American Genetic Association

Journal of Heredity, 2017, 53–62
doi:10.1093/jhered/esw059
Original Article
Advance Access publication September 14, 2016

OXFORD

Symposium Article

Dynamic Sequence Evolution of a Sex-Associated B Chromosome in Lake Malawi Cichlid Fish

Frances E. Clark, Matthew A. Conte, Irani A. Ferreira-Bravo, Andreia B. Poletto, Cesar Martins, and Thomas D. Kocher

From the Department of Biology, University of Maryland, College Park, Maryland 20742 (Clark, Conte, and Kocher); Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland 20742 (Ferreira-Bravo); and Departamento de Morfologia, Instituto de Biociências, UNESP—Universidade Estadual Paulista, Botucatu, SP, Brazil (Poletto and Martins).

Address correspondence to T. D. Kocher at the address above or e-mail: tdk@umd.edu

Received January 27, 2016; First decision March 21, 2016; Accepted August 26, 2016.

Corresponding Editor: Jill Pecon-Slattery

Abstract

B chromosomes are extra chromosomes found in many species of plants, animals, and fungi. B chromosomes often manipulate common cellular processes to increase their frequency, sometimes to the detriment of organismal fitness. Here, we characterize B chromosomes in several species of Lake Malawi cichlid fish. Whole genome sequencing of *Metriaclima* zebra "Boadzulu" individuals revealed blocks of sequence with unusually high sequence coverage, indicative of increased copy number of those sequences. These regions of high sequence coverage were found only in females. SNPs unique to the high copy number sequences permitted the design of specific amplification primers. These primers amplified fragments only in *Metriaclima lombardoi* individuals that carried a cytologically identified B chromosome (B-carriers), indicating these extra copies are located on the B chromosome. These same primers were used to identify B-carrying individuals in additional species from Lake Malawi. Across 7 species, a total of 43 B-carriers were identified among 323 females. B-carriers were exclusively female; no B chromosomes were observed in the 317 males surveyed from these species. Quantitative analysis of the copy number variation of B-specific sequence blocks suggests that B-carriers possess a single B chromosome, consistent with previous karyotyping of *M. lombardoi*. A single B chromosome in B-carriers is consistent with 2 potential drive mechanisms: one involving nondisjunction and preferential segregation in a mitotic division prior to the germ-line, and the other involving preferential segregation during meiosis I.

Subject area: Genomics and gene mapping
Key words: nondisjunction, preferential segregation, selfish genetic element, supernumerary chromosome

Each species has a typical standard set of chromosomes, referred to as the A chromosomes. In 1907, Wilson discovered supernumerary chromosomes or B chromosomes (Wilson 1907). B chromosomes (Bs) are nonessential chromosomes which are not found in every individual, and that typically do not follow Mendelian inheritance (Jones 1991; Camacho et al. 2000). In natural populations, individuals may carry from 1 to as many as 34 Bs (Randolph 1941; Houben et al. 2013). B chromosomes may also vary in number among cells

of a single organism and often differ between gonadal and somatic tissues (Burt and Trivers 2008).

B chromosomes are thought to arise from the A chromosomes (Burt and Trivers 2008; Martis et al. 2012; Houben et al. 2013). Fluorescent in situ hybridization (FISH) has revealed that Bs often share homologous sequence with at least one A chromosome (Martis et al. 2012; Silva et al. 2014). Studies of the rye (*Secale cereal)* B chromosome suggest that it was created by translocations and unbalanced segregation after a segmental or whole-genome duplication (Martis et al. 2012). The B chromosome in *Drosophila albomicans* was produced as a byproduct of a chromosomal fusion between the ancestral third autosome and the ancestral sex chromosome (Zhou et al. 2012). Another evolutionary model in a Lake Victorian cichlid suggests that B chromosomes can arise as a segmental duplication of an A chromosome that includes a centromere but relatively few genes so that it avoids negative selective pressures arising from unbalanced gene dosage. This proto-B might then undergo an internal duplication, producing an isochromosome with 2 nearly identical arms (Valente et al. 2014). Each of these models also includes subsequent accumulation of A sequences, transferred to the B by nonhomologous recombination, hitchhiking on transposons, or reverse transcription. Once these A sequences are inserted onto the B, they can undergo duplication and reach high copy number on the B. Most of these sequences eventually undergo decay because they experience little, if any, purifying selection. Regardless of the origin of the proto-B, the continued accumulation and subsequent duplication of sequences leads to a high repeat content on B chromosomes (Martis et al. 2012; Zhou et al. 2012; Valente et al. 2014).

B chromosomes are typically univalent and require special drive mechanisms to be maintained in populations (Jones et al. 2008). Drive can occur during nuclear divisions before, during, or after meiosis (Jones 1991; Burt and Trivers 2008). Cytological studies have revealed numerous types of drive in plants and animals (Jones and Rees 1982; Jones 1991; Burt and Trivers 2008). B chromosomes often take advantage of pre-existing meiotic and mitotic machinery to increase their rate of transmission. Mechanisms of drive are often limited to one sex, most frequently the female (Burt and Trivers 2008; Camacho et al. 2011). However, the molecular basis of these drive mechanisms remains largely unknown, with the notable exception of the rye B chromosome (Banaei-Moghaddam et al. 2012).

The majority of known drive mechanisms involve nondisjunction, or preferential segregation, or both (Beukeboom 1994). Normally, when a cell divides, chromosome pairs (homologous chromosomes or sister chromatids) will separate from one another (in meiosis I, and in anaphase/meiosis II, respectively) into different daughter cells. Nondisjunction causes the pair of homologous chromosomes (meiosis I), or 2 sister chromatids (meiosis II or mitosis), to end up in the same daughter cell. It is important to note that nondisjunction does not increase the overall number of Bs in the population; a parent cell with a single B chromosome still produces 2 daughter cells with a total of 2 B chromosomes between them. If those 2 daughter cells are gametes, or have an equal chance of producing gametes, then the number of B chromosomes will not increase in frequency in the next generation. The only difference is that those 2 Bs are contained in a single cell, rather than 2. In order for the population frequency of the B chromosome to increase via nondisjunction at a meiotic or mitotic division, the cell with 2 Bs must outcompete the cell without Bs or have a better chance of ending up in the offspring, by either preferential segregation or a B-induced increase in cell division within the germline (Burt and Trivers 2008).

Preferential segregation is the increased likelihood of a chromosome or pair of chromosomes segregating into a specific daughter cell. It can be particularly important during female meiosis, where a B chromosome risks ending up in a polar body rather than the egg nucleus. Preferential segregation can skew the typical 50:50 ratio of inheritance and increase the transmission of a B (or nondisjoined Bs) to the cells that are more likely to produce offspring. This is seen in the grasshopper *Calliptamus palaestinesis*, where most of the somatic tissue has a single B in each cell, but the germ-line cells possess either 2 B chromosomes (2B) or zero B chromosomes (NoB). The ratio between the two is 15:1, respectively (Nur 1963; Jones 1991; Burt and Trivers 2008). 2B cells are more prevalent because of the nondisjunction followed by preferential segregation towards the germ-line during a mitotic division. Because the process is not 100% effective, there are still some NoB germ-line cells. Preferential segregation also occurs during meiosis in the lily, *Lilium callosum*, and the grasshopper, *Myrmeleotettix maculatus* (Kayano 1957; Jones 1991; Burt and Trivers 2008).

