

Review

How Next-Generation Sequencing Has Aided Our Understanding of the Sequence Composition and Origin of B Chromosomes

Alevtina Ruban , Thomas Schmutzer, Uwe Scholz  and Andreas Houben * 

Leibniz Institute of Plant Genetics and Crop Plant Research Gatersleben, 06466 Seeland, Germany; ruban@ipk-gatersleben.de (A.R.); schmutzr@ipk-gatersleben.de (T.S.); scholz@ipk-gatersleben.de (U.S.)

* Correspondence: houben@ipk-gatersleben.de; Tel.: +49-(0)-39482-5486

Received: 28 August 2017; Accepted: 24 October 2017; Published: 25 October 2017

Abstract: Accessory, supernumerary, or—most simply—B chromosomes, are found in many eukaryotic karyotypes. These small chromosomes do not follow the usual pattern of segregation, but rather are transmitted in a higher than expected frequency. As increasingly being demonstrated by next-generation sequencing (NGS), their structure comprises fragments of standard (A) chromosomes, although in some plant species, their sequence also includes contributions from organellar genomes. Transcriptomic analyses of various animal and plant species have revealed that, contrary to what used to be the common belief, some of the B chromosome DNA is protein-encoding. This review summarizes the progress in understanding B chromosome biology enabled by the application of next-generation sequencing technology and state-of-the-art bioinformatics. In particular, a contrast is drawn between a direct sequencing approach and a strategy based on a comparative genomics as alternative routes that can be taken towards the identification of B chromosome sequences.

Keywords: B chromosome; supernumerary chromosome; evolution; next generation sequencing

1. Recent Discoveries Related to the Origin and Evolution of B Chromosomes

The origin and evolution of the B chromosomes, which appear to make a non-essential contribution to the overall genome, have puzzled cytogeneticists for over a century. These generally smaller than standard (A) chromosomes are transmitted in a higher than expected frequency, leading to a rise in their number from one generation to the next, in a process termed “drive” [1,2]. The quantum improvement in DNA sequencing power achieved by so-called next-generation sequencing (NGS), along with associated analytical methodologies, now allows for a rigorous investigation of the nucleotide composition of the B chromosomes. The result will finally provide an unequivocal answer to whether or not they harbor genes, whether they affect the function of the genome and how they originated. A number of B-chromosome-carrying species, representing a broad range of taxa, have been targeted in recent years to address these issues. The outcome of applying NGS and extensive cytogenetic analyses has been that the B chromosomes, despite their being non-essential, have been shown to share much in common with A chromosomes, and that they evolved in the various taxa in comparable ways.

Among the plant B chromosome carriers, the major focus on the sequence composition of B chromosomes has been on rye (*Secale cereale*). Based on molecular clock calculations, it has been estimated that rye B chromosomes originated approximately 1.1–1.3 million years ago, 0.4–0.6 million years after the formation of the *Secale* genus [3]. Analysis of flow-sorted B chromosomes has shown that they harbor a substantial amount of A-chromosome-derived DNA sequences. On the basis of these sequences, the B chromosomes represent a multichromosomal mosaic, with the two A chromosomes 3R and 7R making the largest contribution. The distribution of repetitive DNA

along the B chromosome is largely similar to that found in the A chromosomes, although certain transposable elements are either noticeably rarer or noticeably more abundant in the B chromosomes than in the A chromosomes. Two repetitive sequences, arranged as tandem repeats, were shown some time ago to be B-chromosome-specific, but NGS has now uncovered a further nine sequences that appear to be strongly enriched in the B chromosomes—these are most likely tandemly arranged, and are concentrated either in the non-disjunction control region or in the pericentromere. Some of B-specific repeats are transcribed in a tissue-type specific manner [4]. Other sequences have clearly been derived from organellar (plastids and mitochondria) genomes, which is similarly the case for the B chromosomes of the grass species (and wheat progenitor) *Aegilops speltoides* [5]. As a result, it has been proposed that the rye B chromosomes arose in a stepwise manner, possibly as a by-product of evolutionary rearrangements of the A chromosome complement. The suggestion is that the progenitor of the B chromosome arose in conjunction with a segmental or whole genome duplication event, during which segments of several A chromosomes coalesced. An independent mode of evolution of the B chromosomes requires that they no longer are able to associate meiotically with their A chromosome progenitor(s). The prediction flowing from this scenario is that B chromosomes are more likely to have arisen in taxa which have experienced major karyotypic rearrangements [3]. The B chromosomes found in wild and cultivated rye populations of diverse geographical origin are structurally highly conserved, an observation which has been taken to suggest that, despite their rapid initial evolution, once they had become established, their rate of further structural change and their accumulation of repetitive sequence became greatly attenuated. Intriguingly, however, the level of nucleotide polymorphism appears to be much higher in B chromosome genic sequences than in their A chromosome homologs [6]. Some of these genic sequences are actively transcribed and their transcripts may well be functional [7]: for example, a copy of *AGO4B* residing on a B chromosome is transcribed and has been shown to possess RNA slicer activity, at least in vitro [8].

Uniquely among the many taxa that carry B chromosomes, only fungal species harbor definitively functional B chromosomes; in some cases, they even endow a selective advantage on the host. Some are known to harbor genes encoding the virulence function, which allows for host colonization, and others genes underlying potentially adaptive traits [9,10]. The single-molecule real-time (SMRT) sequencing of a strain of the fungal pathogen *Fusarium poae* carrying at least one B chromosome has established that the A and B chromosomes differ in their content—specifically, the former harbor few transposable elements and no gene duplications, while up to 25% of the latter's sequence is composed of transposable elements, and gene duplications are frequent [11]. Similarly, the B chromosome sequence of *Nectria haematococca*, a fungal pathogen belonging to the *Fusarium solani* species complex, comprises a higher proportion of repetitive DNA than does that of the A chromosome complement; its GC content is lower, and it includes both single copy and duplicated genes [12]. The understanding is that fungal B chromosomes represent a part of the genome able to evolve faster than the standard chromosome complement, thereby permitting the rapid development of pathogenicity without disturbing the core genome (see review by Croll and McDonald [13]).

An NGS-enabled comparison of the genomic sequences of 0B and 4B males of the grasshopper species *Eyprepocnemis plorans*, along with the assembled transcriptomes of 0B and 1B females, has revealed ten B chromosome protein-encoding genes, four of which are complete and six truncated [14]. The abundance of transcript derived from half of these genes was significantly higher in the B chromosome carriers, and in some cases, the increase in abundance could be correlated with the number of B chromosome copies present. A gene ontology analysis has suggested that these B-chromosome-encoded genes are predominantly involved in the regulation of cell division, but it has not been established as yet whether the transcripts generated from the B chromosome gene copies are functional.

