNA; mobile element; organelle DNA; chromosome evolution

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

Introduction

Supernumerary B chromosomes (Bs) are not required for the normal development of organisms and are assumed to represent a specific type of selfish genetic elements. As a result, Bs follow their own species-specific evolutionary pathways. Several recent studies have confirmed the widely accepted view that Bs are derived from their respective A chromosome complement [1-6]. Although Bs may vary in structure and chromatin properties in a species-specific way, their *de novo* formation is probably a rare event, because the occurrence of similar B chromosome variants within related species suggests that they arose from a single origin either from the same or from a related species [7,8].

The evolution of B chromosomes' (Bs) architecture has historically been of interest mainly at the cytogenetic level, with a recent focus to more molecular and genomic level studies (reviewed in [9-11]). Here we focus our review on the impact of repetitive DNA on the evolution of plant B chromosomes.

Methods and tools to characterize the high-copy DNA composition of B chromosomes - past and future

Early investigations about the DNA composition of Bs were based on gradient density centrifugation and renaturation kinetics. In rye, these experiments showed that the ratio and heterogeneity of repeats in Bs did not differ from As, although a slight increase in cytosine and guanine content was observed in the DNA from B-carrier plants [12]. Buoyant densities of DNA from maize plants with and without Bs were found to be similar [13]. Later, [14] introduced the use of comparative restriction endonuclease digestion of genomic DNA of 0B and +B organisms of several *Glossina* species to characterize Bs. Next generation sequencing (NGS) in combination with bioinformatics paved the way to accelerate studies on B chromosomes regarding their sequence composition, gene content and evolution in many ways. In principle, NGS-based methods for the identification of B-specific sequences can be classified in two strategies [10]. For the first strategy, DNA reads of microdissected or flow-sorted B chromosomes are used. Chromosome flow sorting usually produce larger amounts of DNA than microdissection. The most severe disadvantages of sequencing microdissected probes are the low amount of template DNA,

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

contamination by surrounding material, and PCR amplification bias. In consequence, *in silico* purification of produced sequence reads is recommend. However, recent advantages in single cell analysis demonstrated that even the DNA of a single haploid nucleus is sufficient for NGS analysis [15]. The second strategy requires sequencing of two whole genome datasets of the same species, one containing Bs (+B) and one without Bs (0B). It is an indirect approach because B-derived sequences are compared against the 0B-derived sequences as additional step. This approach identifies B-candidate sequences where the ratio of aligned sequences is significantly increased in the +B dataset compared to the 0B dataset.

Identification of B chromosome-enriched sequences like satellite repeats, mobile elements or organelle-derived sequences by similarity-based clustering of next generation sequence reads was achieved using the RepeatExplorer software [16,17] for rye [1] and *Plantago lagopus* [18]. With the RepeatExplorer and RepeatMasker programs the satellite DNA composition of the migratory locust (*E. plorans*) B chromosomes were determined [19]. The 'satellitome' in the grasshopper *E. monticola* consists of 27 satellite DNAs [20], less than half than in the migratory locust, where [19] found 62.

The RepeatExplorer pipeline employs graph representation of read similarities to identify clusters of frequently overlapping reads representing various repetitive elements or their parts [21]. In addition, it provides information about repeat quantities, information about cluster connections via paired-end sequence reads used to identify repeats split between multiple clusters, and outputs from BLASTN and BLASTX similarity searches to custom databases of repetitive elements and repeat-encoded conserved protein domains that aid in repeat annotation.

The *in silico* 'coverage ratio analysis' relies on a read alignment analysis performed for each of the 0B and +B datasets that subsequently are investigated for differences in the read coverage ratio. The strategy was used to determine the B chromosome sequence content of the cichlid *Astatotilapia latifasciata* [3]. Coverage ratio analysis revealed that the B chromosome contains thousands of sequences that have been duplicated from almost all standard chromosomes of this species, although most B-

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

located genes are fragmented. Subsequent sequence analysis of microdissected *A. latifasciata* Bs confirmed this conclusion.

The third comparative approach is a *k*-mer based analysis termed '*k*-mer frequency ratio analysis'. Similar to coverage ratio analysis, it allows the analysis of 0B and +B sequence sets and investigates the differences in the *k*-mer frequency ratio. The web-based tool Kmasker [22] can be applied to run a *k*-mer frequency ratio analysis. The advantages and disadvantage of the different strategies are further discussed in Ruban, Schmutzer, Scholz and Houben [10].

Accumulation of satellite DNA

Because of the high degree of evolutionary conservation and the high copy number, 5S and 45S rDNA satellite repeats have been used as *in situ* hybridization probes for the analysis of Bs since long. The involvement of rDNA satellites in the evolution of plant Bs does not appear to be accidental, because rDNA loci have been detected on Bs of many species of plants [e.g. *Crepis capillaris* (Maluszynska and Schweizer, 1989), *Brachycome dichromosomatica* (Donald et al., 1997; Donald et al., 1995), *Aegilops* (Friebe et al., 1995), and animals [e.g. *Haplochromis obliquidens* (Poletto et al., 2010) and *Eyprepocnemis plorans* (Lopez-Leon et al., 1994)]. Although transcription of B chromosome-located rDNA was long believed not to occur (Ishak et al., 1991), more recent studies have shown that indeed they are expressed, for instance the B-located 45S rRNA genes of the plant *C. capillaris* (Leach et al., 2005) and the grasshopper *Eyprepocnemis plorans* (Ruiz-Estevez et al., 2012). In contrast, the ribosomal RNA genes specific to the B chromosomes in *Brachycome dichromosomatica* are not transcribed [23]. Differences in posttranslational histone modifications, such as histone methylation or acetylation, between A and B chromosomes, which may be responsible for differences in gene activity, have been demonstrated [24-28]. Another possibility is that suppression may occur because of nucleolar dominance, so that the rRNA genes on the A chromosomes are active at the expense of those on the B chromosomes [23]. The inactivity of B chromosome rDNA might explain the presence of multiple ITS2 sequences, since homogenization of rDNA spacers is thought to occur only in transcribed regions. Concerted evolution is a feature of rDNA repeats [29], but mechanisms that control it may not include non-

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

transcribed rDNA sequences [30,31]. Since no sequence homogenization occurs between A and B chromosomes' rDNAs, and since Bs are not constitutively active, one might expect further sequence drift of B-located rDNA sequences. An increase in the number of different members of the already heterogeneous population of ITS sequences could be the evolutionary consequence, as has been postulated for the B chromosome-like paternal sex ratio chromosome of the wasp *Trichogramma kaykai* [32].

In the herb *Plantago lagopus* L. the B chromosome is the product of a spontaneous amplification process of 5S ribosomal DNA derived repeats (Dhar et al., 2002; Kumke et al., 2016). Interestingly, amplification of satellite DNA has been used for the formation of engineered mammalian chromosomes ('satellite-DNA-based-artificial-chromosomes' [33]). In contrast to the situation with animals, the molecular mechanism of sequence amplification in plants is poorly understood. However, except for tobacco [34] no amplification-stimulating DNA elements from plants have been identified thus far. Alternatively, B chromosomal rDNA sites could be a consequence of the reported mobile nature of rDNA [35]; Bs may be the preferred "landing sites" because of B chromosomes' inactivity. Increasing evidence also indicates that rDNA can change position within the genome without corresponding changes in the surrounding sequences [36-38]. Beside rDNA, Bs accumulate chromosome specific satellite DNA (satDNA) which are listed below in a species-specific way.

Rye

The first plant B chromosome-specific satellite repeat has been identified in rye. Comparative restriction digestion of 0B and +B genomic DNA resulted in the isolation of the high-copy repeats E3900 and D1100, which seem to have *de novo* evolved on rye Bs [39-41]. Both repeats have classical features of satDNA as being tandemly repeated, although atypical features as transcriptional activity and euchromatic histone modifications have been reported for these repeats [7,28]. They also show an atypical repeat unit size with 1.1 and 3.9 kb, for D1100 and E3900, respectively [39,40]. These two repeats seem to have been assembled from fragments of a variety

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

of sequence elements [41]. Both contain fragments of mobile elements most likely generated by chromosomal rearrangements, although no protein domain responsible for the autonomous mobility has been found [41].