Table 1 describes the expected outcomes for various combinations of nondisjunction and preferential segregation, occurring at various times during premeiotic mitosis, meiosis I, and meiosis II. For simplicity, this table assumes each mechanism works with perfect efficacy, though it is unlikely such mechanisms in nature are perfect. Knowing the distribution of B chromosomes within families or among individuals of a population, may allow us to distinguish potential drive mechanisms.

B chromosomes have been identified in both African and South American cichlids (Feldberg and Bertollo 1984; Feldberg et al. 2004; Poletto et al. 2010a, 2012a, Pires et al. 2015). *Astatotilapia latifasciata*, an African species from Lake Nawampasa in the Lake Victoria basin, carried either 1 or 2 metacentric B chromosomes in 38 of 96 individuals, both male and female. All of the kidney cells analyzed from B-carrying individuals contained a B chromosome, suggesting mitotic stability. In individuals with 2 B chromosomes, the Bs did not appear to pair during meiosis. Instead, they formed ring-like univalents, consistent with their isochromosomal structure (Poletto et al.

Table 1. Possible drive mechanisms utilizing nondisjunction and preferential segregation

Mitosis	Meiosis I	Meiosis II	Outcome
ND			0B, 1B
ND and PS			1B
	PS		1B
		ND	0B, 2B
		ND and PS	0B, 2B
ND	ND		0B, 2B
ND	ND and PS		0B, 2B
ND		ND	0B, 2B
ND		ND and PS	0B, 2B
	PS	ND	0B, 2B
	PS	ND and PS	2B

Possible combinations of nondisjunction (ND) and preferential segregation (PS), with the corresponding outcomes depicted. The possible drive mechanisms, combinations of ND and PS, are indicated in rows. Columns indicate the time when ND or PS occurs: during mitosis, meiosis I or meiosis II. The final column, "Outcome," indicates the expected type of offspring, that is, the number of B chromosomes they possess. Where 2 types of expected offspring are denoted, both types of offspring would be produced. Drive mechanisms considered here utilize combinations of nondisjunction and preferential segregation during at least one of these 3 cellular divisions. Each mechanism is assumed to be perfect (it produces the indicated result 100% of the time). All outcomes are based on individuals that have a single B chromosome and experience the type of drive listed. For further explanation of each drive mechanism, please see the Supplementary Files: Drive Mechanisms.

2010b, 2012b). B chromosomes were subsequently found in each of 12 cichlid species analyzed from Lake Victoria (Yoshida et al. 2011; Kuroiwa et al. 2014). Two morphologically distinct Bs that share repetitive sequences were found in Lake Victoria, where they were found in both sexes in most species (Fantinatti et al. 2011; Yoshida et al. 2011). In one species, *Lithochromis rubripinnis*, all of the B-carrying individuals were female, but not all females carried the B (Yoshida et al. 2011).

The sequence of a B chromosome from *A. latifasciata* has been further examined by whole genome sequencing (Valente et al. 2014). *Astatotilapia latifasciata* individuals with 2 Bs and without B chromosomes were sequenced and the resulting reads were aligned to a closely related reference genome. Thousands of regions across the genome showed significantly higher coverage in the individual with the B chromosome. These regions representing B chromosome sequences are referred to as "B chromosome blocks" or "B blocks." Copy number for several of these B blocks (estimated by qPCR) was tightly correlated with the B chromosome numbers determined by karyotype. These data support the idea that large portions of the B chromosome originate from A chromosome material, and many of these sequences are found in high copy numbers on the B. Whole-genome sequencing data also revealed that the reference genome assembly of *Pundamilia nyererei* (Brawand et al. 2014) from Lake Victoria contained a B chromosome highly similar to the B chromosome of *A. latifasciata* (Valente et al. 2014). Analysis of the *P. nyererei* transcriptome data (Brawand et al. 2014) also revealed B chromosome-specific transcription of several genes in multiple tissues (Valente et al. 2014).

B chromosomes were also identified in *Metriaclima lombardoi*, a cichlid species from Lake Malawi (Poletto et al. 2010a, 2012a). Divergence of the cichlid flocks of Lake Victoria and Lake Malawi occurred no longer than 8 million years ago (MYA) and all cichlids within Lake Malawi share a common ancestor no more than 1 MYA (Sturmbauer et al. 2001). Of 22 *M. lombardoi* individuals analyzed, 9 females carried a single large B chromosome, but no males were found with a B. The individuals examined were collected from the aquarium trade in Brazil and the Tropical Aquaculture Facility at the University of Maryland. Thus, it is unclear whether the frequency of Bs in this stock accurately reflects the frequency in wild populations.

The purpose of the present study was to quantify the number of B chromosomes among individuals of *B-carrying* Lake Malawi species. We identified B-carriers in *M. lombardoi* and 6 additional Lake Malawi species not previously known to possess B chromosomes. We found that all B-carrying individuals were female, and carried a single B chromosome, which allowed us to narrow the possible mechanisms of B chromosome drive for these species. We also found surprising variation in the copy number of individual B blocks, suggesting a high rate of structural mutation.

Materials and Methods

Animals, Chromosome Preparation, and Genotyping

All procedures involving live animals were approved by and conducted in accordance with the University of Maryland IACUC under Protocol #R-10–73. The 18 male and 18 female *M. lombardoi* used for cytogenetic analysis were obtained from stocks maintained at the Tropical Aquaculture Facility at the University of Maryland and the aquarium trade in Brazil. All stocks were originally sourced from Lake Malawi, Africa. Individuals were euthanized using tricaine methanesulfonate (MS-222) and inspected for testes or ovaries to confirm sex. Mitotic chromosome preparations were obtained from kidney tissue according to (Bertollo et al. 1978), with modifications (Poletto et al. 2010a). DNA was extracted from kidney tissue for these karyotyped individuals using standard phenol chloroform methods. Purified genomic DNA was quantified on a BioTek FLx800 using Pico-green and normalized to a concentration of 0.5 ng/μL. The resulting DNA samples were used in PCR and qPCR analysis.

Male and female *M. lombardoi, Metriaclima zebra* "Boadzulu," *Metriaclima greshakei, Metriaclima mbenji, M. zebra* "Nkhata Bay," *Labeotropheus trewavasae*, and *Melanochromis auratus* fin clips were collected from the wild in 2005, 2008, 2012, and 2014. DNA was extracted from fin tissue using standard phenol chloroform methods. Purified genomic DNA was quantified on a BioTek FLx800 using Pico-green and normalized to a concentration of 0.5 ng/μL. These samples were used for sequencing, PCR and qPCR analysis only, not cytogenetic analysis.

Sequencing

DNA libraries were prepared from the pooled DNA of 20 male or 20 female *M. zebra* "Boadzulu" individuals. The TruSeq DNA sample preparation kit ver.2 rev.C (Illumina) was used for library construction. Libraries were sheared to an average size of 500 bp and paired-end reads (100 bp in length) were obtained using Illumina's HiSeq 1500 platform. Raw sequencing reads were evaluated with FastQC (Babraham Bioinformatics) to ensure that libraries were of good quality. Reads were aligned to the unmasked *M. zebra* reference genome ("M_zebra_v0" available at www.bouillabase.org, Brawand et al. 2014) using Bowtie2 version 2.02 with the "--very-sensitive" preset of parameters (Langmead and Salzberg 2012).