A sequence analysis of the satellite repeat fraction present in the grasshopper *Eumigus monticola* genome has suggested that one of the autosomes contributed the most for B chromosome formation [15]. Selected satellite repeats and 5S rDNA showed a similar distribution in the proximal third of autosome

S8 and the B chromosome. Two repetitive families, which were considered, on the basis of in situ hybridization, to be B-chromosome-specific, were represented by a high copy number. A bioinformatic analysis concluded, however, that both were in fact also to be found in the A genome complement, although at a density which was too low to detect cytogenetically. The observation was taken as supportive of the intraspecific origin of the B chromosomes. Such conclusions remain provisional however, given the dynamic behavior of satellite repeats. Ruiz-Ruano et al. [16] have also analyzed the repetitive DNA content of micro-dissected B chromosomes carried by the migratory locust *Locusta migratoria*, and demonstrated a substantial difference between the proportions of the B and A chromosome sequence represented by repetitive DNA—respectively, 94.9% and 64.1%. Six different satellite repeats were located on the B chromosome, whereas only one member of the A chromosome complement harbored all of these satellite sequences. On this basis, this chromosome, along with a second autosome, which shares histone gene sequences with the B chromosome, have been proposed as the putative donors of the B chromosome sequences. A further feature of this B chromosome is a 17 Kbp segment composed of 29 distinct transposable elements, indicative of the occurrence of multiple insertion events within this region.

The most drastic impact of B chromosomes documented in the literature relates to the jewel wasp (*Nasonia vitripennis*), in which the B chromosomes (also referred to as the “paternal sex ratio” (PSR) chromosomes) are transmitted exclusively via the sperm, and act to eliminate one set of A chromosomes during the zygote’s first mitosis [17,18]. As a result, a female zygote is converted into a male embryo (see review by Aldrich and Ferree [19]). The PSR-induced elimination of the paternal A chromosomes is regulated by post-translational modifications to the histones associated with the sperm’s chromatin [20]. A transcriptomic analysis of the *N. vitripennis* testis has identified a number of PSR-specific transcripts, which may either encode a functional protein or may represent long non-coding RNA [17]. As yet it remains unclear both how these transcripts relate to the key chromatin modifications and what the nature of the controlling mechanism may be.

A plausible example of the de novo formation of a B chromosome has recently come to light in *Drosophila melanogaster*. Although the presence of B chromosomes has been documented in the *Drosophila* genus since at least 1980, they were first noted in *D. melanogaster* karyotype as recently as 2014 in an established stock containing the *mtrm*¹²⁶ allele of the *matrimony* (*mtrm*) gene. Importantly, no B chromosomes had been identified either in the stock from which the *mtrm*¹²⁶ mutant line was created, or in stocks bearing different *mtrm* alleles [21]. The implication was that a B chromosome had formed over the course of the ten-year period of the stock’s maintenance. These particular B chromosomes were highly heterochromatic and resembled chromosome 4 with respect to the arrangement of certain heterochromatin-related satellite repeats. Their presence has been associated with the meiotic non-disjunction of achiasmate copies of chromosome 4 in females. The B chromosomes do not apparently possess a strong drive mechanism and are thought to be mitotically unstable [21]. While these de novo formed B chromosomes appear to offer an appropriate model for revealing the origin and evolution of B chromosomes, as yet little information has been gathered concerning their sequence composition.

The acquisition of genomic sequence obtained from individual cichlid fish (*Astatotilapia latifasciata*) either carrying or not carrying B chromosomes, as well as that of micro-dissected B chromosome sequence has permitted some clarification regarding the origin and evolution of the B chromosomes present in this species [22]. The sequence data have suggested that a proto B chromosome formed before the diversification of the main lineages of the Lake Victoria population, induced by segmental duplications occurring within the autosomes. Three different A chromosomes appear to have provided most of the material making up the B chromosome, but there remains a level of similarity to most of the A chromosomes. The development of a proto B chromosome appears to reflect an accumulation of A chromosome-derived fragments, followed by a burst of sequence amplification and the establishment of a drive mechanism. Besides the large proportion of repetitive DNA present on the B chromosomes (larger than that on the A chromosomes), a number of genic sequences are also present, although most

of these are gene fragments. The few genes remaining intact present are likely either to have been transferred rather recently to the B chromosomes, or to possess functions important for the maintenance and transmission of the B chromosomes. Among the genic sequences detected which have retained a similar structure to their A chromosome homologs, several are associated with the process of cell division, namely the structure of the kinetochore, recombination, cell cycle progression and microtubule organization. Valente et al. [22] have suggested that those that appear to be transcribed are likely involved in the transmission of the B chromosomes. The PCR amplification of some B chromosome sequences present in the cichlid fish *Metriaclicma zebra* and six other Lake Malawi cichlids has revealed a link between the B chromosomes and the sex of the zygote [23], which has been reported to be the case as well for a number of species [24]. The underlying mechanism of this association is either an effect of drive, which results in one of the sexes carrying a higher number of B chromosomes, or represents the outcome of a mechanism that ensures that B chromosomes are more frequent in males or females in which it drives [23].

A high-quality assembly of the domestic dog genome sequence was combined with a comparative cytogenetics approach by Becker et al. [25] to address aspects of genome architecture in a selection of canid species. As part of the study, the sequence composition of the B chromosomes of the common fox (*Vulpes vulpes*) and two subspecies of raccoon dog (*Nyctereutes procyonoides*) was revealed in some detail. In addition to the identification of additional copies of the proto-oncogene *cKIT* [26], the presence of other genomic regions shared by A and B chromosomes was successfully demonstrated. The inference was that the canid B chromosomes likely arose in a single ancestral species as a byproduct of a genome rearrangement event(s), which led to speciation in the Canidae. The canid B chromosomes represent a pool of duplicated sequences, including cancer-associated genes, many of which are associated with chromosome breakpoints [25,27].

The Siberian roe deer (*Capreolus pygargus*) and grey brocket (*Mazama gouazoubira*) are both B-chromosome-carrying members of the Cetartiodactyla [28]. In contrast to the evolution of the canid B chromosomes, in this case, the B chromosomes seem likely to have originated independently. The acquisition of sequence from flow-sorted B chromosomes has enabled a demonstration that both harbor mainly repetitive DNA, with some representation of autosomal sequences undergoing pseudogenization and of amplified non-repetitive sequences. However, both the composition of the repetitive DNA and the spectrum of A-chromosome-derived sequences present differed greatly between the two species. Those in *C. pygargus* harbored at least two duplicated A chromosome regions containing three genes, and the level of heterozygosity and the number of haplotypes was high. In contrast, the *M. gouazoubira* B chromosomes were relatively homogeneous. There were 26 duplicated regions, harboring 34 intact and 21 partial gene sequences. The presence of both the proto-oncogenes *cKIT* and *RET* [25] in the *M. gouazoubira* B chromosomes suggests that the A chromosome genomic regions that become involved in B formation are not random [28].