NGS studies on rye Bs have significantly increased our knowledge about B-specific repeat accumulation and evolution [1,7,42,43]. Since the sequencing of rye B, several B-specific repeats have been identified, mostly being satDNA [1,43]. Furthermore, these studies have shown that rye Bs are descended from rearrangements of the rye standard A chromosomes 3RS and 7R, with subsequent accumulation of repeats and genic fragments from other A chromosomal regions, as well as insertions of organellar DNA [1]. *In silico* identification of the high-copy sequence fraction revealed several B-specific repeats [43]. An accumulation of B-enriched satellites was found mostly in the nondisjunction control region of the B, which is transcriptionally active and late-replicating. All B-enriched sequences are not unique to the B but are also present in other *Secale* species, suggesting the origin of the B from As of the same genus [43]. Moreover, while it was shown that Bs contain a similar proportion of repeats as the A chromosomes, the two differed significantly in composition. This was due to the accumulation of B-specific satellite repeats, mostly in the nondisjunction control region at the terminal part of the long arm, as well as in the extended pericentromere [1,7,43]. The high-copy composition of the rye B seems even more conserved than that of the A chromosomes as different hybridization patterns were found for the repeats ScCl11, Sc36c82 and Sc55c1 on As of different accessions but not on Bs [7], suggesting an important role for the maintenance of the B typical structure. An overall scheme of the rye B repeat composition is shown in Figure 1A.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2018 doi:10.20944/preprints201809.0470.v1

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

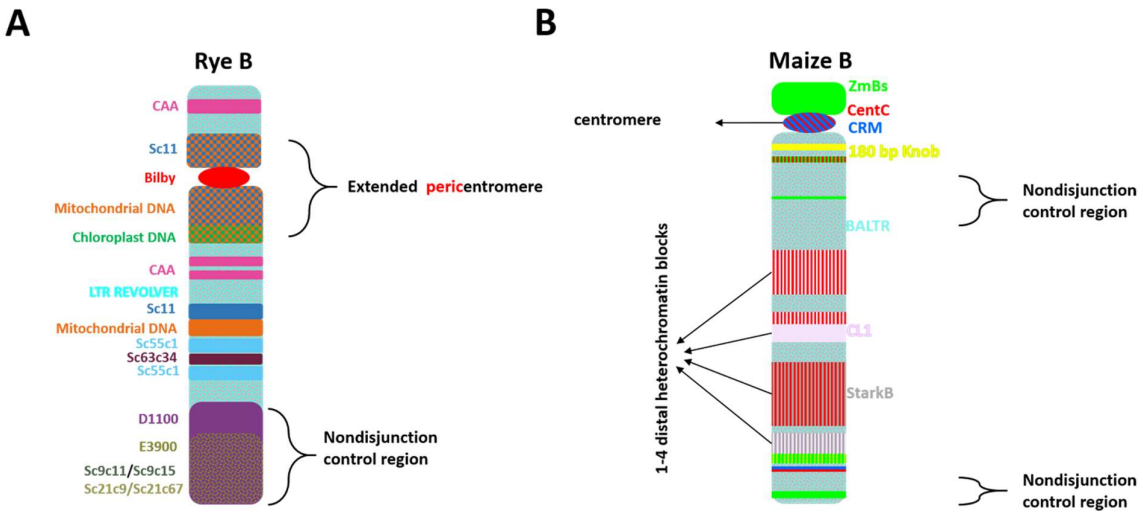

Figure 1. Model for the distribution of (A) rye and (B) maize B chromosome-enriched sequences. While on rye Bs only one region is required for nondisjunction control [44], on maize Bs at least two B chromosome long arm regions that act in trans must be present in the same cell for nondisjunction of the centromere: the very distal tip [45]) and a site in the proximal euchromatin [46].

Maize

The maize (*Zea mays*) B is also well studied and several works have been conducted aiming to understand its composition and accumulation mechanisms. Similar to rye Bs, maize Bs are also characterized by chromosome-type specific repeats [47-51]. The maize B centromere region contains a repetitive element ZmBs that is not present on the A chromosomes, which also shares homology over 90 bp with the maize knob sequence [47]. Later an approximately 700-kb domain that consists of all three previously described repeats (ZmBs repeat, CentC, and CRM) was described that is localised at the core centromere of the maize B and show enhanced association with CENH3 [52].

Two of the elements that are specific to the maize B chromosome are organized in long tandem arrays with repeat units of similar size. The ZmBs repeat, ~1400 bp, is located in and around the B centromere as well as near the tip of the B long arm [47,48]. The CL-1 repeat, ~1500 bp, is present in the first three heterochromatic blocks

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2018 doi:10.20944/preprints201809.0470.v1

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

of the long arm [50,51]. Neither repeat has homology to any known open reading frames or sequences located on As. However, transposition of a retrotransposon and a MITE element involved in the genesis of the CL-1 repeat was detected [51] (Figure 1B).

Brachycome dichromosomatica

The daisy *Brachycome dichromosomatica* has Bs of two different types, the larger Bs are somatically stable whereas the smaller, or micro Bs are somatically unstable. Both types of Bs contain clusters of 45S rDNA.

The large B carries a B-specific tandem repeat (Bd49) that is located mainly at the centromere [53,54]. Multiple copies of sequences related to this repeat are present on the A chromosomes of related species without Bs, whereas only a few copies exist in the As of *B. dichromosomatica* [54]. An isolated Bd49 clone was composed entirely of a tandem array of the repeat unit. However, in other clones the Bd49 repeats were linked to, or interspersed with, sequences that were repetitious and distributed elsewhere on the A and B chromosomes. One such repetitious flanking sequence had similarity to retrotransposon-like sequences and a second was similar to chloroplast DNA [55].

The micro Bs share DNA sequences with the As and the larger Bs, and they also have B-specific repeats (Bdm29 and Bdm54) [56,57]. Bdm29 is highly methylated and after in situ hybridization labelled the entire micro Bs. The Bdm29 is AT-rich with an insert of 290 bp long containing no significant subrepeats. A high number of Bdm29-like sequences was also found in the larger Bs of *B. dichromosomatica* and in other Bs within the genus, suggesting that the Bdm29 sequence is highly conserved and widespread [56]. The Bdm54 repeat is AT-rich with an insert of 477 bp long containing four copies of a subrepeat unit (TCGAAAAGTTCGAAG) as well as three perfect and four degenerate copies of a second short repeat (AGTTCGAA) that are embedded in the first unit [57]. Some micro B-located repeats have been shown to occur as clusters on the A chromosomes in a proportion of individuals within a population [58]. The observation that the genomic organization of the micro B is unlike anything found

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2018 doi:10.20944/preprints201809.0470.v1

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

on the A chromosomes precludes their origin by simple excision from an A chromosome and also indicates that micro Bs do not integrate directly into the A complement to form polymorphic heterochromatic segments [57].

Aegilops speltoides

The *Aegilops* Bs are also known to accumulate several repetitive sequences being characterized by a number of A chromosome-localized repeats like Spelt1, pSc119.2 tandem repeats, 5S rDNA and Ty3-gypsy retroelements [59-62], as well as organelle-derived DNA [63]. However, no B-specific repeat has been found in this species yet.

Plantago lagopus

In *Plantago lagopus*, a weed of the Mediterranean region, a B-specific satellite PLsatB was identified originated from sequence amplification including 5S rDNA fragments [64]. The satellite repeat PLsatB makes up 3.3% of the 1B genotype but only 0.09% of the 0B genotype [18]. *In situ* hybridization with the B-repeat revealed an almost uniform labelling of the Bs at meiotic metaphase I, while extended mitotic prometaphase chromosomes showed a more clustered distribution of the hybridization signals. In any case, no signals were detectable on the A chromosomes. Although PLsatB has evolved from 5S rDNA sequences, a 5S rDNA probe was not sufficient to label the Bs by FISH [18].