Read coverage was initially compared between males and females in a genome browser and blocks of sequence with 10-fold or higher coverage difference in females versus males were found. These blocks, similar in pattern (length of blocks, amount of increase in coverage and variability in coverage across block) to those found in *A. latifasciata* (Valente et al. 2014), are referred to as B chromosome blocks or B blocks. B chromosome blocks were then defined by determining significant coverage differences between male and female pools using the following protocol. The samtools "mpileup" command (version 0.1.18) was used with parameters "-A -q 10" to include anomalous read pairs (common in B blocks) and to filter alignments with mapping quality (mapQ) less than 10 (Li et al. 2009). The mpileup output was streamed into the VarScan "copy-number" command (version 2.3.6) using the "--mpileup 1 --min-segment-size 100 --max-segment-size 50000" parameters as well as an appropriate "--data-ratio" parameter to account for differences in library coverage in the male to female comparison. The VarScan "copycaller" command was then used with the following parameters "--min-coverage 10 --amp-threshold 0.2 --del-threshold 0.2" (Koboldt et al. 2012).

Primer Design and PCR

Primers were designed using *Metriaclima zebra* "Boadzulu" sequence. SNPs identified within the B blocks were incorporated into the primers so that the primers would distinguish between homologous A and B sequence and amplification would be B-specific. SNPs were incorporated such that 1–3 B-specific SNPs were present in at least one primer, forward or reverse, for each pair [See Supplementary Information file on Primers]. Primer3, version 0.4.0, was used to calculate the expected melting temperature and evaluate the primer

sequences (Untergrasser et al. 2012). PCR reactions contained 5 µL of Life Technologies' Dream Taq, 0.5 µL (10 µM/L) forward primer, 0.5 µL (10 µM/L) reverse primer, 3 µL water and 1 µL (0.5 ng/µL) DNA. PCR products were separated on 2% agarose gels.

Quantitative Real-Time PCR

Real-time amplifications were recorded on a Roche LightCycler LC480 thermocycler. Quantitative real-time PCR (qPCR) reactions contained 10 µL of Life Technologies' Maxima SYBR Green/ROX, 1 µL (10 µM/L) of forward primer, 1 µL (10 µM/L) of reverse primer, 2 µL water and 6 µL (0.5 ng/µL) DNA. Each sample was amplified with 3 technical replicates, the average of which was used for all future calculations. Starting template quantities (T) were calculated using the following equation:

$$T = \frac{1}{E^{C_t}} \quad (1)$$

where E is the PCR efficiency of the primer set used and C_t is the critical cycle number from the qPCR reaction. Relative copy number was calculated using a control primer set. The control primer set amplifies the single copy cichlid SWS1 (UV) opsin locus which is present in the A genome, but not the B chromosome. Relative copy number was calculated by using the ratio of starting template of each B block primer pair over the control primer pair. Hierarchical cluster analysis of individuals by sequence copy number was performed using IBM SPSS Statistics 23 software.

Data Availability

We have deposited the primary data underlying these analyses as follows:

-DNA sequences: Illumina sequencing files are available from the NCBI BioProject PRJNA266314
-Primer sequence uploaded as online Supplementary Information.

Results

Identification of High Coverage Blocks

Pools of 20 wild-caught male and female *Metriaclima zebra* "Boadzulu" were sequenced and the resulting reads were aligned to the *M. zebra* reference genome. An example of the coverage differences between male and female pools is shown in Figure 1. Males and females show similar coverage across most of the genome, roughly 25.5×, but as shown in Figure 1, there are blocks of sequence with much higher coverage in females versus males. There are thousands of these short blocks of high sequence coverage, distributed across all linkage groups. The blocks were found exclusively in females. The female-limited presence of these blocks is consistent with the karyotypic detection of Bs in females, but not males, of *M. lombardoi* (Poletto et al. 2010a). These blocks are similar in pattern (i.e., length, coverage increase, variability of coverage across block) to those found in *A. latifasciata* (Valente et al. 2014), but their identities and locations are different (data not shown). There is almost no overlap among the sequence blocks found in the 2 species. We hypothesize that these blocks represent repetitive B chromosome sequence, and we test this idea below.

High Coverage Sequence Blocks Correspond to the Presence of B Chromosomes

High coverage block sequences from *M. zebra* "Boadzulu," located in 11 separate scaffolds of the *M. zebra* A genome reference assembly, were used to design a total of 21 primers sets for PCR amplification (See S1_File on Primers). Sequence corresponding to the B was distinguishable from A sequence by high frequency SNPs found within the high coverage blocks of female sequences but not found in the sequences from the male pool. These SNPs were incorporated into the forward primer, the reverse primer, or both (See S1_File on Primers). While the sequences appear to be continuous on the A chromosomes, the homologous sequences on the B may have undergone structural rearrangement, preventing efficient amplification with some primer sets. Of the 21 primer sets designed, 7 amplified the expected fragments, 6 amplified products too large to use for qPCR, 4 amplified sequence not specific to the B chromosome, 3 amplified a complex set of fragments, and 1 failed to amplify any sequence at all. Five primer sets amplifying the fragments of the expected size were selected for further analyses.

We next tested the hypothesis that the high-coverage blocks correspond to B chromosome sequence. We used the 5 primer sets to amplify DNA from *M. lombardoi* individuals that had been karyotyped and found to either carry the B (N = 8) or not to carry a B (N = 5). The karyotypes of a female with a B chromosome, a female without a B chromosome and a male without a B chromosome are shown in Figure 2. A subsample of the PCR data is shown in Figure 3. Amplification of all 5 primer sets was observed in each karyotyped individual with a B. No amplification was observed in any of the karyotyped individuals without a B. These data demonstrate that the primer sets are B-specific. The high coverage sequence blocks are hereinafter referred to as B blocks.

Prevalence of B Chromosomes in Lake Malawi Cichlids

Because these primer sets are B-specific, we can use them to assay the presence/absence of B chromosomes in additional individuals

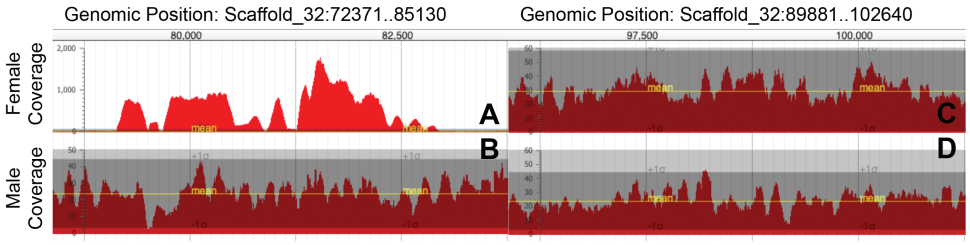

Figure 1. Read coverage of female and male *Metriaclima zebra* "Boadzulu" at a B block. Read coverage at 2 locations on scaffold_32 for female (**A** and **C**) and male (**B** and **D**) pooled samples of 20 *Metriaclima zebra* "Boadzulu" individuals. The location in plots **A** and **B** (72 371–85 130 bp on Scaffold 32) includes a B block, visible in plot **A**. Plots **C** and **D** represent sequence without a B block. Reads were aligned to the *M. zebra* reference genome. *Note:* the *y* axis differs between plots.

collected from the wild whose karyotypes are unknown. B chromosomes were identified in 6 additional species: *Metriaclima zebra* "Boadzulu," *M. zebra 'Nkhata Bay'*, *M. greshakei*, *M. mbenji*, *L. trewavasae* and *M. auratus* (Table 2). Together with the previously published identification of B chromosomes in *M. lombardoi*, B chromosomes now have been found in a total of 7 species of Lake Malawi cichlid. In all 7 of these species, B chromosomes have been found only in females.

