

The definitive version is available at:

[http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1469-8137](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1469-8137)

DOI: <https://doi.org/10.1111/nph.14723>

## Letters

# B-chrom: a database on B-chromosomes of plants, animals and fungi

### Introduction

We present here B-chrom ([www.bchrom.csic.es](http://www.bchrom.csic.es)), an online database with comprehensive information on B chromosomes (Bs) for plants, animals and fungi. Data have been extracted from 3041 sources published between 1907 and July 2016. There are 5760 entries corresponding to 2828 species. Besides presence, number or range of B chromosomes, the database provides information on chromosome number, ploidy level and genome size when available. After an extensive publication search strategy and data mining, the content of the database has been analysed statistically and significant positive correlations, though faint, have been found between the average number of Bs, chromosome number and ploidy level. Some of the species with the highest number of Bs in plants reproduce asexually, which may be related to an accumulation of these *selfish* genomic elements. Despite the increased interest in B chromosome research in recent years (20% of the data in B-chrom was published in the last decade) there is still limited knowledge with respect to global biodiversity. The database is the first step to systematize information on Bs, and we expect that it will be used by scientists interested in cytogenetics for data-mining or comparative studies involving B chromosomes.

B chromosomes, also known as supernumerary or accessory chromosomes, are additional and non-essential constituents of karyotypes. Their most significant traits are that: (1) they may be present in some, but not all, individuals of a population, or in certain cells of the same individual but not in all; and (2) they fail to recombine with A chromosomes during meiosis (Jones *et al.*, 2007; Houben *et al.*, 2013). B chromosomes are usually, but not always, smaller than A chromosomes and they are a sometimes overlooked source of intraspecific genome size variation (e.g. natural rye plants can have from zero to four Bs, which can clearly affect the constancy of its *C*-value; Jones, 1976). They are usually considered useless, but both favourable (e.g. antibiotic resistance in the fungus *Nectria haematococca*, Coleman *et al.*, 2009; selective environmental advantages in *Allium schoenoprasum*, Holmes & Bougourd, 1989) and adverse (e.g. lower vigour and impaired fertility in *Dactylis glomerata* hybrids, Williams & Barclay, 1968; parasite-like behaviour of Bs in *Eyprepocnemis plorans*, Camacho *et al.*, 2003) effects of Bs have been described.

The presence of Bs is not associated with any phenotype in most cases (Valente *et al.*, 2017), particularly when the number of Bs is

low. The maximum number of Bs tolerated is variable across species and it is counterbalanced by their potential negative consequences, particularly on fertility and vigour (Houben, 2017). In genera displaying apomixis, the presence of Bs is well documented, such as in *Boechera* where these may be involved in the genetic control of apomixis (Kantama *et al.*, 2007; Mandáková *et al.*, 2015). There are also interesting relationships between Bs and sex chromosomes: Sharbel *et al.* (1998) proposed a sex chromosome as the ancestor of the B chromosome in a frog species; conversely, in some cichlid fish species it was hypothesized that a portion of sex chromosomes is derived from Bs (Yoshida *et al.*, 2011).

B chromosomes were first observed in insects from the genus *Metapodius* (now *Acantocephala*) (Wilson, 1907). In plants, they were discovered in crops from the genus *Secale* (Gotoh, 1924). Since then, thousands of reports have steadily increased our knowledge of the distribution and features of Bs across life on Earth. B chromosomes have been classically considered nonfunctional. However, it is not until recently that active genes have been found in Bs. For example, ribosomal RNA (rRNA) genes may play a role in the evolution of Bs, as these have been detected on Bs of many plant and animal species. In *Plantago lagopus*, a new B chromosome arose from the extensive amplification of 5S rRNA genes (Dhar *et al.*, 2002). Also, transcription of rRNA genes was the first molecular evidence of gene activity in Bs, again both in plants and animals (Leach *et al.*, 2005; van Vugt *et al.*, 2005). Later, genes from other multigene families such as H1, H3 and H4 histones, as well as U2 snRNA, transposable elements and satellite DNA, have also been documented as components of Bs (Valente *et al.*, 2017).

There is no general ubiquitous mechanism for the evolutionary origin of Bs. Most likely, there are several possible origins; the most widely accepted is that they are derived from the A chromosome complement (Houben *et al.*, 2013). After unbalanced or asymmetric translocation, small centric fragments can also become Bs (Jones & Rees, 1982). Some evidence also suggests that Bs can act as diploidizing agents after a polyploidization process (Jones & Houben, 2003).