Cestrum

Species of *Cestrum* have shown large diversity in the accumulation and distribution of repetitive DNA families [65], and Bs have been described in seven species: *C. strigilatum*, *C. diurnum*, *C. parqui* × *C. aurantiacum*, *C. intermedium*, *C. parqui, C. euanthes* and *C. nocturnum* [66-69]. Some of these repeats have already been identified and associated with B chromosomes [66]. In the hybrid *C. parqui* × *C. aurantiacum*, for instance, the B chromosome contains 35S and 5S rDNA and SSR AT-rich motifs [68]. Sequences of rDNA were also identified in Bs of *C. parqui*, *C. euanthes* and *C. nocturnum* [69]. In *C. intermedium* and *C. strigilatum*, besides C-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2018 doi:10.20944/preprints201809.0470.v1

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

Giemsa+/CMA+/DAPI+ bands [66], the Bs also display hybridization signals with the Gypsy-like retrotransposon probe but not with rDNA probes [66]. Some types of repetitive DNA were identified in A and B chromosomes in *C. strigilatum* and in species of this plant group, such as AT-rich SSR, 35S and 5S rDNA, C-Giemsa and C-CMA/DAPI bands and retrotransposons [70].

Crepis capillaris

In *Crepis capillaris in situ* hybridization of cells with labelled DNA derived from microdissected Bs confirmed that the B is composed mainly of sequences also present in the A chromosomes, but lacks the main repeats located on A chromosomes [71,72]. No B-specific repeat has been found. The highly abundant repeat B134 shows repeating units with a sequence similarity range from 69% to 90% and characterized by its richness in (CA)n repeats. Members of this family are dispersed throughout the A and B chromosomes but are more concentrated in the pericentromeric heterochromatin of the B, indicating that the molecular organization of B heterochromatin is different from that of the As. B-located B134 repeats also have diverged from those on the As [72].

B chromosome-specific accumulation of organelle DNA

Chloroplast and mitochondrial DNA sequences are frequently transferred into the nuclear genome in plants. The endosymbiotic transfer is usually dependent on recombination-based insertions of large fragments of organellar DNA into the nucleus, resulting in either nuclear insertions of plastid DNA (NUPTs) or nuclear insertions of mitochondrial DNA (NUMTs) [73]. Transfer of organelle DNA to the nucleus is a well-known phenomenon (reviewed in [74-77] and nuclear insertions of plastid DNA (NUPTs) and mitochondrial DNA (NUMTs) have been shown to be involved in the origin of new functional genes [76].

Large insertions of organelle DNA were found on the rye B chromosomes [1]. While plastid DNA is not absent from the As, the B localized sequences are considerably larger. They most likely stem from several independent insertion events. However,

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

why they are organized as clustered regions on Bs is still not well known. Between Bs from different geographical origin there is no difference in organelle DNA content or distribution except for a pericentric inversion after FISH [7]. In contrast, the B chromosomes of *Ae. speltoides* of different accessions showed differences in abundance and location of organelle DNA [63]. However, all tested accessions did show accumulation in the B chromosomes over A chromosomes. When comparing A- and B-derived organelle reads to the original organelle genome sequence from wheat, additionally to the accumulation, the reads from the rye B show less similarity to the source genome [1]. The inserts on the B therefore accumulated more mutations than A located inserts. It is also noticeable that the strongest accumulation happened in the pericentromere for rye Bs, and in the distal chromosome arms for *Ae. speltoides*. Pericentric insertions of organelle DNA have also been shown in rice A chromosomes [78]. Pericentromeric regions generally contain few functional genes, and this low gene density may facilitate the repeated integration of the organelle derived DNA [78]. Alternatively, consistent with the rapid evolution of centromeres [79] after sequence integration, subsequent amplification of these sequences might have occurred within this region. The confinement to mostly one region hints at two different possibilities for enrichment in mitochondrial DNA: directed repeated insertion into the surroundings of this area, or few insertions with subsequent local amplification of these sequences. Although the second explanation seems more likely, the diverse nature of the organelle sequence reads suggests many independent insertion events instead of many copies derived from one incorporation. Interestingly it has also been shown that organelle DNA often integrates several fragments into one location [76]. This would fit well with data indicating many events rather than amplification of just one insertion [1].

What mechanism could account for the accumulation of organellar DNA in B chromosomes? It is known that environmental stresses increase the incorporation of organelle DNA into the nucleus [77]. As the Bs, especially in higher numbers, can be considered as stress factors to the cell [80], their presence might increase the basal rate of DNA transfer. Transfer of organellar DNA to the nucleus is very frequent [74,81,82], but most of the “promiscuous” DNA is also rapidly lost again via a counterbalancing removal process [83]. If this expulsion mechanism is impaired in B

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

chromosomes, then the high turnover rates that prevent such sequences on the A chromosomes from accumulating and degrading would be absent and allow for sequence decay. Thus, the dynamic equilibrium between frequent integration and rapid elimination of organellar DNA could be imbalanced for B chromosomes. This hypothesis is supported by higher divergence of B-derived NUMT reads compared to A-derived NUMT reads in *S. cereale* [1]. Future analyses of other B-bearing species are needed to address the question whether organelle-to-nucleus DNA transfer is an important mechanism that drives the evolution of B chromosomes. Another possibility for the accumulation is the dependence on double strand breaks (DSBs). If B chromosomes are more prone to DSB, this could aid in the insertion of random available DNA [63]. Organelle DNA on Bs might be underreported in genomic studies, since organelle DNA is generally filtered out during sequence analysis, due to contamination with DNA extracted from organelles. The presence of large insertions of organelle DNA on B chromosomes might be a plant specific phenomenon, as there have not been any reports yet of animal B chromosomes with mitochondrial insertions.

B chromosome-specific accumulation of transposable elements

Recent works have confirmed what has been proposed, that B chromosomes accumulate DNA from various sources existing as amalgamations of mixed repeats and single-copy regions. In such scenario Bs would provide a safe haven for the accumulation and spread of transposable elements (TEs). This has been suggested as a mechanism through which some of the variability in mammalian Y chromosomes have arisen, as random insertions of transposable DNA into different regions of the Y chromosome would result in elements differing with respect to DNA composition and structure [73,84].

Although TEs compose the vast majority of repetitive DNA in eukaryotic genomes, very few studies have been conducted on plants carrying Bs. Thus, little is known about TE accumulation mechanisms on Bs. Evidence that Bs can provide an ideal target for transposition of TEs comes from works on the retrotransposon NATE (*Nasonia* Transposable Element) which has been described from the PSR element of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2018 doi:10.20944/preprints201809.0470.v1

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

N. vitripennis [85,86]. B-specific accumulation of Ty3/gypsy retrotransposons has been also reported for the fish *Alburnus alburnus* (L.) [87]. In plants, the rye Bs show B-specific accumulation of LTR families [1,43]. In maize B was found a retrotransposon derived repeat specifically located on heterochromatic domains, although this repeat lacks the autonomous domains being characterized as non-autonomous chimeric element [48]. A retrotransposon has also been invoked in the transposition of chloroplast DNA into the repeat element Bd49 of the B chromosomes of *B. dichromosomatica* [55]. Thus, insertion of such elements may therefore be responsible for the generation of structural variability in Bs [88].

Rye

Although most repeats are similarly distributed along As and Bs of rye, several transposons are either amplified or depleted on the B chromosome. For instance, the ancient retroelement Sabrina, abundant in all Triticeae and transcriptionally inactive in rye [89], is highly accumulated on As but less abundant on Bs. In contrast, the active element Revolver, as well as the predicted Copia transposon Sc36c82 seem to be more amplified on the Bs [43]. The accumulation of active elements might have its cause in the lack of selection pressure on Bs, where the integration of a mobile element does not interrupt functional genes. Meiotic crossing-over has been proposed to remove (hemizygous) mobile elements [90]. Rye Bs pair frequently with each other and themselves in pachytene [91], but bivalents are less connected by chiasmata than the As [92]. Thus, reduced crossing-over might facilitate retroelement accumulation, as proposed for plant Y chromosomes [93,94].