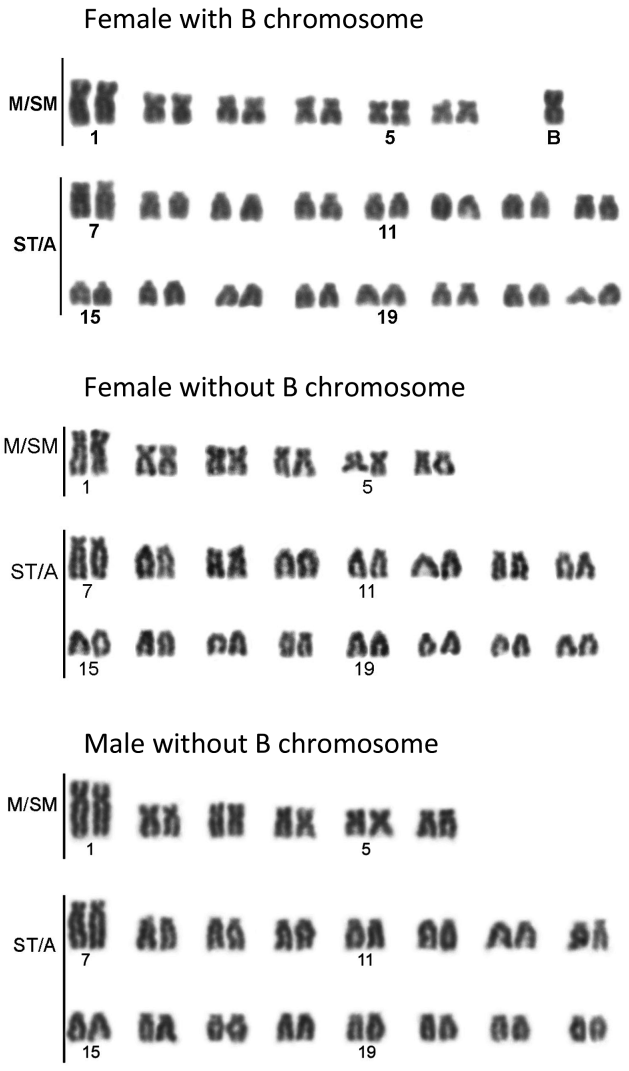

Figure 2. Karyotypes. Giemsa-stained karyograms of an *Metriaclima lombardoi* female B-carrier, an *M. lombardoi* female without a B and an *M. lombardoi* male without a B.

Copy Number Variation of B Chromosome Sequences

Next we wanted to know if all individuals carried a single B, as observed karyotypically in *M. lombardoi*. We reasoned that if individuals had 2 B chromosomes they should have twice as much of the B specific sequences as individuals with a single B chromosome. We performed quantitative PCR on the DNAs from 7 of the 8 karyotyped individuals shown to have the B, as well as B-carrying individuals identified by PCR from the population samples. We used qPCR to quantify the copy number of each of the 5 B block repeats studied above in each individual (see Methods section for details) (See S2_File on qPCR Standard Error). If individuals possessed 2 or 3 B chromosomes, we expected that the copy number of each B block repeat would roughly double or triple, respectively, compared to individuals known to carry a single B. Individuals with the same number of Bs should consistently cluster in pairwise comparisons of copy number.

Figure 4 shows the variation in copy number from each B block amplified, organized by population and then left to right by increasing average copy number. Table 3 lists the mean copy number and standard deviation for each species. Each individual used in Figure 4 and Table 3 has a B chromosome as determined through PCR. Some B chromosome repeats (corresponding to the sequence from B blocks on scaffold_13, scaffold_58, and scaffold_14) tend to show little copy number variation (CNV), while other regions of the B chromosome (corresponding to the sequence from B blocks on scaffold_23 and scaffold_32) show much higher CNV between individuals. Sequences with a higher average copy number show higher absolute variation in copy number than sequences with lower average copy number. Within species, however, the range of copy number rarely exceeds 2-fold and appears to be a single cluster, consistent with the idea that all individuals carry a single B chromosome. A single individual (sample ID: 2005–0995) from the *M. zebra* "Boadzulu" population appears to be an outlier for each B block sequence (most easily observed in Figure 4 for repeat sequences 14 and 23). This individual does not possess a copy of sequence 13 (data point not shown in Figure 4). The copy number of each sequence in this individual is not only smaller than any other *M. zebra* "Boadzulu" individual, but it is smaller than all of the karyotyped *M. lombardoi* individuals, which are known to carry a single B chromosome. This suggests this individual (2005–0995) possesses only a fragment of the B chromosome.

Figure 5 shows pairwise comparisons between each pair of B block repeats, organized in a half-matrix fashion. Figure 5 includes

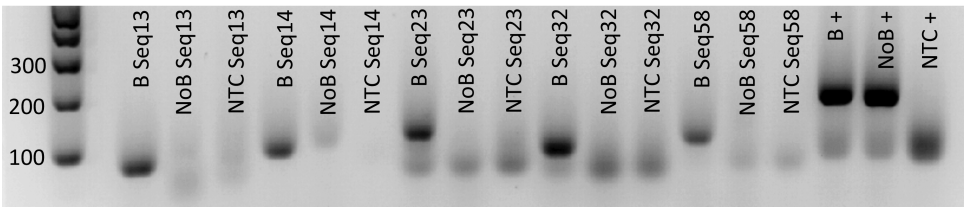

Figure 3. B-specific amplification. Agarose gel (2%) of PCR product resulting from amplification with either control or B-specific primers. The control primer (+) set amplifies the single copy cichlid SWS1 (UV) opsin locus which is present in the A genome, not the B chromosome. Amplification with the control primer is a positive control to indicate amplifiable DNA. DNA from 2 individuals was used, 1 female known cytogenetically to carry a B chromosome (B) and 1 female known cytogenetically to not carry a B (NoB). Amplification of a nontemplate control (NTC), containing no DNA, was also used for each primer set as a negative control.

data from individuals genotyped as having the B chromosome through PCR as well as the 7 *M. lombardoi* individuals shown to have the B chromosome through cytogenetic analysis (as indicated in the legend). Several patterns are apparent. First, individuals of the same species tend to cluster. Second, there appear to be structural (copy number) differences among the B chromosomes of different species. Third, within species, there is no apparent correlation of the copy numbers for different sequence blocks. There are a few individuals that appear to be outliers with respect to the cluster of individuals for that species. However, these individuals are not consistent outliers for each of the B block repeat classes. Thus, there is CNV of individual B block repeats among B chromosomes, but there is no evidence for correlation of CNV across loci, as would be expected if there was variation in the number of B chromosomes among individuals.