Chromosome-specific probes generated from flow-sorted B chromosomes of the rodent *Holochilus brasiliensis* have been exploited for in situ hybridization in *Oryzomyini* spp. [29]. These experiments have revealed some common sequences, referred to as the “anonymous *Oryzomyini* shared heterochromatic region” (OSHR), which are found in 12 of the 15 species analyzed. The OSHR is thought to have arisen 3.0–7.8 million years ago on the sex chromosomes of an ancestral species, spreading later to some autosomes as well as to established B chromosomes through the action of transposable elements. An independent evolution of B chromosomes in the genus *Oryzomyini* has been proposed by Ventura et al. [29]. As yet, the origin of the heterogeneous B chromosomes remains unclear. The proposition is that, in rodents, both the autosomes and the sex chromosomes have contributed to the evolution of B chromosomes [30]. The involvement of sex chromosomes has also been suggested in the frog species *Leiopelma hochstetteri* [31].

2. The Acquisition of Sequences Enriched in B Chromosome

Differences between the A and B chromosomes with respect to their size, structure, and pattern of meiotic pairing behavior offer the opportunity to isolate the B chromosomes via micro-dissection. In plant species where it is difficult to synchronize mitotic division across many cells, advantage can be taken of the natural synchrony associated with meiosis, particularly in the anthers, where large numbers of pollen mother cells passage through meiosis simultaneously. In the earliest reported use of micro-dissection to obtain B chromosome-specific sequences, Sandery et al. [32] attempted to clone into lambda phage DNA obtained from a very large number of rye B chromosomes; although the approach was rather inefficient. The introduction of PCR was responsible for a quantum leap in efficiency, and this technology lies behind most current protocols for chromosome micro-dissection and the subsequent handling of the DNA [33–36]. Successful in situ painting of B chromosomes (e.g., rye [37], *Brachycome dichromosomatica* [38]) with labelled DNA generated after microdissection was possible because of the enrichment of chromosome-specific repetitive sequences, rather than the chromosome specific low- and single-copy sequences. A list of B chromosomes successfully isolated via micro-dissection is given in Table 1. An alternative route to acquiring chromosome-specific DNA takes advantage of the power of flow-sorting to separate chromosomes on the basis of their size [39]. A major advantage of this approach is that it isolates orders of magnitude higher numbers of chromosomes than is feasible using micro-dissection. The resulting DNA can be amplified using either degenerate oligonucleotide primed PCR [40] or Phi29 multiple displacement amplification [41]. The latter technique is more effective where longer amplicons (5–30 Kbp) are preferred [42]. Species for which flow-sorting has been successfully used to purify B chromosome DNA are listed in Table 2.

Table 1. Isolation of B chromosomes by microdissection.

Species	Reference
Rye (<i>Secale cereale</i>)	[32,37,43]
Hawks beard (<i>Crepis capillaris</i>)	[44]
<i>Brachycome dichromosomatica</i>	[38]
Maize (<i>Zea mays</i>)	[45]
Dung beetle (<i>Dichotomius sericeus</i>)	[46]
Locust (<i>Locusta migratoria</i>)	[16,47]
Grasshopper (<i>Podisma kanoi</i>)	[48]
Grasshopper (<i>Podisma sapporensis</i>)	[49]
Grasshopper (<i>Abracris flavolineata</i>)	[50]
Hylid frog (<i>Hypsiboas albopunctatus</i>)	[51]
Fish (<i>Moenkhausia sanctaefilomenae</i>)	[52]
Fish (<i>Astyanax scabripinnis</i>)	[53]
Cichlid fish (<i>Astatotilapia latifasciata</i>)	[22]
Fishes genus <i>Astyanax</i> : <i>A. paranae</i> , <i>A. fasciatus</i> , <i>A. bockmanni</i>	[54]
Fish (<i>Prochilodus lineatus</i>)	[55]
Pacific giant salamander (<i>Dicamptodon tenebrosus</i>)	[56]
Korean field mouse (<i>Apodemus peninsulae</i>)	[57]

Table 2. Isolation of B chromosomes by flow sorting.

Species	Reference
Rye (<i>Secale cereale</i>)	[3,4]
Red fox (<i>Vulpes vulpes</i>)	[26]
Oryzomyines (<i>Holochilus brasiliensis</i>)	[29]
Siberian roe deer (<i>Capreolus pygargus</i>)	[28]
Grey brocket deer (<i>Mazama gouazoubira</i>)	[28]

3. The In Silico-Based Identification of B Chromosome-Enriched Sequences

Various strategies have been elaborated to identify B chromosome sequences from NGS-acquired data. This section summarizes the differences between the direct and indirect (comparative) methods (Figure 1).