The study of Bs started with classical karyological methods and now even -omics approaches are being used for this research (Valente *et al.*, 2017). Thousands of reports on Bs have been produced during the last two centuries, yet an overview of its distribution across the tree of life is still missing, as well as a clear understanding of their role. As more and more organisms have been found to harbour accessory genetic materials over the last decades, the number of literature reviews on Bs has increased substantially, including overarching examinations (Borisov, 2014), revisions based on taxonomic groups (e.g. Jones *et al.* (2007) and Datta *et al.* (2016) for plants, Palestis *et al.* (2010) for orthopterans, Makunin *et al.* (2014) for mammals, or Galazka & Freitag (2014) and Stukenbrock & Croll (2014) for fungi), as well as reviews analysing their putative role in genetics and evolution (Camacho *et al.*, 2000;

Banaei-Moghaddam *et al.*, 2015). The phylogenetic diversity within and between taxa harbouring Bs suggests that this polyphyletic phenomenon is a rather common one, with their complex characteristics and dynamics still poorly understood from a systemic perspective (Valente *et al.*, 2017). There are also periodical international conferences devoted to Bs (i.e. *B-chromosome Conference*, the last one held in Gatersleben, Germany, in 2014).

Jones & Díez (2004) compiled a comprehensive database which included any report on the presence of Bs published between 1907 and 1994. However, the database has not been updated and many reports on Bs have been released since 1994. Currently, there is a wealth of information available but sometimes the access is difficult, as this can be published in a variety of local or national journals. Taking advantage of the current and powerful literature search engines, the main purpose of this work was constructing a new resource, which included extensive data on presence and numbers of Bs in plants, animals and fungi, and making it available online. The number of publications which include information on Bs is remarkable, and interest remains high (Fig. 1). To our knowledge, there are no online initiatives offering such information and embracing the three largest biological kingdoms. We expect that this comprehensive and updated catalogue of species presenting Bs will contribute to the understanding of these 'ultimate genome parasites' (Jones *et al.*, 2007) allowing the analysis of their distribution across the tree of life.

## Materials and Methods

### Information sources

In order to obtain the data, a search strategy was created to retrieve scientific documents which included reports on Bs. The online bibliographic databases used were Scopus (<https://www.scopus.com/>), Web of Science (WOS, <https://apps.webofknowledge.com/>), SciELO (<http://www.scielo.org/>), Directory of Open Access Journals (DOAJ, <https://doaj.org/>) and Google Scholar (<https://scholar.google.es/>). Searches were limited to a specific time period (1995–July 2016) since the information provided by the 'B chromosome database' (Jones & Díez, 2004) comprising years

1907–1994 would be incorporated in B-chrom. Database queries were conducted from January to July 2016. The search strategy was adjusted to the different interfaces of each bibliographic database. In particular, 'B chromosome' was searched in the fields *Title*, *Abstract* and *Keywords* both in Scopus and WOS. However, in Google Scholar, searches can only be made using two fields: *In the title of the article* or *Anywhere in the article*. We selected the former due to the amount of noise generated by the latter. With regard to SciELO and DOAJ, the search was performed in all fields given the considerably lower number of results obtained with the previous search strategy. Scopus was our choice for preparing the initial corpus of publications due to its wider coverage of documents as compared with WOS, which indexes fewer publications (Mongeon & Paul-Hus, 2016). When the relevance of certain documents was not clear, we looked for the presence of other related keywords such as 'supernumerary chromosome', 'accessory chromosome' or 'selfish chromosome'. Apart from the presence of these keywords, we also evaluated where they appeared (title, abstract, keywords, etc.) to assess their relevance. As searches were performed, results were downloaded in CSV format, except in the case of DOAJ, where this option was not available. Subsequently, we used Google Sheets for analysing the results of our searches. We discarded duplicates and not relevant documents and we obtained the complete text of all the documents we had access to. Papers without at least a summary in English were excluded. From this preliminary corpus of documents (2837 publications), we created a bibliographic database using ZOTERO (<https://www.zotero.org/>). The references from 'B chromosome database' (our starting point database) were manually added to that one. During this process, errors were corrected (spelling mistakes in author names, wrong publication years, etc.) and DOI identifiers were included (or URLs when DOIs were not available) to complete the bibliographic references.

In its final form, B-chrom includes data from 3041 references (mostly journal articles, but also books and meeting proceedings): 2410 coming from Jones & Díez (2004) and 631 obtained from the five bibliographic databases analysed (meaning that 22% of the newly retrieved documents had relevant information). Note that the initial work by Jones & Díez (2004) covered