Table 2. Individuals genotyped for B chromosome

Population	Females with B/ total females	Males with B/ total males
Metriaclima lombardoi	10/93	0/43
Metriaclima zebra "Boadzulu"	21/49	0/30
Metriaclima greshakei	3/26	0/47
Metriaclima mbenji	1/27	0/33
Labeotropheus trewavasae	3/36	0/101
Melanochromis auratus	2/12	0/12
M. zebra "Nkhata Bay"	3/80	0/51

The number of individuals, from each population, that were shown to carry a B chromosome via amplification with B-specific primers. Individuals were initially genotyped using primers for sequence 32 (Seq32), which produce the strongest amplification. Positive amplification was then confirmed by amplification for the other 4 B-specific sequences. Of the 43 B-carrying individuals genotyped, 1 individual (sample ID: 2005–0995) amplified with 4 of the 5 primer sets. The other 42 amplified with all 5 primer sets used.

To further test for correlations in repeat number among individuals we performed a hierarchical cluster analysis of the B block copy numbers. Three groups emerged (Figure 6). One cluster contains all of the samples of *L. trewavasae* and *M. mbenji*. This cluster appears to reflect species-specific differences in the copy number of Seq32 (Figure 5). The 2 remaining groups each contain individuals that have been karyotyped and shown to have a single B. We conclude that all individuals in each group have a single B chromosome.

Intergenerational Variation in Copy Number

We then examined variation in copy number of B block sequences among siblings of karyotyped, lab-reared *M. lombardoi*. Since the offspring inherit the B exclusively from their mother, any sibling variation in sequence of the B must have arisen in a single generation. Quantitative PCR on the DNA from these individuals allowed us to quantify sequence CNV among siblings (Table 4). Individuals 58 and 59 are from family A002, and individuals 77–81 are from family A024. For most sequences, the copy number of each B block is consistent across individuals of a family, but there was considerable

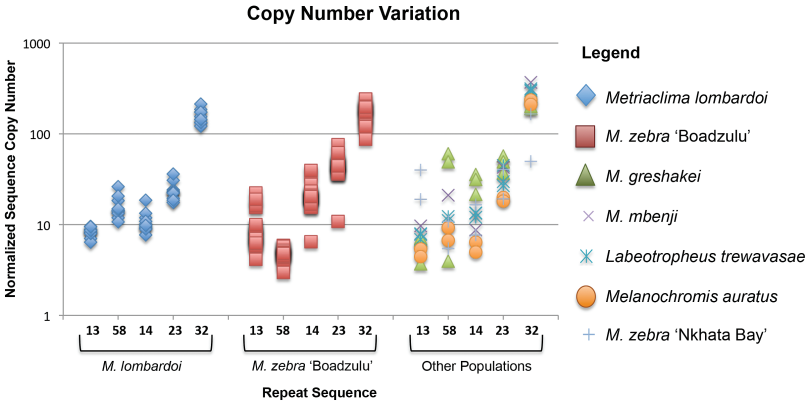

Figure 4. Copy number variation. Copy number of each B block repeat is shown in 3 groups; *Metriaclima lombardoi, Metriaclima zebra* "Boadzulu" and all other populations. The *y* axis depicts copy number and the *x* axis depicts the specific B block repeat analyzed. B block repeats are ordered left to right (for each of the 3 groups) by increasing average copy number. *Note:* the copy number of sequence 13 is not shown for 1 *M. zebra* "Boadzulu" individual (sample ID: 2005–0995), because the individual did not carry a copy of the sequence.

Table 3. Copy number for each B-specific sequence

Population	*N*	Seq13	Seq14	Seq23	Seq32	Seq58
Metriaclima lombardoi	10	9.21 (2.08, 22.6%)	10.73 (2.58, 24.0%)	23.3 (4.69, 20.1%)	155.98 (36.5, 23.4%)	15.22 (4.18, 27.5%)
Metriaclima zebra Boadzulu	21	9.11 (5.13, 56.3%)	21.17 (6.88, 32.5%)	44.66 (12.21, 27.3%)	169.34 (40.74, 24.1%)	4.86 (0.73, 15.0%)
Metriaclima greshakei	3	5.85 (1.89, 32.3%)	29.42 (7.1, 24.1%)	47.49 (10.15, 21.4%)	214.82 (36.76, 17.1%)	37.73 (29.83, 79.1%)
Metriaclima mbenji	1	9.67 (—)	8.7 (—)	42.92 (—)	369.27 (—)	21.11 (—)
Labeotropheus trewavasae	3	7.53 (0.59, 7.8%)	12.9 (0.76, 5.9%)	33.67 (9.5, 28.2%)	303.18 (9.01, 3.0%)	10.46 (2.21, 21.1%)
Melanochromis auratus	2	4.9 (0.64, 13.1%)	5.65 (1.01, 17.9%)	18.98 (1.1, 5.8%)	222.55 (16.64, 7.5%)	7.93 (1.75, 22.1%)
M. zebra Nkhata Bay	3	29.36 (14.61, 49.8%)	12.06 (6.18, 51.2%)	29.73 (14.91, 50.2%)	107.64 (82.11, 76.3%)	8.15 (3.85, 47.2%)

The mean copy number followed by the standard deviation and the relative standard deviation (in parentheses) for each population and each B-specific primer set. The value *N* represents the number of individuals. There was a significant difference in CNV between sequences, $F(4, 25) = 6.99$, $P \leq 0.001$. Standard deviation is not listed for *Metriaclima mbenji* as only a single individual from this population was found to have a B.