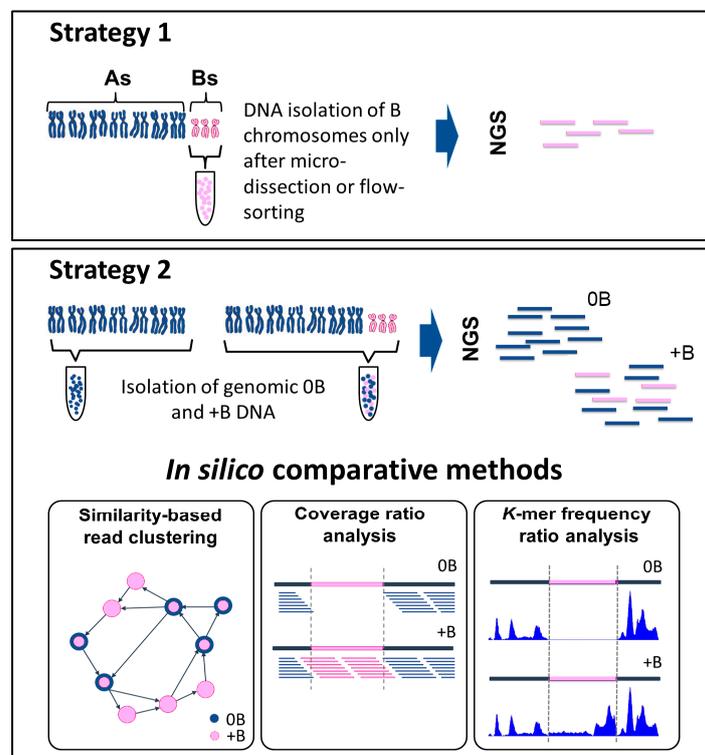


Figure 1. Direct and indirect methods used to identify B chromosome sequences using next generation sequencing (NGS). Strategy 1: the direct method. This approach requires a prior step, in which the B chromosomes are isolated either by micro-dissection or by flow-sorting. Strategy 2: the indirect method. This method requires the acquisition of sequence data from both an individual carrying a B chromosome(s) (+B dataset) and a related individual lacking any B chromosome(s) (0B dataset). The two datasets are compared using three alternative methods. In “similarity-based read clustering”, a graphically based analysis is performed using, for example, the RepeatExplorer pipeline. Sequence information is transformed into graphical structures (vertices correspond to sequence reads and edges characterize the overlap between reads). Differences (presence/absence of sequence reads) in the 0B and +B datasets affect the clusters, and are used to distinguish B chromosome sequences. The two-colored circles indicate reads containing sequences from 0B and +B probes. The “coverage ratio analysis” requires an initial alignment of reads, using an alignment pipeline such as Burrows-Wheeler Alignment tool (BWA). Differences in the read coverage ratio indicate B chromosome-derived candidate regions. The pink section illustrates an example of a putative candidate region, which features the absence of reads in the 0B dataset and their presence in the +B dataset. In the “k-mer frequency ratio analysis” approach, a program such as the Kmasker pipeline identifies differences in the k-mer frequency ratio. The illustration shows an example of a B chromosome segment (shown in pink) in which the k-mer frequency is low or zero in the 0B dataset, but high in the +B dataset. Both the coverage ratio and k-mer frequency ratio analyses, but not the similarity-based read clustering approach, require a reference sequence.

3.1. Strategy 1—The Direct Route: Isolating, then Sequencing Micro-Dissected or Flow-Sorted B Chromosomes

Once B chromosomes have been isolated by either micro-dissection (Table 1) or flow-sorting (Table 2), it is possible to derive their nucleotide content by standard DNA sequencing approaches.

The benefit of this direct method is that there is an a priori assurance that most of the sequences generated are harbored by a B chromosome complement. Employing sufficient sequencing depth, in conjunction with the deployment of advanced bioinformatic tools such as the “targeted chromosome-based cloning via long-range assembly” method [58] can generate sequence assemblies of high quality. Data acquired from a low sequencing depth experiment cannot produce sufficient sequence coverage to allow for a reliable assembly. The major problem encountered with sequencing DNA from micro-dissected material is the noise generated by contamination from non-target chromosomes, from non-target species and from PCR amplification bias. Thus, sequence reads should always be tested (where possible) against reference genome sequences. Here, high specificity and sequence uniqueness is required to identify B chromosome-specific fragments.

Similar to the micro-dissection approach, flow-sorted chromosomes offer a significant reduction in sample complexity, since a specific chromosome can be purified for sequencing. An effective method of sequencing flow-sorted material platform is the so-called “Chicago Hi-C scaffolding” approach, since it requires only small amounts of template DNA [59]. The ability to assemble long sequence scaffolds aids in assessing co-linearity and synteny between B and A chromosomes, and in addressing the origin of B chromosomes sequences. The major limitation encountered with flow-sorting is the difficulty of discriminating between B and fragmented A chromosomes. Measurable progress has been made in recent years towards minimizing this source of contamination [60].

3.2. Strategy 2—The Indirect Route: Comparing Whole Genome Sequence Acquired from Individuals Carrying and Not Carrying B Chromosomes

Inferring a B chromosome location for a given sequence from whole genome sequence data requires a comparison between datasets from a pair of (preferably related) accessions, one of which carries one or more B chromosomes (+B) and the other does not (0B). In principle, the approach identifies peaks where the ratio of aligned sequences is significantly higher in the +B dataset than in the 0B dataset. These regions are identified as putative candidates that are enriched in B chromosome sequences. Here, three different methods have been suggested to identify B chromosome-enriched sequences. The use of several independent +B and 0B identification methods helps to reduce the number of false positives.

3.2.1. Similarity-Based Read Clustering

B chromosome-enriched sequences, such as satellite DNA, retrotransposons, and organelle-derived sequences, can be identified by the similarity-based clustering of NGS reads, as attempted by the RepeatExplorer pipeline, which identifies clusters of frequently overlapping reads, and interprets these as parts of repetitive elements [61]. In addition, the pipeline estimates copy numbers, based on the frequency of duplicate reads. It is able to connect adjacent sequence clusters via the use of paired-end sequence reads. Furthermore, it performs BLAST nucleotide and protein sequence (BLASTN and BLASTX) similarity searches [62] against specialized databases of repetitive elements and repeat-encoded conserved protein domains, which supports the annotation of repetitive elements. To reveal the presence of repetitive elements on a B chromosome, the analysis can be run in a comparative mode, performing a simultaneous clustering of reads from the +B and 0B samples. The structure of the clusters can be investigated using the SeqGrapple program [61]. The approach has been applied with some success in both rye [3] and *Plantago lagopus* [63].

3.2.2. Coverage Ratio Analysis

The “coverage ratio analysis” can be performed by mapping genomic reads against a reference genome [22], as is cited in the manuscript. However, it could be also performed by mapping genomic reads against a reference transcriptome as performed by Navarro-Dominguez et al. [14]. The method works by aligning the +B and 0B dataset, looking for differences in the sequence read coverage ratio (Figure 1). Alignment software such as Burrows-Wheeler Alignment tool (BWA) [64] and Bowtie [65]

can be used to construct sequence alignment/maps [66]. Subsequently, the constructed SAM/BAM files are investigated for regions with different numbers of aligned reads. The B chromosome sequence content of the cichlid fish *A. latifasciata* was determined from high coverage whole genome sequence (acquired with an Illumina HiSeq platform, San Diego, CA, USA) of individuals with and without the B chromosomes, and the reads were mapped onto a reference genome—in this case, that of the related cichlid species *M. zebra* [22]. The coverage ratio analysis revealed that the B chromosomes contain thousands of sequences which have copies on almost every A chromosome. Although most of the genic sequences on the B chromosomes have been fragmented, a few do appear to be intact. Subsequent sequence analysis of micro-dissected *A. latifasciata* B chromosomes has confirmed this conclusion [22].

3.2.3. *k*-mer Frequency Ratio Analysis

A third possible approach is referred to “*k*-mer frequency ratio analysis.” Here, the critical variable is the *k*-mer frequency ratio (Figure 1). A *k*-mer is defined as a sequence fragment of length *k*. The method relies on the construction of a set of such *k*-mer indices covering all sequence motifs occurring in the dataset. Two programs designed to perform this task are Tallymer [67] and Jellyfish [68]. The Kmasker tool [69] can be applied to run the *k*-mer frequency ratio analysis. In addition to its core functionality of masking repetitive elements and identifying low copy sequences, Kmasker can also be used to design both probes for in situ hybridization [70] and single nucleotide polymorphism markers.

The approach was applied in the carnivorous plant species *Genlisea* to study its divergent genome size evolution [71]. In this regard, when comparing the two-sister species *Genlisea nigrocaulis* and *Genlisea hispidula* in their repeat composition, the approach revealed sequence candidates that were involved in the genome size expansion, which is a similar experiment as comparing 0B and +B datasets.