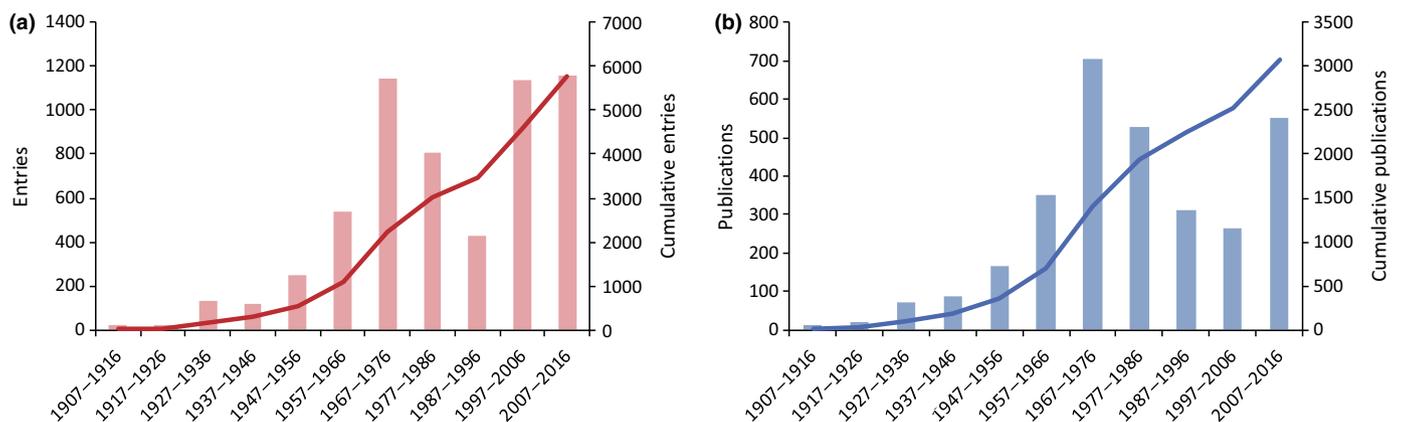


Fig. 1 Number of (a) entries and (b) publications reporting B chromosomes present in B-chrom over 11 successive 10-yr periods, between 1907 and 2016. The right axis indicates cumulative data both in (a) individual entries and (b) number of publications.

1906 to 1994, while our searches encompassed only from 1995 to July 2016.

### Data mining

The information was manually extracted from each source publication and when available (in most cases), the presence and number of Bs were visually checked in the figures. Data were introduced in a Google Sheet filling the following fields for each entry: (1) kingdom, (2) phylum or (sub)division, (3) popular name of the containing group, (4) class, (5) order, (6) family, (7) genus, (8) specific epithet, (9) complete species name, (10) ploidy level, (11) somatic chromosome number ( $2n$ ), (12) B-chromosomes (see later), (13) complete citation reference in APA format, and (14) DOI or URL where the source publication is available. For angiosperm plants, there was an additional category, (15) monocots or eudicots. Each entry and each publication had unique identification numbers. The information available from the work by Jones & Díez (2004) until 1994 and the IPCN compilations from years 1994–2006 (Goldblatt & Johnson, 1998, 2000, 2003, 2006, 2010) were imported to the spreadsheet and formatted according to the earlier mentioned structure.

Data on Bs are displayed as either: presence, with the letters 'Bs', and a specific number (e.g. 1, 2, 8) or as a range (e.g. 2–6), depending on the information provided by the source publication. When the number of Bs referred to the gametic cells it was indicated with the letter 'G' after the number of Bs. Also, when Bs were found either in male or in female individuals it was shown by the letters 'm' and 'f', respectively. When the number of Bs was found at the haploid level we used the letter 'n' (the case of some liverworts). Finally, the letters 'PRS' indicate Paternal Sex Ratio, types of Bs occurring in certain arthropods (Werren & Stouthamer, 2003).

### Database web environment and construction

Release 1.0 of B-chrom was launched in March 2017. Access to the information is made easy through a search box in which queries per genus or per species can be inserted. Additionally, data can be browsed by the largest groups present in the database (eudicots, monocots, gymnosperms, fungi, fish, insects and mammals) directly from the home page or from the 'Browse' tab. Data are returned in customizable tables in which users can select to display their desired options. Search results are also downloadable as CSV files from the institutional repository Digital CSIC. The default settings of a regular search include family, genus, species name, ploidy level, chromosome number, information on Bs (presence, number or range) and complete citation reference linked to its DOI or URL. Additionally, the tab 'Publications' offers the complete reference list of the publications retrieved for data mining, while the tab 'Links' provides directions to other databases with cytogenetic data, as well as other useful web resources. Finally, the tab 'Contact' is intended to foster communication with researchers interested in providing new data or in making any comments or corrections to the database.

The database structure was created in the MySQL server and hosted in [www.bchrom.csic.es](http://www.bchrom.csic.es). The initial Google Sheet in which

the data were compiled was imported to a CSV file. The website uses LARAVEL v.5.3 (<https://laravel.com/>; for PHP 5 developments) and BOOTSTRAP v.3.3.7 (with HTML, CSS and JS; <http://getbootstrap.com/>) open source frameworks.