Based on these observations Klemme, Banaei-Moghaddam, Macas, Wicker, Novak and Houben [43] proposed a model for the selective accumulation of transposable elements on rye Bs: “In the Triticeae ancestor, Sabrina was transposing and spread over the entire genome. After inactivation of Sabrina [89] before or during speciation of rye, the B was formed from the As with Sabrina still present. The newly evolving elements such as Revolver then became active and transposed throughout the rye genome. The dispensable nature of the B and the lack of selective pressure allowed for stronger

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

accumulation of Revolver on the B, even further diluting the remnants of inactive elements which can no longer increase copy number."

Additionally, other retroelements have been also found to be accumulated on rye Bs, for instance, the Copia elements Sc11c32 and Sc11c927, similar to the centromeric sequences Bilby and Sc11 (Figure 1A), expanded more in the extended pericentromere of the B than in those of As [44]. The presumed ancestral elements are still detectable in subterminal positions on As [7,43].

Maize

B-specific accumulation of retrotransposon-derived sequences has also been reported for the Stark B element. This repeat was formerly identified as B-specific sequence family with a relationship to the PREM-1 family of maize retroelements, which are preferentially transcribed in pollen [49]. This B-specific element was later characterized as StarkB and is composed of fragments from the A genome as well as B-specific sequences. The StarkB element is much larger than the other B-specific elements and is not present in large tandem arrays. Different copies of StarkB varied by small insertions, deletions, and duplications as well as single-nucleotide polymorphisms. Using the LTR divergence of retroelements interrupting the B-specific sequences, the minimum age of the StarkB repeat array and, by inference, of the B chromosome, was estimated to be 2 million years [48].

StarkB is distributed throughout the third and fourth blocks of heterochromatin and is not arranged in arrays, in contrast to the other maize B chromosome-specific repeats, ZmBs and CL-1. It is composed of repetitive sequences known from the A genome as well as novel sequences unique to the B chromosome. StarkB elements differ by insertions and deletion and are frequently interrupted by retrotransposons. By determining the divergence of the LTR of retrotransposons inserted into B chromosome-specific sequences, it was possible to place a minimum age on the B chromosome. The formation of the StarkB element reflects a process that generates large amounts of DNA by creating and amplifying novel sequences on the B

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

chromosome. Furthermore, StarkB was found to be expressed. Therefore, the process that contributed to a large portion of the B chromosome combined pre-existing coding regions to produce novel transcripts [48]. Moreover, because the two blocks of heterochromatin that contain StarkB have persisted over many generations, their presence may play a role in B chromosome transmission through involvement in accumulation mechanisms of some type or by providing an adequate amount of chromatin for efficient chromosome transmission [48].

Furthermore, the maize B is also known to share the same centromere-specific retrotransposons (CRM elements) with the As [52]. Another retroelement called BALTR1 (B and A LTR element 1) was also found to be hybridized throughout the A and B genome but showed enrichment near centromeres and on the B long arm [48]. The BALTR1 element was 7892 bp in length with LTR that were 1159 and 1173 bp, and possessed the internal coding sequences and the LTRs. The LTR of BALTR1 did not show similarity to sequences in public databases. Because the coding region was most similar to the CRM family of retrotransposons, this novel retrotransposon is likely a member of the Ty3/gypsy class [48]. An overview of maize B-enriched repeats is shown in Figure 1B.

Rapid evolution of repeats on Bs

The accumulation of repeats accompanies the evolution of Bs in several plant species, suggesting that the Bs most likely represents a suitable chromosomal context for satellite expansion. But why do Bs often accumulate repeats?

There are at least two possible pathways for Bs to undergo repeat accumulation. One likely source for the accumulation of repeats on Bs comes from studies conducted on rye where a restriction of meiotic recombination in Bs was observed, which showed a variation in the frequency of bivalents formation among different genotypes [95,96]. This restriction of recombination can be considered as starting point for the independent evolution of Bs. The presence of fast-evolving repetitive sequences, could predispose a nascent B to undergo further rapid structural modifications required to establish and amplify new B-specific repeats. Additionally, Bs are under

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

reduced selective pressure due to their non-essential nature, as far as it does not affect its accumulation mechanism. Such chromosomal environment may be considered as a safe haven for non-coding fast-evolving sequences, such as promiscuous DNA as satDNA, organelle DNA and TEs, which are frequent genome hitchhikers able to settle in non-recombining regions. Thus, taking in account the features of Bs, these two pathways may explain the great diversification of repeats found on Bs.

Recent papers showed that both the expansion and shrinkage of the Y chromosome is influenced by sex-specific regulation of repetitive DNA. It was proposed that the dynamics of Y chromosome formation is an interplay of genetic and epigenetic processes [73]. Eukaryotic genomes contain a large proportion of repetitive DNA sequences, mostly TEs and tandem repeats. These repetitive sequences often colonize specific chromosomal (Y or W chromosomes, B chromosomes) or subchromosomal (telomeres, centromeres) niches. Sex chromosomes, especially non-recombining regions of the Y chromosome, are subject to different evolutionary forces compared with autosomes. In non-recombining regions of the Y chromosome repetitive DNA sequences are accumulated, representing a dominant and early process forming the Y chromosome, probably before genes start to degenerate [97]. A similar mechanism may take place in Bs, since it is known that, at least for rye, the frequency of meiotic recombination and proper bivalent formation may vary greatly among different genotypes as discussed above.

Because Bs are non-essential they are supposed to have a neutral situation in host genomes and thus it is expected to observe high B sequence variability among different samples. The origin of B structural variants is mostly attributed to be of monophyletic origin from a unique type of ancestral B chromosome which afterwards diverged in different types through generations [7,8,98]. Indeed, there are several cases of B polymorphisms whether numerical or structural [99,100]. For the B chromosome of the grasshopper *E. plorans* a large variety of structural variants has been demonstrated among many populations [99]. In plants, B polymorphisms have been mainly attributed to numerical polymorphisms [100,101]. Although, in a few cases

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

B structural variants in natural populations have been identified e.g. *B. dichromosomatica* [98], *Ae. speltoides* [63] and *S. autumnalis* [102].

Evolutionary aspects of repeat accumulation

The non-disjunction region of the rye B containing the repetitive elements E3900 and D1100 have unusual properties including a predisposition to instability [41]. There are striking similarities between maize StarkB and the rye B element, E3900. Both elements are composed of a complex mixture of unique sequences and sequences shared with the respective A genome. Both contain many sites of small duplications and regions of low complexity, composed of single nucleotides or simple satellite repeats. Both are found in clusters intermixed with other B-specific sequences near the end of the chromosome arm but not organized in long tandem arrays [41,48]. These common features may reflect similarities in how these elements formed and expanded, or requirements for B survival. It has been suggested that alterations to the E3900/D1100 region could influence the degree of chromatin packaging and affect the efficiency of the rye B transmission [41]. The presence of a region on the B itself that could modulate transmission frequency would help ensure the long-term survival of the chromosome. When the number of Bs in a population is excessively high, so that the hosts fitness is adversely affected, B variants with lower transmission would be selected [48]. That may explain why a variation of genotypes with different transmission rates was observed in rye Bs [95,96].