3.3. Benefits and Merits of Indirect and Direct Strategies

The major advantage of the indirect over the direct strategy lies in its not requiring a technical intervention (micro-dissection or flow-sorting), which not only incurs cost, but also introduces an unavoidable degree of contamination by off-target material. While most of the unwanted sequence can be excluded using bioinformatics approaches, this further intervention adds yet another intermediate step. Nevertheless, the direct approach gains from the fact that the bulk of the sequence acquired is relevant, while in the indirect approach, the opposite is the case, since most of the sequence acquired originates from the A chromosome complement or from the organellar genomes. Contamination in the template acquired by micro-dissection is likely to derive from off-target species (microorganisms, human) rather than from the host, whereas for the flow-sorted template, the major source of contaminating DNA is likely to be the host’s A chromosome complement and/or organellar DNA. Where a reference genome sequence has been established, much of the contamination should be identifiable using homology searches, except for sequences that are shared between the B and A chromosomes. This is less obviously the case for a template acquired from micro-dissected chromosomes, as in this case, the source of the contamination is unknown. The challenge for the indirect method is to set an appropriate threshold that minimizes type I error, while still retaining a sufficient number of sequences. Defining this threshold depends on the sequencing depth, the sequence diversity to reference genome sequence and the probability of assembly error. Thus, all sequences identified via the indirect route are associated with a level of uncertainty. In general, the indirect approach is most effective for the discovery of sequences that are abundant on the B chromosomes. In some situations, technical considerations can suggest one method as more suitable than the alternative. For instance, where it is not possible to boost the number of somatic cells undergoing mitosis, flow-sorting becomes inefficient. Similarly, micro-dissection is difficult to carry out where the target chromosome cannot be readily identified on the basis of its morphology. Combining direct and indirect approaches can be an effective strategy, since the outcome of one can be used to validate the outcome of the other.

3.4. An In Silico Method Used to Identify B Chromosome Sequences

One way of assigning the origin of specific sequences to the B chromosome is to make use of synteny between closely related species, a phenomenon whereby interspecific gene order is maintained, at least within relatively short genomic segments. The “genome zipper approach” [72], which exploits this conservation of gene order, has been used in a number of plant species to order and structure NGS sequences [3,72]. As demonstrated in rye [3], the “genome zipper approach” can be extended to B chromosome sequences, once candidate sequences have been identified by a BLASTN analysis against an appropriate reference genome sequence.

4. Conclusions

Combining NGS with state-of-the-art bioinformatics is providing new ways of identifying sequences specific to B chromosomes, revealing a wealth of molecular data relevant for the study of their origin and evolution. Based on sequence data obtained from animal, plant and fungal B chromosomes, the present consensus is that the B chromosomes are composed of duplicated segments derived from potentially multiple A chromosomes, with the addition of some organellar DNA (see review by Houben et al. [73]). Some B chromosomes contain paralogs of A-chromosome-located genes, either as intact or as degenerate sequences. Genic sequences on the B chromosomes do make some contribution to the host transcriptome [74,75]. B-chromosome-specific repeats tend to be derived from the amplification of A chromosome coding and non-coding sequences [76]. The similarities between the B chromosomes and both the so-called “double minute” chromosomes and homogeneously staining regions has suggested that these structures were formed in a comparable manner [76]. There may be parallels between B chromosomes and marker chromosomes in tumor tissue formed by chromothripsis, a process in which several distinct chromosomal regions simultaneously fragment and subsequently are imperfectly reassembled [77]. Human small supernumerary marker chromosomes may serve as an appropriate model for the early evolution of the B chromosomes [78], although these do not share the drive mechanism characteristic of the B chromosomes. Taking into account the growing number of species for which B chromosome-located genic sequences with possible functions have been reported, B chromosomes cannot be considered as “genetically inert” any more. However, their physiological importance still remains at best sketchily understood.

Modern sequencing and bioinformatics methods can be expected to shed new light on the B chromosomes and thereby improve our knowledge of their genomic dynamics. A detailed understanding of the workings of the (peri)centromere will be needed before the mechanistic basis of their characteristic drive can be unraveled. Further progress in RNA sequencing technology will allow for a more rounded picture of the effect on the transcriptome of the B chromosomes to be generated. Additional analysis of the B chromosomes can be expected to provide exciting information relevant to the rapid genome changes that can occur in higher eukaryotes.

Acknowledgments: We thank the Deutsche Forschungsgemeinschaft (HO 1779/261 and SCHO 1420/2-1) for the financial support. Uwe Scholz acknowledges support from the German Ministry of Education and Research (BMBF) for grant 031A536 “de.NBI”.

Author Contributions: A.R., T.S., U.S., and A.H. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Jones, R.N. B-chromosome drive. *Am. Nat.* **1991**, *137*, 430–442. [[CrossRef](#)]
2. Houben, A. B chromosomes—A matter of chromosome drive. *Front. Plant Sci.* **2017**, *8*. [[CrossRef](#)] [[PubMed](#)]