### Statistical analyses

The Taxonomic Name Resolution Service v.4.0 (accessed 15 September 2016) (<http://tnrs.iplantcollaborative.org>) was used for correcting and standardizing plant names in order to avoid ambiguous, superfluous or incorrect names resulting in mismatched or unwitting duplication of records (Boyle *et al.*, 2013) followed an approach applied previously (García *et al.*, 2017). When ploidy levels and/or chromosome numbers were not indicated in the source publication, data have been extracted from the Chromosome Counts Database (CCDB) (accessed 15–25 November 2016) (<http://ccdb.tau.ac.il>) (Rice *et al.*, 2015). Genome size data have been obtained through the Plant DNA C-values database (accessed 15–25 November 2016) (<http://data.kew.org/cvalues>), the GSAD database (accessed 15–25 November 2016) (<http://www.asteraceagenomesize.com>) and the Animal Genome Size database (accessed 15–25 November 2016) (<http://www.genomesize.com>). If a species had a different number of Bs, chromosome number or ploidy level that differed between or within publications, we treated each difference as a separate entry. Statistical analyses were performed with RSTUDIO, v.0.98.1078, a user interface for R (<http://www.rstudio.com>). Duplicates were removed from the dataset before all analyses. Since datasets were not normally distributed, we performed the non-parametric Spearman rank correlation for three analyses: (1) number of Bs vs chromosome number ( $2n$ ); (2) number of Bs vs ploidy level; (3) number of Bs vs genome size (2C). These analyses have been performed at different taxonomic levels (see Supporting Information Tables S1–S4). Some assumptions were made: when there was a range of Bs for a given entry we have used the average value, and when the number of Bs was not specified, we have assumed it was one, for calculations.

## Results and Discussion

### Coverage

It is difficult to estimate in how many species Bs may be present because the representation in the dataset is highly biased for the reasons explained earlier. Besides, most estimates refer only to plants. Darlington & Wylie (1955) listed chromosome numbers of over 17 000 species of flowering plants of which 0.8% had Bs while Fedorov (1969) estimated 1.1% of plant species had Bs. More recently, Levin *et al.* (2005) reported Bs in 8% of monocots and 3% of eudicots (c. 4% of angiosperms). In the CCDB (Rice *et al.*, 2015), one of the most recent resources providing chromosome numbers for plants, data are available for 77 958 species. Considering that we have assembled information on Bs for 2087 plant species, we can estimate that 2.68% of species with known chromosome numbers have Bs. In animals it is more difficult to make this calculation since chromosome counts are more scattered

in this kingdom and there are no available lists as in plants, to our knowledge. In the same line, it is impossible to know if Bs are more frequent in animals, plants or fungi, since the high frequency of Bs in certain groups better reflects the intensity and technical ease with which each group has been studied (Camacho *et al.*, 2000). In fact, the effect of study intensity of certain groups on the presence of Bs can be huge, as pointed out by Palestis *et al.* (2004), who found that species less studied had 12-fold lower Bs frequency than the most studied in the database of mammalian karyotypes. Levin *et al.* (2005) also reported a significant correlation between the proportion of Bs across angiosperms and the study effort. More factors certainly bias the assessment of species with Bs; that is most chromosome counts are based on a single or few individuals, most species have not been assessed cytologically, and the phylogenetic relationships between species should also be taken into account to evaluate the frequency of Bs. Current assessments of biodiversity estimate that there may be 7 000 000 species of animals, 400 000 of plants and 1 500 000 of fungi (based on Chapman, 2009) therefore the correct figure of species with Bs is certainly much larger. Yet it is astonishing that some groups lack a single representation in the database, besides considerable cytogenetic knowledge, as is the case of birds (e.g. Tegelström & Rytman, 1981; Ellegren, 2010), with > 18 000 species estimated (Barrowclough *et al.*, 2016). Interestingly, most birds present small DNA contents and a narrow range of its variation, indicating that a tight control of genome size could explain the absence of Bs. Vujošević & Blagojević (2004) also related the absence of Bs in birds with their small genome sizes. Another group of animals in which Bs have not been detected to date is phylum Cnidaria (9000 estimated species), the group containing corals, sea anemones and jellyfish, although their chromosome counts are very scarce too. In plants, algae have no reports on Bs, despite the size of the group (*c.* 12 000 species described) and karyological knowledge (e.g. Austin, 1956; Kapraun, 1993; Lewis, 1996; Muravenko *et al.*, 2001). Ferns and

mosses (with 15 000 estimated species each) are poorly represented, with only seven and 10 species with data on Bs, respectively, despite being well-known karyologically, with reports on 45 pteridophyte families and 120 bryophyte families in the CCDB (Rice *et al.*, 2015). Regarding flowering plants, it is surprising that some well-known families with available chromosome counts do not present any reports on Bs, such as Apocynaceae (1500 species, with 805 counts in the CCDB), Berberidaceae (700 species, with 156 counts in the CCDB), Cactaceae (1700 species, with 698 counts in the CCDB), or Ericaceae (4250 species, with 700 counts in the CCDB).