As a general model for the accumulation and evolution of repeats on plant Bs we propose the following: (1) A proto-B chromosome derived as a result of multiple translocations and duplications of A chromosome fragments and therefore show a similar sequence composition as the original A chromosome fragments. The presence of a chromosome fragment processing a functional centromere to assure mitotic and meiotic segregation is essential. (2) Gene silencing/erosion followed by restriction of meiotic recombination and reduced selective pressure triggers a B-specific repeat accumulation due to non-essential nature of the B chromosome. At this point the formation of a B chromosome-specific drive/accumulation mechanism is essential for the survival of the B chromosome. (3) Mature Bs are characterized by B-specific repeats and a tolerable impact of the Bs on the fitness of the host organism.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 September 2018 doi:10.20944/preprints201809.0470.v1

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

Unless the B is eliminated, it could constantly accumulate sequences and change its structure (neutral selection) and enhance its drive mechanism (positive selection) along its evolution. Furthermore, it is important to notice that the B centromeres at least in rye and maize show specific features to assure chromosome drive via nondisjunction at first or second pollen grain mitosis, respectively (reviewed in Houben [103]). Thus, the evolution of B-specific (peri)centromeric properties seems to be a key step in the evolution of Bs. Figure 2 shows a diagrammatic summary for the proposed model of repeat accumulation and evolution of Bs.

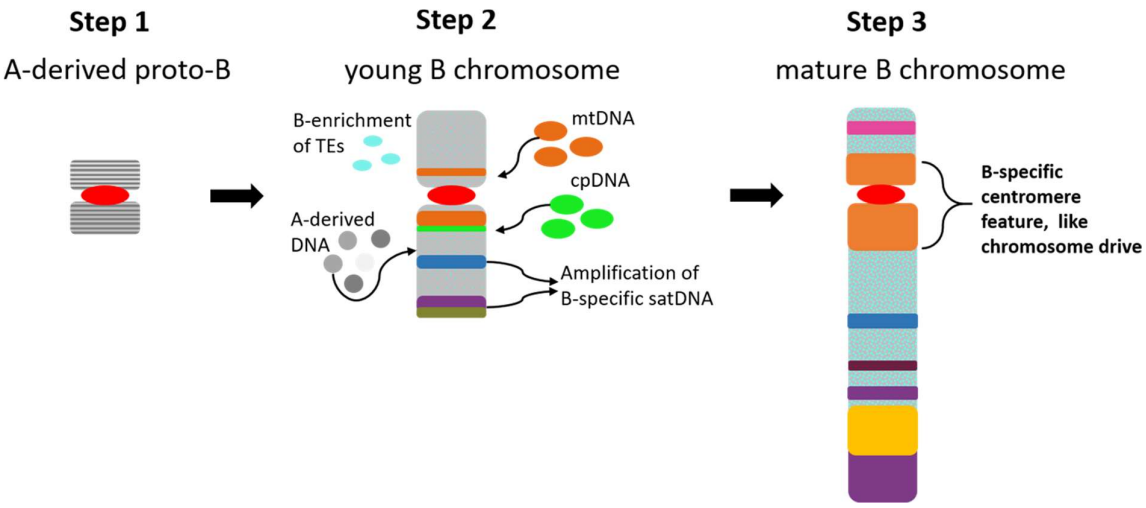

Figure 2. Model for the repeat accumulation and evolution of plant B chromosomes. (1) Proto-B chromosome derived as a result of multiple translocations and duplications of A chromosome fragments (harboring a functional centromere). (2) Gene erosion/silencing followed by restriction of meiotic recombination triggers B-specific repeat accumulation. (3) Mature Bs could achieve a high degree of B-specific repeats, an efficient drive mechanism and a tolerable impact on the fitness of the host organism.

Concluding remarks and future perspectives

As discussed above, Bs are expected to undergo reduced selective pressure due to their non-essential nature and, thus, exhibit larger variation than A chromosomes. Although it is remarkable to see that in fact some Bs have a relatively conserved structure with rare occurrence of structural variants across plant species [7,98]. In

Peer-reviewed version available at Genes 2018, 9, 515; doi:10.3390/genes9100515

some cases, it is remarkable to see that the variation in the As might be even higher than that observed in the Bs, as for instance in *B. dichromosomatica* [98] and for the B-specific repeat ScCl11 in rye [7]. This suggests the existence of a control mechanism for the maintenance of a standard B structure. Although Bs seem always to have an A-derived architecture, they are frequently found to have accumulated several B-specific repeats and, at least in some cases, insertions of organellar DNA. Furthermore, they may show notable variation in their quantitative repeat composition. These unique features observed on Bs highlight on the one hand the origin of Bs from As, and on the other a different evolutionary pathway of As and Bs. It seems that the B acts like a “genomic sponge” which collects and maintains sequences of diverse origins.

Acknowledgments: AM is financially supported by CNPq/FAPEAL DCR fellowship. We thank the Deutsche Forschungsgemeinschaft DFG (HO1779/26-1) for financial support.

Author Contributions: AM and AH designed and wrote the MS. SK wrote the MS.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

Figure legends

Figure 1.

Model for the distribution of (A) rye and (B) maize B chromosome-enriched sequences. While on rye Bs only one region is required for nondisjunction control [44], on maize Bs at least two B chromosome long arm regions that act in trans must be present in the same cell for nondisjunction of the centromere: the very distal tip [45]) and a site in the proximal euchromatin [46].

Figure 2.

Model for the repeat accumulation and evolution of plant B chromosomes. (1) Proto-B chromosome derived as a result of multiple translocations and duplications of A chromosome fragments (harboring a functional centromere). (2) Gene

erosion/silencing followed by restriction of meiotic recombination triggers B-specific repeat accumulation. (3) Mature Bs could achieve a high degree of B-specific repeats, an efficient drive mechanism and a tolerable impact on the fitness of the host organism.

References

1. Martis, M.M.; Klemme, S.; Banaei-Moghaddam, A.M.; Blattner, F.R.; Macas, J.; Schmutzer, T.; Scholz, U.; Gundlach, H.; Wicker, T.; Simkova, H., et al. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences of the USA* **2012**, *109*, 13343-13346, doi:10.1073/pnas.1204237109.
2. Silva, D.M.; Pansonato-Alves, J.C.; Utsunomia, R.; Araya-Jaime, C.; Ruiz-Ruano, F.J.; Daniel, S.N.; Hashimoto, D.T.; Oliveira, C.; Camacho, J.P.; Porto-Foresti, F., et al. Delimiting the origin of a B chromosome by FISH mapping, chromosome painting and DNA sequence analysis in Astyanax paranae (Teleostei, Characiformes). *PloS one* **2014**, *9*, e94896, doi:10.1371/journal.pone.0094896.
3. Valente, G.T.; Conte, M.A.; Fantinatti, B.E.; Cabral-de-Mello, D.C.; Carvalho, R.F.; Vicari, M.R.; Kocher, T.D.; Martins, C. Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Molecular biology and evolution* **2014**, *31*, 2061-2072, doi:10.1093/molbev/msu148.
4. Croll, D.; McDonald, B.A. The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathog* **2012**, *8*, e1002608, doi:10.1371/journal.ppat.1002608.
5. Navarro-Domínguez, B.; Ruiz-Ruano, F.J.; Cabrero, J.; Corral, J.M.; López-León, M.D.; Sharbel, T.F.; Camacho, J.P.M. Protein-coding genes in B chromosomes of the grasshopper *Eyprepocnemis plorans*. *Scientific Reports* **2017**, *7*, 45200, doi:10.1038/srep45200.
6. Ruiz-Ruano, F.J.; Cabrero, J.; Lopez-Leon, M.D.; Camacho, J.P. Satellite DNA content illuminates the ancestry of a supernumerary (B) chromosome. *Chromosoma* **2016**, 10.1007/s00412-016-0611-8, doi:10.1007/s00412-016-0611-8.
7. Marques, A.; Banaei-Moghaddam, A.M.; Klemme, S.; Blattner, F.R.; Niwa, K.; Guerra, M.; Houben, A. B chromosomes of rye are highly conserved and accompanied the development of early agriculture. *Annals of botany* **2013**, *112*, 527-534, doi:10.1093/aob/mct121.
8. Muñoz-Pajares, A.; Martinez-Rodriguez, L.; Teruel, M.; Cabrero, J.; Camacho, J.P.M.; Perfectti, F. A single, recent origin of the accessory B chromosome of the grasshopper *Eyprepocnemis plorans*. *Genetics* **2011**, *187*, 853-863, doi:DOI 10.1534/genetics.110.122713.
9. Valente, G.T.; Nakajima, R.T.; Fantinatti, B.E.; Marques, D.F.; Almeida, R.O.; Simoes, R.P.; Martins, C. B chromosomes: from cytogenetics to systems biology. *Chromosoma* **2017**, *126*, 73-81, doi:10.1007/s00412-016-0613-6.
10. Ruban, A.; Schmutzer, T.; Scholz, U.; Houben, A. How Next-Generation Sequencing Has Aided Our Understanding of the Sequence Composition and Origin of B Chromosomes. *Genes (Basel)* **2017**, *8*, doi:10.3390/genes8110294.
11. Makunin, A.I.; Dementyeva, P.V.; Graphodatsky, A.S.; Volobouev, V.T.; Kukekova, A.V.; Trifonov, V.A. Genes on B chromosomes of vertebrates. *Molecular cytogenetics* **2014**, *7*, 99, doi:10.1186/s13039-014-0099-y.
12. Timmis, J.N.; Ingle, J.; Sinclair, J.; Jones, R.N. Genomic quality of rye B chromosomes. *Journal of Experimental Botany* **1975**, *26*, 367-378.
13. Chilton, M.D.; Mccarthy, B.J. DNA from Maize with and without B-Chromosomes - Comparative Study. *Genetics* **1973**, *74*, 605-614.