3. 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. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13343–13346. [[CrossRef](#)] [[PubMed](#)]
4. Klemme, S.; Banaei-Moghaddam, A.M.; Macas, J.; Wicker, T.; Novák, P.; Houben, A. High-copy sequences reveal distinct evolution of the rye B-chromosome. *New Phytol.* **2013**, *199*, 550–558. [[CrossRef](#)] [[PubMed](#)]
5. 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. [[CrossRef](#)] [[PubMed](#)]
6. Marques, A.; Klemme, S.; Guerra, M.; Houben, A. Cytomolecular characterization of de novo formed rye B chromosome variants. *Mol. Cytogenet.* **2012**, *5*. [[CrossRef](#)] [[PubMed](#)]
7. Banaei-Moghaddam, A.M.; Meier, K.; Karimi-Ashtiyani, R.; Houben, A. Formation and expression of pseudogenes on the B chromosome of rye. *Plant Cell* **2013**, *25*, 2536–2544. [[CrossRef](#)] [[PubMed](#)]
8. Ma, W.; Gabriel, T.S.; Martis, M.M.; Gursinsky, T.; Schubert, V.; Vrána, J.; Doležel, J.; Grundlach, H.; Altschmied, L.; Scholz, U.; et al. Rye B chromosomes encode a functional Argonaute-like protein with in vitro slicer activities similar to its A chromosome paralog. *New Phytol.* **2016**, *213*, 916–928. [[CrossRef](#)] [[PubMed](#)]
9. Han, Y.; Liu, X.; Benny, U.; Kistler, H.C.; VanEtten, H.D. Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. *Plant J.* **2001**, *25*, 305–314. [[CrossRef](#)] [[PubMed](#)]
10. Harimoto, Y.; Hatta, R.; Kodama, M.; Yamamoto, M.; Otani, H.; Tsuge, T. Expression profiles of genes encoded by the supernumerary chromosome controlling AM-toxin biosynthesis and pathogenicity in the apple pathotype of *Alternaria alternata*. *Mol. Plant Microbe Interact.* **2007**, *20*, 1463–1476. [[CrossRef](#)] [[PubMed](#)]
11. Vanheule, A.; Audenaert, K.; Warris, S.; van de Geest, H.; Schijlen, E.; Höfte, M.; De Saeger, S.; Haesaert, G.; Waalwijk, C.; van der Lee, T. Living apart together: Crosstalk between the core and supernumerary genomes in a fungal plant pathogen. *BMC Genom.* **2016**, *17*, 670. [[CrossRef](#)] [[PubMed](#)]
12. Coleman, J.J.; Rounsley, S.D.; Rodriguez-Carres, M.; Kuo, A.; Wasmann, C.C.; Grimwood, J.; Schmutz, J.; Taga, M.; White, G.J.; Zhou, S.; et al. The genome of *Nectaria haematococca*: Contribution of supernumerary chromosomes to gene expansion. *PLoS Genet.* **2009**, *5*, e1000618. [[CrossRef](#)] [[PubMed](#)]
13. Croll, D.; McDonald, B.A. The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathog.* **2012**, *8*, e1002608. [[CrossRef](#)] [[PubMed](#)]
14. 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*. *Sci. Rep.* **2017**, *7*, 45200. [[CrossRef](#)] [[PubMed](#)]
15. 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* **2017**, *126*, 487–500. [[CrossRef](#)] [[PubMed](#)]
16. Ruiz-Ruano, F.J.; Cabrero, J.; Lopez-Leon, M.D.; Sanchez, A.; Camacho, J.P.M. Quantitative sequence characterization for repetitive DNA content in the supernumerary chromosome of the migratory locust. *Chromosoma* **2017**. [[CrossRef](#)] [[PubMed](#)]
17. Akbari, O.S.; Antoshechkin, I.; Hay, B.A.; Ferree, P.M. Transcriptome profiling of *Nasonia vitripennis* testis reveals novel transcripts expressed from the selfish B chromosome, paternal sex ratio. *G3* **2013**, *3*, 1597–1605. [[CrossRef](#)] [[PubMed](#)]
18. Nur, U.; Werren, J.H.; Eickbush, D.G.; Burke, W.D.; Eickbush, T.H. A “selfish” B chromosome that enhances its transmission by eliminating the paternal genome. *Science* **1988**, *240*, 512–514. [[CrossRef](#)] [[PubMed](#)]
19. Aldrich, J.C.; Ferree, P.M. Genome silencing and elimination: Insights from a “selfish” B chromosome. *Front. Genet.* **2017**, *8*, 50. [[CrossRef](#)] [[PubMed](#)]
20. Aldrich, J.C.; Leibholz, A.; Cheema, M.S.; Ausió, J.; Ferree, P.M. A ‘selfish’ B chromosome induces genome elimination by disrupting the histone code in the jewel wasp *Nasonia vitripennis*. *Sci. Rep.* **2017**, *7*, 42551. [[CrossRef](#)] [[PubMed](#)]
21. Bauerly, E.; Hughes, S.E.; Vietti, D.R.; Miller, D.E.; McDowell, W.; Hawley, R.S. Discovery of supernumerary B chromosomes in *Drosophila melanogaster*. *Genetics* **2014**, *196*, 1007–1016. [[CrossRef](#)] [[PubMed](#)]

22. Valente, G.T.; Conte, M.A.; Fantinatti, B.E.A.; 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. *Mol. Biol. Evol.* **2014**, *31*, 2061–2072. [[CrossRef](#)] [[PubMed](#)]
23. Clark, F.E.; Conte, M.A.; Ferreira-Bravo, I.A.; Poletto, A.B.; Martins, C.; Kocher, T.D. Dynamic sequence evolution of a sex-associated B chromosome in Lake Malawi cichlid fish. *J. Hered.* **2017**, *108*, 53–62. [[CrossRef](#)] [[PubMed](#)]
24. Camacho, J.P.M.; Schmid, M.; Cabrero, J. B chromosomes and sex in animals. *Sex. Dev.* **2011**, *5*, 155–166. [[CrossRef](#)] [[PubMed](#)]
25. Becker, S.E.; Thomas, R.; Trifonov, V.A.; Wayne, R.K.; Graphodatsky, A.S.; Breen, M. Anchoring the dog to its relatives reveals new evolutionary breakpoints across 11 species of the Canidae and provides new clues for the role of B chromosomes. *Chromosom. Res.* **2011**, *19*, 685–708. [[CrossRef](#)] [[PubMed](#)]
26. Graphodatsky, A.S.; Kukekova, A.V.; Yudkin, D.V.; Trifonov, V.A.; Vorobieva, N.V.; Beklemisheva, V.R.; Perelman, P.L.; Graphodatskaya, D.A.; Trut, L.N.; Yang, F.; et al. The proto-oncogene C-KIT maps to canid B-chromosomes. *Chromosom. Res.* **2005**, *13*, 113–122. [[CrossRef](#)] [[PubMed](#)]
27. Yudkin, D.V.; Trifonov, V.A.; Kukekova, A.V.; Vorobieva, N.V.; Rubtsova, N.V.; Yang, F.; Acland, G.M.; Ferguson-Smith, M.A.; Graphodatsky, A.S. Mapping of *KIT* adjacent sequences on canid autosomes and B chromosomes. *Cytogenet. Genome Res.* **2007**, *116*, 100–103. [[CrossRef](#)] [[PubMed](#)]
28. Makunin, A.I.; Kichigin, I.G.; Larkin, D.M.; O'Brien, P.C.; Ferguson-Smith, M.A.; Yang, F.; Proskuryakova, A.A.; Vorobieva, N.V.; Chernyaeva, E.N.; O'Brien, S.J.; et al. Contrasting origin of B chromosomes in two cervids (Siberian roe deer and grey brocket deer) unravelled by chromosome-specific DNA sequencing. *BMC Genom.* **2016**, *17*, 618. [[CrossRef](#)] [[PubMed](#)]
29. Ventura, K.; O'Brien, P.C.M.; do Nascimento Moreira, C.; Yonenaga-Yassuda, Y.; Ferguson-Smith, M.A. On the origin and evolution of the extant system of B chromosomes in Oryzomyini radiation (Rodentia, Sigmodontinae). *PLoS ONE* **2015**, *10*, e0136663. [[CrossRef](#)] [[PubMed](#)]
30. Rubtsov, N.B.; Karamysheva, T.V.; Andreenkova, O.V.; Bochkarev, M.N.; Kartavtseva, I.V.; Roslik, G.V.; Borissov, Y.M. Comparative analysis of micro and macro B chromosomes in the Korean field mouse *Apodemus peninsulae* (Rodentia, Murinae) performed by chromosome microdissection and FISH. *Cytogenet. Genome Res.* **2004**, *106*, 289–294. [[CrossRef](#)] [[PubMed](#)]
31. Sharbel, T.F.; Green, D.M.; Houben, A. B-chromosome origin in the endemic New Zealand frog *Leiopelma hochstetteri* through sex chromosome devolution. *Genome* **1998**, *41*, 14–22. [[CrossRef](#)] [[PubMed](#)]
32. Sandery, M.J.; Forster, J.W.; Macadam, S.R.; Blunden, R.; Jones, R.N.; Brown, D.M. Isolation of a sequence common to A- and B-chromosomes of rye (*Secale cereale*) by microcloning. *Plant Mol. Biol. Rep.* **1991**, *9*, 21–30. [[CrossRef](#)]
33. Houben, A. Chromosome microdissection and utilization of microisolated DNA. In *Plant Cytogenetics. Plant Genetics and Genomics: Crops and Models*; Bass, H., Birchler, J., Eds.; Springer: New York, NY, USA, 2011; Volume 4, pp. 257–270.
34. Kosyakova, N.; Liehr, T.; Al-Rikabi, H.A.B. FISH-microdissection. In *Fluorescence In Situ Hybridization (FISH). Application Guide*, 2nd ed.; Liehr, T., Ed.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 81–100.
35. Zhang, Y.X.; Deng, C.L.; Hu, Z.M. The chromosome microdissection and microcloning technique. *Methods Mol. Biol.* **2016**, *1429*, 151–160. [[CrossRef](#)] [[PubMed](#)]
36. Zhou, R.N.; Hu, Z.M. The development of chromosome microdissection and microcloning technique and its applications in genomic research. *Curr. Genom.* **2007**, *8*, 67–72. [[CrossRef](#)]
37. Houben, A.; Kynast, R.G.; Heim, U.; Hermann, H.; Jones, R.N.; Forster, J.W. Molecular cytogenetic characterisation of the terminal heterochromatic segment of the B-chromosome of rye (*Secale cereale*). *Chromosoma* **1996**, *105*, 97–103. [[CrossRef](#)] [[PubMed](#)]
38. 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. [[CrossRef](#)] [[PubMed](#)]
39. Dolezel, J.; Kubalaková, M.; Paux, E.; Bartos, J.; Feuillet, C. Chromosome-based genomics in the cereals. *Chromosom. Res.* **2007**, *15*, 51–66. [[CrossRef](#)] [[PubMed](#)]
40. Telenius, H.; Carter, N.P.; Bebb, C.E.; Nordenskjold, M.; Ponder, B.A.J.; Tunnacliffe, A. Degenerate oligonucleotide-primed PCR—General amplification of target DNA by a single degenerate primer. *Genomics* **1992**, *13*, 718–725. [[CrossRef](#)]