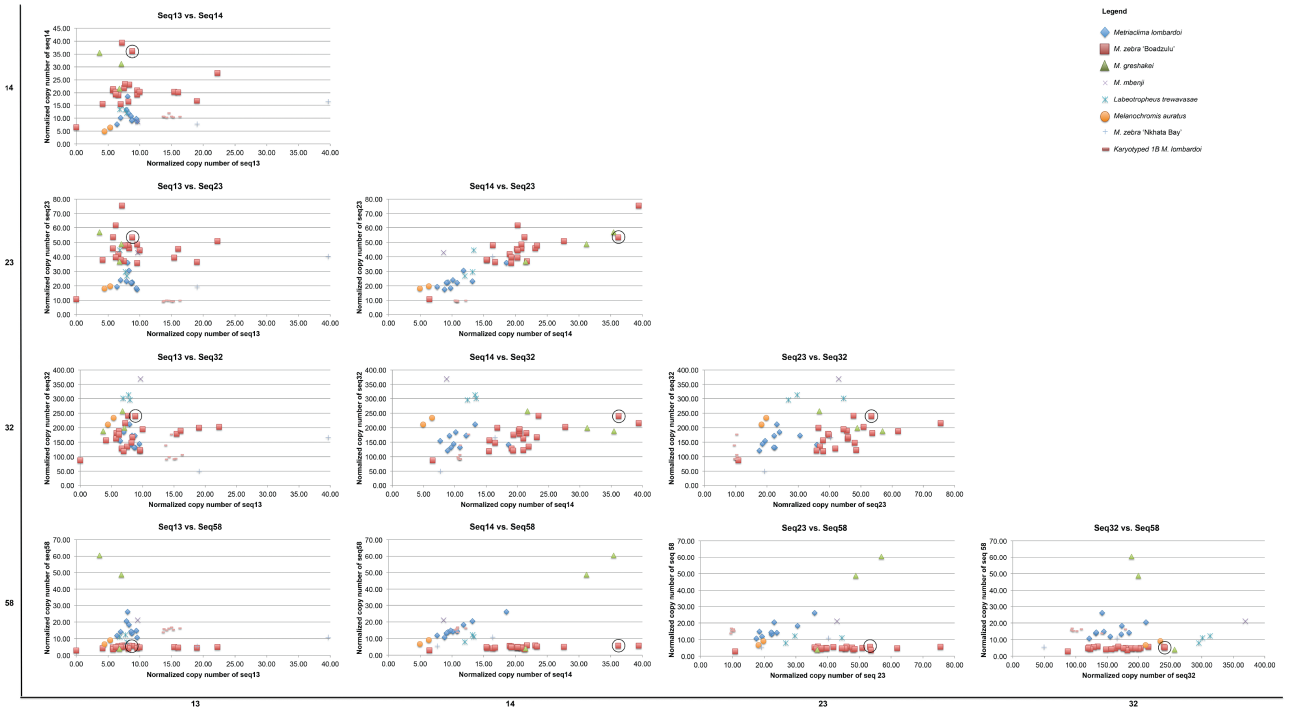

Figure 5. Two-way plots of copy number. Two-way plots arrayed in a half-matrix portray the copy number of a B block repeat on each axis. The B block repeats depicted are identified on each graph as well as along the half-matrix axes. The same individual has been circled in each graph to demonstrate that while an individual may be an outlier for one B block repeat, it is not an outlier for other B block repeats.

variation in copy number of Seq32. The variation at each locus is independent of the other loci. Since these individuals had been karyotyped and shown to possess only one B, these data reinforce the idea that minor variations in copy number reflect structural variation, not differences in the number of B chromosomes.

Discussion

Insights into the Mechanism of B Chromosome Drive

This study demonstrates that 1) Malawi cichlid B chromosomes occur at appreciable frequency in populations of at least 6 Malawi cichlid species, 2) that all B-carriers are female, and 3) that carriers have only a single B in their somatic tissue. This is in contrast to the B chromosomes found in Lake Victoria cichlids, which are found in both males and females, and in up to 3 copies per individual (Poletto et al. 2010a, 2010b; Yoshida et al. 2011). It is likely that these differences in B chromosome distribution reflect differences in the underlying mechanisms of B chromosome drive.

A drive mechanism may act by increasing the number of B chromosomes that individuals carry, or by increasing the frequency of carriers, or both. The mechanism of drive will determine the frequency of Bs in tissues, individuals and populations. The efficiency of the mechanism will also contribute to the frequency of B chromosomes, but for simplicity we consider only perfect mechanisms. Only 3 combinations of nondisjunction and preferential segregation will produce a population with only 1 B among carriers (Table 1). Only 2 of these combinations actually produce drive. If the Lake Malawi B chromosome uses nondisjunction or preferential segregation to drive, it would either use a combination of both during a pivotal mitotic division, or preferential segregation during meiosis I to avoid the polar body. The latter indicates a drive mechanism specific to females, meaning the B chromosome has a higher fitness in females. This may explain its female-specific presence.

While Bs have been shown to be mitotically unstable, or to employ nondisjunction and preferential segregation during mitotic divisions (Nur 1969; Kayano 1971), we are not aware of any such examples where mitotic nondisjunction is controlled in such a manner that it occurs precisely once during development. If the processes of nondisjunction and preferential segregation are recurring in mitosis, it would lead to higher variation of B chromosome number in the gonads. Furthermore, we can see no reason to expect that one mitotic division during development would be so distinct as to allow for this precise control. For this reason, we suggest that preferential segregation during meiosis I is the most likely scenario for the drive mechanism of the B chromosome in Lake Malawi cichlids.

Sex-Specific B Chromosome

We found a complete association between B chromosome presence and the female sex. This differs from the B chromosomes of most Lake Victoria cichlids, which are found in both males and females. Associations between B chromosomes and sex have long been recognized in many species (Camacho et al. 2011). Many B chromosomes drive in one sex, but not the other (Burt and Trivers 2008). While most B chromosome systems exhibit similar frequencies of B-carriers in both sexes, some have a higher frequency in one sex or the other, and others lack B-carriers in one sex altogether. Among the characid fishes, *Astyanax scabripinnis*, demonstrates a system in which B chromosomes are found more frequently in females. Additionally, several intersex individuals were identified, all of which possessed B chromosomes (Vicente et al. 1996; Neo et al. 2000). In one population of the characid fish *Moenkhausia sanctaefilomenae*,

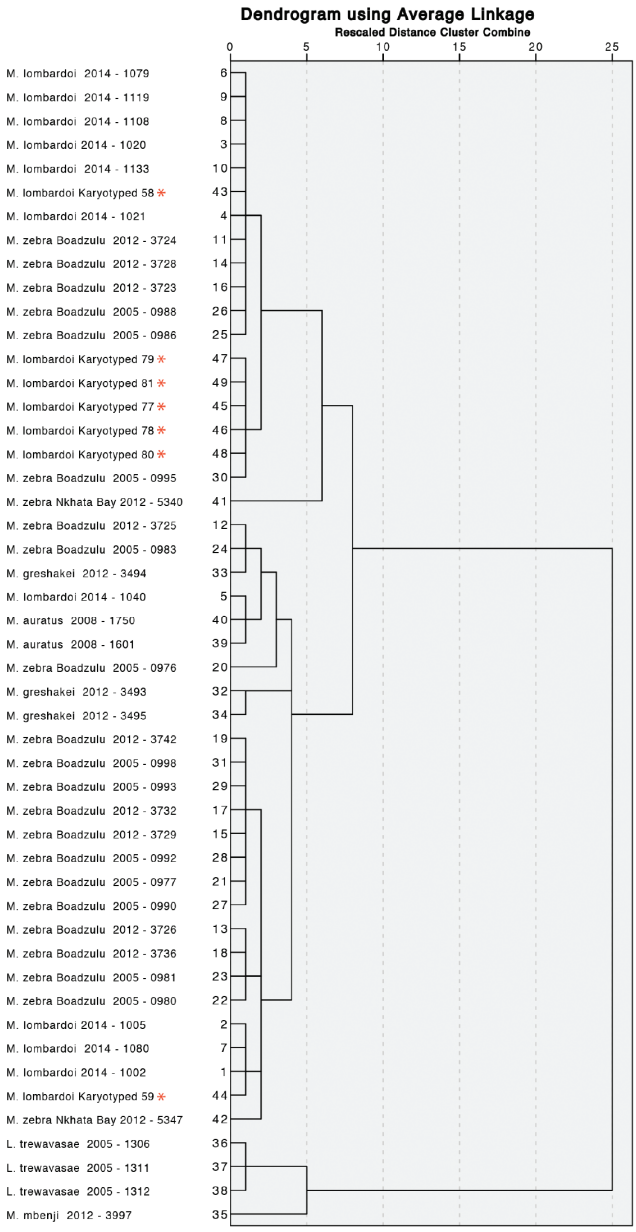

Figure 6. Cluster analysis dendrogram. Hierarchical clustering of samples by the CNVs of sequences Seq13, Seq14, Seq23, Seq32, and Seq58. Species and sample ID listed to the left. Karyotyped *Metriaclima lombardoi* individuals, known to have 1 B chromosome, are marked with a red asterisk (*).