Data have been extracted from 3041 papers published between 1907 and 2016. Fig. 1 shows the data and publication distribution divided into periods of 10 yr. The latter period (2007–2016) is the most productive in terms of individual B-chromosome assessments (1157 entries, or 20.10% of the whole dataset) and the second in which more papers on Bs were published (553, after the period 1967–1976 with 704 publications). These figures illustrate that the interest in cytogenetics remains high amongst the scientific community, as found for other recently updated or released cytogenetic databases like the Plant rDNA database (García *et al.*, 2014) or the Animal rDNA database (J. Sochorová, S. García, F. Gálvez, R. Symonova & A. Kovařík, unpublished), in which recent years also tend to be the most productive. Table 1 lists the journals which have published more documents containing information on Bs.

The 5760 entries available in the database correspond to 2828 eukaryotic species which have been reported to present Bs in their genomes, of which 73.56% (2087 species) were plants (53.20% monocots and 46.80% eudicots), 25.95% (736 species) animals and only 0.49% (14 species) fungi, excluding duplicates. With respect to the previous database on Bs (Jones & Díez, 2004), our database represents an increase of 65% in the number of entries and of 61.5% in the number of species (3484 entries and 1757 species in

**Table 1** The 20 journals (including publisher information) providing most articles to B-chrom

	Journals	Publisher	Articles
1	<i>Chromosoma</i>	Springer Nature	201
2	<i>Cytologia</i>	Japan Mendel Society	176
3	<i>Caryologia</i>	Taylor & Francis	142
4	<i>Heredity</i>	Springer Nature	138
5	<i>Hereditas</i>	John Wiley & Sons	129
6	<i>Genetica</i>	Springer Nature	99
7	<i>Maize Genetics Cooperation Newsletter</i>	Iowa State University	90
8	<i>American Journal of Botany</i>	Botanical Society of America	77
9	<i>Canadian Journal of Genetics and Cytology</i>	Canadian Science Publishing	61
10	<i>Genetics</i>	Genetics Society of America	60
11	<i>Cytogenetic and Genome Research</i>	Karger Medical and Scientific Publishers	58
12	<i>Chromosome Research</i>	Springer Nature	54
13	<i>Genome</i>	Canadian Science Publishing	42
14	<i>Plant Systematics and Evolution</i>	Springer Nature	38
15	<i>The Japanese Journal of Genetics</i>	Genetics Society of Japan	28
16	<i>Canadian Journal of Botany</i>	Canadian Science Publishing	27
17	<i>Current Science</i>	Current Science Association	27
18	<i>Journal of Heredity</i>	Oxford University Press	26
19	<i>TAG Theoretical and Applied Genetics</i>	Springer Nature	26
20	<i>Tsitologija</i>	Department of Biological Sciences of the Russian Academy of Sciences	26

the ‘B chromosome database’). In total, data are available for 311 families (119 of plants, 185 of animals and seven of fungi), 1095 genera (635 of plants, 450 of animals and 10 of fungi). In animals, the sample is mostly composed of insects (52.28%), mammals (17.83%) and ray-finned fish (Actinopterygii) (16.28%). This is not surprising given the relative abundance, potential interest or suitability for cytogenetic studies of these groups. Indeed, insects account for the most diverse and abundant group of animals, mammals have been subject to deep cytogenetic studies, particularly in some model species like mice, and ray-finned fish are the dominant class of vertebrates, accounting for most (> 95%) extant fish species. In plants, the representation is again biased by economic interest (crops) or relative abundance, with Poales (21.76%), Asterales (20.75%) and Asparagales (15.72%) the most represented orders, which is in accordance with previous findings (Houben *et al.*, 2013). As for fungi, data are only available for 14 species from several orders from the phylum Ascomycota, which is its largest phylum. Model organisms such as maize, rye and the grasshopper *Eyprepocnemis plorans* are the species which have been studied most times individually (with 309, 210 and 81 independent reports, respectively).

Organisms harbouring Bs in our database present ploidy levels ranging from 1 (some fungi) to 22, although the range is narrower in animals (2–6) than in plants (1–22). The most common ploidy level is the diploid, accounting for 66% of the database, followed by the tetraploid (11%) (Fig. 2). Chromosome numbers range from 4 to 720, with animals having again a narrower range (6–150). In 13.30% of entries only presence (Bs) is recorded, and in 25.51% only one B-chromosome is found (the modal value), the remaining ranging from 1 to 50. The average number of Bs is 2.4, being one to two Bs the most common reports (44.31% of the database entries), followed by those that state only ‘presence’ of Bs (17.98%).