41. Dean, F.B.; Nelson, J.R.; Giesler, T.L.; Lasken, R.S. Rapid amplification of plasmid and phage DNA using Phi29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* **2001**, *11*, 1095–1099. [[CrossRef](#)] [[PubMed](#)]
42. Šimková, H.; Svensson, J.T.; Condamine, P.; Hříbová, E.; Suchánková, P.; Bhat, P.R.; Bartoš, J.; Šafář, J.; Close, T.J.; Doležel, J. Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. *BMC Genom.* **2008**, *9*, 294. [[CrossRef](#)] [[PubMed](#)]
43. Long, H.; Qi, Z.X.; Sun, X.M.; Chen, C.B.; Li, X.L.; Song, W.Q.; Chen, R.Y. Characters of DNA constitution in the rye B chromosome. *J. Integr. Plant Biol.* **2008**, *50*, 183–189. [[CrossRef](#)] [[PubMed](#)]
44. Jamilena, M.; Garrido-Ramos, M.; Rejon, M.R.; Rejon, C.R.; Parker, J.S. Characterisation of repeated sequences from microdissected B chromosomes of *Crepis capillaris*. *Chromosoma* **1995**, *104*, 113–120. [[CrossRef](#)] [[PubMed](#)]
45. Cheng, Y.M.; Lin, B.Y. Cloning and characterization of maize B chromosome sequences derived from microdissection. *Genetics* **2003**, *164*, 299–310. [[PubMed](#)]
46. Amorim, I.C.; Milani, D.; Cabral-de-Mello, D.C.; Rocha, M.F.; Moura, R.C. Possible origin of B chromosome in *Dichotomius sericeus* (Coleoptera). *Genome* **2016**, *59*, 575–580. [[CrossRef](#)] [[PubMed](#)]
47. Teruel, M.; Cabrero, J.; Montiel, E.E.; Acosta, M.J.; Sanchez, A.; Camacho, J.P. Microdissection and chromosome painting of X and B chromosomes in *Locusta migratoria*. *Chromosom. Res.* **2009**, *17*, 11–18. [[CrossRef](#)] [[PubMed](#)]
48. Bugrov, A.G.; Karamysheva, T.V.; Perepelov, E.A.; Elisaphenko, E.A.; Rubtsov, D.N.; Warchalowska-Sliwa, E.; Tatsuta, H.; Rubtsov, N.B. DNA content of the B chromosomes in grasshopper *Podisma kanoi* Storozh. (Orthoptera, Acrididae). *Chromosom. Res.* **2007**, *15*, 315–325. [[CrossRef](#)] [[PubMed](#)]
49. Bugrov, A.G.; Karamysheva, T.V.; Pyatkova, M.S.; Rubtsov, D.N.; Andreenkova, O.V.; Warchalowska-Sliwa, E.; Rubtsov, N.B. B chromosomes of the *Podisma sapporensis* Shir. (Orthoptera, Acrididae) analysed by chromosome microdissection and FISH. *Folia Biol.* **2003**, *51*, 1–11.
50. Menezes-de-Carvalho, N.Z.; Palacios-Gimenez, O.M.; Milani, D.; Cabral-de-Mello, D.C. High similarity of U2 snDNA sequence between A and B chromosomes in the grasshopper *Abracris flavolineata*. *Mol. Genet. Genom.* **2015**, *290*, 1787–1792. [[CrossRef](#)] [[PubMed](#)]
51. Gruber, S.; Diniz, D.; Sobrinho-Scudeler, P.; Foresti, F.; Haddad, C.; Kasahara, S. Possible interspecific origin of the B chromosome of *Hypsiboas albopunctatus* (Spix, 1824) (Anura, Hylidae), revealed by microdissection, chromosome painting, and reverse hybridisation. *Comp. Cytogenet.* **2014**, *8*, 17–29. [[CrossRef](#)] [[PubMed](#)]
52. Scudeler, P.E.; Diniz, D.; Wasko, A.P.; Oliveira, C.; Foresti, F. Whole chromosome painting of B chromosomes of the red-eye tetra *Moenkhausia sanctaefilomenae* (Teleostei, Characidae). *Comp. Cytogenet.* **2015**, *9*, 661–669. [[CrossRef](#)] [[PubMed](#)]
53. Vicari, M.R.; de Mello Pistune, H.F.; Castro, J.P.; de Almeida, M.C.; Bertollo, L.A.; Moreira-Filho, O.; Camacho, J.P.; Artoni, R.F. New insights on the origin of B chromosomes in *Astyanax scabripinnis* obtained by chromosome painting and FISH. *Genetica* **2011**, *139*, 1073–1081. [[CrossRef](#)] [[PubMed](#)]
54. De A. Silva, D.M.; Daniel, S.N.; Camacho, J.P.; Utsunomia, R.; Ruiz-Ruano, F.J.; Penitente, M.; Pansonato-Alves, J.C.; Hashimoto, D.T.; Oliveira, C.; Porto-Foresti, F.; et al. Origin of B chromosomes in the genus *Astyanax* (Characiformes, Characidae) and the limits of chromosome painting. *Mol. Genet. Genom.* **2016**, *291*, 1407–1418. [[CrossRef](#)] [[PubMed](#)]
55. Voltolin, T.A.; Laudicina, A.; Senhorini, J.A.; Bortolozzi, J.; Oliveira, C.; Foresti, F.; Porto-Foresti, F. Origin and molecular organization of supernumerary chromosomes of *Prochilodus lineatus* (Characiformes, Prochilodontidae) obtained by DNA probes. *Genetica* **2010**, *138*, 1133–1139. [[CrossRef](#)] [[PubMed](#)]
56. Brinkman, J.N.; Sessions, S.K.; Houben, A.; Green, D.M. Structure and evolution of supernumerary chromosomes in the Pacific giant salamander, *Dicamptodon tenebrosus*. *Chromosom. Res.* **2000**, *8*, 477–485. [[CrossRef](#)]
57. Karamysheva, T.V.; Andreenkova, O.V.; Bochkarev, M.N.; Borissov, Y.M.; Bogdanchikova, N.; Borodin, P.M.; Rubtsov, N.B. B chromosomes of Korean field mouse *Apodemus peninsulae* (Rodentia, Murinae) analysed by microdissection and FISH. *Cytogenet. Genome Res.* **2002**, *96*, 154–160. [[CrossRef](#)] [[PubMed](#)]
58. Thind, A.K.; Wicker, T.; Simkova, H.; Fossati, D.; Moullet, O.; Brabant, C.; Vrana, J.; Dolezel, J. Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. *Nat. Biotechnol.* **2017**, *35*, 793–796. [[CrossRef](#)] [[PubMed](#)]