B chromosomes are found in males, but not females (Camacho et al. 2011). B chromosomes in the fairy shrimp *Branchipus schaefferi*, are also found solely in males and the number of B chromosomes is associated with the sex ratio of the population (Beladjal et al. 2002; Burt and Trivers 2008). In some Orthoptera, drive has been observed to occur in females, not males. Yet an interesting mechanism has evolved in male B-carriers. The B chromosome associates and segregates with the X chromosome during male meiosis, ensuring that the B chromosome is more frequently transmitted to females where it has the benefit of drive (Burt and Trivers 2008). In each of these examples, it is believed that the difference in B-carrier sex ratio is a result of the drive mechanism of the B. Either the drive mechanistically results in one sex having more B chromosomes, or a mechanism

Table 4. Family member copy number

Sample ID	Family	Seq13	Seq14	Seq23	Seq32	Seq58
58	A002	13.5	10.6	9.1	139.3	13.9
59	A002	14.4	12.0	9.9	176.6	16.2
77	A024	15.1	10.4	9.3	93.8	15.4
78	A024	16.1	10.6	9.9	106.7	16.4
79	A024	14.8	10.7	9.3	92.2	16.8
80	A024	14.0	10.3	10.1	99.7	15.2
81	A024	13.6	10.6	9.8	91.1	15.8

This table lists the B block repeat copy number for each primer set for the individuals of 2 families. Columns 1 and 2 indicate Sample ID and family, respectively. All individuals included in this table have been karyotyped and possess a single B chromosome.

evolves to ensure the B is more frequently found in the sex in which it drives. The female-specific presence of B chromosomes in Lake Malawi cichlids may be a result of the drive mechanism employed, or a secondary mechanism (i.e., B chromosome sex determination) that benefits the B chromosome and its mechanism of drive. If drive is accomplished through preferential segregation in meiosis I, it is only effective in females, which produce polar bodies. If the B chromosome acquired a female sex determining sequence, thus ensuring its transmission only to females, it would increase the fitness of the B chromosome. There are certainly other hypotheses (i.e., male lethality) that would explain the female-specific B presence. An investigation of these hypotheses and a better understanding of what causes this correlation between B presence and the female sex is necessary to fully understand drive of this B chromosome and its corresponding evolutionary impacts.

Dynamics of Repeated Sequences on the B Chromosome

The variation we identified by qPCR shows that the copy number of a sequence on the B can change quickly. Not only can sequence copy number vary among individuals in the same population, but it also varies among siblings. It is unclear whether this variation is produced during meiosis or mitosis. It is interesting to speculate what duplication mechanism could bring about the copy number changes in a single generation if there is only a single B chromosome present in the cell. Poletto et al. (2010a) and Valente et al. (2014) suggest that the Lake Victoria B is an isochromosome (Poletto et al. 2010b; Valente et al. 2014). While univalent, the 2 chromosome arms can associate with one another and potentially undergo recombination in meiosis. Unequal crossovers between the chromosome arms of sister chromatids may contribute to CNV among siblings. Various other duplication and deletion mechanisms (long-range slippage, break- induced replication, single strand annealing) may play a role in CNV, but it is difficult to distinguish these mechanisms without knowing the size of the duplications and their arrangement on the B chromosome. Alternatively, a drive mechanism that utilizes nondisjunction and preferential segregation during mitosis would result in 2B tissue prior to meiosis. During prophase I of meiosis, these 2 B chromosomes could then function as homologous chromosomes and undergo unequal crossover. Any of these mechanisms could produce the variation in copy number between single generations seen in our data.

We propose the following model of B sequence evolution in Lake Malawi cichlids. Once established, the B chromosome experiences a continuous bombardment of sequences derived from the A genome,

through translocation, transposition, and reverse transcription. Once on the B, these sequences are frequently duplicated, leading to a highly repetitive DNA sequence. It is unclear whether related B block repeats remain tandemly arrayed, or become dispersed throughout the B by structural rearrangements such as inversions. Localization of B block repeats via FISH or long read sequencing may resolve this. Because the B chromosome is inherited clonally, mutations (single nucleotide polymorphisms, duplications, deletions, and structural rearrangements) are expected to accumulate rapidly.

If recombination is an important force of sequence evolution for the B, the location of the sequence along the chromosome arm may be important in determining the frequency and size of mutations. Sequences located in the middle of each arm may experience higher rates of recombination, and thus more rapid changes in copy number, than sequences near the centromere. We have shown that the amount of variation in copy number can vary between different B block repeats. It would be interesting to look for correlations between the chromosomal location of a sequence and the rate of change in copy number. Alternatively, copy number may be a factor in copy number variation. B blocks present in higher copy number may be more likely to undergo unequal crossing over or slipped strand mispairing.

If our model of B sequence evolution is correct, then it may be difficult to determine whether a B chromosome arose from a particular chromosome in the A set. The presence of a few homologous sequences found prominently on the B might be evidence of origin from a single or a few A chromosomes. Alternatively, this could merely indicate the success of those sequences in colonizing an already existing B, and not reflect the origin of the B itself. The abundance of a B block repeat also may not be related to the length of time it has been on the B. Other methods are required to confirm the homology of the B with portions of the A genome. Given the rapid turnover of sequences on the B, it may prove impossible to determine from which A chromosome the B was originally derived.

Conclusion

We developed a PCR assay that allowed for efficient genotyping of 640 individual Lake Malawi cichlids for the presence of B chromosomes. We describe the presence of B chromosomes in 6 additional species, present only in female individuals in each species. A combination of qPCR and cluster analysis was used to demonstrate that individuals carry no more than one B chromosome. Only 2 drive mechanisms, which utilize nondisjunction or preferential segregation, are consistent with the presence of a single B. We suggest that preferential segregation during meiosis I is the most likely mechanism. CNV of the B block repeats provides insight into the variability among these clonally inherited chromosomes. We have outlined a basic model of B sequence evolution, but given the rapid turnover of B chromosome sequences, it may prove difficult to characterize the earliest stages of divergence of the B chromosome from the A chromosome set.

Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding

This work was supported by the National Science Foundation (DEB-1143920) along with the American Genetic Association's 2015 Special Event Award.

Acknowledgments

We thank Dr Karen Carleton for her insight in qPCR design and analysis, and William Gammerdinger, Sri Pratima Nandamuri, and Daniel Escobar Camacho for their continued input in both hypotheses testing and troubleshooting. All procedures involving live animals were approved by and in accordance with University of Maryland IACUC under Protocol #R-10–73.

References

Babraham Bioinformatics FastQC [Internet]. Cambridge (UK): Babraham Institute. Available from: http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/

Banaei-Moghaddam AM, Schubert V, Kumke K, WeiB O, Klemme S, Nagaki K, Macas J, Gonzalez-Sanchez M, Heredia V, Gomez-Revilla D, *et al*. 2012. Nondisjunction in favor of a chromosome: the mechanism of rye B chromosome drive during pollen mitosis. *Plant Cell*. 24:4124–4134.