The highest B chromosome numbers were detected in the plant species *Pachyphytum fittkaui* (Crassulaceae) with  $2n = 120 + 50B$ , followed by *Albuca bracteata* (synonym *Ornithogalum caudatum*)

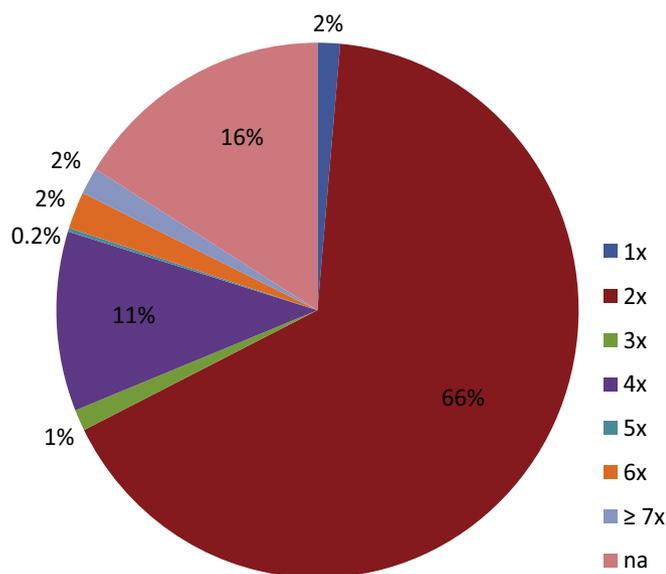


Fig. 2 Distribution of ploidy levels across the database. na, data not available.

(Asparagaceae) with  $2n = 18 + 36B$  and *Zea mays* (Poaceae) with  $2n = 20 + 34B$ . In the three species, vegetative reproduction is well known (for Crassulaceae, Guo *et al.*, 2015; for Asparagaceae, Byers *et al.*, 2014; for maize, Wolff, 1971). Also, in other taxa with a high number of Bs, asexual reproduction has been observed, as in *Fritillaria japonica* (Liliaceae) ( $2n = 22 + 26B$ ) or *Centaurea scabiosa* (Asteraceae) ( $2n = 22 + 20B$ ). Perhaps species that reproduce vegetatively can better tolerate the presence of Bs or genome size accumulation because of the absence of meiosis as a controlling mechanism of additional genomic load, although more data are needed to substantiate this hypothesis.

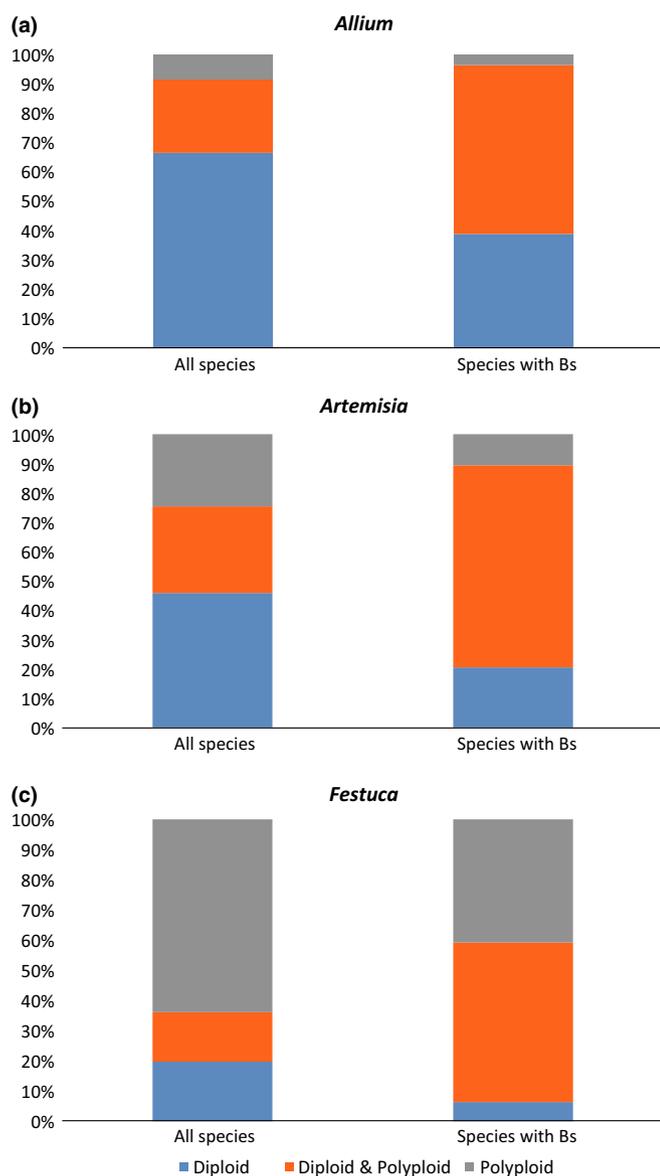
Among animals, the highest B chromosome number is found in the rodent *Apodemus peninsulae* (Muridae) ( $2n = 48 + 30B$ ), followed by the spider *Clubiona japonicola* (Clubionidae) ( $2n = 24 + 28B$ ). [Correction added after first publication 25 July 2017: the preceding sentence has been amended.] Although extreme reports of Bs are found more frequently in plants than in animals, there are more animal than plant species harbouring  $\geq 5$  Bs (23.1% vs 17.3%) or  $\geq 10$  Bs (6.1% vs 3.7%), bearing in mind that our knowledge on Bs in animals is more limited than in plants. Usually, Bs do not exceed the A chromosome complement, however, this happens in 28 species and it can reach up to 5.5-fold of the A chromosome complement as in the case of *Brachycome lineariloba* (Asteraceae) ( $2n = 4 + 22B$ ) (Smith-White & Carter, 1970). Certainly, the presence of Bs can interfere with the normal functioning of cells, including processes of mitosis and meiosis and this may be why an overload of Bs is rarely tolerated, while, one B chromosome is the most common situation.