14. Amos, A.; Dover, G. The distribution of repetitive DNAs between regular and supernumerary chromosomes in species of Glossina (Tsetse) - a 2-step process in the origin of supernumeraries. *Chromosoma* **1981**, *81*, 673-690, doi:Doi 10.1007/Bf00329579.
15. Dreissig, S.; Fuchs, J.; Himmelbach, A.; Mascher, M.; Houben, A. Sequencing of Single Pollen Nuclei Reveals Meiotic Recombination Events at Megabase Resolution and Circumvents Segregation Distortion Caused by Postmeiotic Processes. *Frontiers in plant science* **2017**, *8*, 1620, doi:10.3389/fpls.2017.01620.
16. Novak, P.; Avila Robledillo, L.; Koblizkova, A.; Vrbova, I.; Neumann, P.; Macas, J. TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic acids research* **2017**, *45*, e111, doi:10.1093/nar/gkx257.
17. Novak, P.; Neumann, P.; Pech, J.; Steinhaisl, J.; Macas, J. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* **2013**, *29*, 792-793, doi:10.1093/bioinformatics/btt054.
18. Kumke, K.; Macas, J.; Fuchs, J.; Altschmied, L.; Kour, J.; Dhar, M.K.; Houben, A. Plantago lagopus B Chromosome Is Enriched in 5S rDNA-Derived Satellite DNA. *Cytogenetic and genome research* **2016**, *148*, 68-73, doi:10.1159/000444873.
19. Ruiz-Ruano, F.J.; Lopez-Leon, M.D.; Cabrero, J.; Camacho, J.P. High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Sci Rep* **2016**, *6*, 28333, doi:10.1038/srep28333.
20. Ruiz-Ruano, F.J.; Cabrero, J.; Lopez-Leon, M.D.; Camacho, J.P.M. Satellite DNA content illuminates the ancestry of a supernumerary (B) chromosome. *Chromosoma* **2017**, *126*, 487-500, doi:10.1007/s00412-016-0611-8.
21. Novak, P.; Neumann, P.; Macas, J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC bioinformatics* **2010**, *11*, 378, doi:10.1186/1471-2105-11-378.
22. Schmutzer, T.; Ma, L.; Pousarebani, N.; Bull, F.; Stein, N.; Houben, A.; Scholz, U. Kmasker--a tool for in silico prediction of single-copy FISH probes for the large-genome species Hordeum vulgare. *Cytogenetic and genome research* **2014**, *142*, 66-78, doi:10.1159/000356460.
23. Donald, T.M.; Houben, A.; Leach, C.R.; Timmis, J.N. Ribosomal RNA genes specific to the B chromosomes in Brachycome dichromosomatica are not transcribed in leaf tissue. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **1997**, *40*, 674-681.
24. Carchilan, M.; Kumke, K.; Mikolajewski, S.; Houben, A. Rye B chromosomes are weakly transcribed and might alter the transcriptional activity of A chromosome sequences *Chromosoma* **2009**, *(in press)*.
25. Kumke, K.; Jones, R.N.; Houben, A. B chromosomes of Puschkinia libanotica are characterized by a reduced level of euchromatic histone H3 methylation marks. *Cytogenetic and Genome Research* **2008**, *121*, 266-270.
26. Marschner, S.; Kumke, K.; Houben, A. B chromosomes of B. dichromosomatica show a reduced level of euchromatic histone H3 methylation marks. *Chromosome Res* **2007**, *15*, 215-222.
27. Jin, W.W.; Lamb, J.C.; Zhang, W.L.; Kolano, B.; Birchler, J.A.; Jiang, J.M. Histone modifications associated with both A and B chromosomes of maize. *Chromosome Research* **2008**, *16*, 1203-1214.
28. Carchilan, M.; Delgado, M.; Ribeiro, T.; Costa-Nunes, P.; Caperta, A.; Morais-Cecilio, L.; Jones, R.N.; Viegas, W.; Houben, A. Transcriptionally active heterochromatin in rye B chromosomes. *The Plant cell* **2007**, *19*, 1738-1749.
29. Dover, G. Concerted evolution, molecular drive and natural selection. *Curr Biol* **1994**, *4*, 1165-1166.

30. Lim, K.Y.; Kovarik, A.; Matyasek, R.; Bezdek, M.; Lichtenstein, C.P.; Leitch, A.R. Gene conversion of ribosomal DNA in Nicotiana tabacum is associated with undermethylated, decondensed and probably active gene units. *Chromosoma* **2000**, *109*, 161-172.
31. Dadejova, M.; Lim, K.Y.; Souckova-Skalicka, K.; Matyasek, R.; Grandbastien, M.A.; Leitch, A.; Kovarik, A. Transcription activity of rRNA genes correlates with a tendency towards intergenomic homogenization in Nicotiana allotetraploids. *New Phytol* **2007**, *174*, 658-668.
32. van Vugt, J.; de Nooijer, S.; Stouthamer, R.; de Jong, H. NOR activity and repeat sequences of the paternal sex ratio chromosome of the parasitoid wasp Trichogramma kaykai. *Chromosoma* **2005**, *114*, 410-419.
33. Csonka, E.; Cserpan, I.; Fodor, K.; Hollo, G.; Katona, R.; Kereso, J.; Praznovszky, T.; Szakal, B.; Telenius, A.; deJong, G., et al. Novel generation of human satellite DNA-based artificial chromosomes in mammalian cells. *Journal of Cell Science* **2000**, *113*, 3207-3216.
34. Borisjuk, N.; Borisjuk, L.; Komarnytsky, S.; Timeva, S.; Hemleben, V.; Gleba, Y.; Raskin, I. Tobacco ribosomal DNA spacer element stimulates amplification and expression of heterologous genes. *Nature Biotechnology* **2000**, *18*, 1303-1306.
35. Schubert, I.; Wobus, U. In situ hybridisation confirms jumping nucleolus organizing regions in Allium. *Chromosoma* **1985**, *92*, 143-148.
36. Abirached-Darmency, M.; Prado-Vivant, E.; Chelysheva, L.; Pouthier, T. Variation in rDNA locus number and position among legume species and detection of 2 linked rDNA loci in the model Medicago truncatula by FISH. *Genome* **2005**, *48*, 556-561.
37. Dubcovsky, J.; Dvorak, J. Ribosomal RNA multigene loci: nomads of the Triticeae genomes. *Genetics* **1995**, *140*, 1367-1377.
38. Datson, P.M.; Murray, B.G. Ribosomal DNA locus evolution in Nemesia: transposition rather than structural rearrangement as the key mechanism? *Chromosome Res* **2006**, *14*, 845-857.
39. Sandery, M.J.; Forster, J.W.; Blunden, R.; Jones, R.N. Identification of a family of repeated sequences on the rye B chromosome. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **1990**, *33*, 908-913.
40. Blunden, R.; Wilkes, T.J.; Forster, J.W.; Jimenez, M.M.; Sandery, M.J.; Karp, A.; Jones, R.N. Identification of the E3900 family, a second family of rye B chromosome specific repeated sequences. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **1993**, *36*, 706-711.
41. Langdon, T.; Jenkins, G.; Seago, C.; Jones, R.N.; Ougham, H.; Thomas, H.; Forster, J.W. De novo evolution of satellite DNA on the rye B chromosome. *Genetics* **2000**, *154*, 869-884.
42. Marques, A.; Klemme, S.; Guerra, M.; Houben, A. Cytomolecular characterization of de novo formed rye B chromosome variants. *Molecular cytogenetics* **2012**, *5*, 34-37, doi:10.1186/1755-8166-5-34.
43. Klemme, S.; Banaei-Moghaddam, A.M.; Macas, J.; Wicker, T.; Novak, P.; Houben, A. High-copy sequences reveal distinct evolution of the rye B chromosome. *New Phytol* **2013**, *199*, 550-558, doi:10.1111/nph.12289.
44. Banaei-Moghaddam, A.M.; Schubert, V.; Kumke, K.; Weibeta, O.; Klemme, S.; Nagaki, K.; Macas, J.; Gonzalez-Sanchez, M.; Heredia, V.; Gomez-Revilla, D., et al. Nondisjunction in Favor of a Chromosome: The Mechanism of Rye B Chromosome Drive during Pollen Mitosis. *The Plant cell* **2012**, *24*, 4124-4134, doi:10.1105/tpc.112.105270.
45. Ward, E.J. Nondisjunction: localization of the controlling site in the maize B chromosome. *Genetics* **1973**, *73*, 387-391.
46. Lin, B.Y. Regional control of nondisjunction of the B chromosome in maize. *Genetics* **1978**, *90*, 613-627.
47. Alfenito, M.R.; Birchler, J.A. Molecular characterization of a maize B chromosome centric sequence. *Genetics* **1993**, *135*, 589-597.
48. Lamb, J.C.; Riddle, N.C.; Cheng, Y.M.; Theuri, J.; Birchler, J.A. Localization and transcription of a retrotransposon-derived element on the maize B chromosome. *Chromosome Research* **2007**, *15*, 383-398.