Beladjal L, Vandekerckhove TTM, Muyssen B, Heyrman J, de Caesemaeker J, Mertens J. 2002. B-chromosomes and male-biased sex ratio with paternal inheritance in the fairy shrimp *Branchipus schaefferi* (Crustacea, Anostraca). *Heredity*. 88:356–360.

Bertollo LAC, Takahashi CS, Moreira-Filho O. 1978. Citotaxonomic consideration on *Hoplias lacerdae* (Pisces, Erythrinidae). *Braz J Genet*. 1:103–120.

Beukeboom LW. 1994. Bewildering Bs: an impression of the 1st B-chromosome conference. *Heredity*. 73:328–336.

Brawand D, Wagner CE, Li YI, Manlinsky M, Keller I, Fan S, Simakov O, Ng AY, Lim ZW, Bezault E, *et al*. 2014. The genomic substrate for adaptive radiation in African cichlid fish. *Nature*. 513:375–381.

Burt A, Trivers R. 2008. *Genes in conflict: the biology of selfish genetic elements*. 1st ed. Cambridge: Belknap Press.

Camacho JP, Sharbel TF, Beukeboom LW. 2000. B-chromosome evolution. *Philos Trans R Soc B Biol Sci*. 355:163–178.

Camacho JP, Schmid M, Cabrero J. 2011. B chromosomes and sex in animals. *Sex Dev*. 5:155–166.

Fantinatti BE, Mazzuchelli J, Valente GT, Cabral-de-Mello DC, Martins C. 2011. Genomic content and new insights on the origin of the B chromosome of the cichlid fish *Astatotilapia latifasciata*. *Genetica*. 139:1273–1282.

Feldberg E, Bertollo LAC. 1984. Discordance in chromosome number among somatic and gonadal tissue cells of *Gymnogeophagus balzanii* (Pisces: Cichlidae). *Rev Brasil Genet*. 7:639–645.

Feldberg E, Porto JI, Alves-Brinn MN, Mendonça MN, Benzaquem DC. 2004. B chromosomes in Amazonian cichlid species. *Cytogenet Genome Res*. 106:195–198.

Houben A, Banaei-Moghaddam AM, Klemme S, Timmis JN. 2013. Evolution and biology of supernumerary B chromosomes. *Cell Mol Life Sci*. 71:467–478.

IBM Corp. 2013. *IBM SPSS Statistics for Macintosh*, Version 22.0. Armonk (NY): IBM Corp.

Jones RN, Rees H. 1982. *B chromosomes*. 1st ed. New York: Academic Press.

Jones RN. 1991. B-chromosome drive. *Am Nat*. 137:430–442.

Jones RN, Gonzalez-Sanchez M, Gonzalez-Garcia M, Vega JM, Puertas MJ. 2008. Chromosomes with a life of their own. *Cytogenet Genome Res*. 120:265–280.

Kayano H. 1957. Cytogenetic studies in *Lilium callosum*. III. Preferential segregation of a supernumerary chromosome in EMC's. *Proc Japan Acad*. 33:553–558.

Kayano H. 1971. Accumulation of B chromosomes in the germ line of *Locusta migratoria*. *Heredity*. 27:119–123.

Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 22:568–576.

Kuroiwa A, Terai Y, Kobayashi N, Yoshida K, Suzuki M, Nakanishi A, Matsuda Y, Watanabe M, Okada N. 2014. Construction of chromosome markers from the Lake Victoria cichlid *Paralabidochromis chilotes* and their application to comparative mapping. *Cytogenet Genome Res*. 142:112–120.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 9:357–359.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Burbin R; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. 25:2078–2079.

Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Scholz U, Gundlach H, Wicker T, Simkova H, *et al*. 2012. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proc Natl Acad Sci USA*. 109:13343–13346.

Neo DM, Moreira-Filho O, Camacho JMP. 2000. Altitudinal variation for B chromosome frequency in the characid fish *Astyanax scabripinnis*. *Heredity*. 85:136–141.

Nur U. 1963. A mitotically unstable supernumerary chromosome with an accumulation mechanism in a grasshopper. *Chromosoma*. 14:407–422.

Nur U. 1969. Mitotic instability leading to an accumulation of B-chromosomes in grasshoppers. *Chromosoma*. 27:1–19.

Pires LB, Sampaio TR, Dias AL. 2015. Mitotic and meiotic behavior of B chromosomes in *Crenicichla lepidota*: new report in the family Cichlidae. *J Hered*. 106:289–295.

Poletto AB, Ferreira IA, Cabral-de-Mello DC, Nakajima RT, Mazzuchelli J, Ribeiro HB, Venere PC, Nirchio M, Kocher TD, Martins C, *et al*. 2010a. Chromosome differentiation patterns during cichlid fish evolution. *BMC Genet*. 11:50.

Poletto AB, Ferreira IA, Martins C. 2010b. The B chromosomes of the African cichlid fish *Haplochromis obliquidens* harbour 18S rRNA gene copies. *BMC Genet*. 11:1.

Poletto AB, Ferreira IA, Cabral-de-Mello DC, Nakajima RT, Mazzuchelli J, Ribeiro HB, Venere PC, Nirchio M, Kocher TD, Martins C. 2012a. Chromosome differentiation patterns during cichlid fish evolution: correction. *BMC Genet*. 13:2.

Poletto AB, Ferreira IA, Martins C. 2012b. The B chromosome of the cichlid fish *Haplochromis obliquidens* harbors 18S rRNA genes: correction. *BMC Genet*. 13:3.

Randolph LF. 1941. Genetic characteristics of the B chromosome in maize. *Genetics*. 26:608–631.

Silva DM, Pansonato-Alves JC, Utsunomia R, Araya-Jaime C, Ruiz-Ruano FJ, Daniel SN, Hashimoto DT, Oliveira C, Camacho JPM, Porto-Foresti F, *et al*. 2014. Delimiting the origin of a B chromosome by FISH mapping, chromosome painting and DNA sequence analysis in *Astyanax paranae* (Teleostei, Characiformes). *PLoS One*. 9:e94896

Sturmbauer C, Baric S, Salzburger W, Rüber L, Verheyen E. 2001. Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. *Mol Biol Evol*. 18:144–154.

Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 40:e115.

Valente GT, Conte MA, Fantinatti BE, Cabral-de-Mello DC, Carvalho RF, Vicari MR, Kocher TD, Martins C. 2014. Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Mol Biol Evol*. 31:2061–2072.

Vicente VE, Moreira-Filho O, Camacho JP. 1996. Sex-ratio distortion associated with the presence of a B chromosome in *Astyanax scabripinnis* (Teleostei, Characidae). *Cytogenet Cell Genet*. 74:70–75.

Wilson EB. 1907. The supernumerary chromosomes of Hemiptera. *Science*. 26:870–871.

Yoshida K, Terai Y, Mizoiri S, Aibara M, Nishihara H, Watanabe M, Kuroiwa A, Hirai H, Hirai Y, Matsuda Y, *et al*. 2011. B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. *PLoS Genet*. 7:e1002203.

Zhou Q, Zhu HM, Huang QF, Zhao L, Zhang GJ, Roy SW, Vicoso B, Xuan ZL, Ruan J, Zhang Y, *et al*. 2012. Deciphering neo-sex and B chromosome evolution by the draft genome of *Drosophila albomicans*. *BMC Genomics*. 13:109.