We have analysed the possible relationship between the average number of Bs and polyploidy, chromosome number and genome size at different taxonomic levels (Tables S1–S4) in plants and animals (the small number of available species of fungi does not allow proper statistical analyses). There is a very faint trend to higher number of Bs with higher ploidy levels, chromosome numbers and genome sizes in both plants and animals although the relationship is only significant in eudicots and with chromosome number ( $\rho = 0.082$ ,  $P = 0.003$ ) and ploidy level ( $\rho = 0.080$ ,  $P = 0.005$ ), and in animals but only with chromosome number ( $\rho = 0.127$ ,  $P = 0.0005$ ). Palestis *et al.* (2004) found a positive correlation of Bs with genome size, assessing B-chromosome frequency across angiosperms. In the same line, Trivers *et al.* (2004) found a strong positive correlation between Bs and genome size in British angiosperms, and Levin *et al.* (2005) hypothesized that species with small genomes would have a lower incidence of Bs, as larger genomes would better tolerate additional genetic material. However, to our knowledge the positive (though faint) relationship found between ploidy level and B presence had not been detected in previous works (Jones & Rees, 1982; Trivers *et al.*, 2004; Levin *et al.*, 2005).

Nevertheless, the large and biased sample prevents general conclusions and these relationships are better studied at lower taxonomic levels such as family or genera. In these cases, certain groups behave completely differently from others: for example while in genera *Artemisia*, *Bromus*, *Oryzopsis*, *Lolium*, *Diabrotica* and *Ophris* the number of Bs is positively and significantly correlated with chromosome number, the contrary is true for

genera *Poa*, *Fritillaria*, *Crotalaria*, *Brachycome*, *Cytisus* and *Listera* (see Table S3). It has been hypothesized that there are several mechanisms of origin of Bs and also different selection/environmental pressures may shape the destiny of B-chromosome behaviour depending on the group. Besides, in most groups there is no relationship with Bs, highlighting the independent, perhaps parasitic, nature of these unpredictable genomic components.

We have also studied the possible influence of polyploidy on the presence of Bs in certain genera in which enough data were available (Table S5) to allow a proper analysis. In all three genera (*Allium*, *Artemisia* and *Festuca*, representatives of the most commonly studied families regarding Bs) we have found that species which are only present at diploid level have lower proportions of counts with Bs than species which are present at different ploidy levels (Fig. 3).



**Fig. 3** Comparison between the distribution of ploidy levels in all species vs species with B chromosomes (Bs) in genera (a) *Allium*, (b) *Artemisia* and (c) *Festuca*. Both autopolyploid and allopolyploid taxa may be included in the analyses for the three genera.

We can hypothesize that the mechanism(s) promoting polyploidy and/or chromosome number variability may be related to the ways in which Bs might have arisen. *De novo* origin of Bs has been detected in the complex of *Prospero autumnale* (previously *Scilla autumnalis*) in which these may be by-products of its large-scale chromosomal rearrangements (Jang *et al.*, 2016).

The release of B-chrom has meant a considerable assembling effort, but it is still an initial step on the systematization of data on Bs. The database is envisaged as a long-term project, and in future updates we aim to include other relevant information such as Bs morphology (e.g. visibility and position of centromere), relative size with respect to the A-chromosome complement (e.g. micro- or macro-Bs), known gene content (e.g. rDNA), etc. Furthermore, in future releases, we would like to incorporate data from other information sources which have not been explored e.g. PhD theses or reference lists from the most important books and articles dealing with this topic. The analyses here presented are a reflection of the seemingly unpredictable nature of these ‘passengers in the genome’ (Jones *et al.*, 2007). Yet, as pointed out previously, the database embraces a small fraction of eukaryotic diversity, highlighting some relevant gaps in knowledge. In particular, research should be triggered in unexplored groups such as birds, algae or fungi.