49. Stark, E.A.; Connerton, I.; Bennett, S.T.; Barnes, S.R.; Parker, J.S.; Forster, J.W. Molecular analysis of the structure of the maize B-chromosome. *Chromosome Research* **1996**, *4*, 15-23.
50. Cheng, Y.M.; Lin, B.Y. Cloning and characterization of maize B chromosome sequences derived from microdissection. *Genetics* **2003**, *164*, 299-310.
51. Cheng, Y.M.; Lin, B.Y. Molecular organization of large fragments in the maize B chromosome: indication of a novel repeat. *Genetics* **2004**, *166*, 1947-1961.
52. Jin, W.; Lamb, J.C.; Vega, J.M.; Dawe, R.K.; Birchler, J.A.; Jiang, J. Molecular and functional dissection of the maize B chromosome centromere. *The Plant cell* **2005**, *17*, 1412-1423, doi:10.1105/tpc.104.030643.
53. John, U.P.; Leach, C.R.; Timmis, J.N. A sequence specific to B chromosomes of Brachycome dichromosomatica. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **1991**, *34*, 739-744.
54. Leach, C.R.; Donald, T.M.; Franks, T.K.; Spiniello, S.S.; Hanrahan, C.F.; Timmis, J.N. Organisation and origin of a B chromosome centromeric sequence from *Brachycome dichromosomatica*. *Chromosoma* **1995**, *103*, 708-714.
55. Franks, T.K.; Houben, A.; Leach, C.R.; Timmis, J.N. The molecular organisation of a B chromosome tandem repeat sequence from *Brachycome dichromosomatica*. *Chromosoma* **1996**, *105*, 223-230.
56. Houben, A.; Leach, C.R.; Verlin, D.; Rofe, R.; Timmis, J.N. A repetitive DNA sequence common to the different B chromosomes of the genus Brachycome. *Chromosoma* **1997**, *106*, 513-519.
57. Houben, A.; Verlin, D.; Leach, C.R.; Timmis, J.N. The genomic complexity of micro B chromosomes of *Brachycome dichromosomatica*. *Chromosoma* **2001**, *110*, 451-459, doi:10.1007/s00412-001-0173-1.
58. Houben, A.; Wanner, G.; Hanson, L.; Verlin, D.; Leach, C.R.; Timmis, J.N. Cloning and characterisation of polymorphic heterochromatic segments of Brachycome dichromosomatica. *Chromosoma* **2000**, *109*, 206-213.
59. Raskina, O.; Brodsky, L.; Belyayev, A. Tandem repeats on an eco-geographical scale: outcomes from the genome of Aegilops speltoides. *Chromosome Research* **2011**, *19*, 607-623, doi:10.1007/s10577-011-9220-9.
60. Hosid, E.; Brodsky, L.; Kalendar, R.; Raskina, O.; Belyayev, A. Diversity of long terminal repeat retrotransposon genome distribution in natural populations of the wild diploid wheat Aegilops speltoides. *Genetics* **2012**, *190*, 263-274, doi:10.1534/genetics.111.134643.
61. Belyayev, A.; Raskina, O. Chromosome evolution in marginal populations of Aegilops speltoides: causes and consequences. *Annals of botany* **2013**, *111*, 531-538, doi:10.1093/aob/mct023.
62. Friebe, B.; Jiang, J.; Gill, B. Detection of 5S rDNA and other repeated DNA on supernumerary B chromosomes of *Triticum* species (Poaceae). *Plant Systematics and Evolution* **1995**, *196*, 131-139.
63. Ruban, A.; Fuchs, J.; Marques, A.; Schubert, V.; Soloviev, A.; Raskina, O.; Badaeva, E.; Houben, A. B chromosomes of *Aegilops speltoides* are enriched in organelle genome-derived sequences. *PloS one* **2014**, *9*, e90214, doi:10.1371/journal.pone.0090214.
64. Dhar, M.K.; Friebe, B.; Koul, A.K.; Gill, B.S. Origin of an apparent B chromosome by mutation, chromosome fragmentation and specific DNA sequence amplification. *Chromosoma* **2002**, *111*, 332-340, doi:10.1007/s00412-002-0214-4.
65. Paula, A.A.; Fernandes, T.; Vignoli-Silva, M.; Vanzela, A.L.L. Comparative cytogenetic analysis of *Cestrum* (Solanaceae) reveals different trends in heterochromatin and rDNA sites distribution. *Plant Biosystems* **2015**, *149*, 976-983, doi:http://dx.doi.org/10.1080/11263504.2014.969354.
66. Fregonezi, J.N.; Rocha, C.; Torezan, J.M.; Vanzela, A.L. The occurrence of different Bs in Cestrum intermedium and C. strigilatum (Solanaceae) evidenced by chromosome banding. *Cytogenetic and genome research* **2004**, *106*, 184-188, doi:10.1159/000079285.