59. Putnam, N.H.; O'Connell, B.L.; Stites, J.C.; Rice, B.J.; Blanchette, M.; Calef, R.; Troll, C.J.; Fields, A.; Hartley, P.D.; Sugnet, C.W.; et al. Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Res.* **2016**, *26*, 342–350. [[CrossRef](#)] [[PubMed](#)]
60. Dolezel, J.; Vrana, J.; Capal, P.; Kubalaková, M.; Buresova, V.; Simkova, H. Advances in plant chromosome genomics. *Biotechnol. Adv.* **2014**, *32*, 122–136. [[CrossRef](#)] [[PubMed](#)]
61. Novák, P.; Neumann, P.; Macas, J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinform.* **2010**, *11*. [[CrossRef](#)] [[PubMed](#)]
62. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
63. 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. *Cytogenet. Genome Res.* **2016**, *148*, 68–73. [[CrossRef](#)] [[PubMed](#)]
64. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
65. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
66. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)] [[PubMed](#)]
67. Kurtz, S.; Narechania, A.; Stein, J.C.; Ware, D. A new method to compute *k*-mer frequencies and its application to annotate large repetitive plant genomes. *BMC Genom.* **2008**, *9*, 517. [[CrossRef](#)] [[PubMed](#)]
68. Marçais, G.; Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. *Bioinformatics* **2011**, *27*, 764–770. [[CrossRef](#)] [[PubMed](#)]
69. 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*. *Cytogenet. Genome Res.* **2014**, *142*, 66–78. [[CrossRef](#)] [[PubMed](#)]
70. Aliyeva-Schnorr, L.; Beier, S.; Karafiátová, M.; Schmutzer, T.; Scholz, U.; Doležel, J.; Stein, N.; Houben, A. Cytogenetic mapping with centromeric bacterial artificial chromosomes contigs shows that this recombination-poor region comprises more than half of barley chromosome 3H. *Plant J.* **2015**, *84*, 385–394. [[CrossRef](#)] [[PubMed](#)]
71. Vu, G.T.H.; Schmutzer, T.; Bull, F.; Cao, H.X.; Fuchs, J.; Tran, T.D.; Jovtchev, G.; Pistrick, K.; Stein, N.; Pecinka, A.; et al. Comparative genome analysis reveals divergent genome size evolution in a carnivorous plant genus. *Plant Genome* **2015**, *8*. [[CrossRef](#)]
72. Mayer, K.F.; Martis, M.; Hedley, P.E.; Simkova, H.; Liu, H.; Morris, J.A.; Steuernagel, B.; Taudien, S.; Roessner, S.; Gundlach, H.; et al. Unlocking the barley genome by chromosomal and comparative genomics. *Plant Cell* **2011**, *23*, 1249–1263. [[CrossRef](#)] [[PubMed](#)]
73. Houben, A.; Banaei-Moghaddam, A.M.; Klemme, S.; Timmis, J.N. Evolution and biology of supernumerary B chromosomes. *Cell. Mol. Life Sci.* **2014**, *71*, 467–478. [[CrossRef](#)] [[PubMed](#)]
74. Banaei-Moghaddam, A.M.; Martis, M.M.; Macas, J.; Gundlach, H.; Himmelbach, A.; Altschmied, L.; Mayer, K.F.; Houben, A. Genes on B chromosomes: Old questions revisited with new tools. *Biochim. Biophys. Acta* **2015**, *1849*, 64–70. [[CrossRef](#)] [[PubMed](#)]
75. 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. [[CrossRef](#)] [[PubMed](#)]
76. Makunin, A.; Dementyeva, P.; Graphodatsky, A.; Volobouev, V.; Kukekova, A.; Trifonov, V. Genes on B chromosomes of vertebrates. *Mol. Cytogenet.* **2014**, *7*, 99. [[CrossRef](#)] [[PubMed](#)]

77. Leibowitz, M.L.; Zhang, C.Z.; Pellman, D. Chromothripsis: A new mechanism for rapid karyotype evolution. *Annu. Rev. Genet.* **2015**, *49*, 183–211. [[CrossRef](#)] [[PubMed](#)]
78. Liehr, T.; Mrasek, K.; Kosyakova, N.; Ogilvie, C.M.; Vermeesch, J.; Trifonov, V.; Rubtsov, N. Small supernumerary marker chromosomes (sSMC) in humans; are there B chromosomes hidden among them. *Mol. Cytogenet.* **2008**, *1*, 12. [[CrossRef](#)] [[PubMed](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).