67. Sobti, S.N.; Verma, V.; Rao, B.L.; Pushpangadan, P. In: IOPB chromosome number reports LXV. *Taxon* **1979**, *28*, 627.
68. Sykorova, E.; Lim, K.Y.; Fajkus, J.; Leitch, A.R. The signature of the Cestrum genome suggests an evolutionary response to the loss of (TTTAGGG)n telomeres. *Chromosoma* **2003**, *112*, 164-172, doi:10.1007/s00412-003-0256-2.
69. Urdampilleta, J.D.; Chiarini, F.; Stiefkens, L.; Bernardello, G. Chromosomal differentiation of Tribe Cestreae (Solanaceae) by analyses of 18-5.8-26S and 5S rDNA distribution. *Plant Systematics and Evolution* **2015**, *301*, 1325-1334, doi:https://doi.org/10.1007/s00606-014-1158-x.
70. Vanzela, A.L.L.; de Paula, A.A.; Quintas, C.C.; Fernandes, T.; Baldissera, J.; de Souza, T.B. Cestrum strigilatum (Ruiz & Pavon, 1799) B chromosome shares repetitive DNA sequences with A chromosomes of different Cestrum (Linnaeus, 1753) species. *Comp Cytogenet* **2017**, *11*, 511-524, doi:10.3897/CompCytogen.v11i3.13418.
71. Jamilena, M.; Ruiz Rejon, C.; Ruiz Rejon, M. A molecular analysis of the origin of the *Crepis capillaris* B chromosome. *J Cell Sci* **1994**, *107*, 703-708.
72. Jamilena, M.; Garrido-Ramos, M.; Ruiz Rejon, M.; Ruiz Rejon, C.; Parker, J.S. Characterisation of repeated sequences from microdissected B chromosomes of *Crepis capillaris*. *Chromosoma* **1995**, *104*, 113-120.
73. Hobza, R.; Cegan, R.; Jesionek, W.; Kejnovsky, E.; Vyskot, B.; Kubat, Z. Impact of Repetitive Elements on the Y Chromosome Formation in Plants. *Genes (Basel)* **2017**, *8*, doi:10.3390/genes8110302.
74. Timmis, J.N.; Ayliffe, M.A.; Huang, C.Y.; Martin, W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature reviews. Genetics* **2004**, *5*, 123-135, doi:10.1038/nrg1271.
75. Kleine, T.; Maier, U.G.; Leister, D. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* **2009**, *60*, 115-138, doi:10.1146/annurev.arplant.043008.092119.
76. Lloyd, A.H.; Timmis, J.N. The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome. *Molecular biology and evolution* **2011**, *28*, 2019-2028, doi:10.1093/molbev/msr021.
77. Wang, D.; Lloyd, A.H.; Timmis, J.N. Environmental stress increases the entry of cytoplasmic organellar DNA into the nucleus in plants. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 2444-2448, doi:10.1073/pnas.1117890109.
78. Matsuo, M.; Ito, Y.; Yamauchi, R.; Obokata, J. The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast-nuclear DNA flux. *The Plant cell* **2005**, *17*, 665-675, doi:10.1105/tpc.104.027706.
79. Hall, A.E.; Keith, K.C.; Hall, S.E.; Copenhaver, G.P.; Preuss, D. The rapidly evolving field of plant centromeres. *Current opinion in plant biology* **2004**, *7*, 108-114, doi:10.1016/j.pbi.2004.01.008.
80. Jones, R.N.; Viegas, W.; Houben, A. A century of B chromosomes in plants: so what? *Annals of botany* **2008**, *101*, 767-775, doi:10.1093/aob/mcm167.
81. Huang, C.Y.; Ayliffe, M.A.; Timmis, J.N. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* **2003**, *422*, 72-76, doi:10.1038/nature01435.
82. Sheppard, A.E.; Ayliffe, M.A.; Blatch, L.; Day, A.; Delaney, S.K.; Khairul-Fahmy, N.; Li, Y.; Madesis, P.; Pryor, A.J.; Timmis, J.N. Transfer of plastid DNA to the nucleus is elevated during male gametogenesis in tobacco. *Plant physiology* **2008**, *148*, 328-336, doi:10.1104/pp.108.119107.
83. Sheppard, A.E.; Timmis, J.N. Instability of plastid DNA in the nuclear genome. *PLoS genetics* **2009**, *5*, e1000323, doi:10.1371/journal.pgen.1000323.
84. Graves, J.A. The origin and function of the mammalian Y chromosome and Y-borne genes--an evolving understanding. *BioEssays : news and reviews in molecular, cellular and developmental biology* **1995**, *17*, 311-320, doi:10.1002/bies.950170407.

85. McAllister, B.F. Isolation and characterization of a retroelement from B chromosome (PSR) in the parasitic wasp Nasonia vitripennis. *Insect molecular biology* **1995**, *4*, 253-262.
86. McAllister, B.F.; Werren, J.H. Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. *Chromosoma* **1997**, *106*, 243-253.
87. Ziegler, C.G.; Lamatsch, D.K.; Steinlein, C.; Engel, W.; Schartl, M.; Schmid, M. The giant B chromosome of the cyprinid fish *Alburnus alburnus* harbours a retrotransposon-derived repetitive DNA sequence. *Chromosome Research* **2003**, *11*, 23-35.
88. Camacho, J.P.; Sharbel, T.F.; Beukeboom, L.W. B-chromosome evolution. *Philosophical Transactions of the Royal Society B: Biological Science* **2000**, *355*, 163-178, doi:10.1098/rstb.2000.0556.
89. Shirasu, K.; Schulman, A.H.; Lahaye, T.; Schulze-Lefert, P. A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. *Genome research* **2000**, *10*, 908-915.
90. Charlesworth, B.; Sniegowski, P.; Stephan, W. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **1994**, *371*, 215-220, doi:10.1038/371215a0.
91. Diez, M.; Jimenez, M.M.; Santos, J.L. Synaptic patterns of rye B chromosomes. II. The effect of the standard B chromosomes on the pairing of the A set. *TAG. Theoretical and applied genetics* **1993**, *87*, 17-21, doi:10.1007/BF00223737.
92. Jimenez, G.; Manzanero, S.; Puertas, M.J. Relationship between pachytene synapsis, metaphase I associations, and transmission of 2B and 4B chromosomes in rye. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada* **2000**, *43*, 232-239.
93. Charlesworth, D. Plant sex chromosomes. *Genome dynamics* **2008**, *4*, 83-94, doi:10.1159/000126008.
94. Charlesworth, D. Plant Sex Chromosomes. *Annu Rev Plant Biol* **2016**, *67*, 397-420, doi:10.1146/annurev-arplant-043015-111911.
95. Jimenez, M.M.; Romera, F.; Gonzalez-Sanchez, M.; Puertas, M.J. Genetic control of the rate of transmission of rye B chromosomes. III. Male meiosis and gametogenesis. *Heredity* **1997**, *78*, 636-644.
96. Jimenez, M.M.; Romera, F.; Gallego, A.; Puertas, M.J. Genetic control of the rate of transmission of rye B chromosomes. II. 0Bx2B crosses. *Heredity* **1995**, *74*, 518-523.
97. Kejnovsky, E.; Hobza, R.; Cermak, T.; Kubat, Z.; Vyskot, B. The role of repetitive DNA in structure and evolution of sex chromosomes in plants. *Heredity* **2009**, *102*, 533-541.
98. Houben, A.; Thompson, N.; Ahne, R.; Leach, C.R.; Verlin, D.; Timmis, J.N. A monophyletic origin of the B chromosomes of *Brachycome dichromosomatica* (Asteraceae). *Plant Systematics and Evolution* **1999**, *219*, 127-135.
99. Bakkali, M.; Camacho, J.P.M. The B chromosome polymorphism of the grasshopper *Eyprepocnemis plorans* in North Africa: III. Mutation rate of B chromosomes. *Heredity* **2004**, *92*, 428-433.
100. Lia, V.V.; Confalonieri, V.A.; Poggio, L. B chromosome polymorphism in maize landraces: Adaptive vs. demographic hypothesis of clinal variation. *Genetics* **2007**, *177*, 895-904.
101. Sieber, V.K.; Murray, B.G. Structural and numerical chromosomal polymorphism in natural populations of *Alopecurus* (Poaceae). *Plant Systematics and Evolution* **1981**, *139*, 121-136.
102. Parker, J.S.; Lozano, R.; Taylor, S.; Rejon, M.R. Chromosomal structure of populations of *Scilla autumnalis* in the Iberian Peninsula. *Heredity* **1991**, *67*, 287-297.
103. Houben, A. B Chromosomes - A Matter of Chromosome Drive. *Frontiers in plant science* **2017**, *8*, 210, doi:10.3389/fpls.2017.00210.