## Acknowledgements

The authors would like to thank Paula Bonaventura and María Luisa Gutiérrez (IBB-CSIC-ICUB, Barcelona, Spain) for their help in data mining and Teresa Garnatje (IBB-CSIC-ICUB) for carefully reading the manuscript. The authors acknowledge Sergi Garcia for the creation of the database web environment. The Dirección General de Investigación Científica y Técnica from the Government of Spain (CGL2016-75694-P), the Czech Science Foundation (P506/16/02149J) and the Government of Catalonia (‘Ajuts a grups de recerca consolidats’, 2014SGR514) are acknowledged for funding. S.G. benefits from a ‘Ramón y Cajal’ contract from the Government of Spain (RYC-2014-16608).

## Author contributions

S.G. planned and designed the research. U.D’A., M.P.A-F. and K.B. performed publication search and data collection. G.M.d.X., S.G. and A.K. analysed data. U.D’A., M.P.A-F., K.B. and S.G. wrote the manuscript with significant contributions from A.K. and G.M.d.X.

Ugo D’Ambrosio<sup>1†</sup>, M. Pilar Alonso-Lifante<sup>1†</sup>,  
Karina Barros<sup>1†</sup>, Aleš Kovařík<sup>2</sup>, Gemma Mas de Xaxars<sup>3</sup> and  
Sònia Garcia<sup>1\*</sup>

<sup>1</sup>Institut Botànic de Barcelona (IBB-CSIC-ICUB), Passeig del Migdia s/n, Barcelona 08038, Catalonia, Spain;

<sup>2</sup>Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno CZ-61265, Czech Republic;

<sup>3</sup>Laboratori de Botànica (UB), Unitat Associada al CSIC, Facultat de Farmàcia i Ciències de l’Alimentació, Universitat de Barcelona,

Av. Joan XXIII s.n., Barcelona 08028, Catalonia, Spain  
(\*Author for correspondence: tel +34 93289611;  
email soniagarcia@ibb.csic.es)

†These authors contributed equally to this work.

## References

- Austin AP. 1956. Chromosome counts in the Rhodophyceae. *Nature* 178: 370–371.
- Banaei-Moghaddam AM, Martis MM, Macas J, Gundlach H, Himmelbach A, Altschmied L, Mayer KFX, Houben A. 2015. Genes on B chromosomes: old questions revisited with new tools. *Biochimica et Biophysica Acta (BBA)–Gene Regulatory Mechanisms* 1849: 64–70.
- Barrowclough GF, Cracraft J, Klicka J, Zink RM. 2016. How many kinds of birds are there and why does it matter? *PLoS ONE* 11: e0166307.
- Borisov YM. 2014. B-chromosomes and the plasticity of the species. *Russian Journal of Genetics: Applied Research* 4: 341–350.
- Boyle B, Hopkins N, Lu Z, Garay JAR, Mozzherin D, Rees T, Matasci N, Narro ML, Piel WH, Mckay SJ *et al.* 2013. The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics* 14: 16.
- Byers C, Maughan PJ, Clouse J, Stewart JR. 2014. Microsatellite primers in *Agave utahensis* (Asparagaceae), a keystone species in the Mojave Desert and Colorado Plateau. *Applications in Plant Sciences* 2: 1400047.
- Camacho JPM, Cabrero J, López-León MD, Bakkali M, Perfectti F. 2003. The B chromosomes of the grasshopper *Eyprepocnemis plorans* and the intragenomic conflict. *Genetica* 117: 77–84.
- Camacho JPM, Sharbel TF, Beukeboom LW. 2000. B-chromosome evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 355: 163–178.
- Chapman AD. 2009. *Numbers of living species in Australia and the World, 2<sup>nd</sup> edn.* Canberra, ACT, Australia: Australian Biological Resources Study.
- Coleman JJ, Rounsley SD, Rodríguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S *et al.* 2009. The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genetics* 5: e1000618.
- Darlington CD, Wylie AP. 1955. *Chromosome atlas of flowering plants.* London, UK: George Allen & Unwin.
- Datta AK, Mandal A, Das D, Gupta S, Saha A, Paul R, Sengupta S. 2016. B chromosomes in angiosperm – a review. *Cytology and Genetics* 50: 60–71.
- Dhar MK, Friebe B, Koul AK, Gill BS. 2002. Origin of an apparent B chromosome by mutation, chromosome fragmentation and specific DNA sequence amplification. *Chromosoma* 111: 332–340.
- Ellegren H. 2010. Evolutionary stasis: the stable chromosomes of birds. *Trends in Ecology & Evolution* 25: 283–291.
- Fedorov AA. 1969. *Chromosome numbers of flowering plants.* Leningrad, USSR: Academy of Natural Sciences of the USSR.
- Galazka JM, Freitag M. 2014. Variability of chromosome structure in pathogenic Fungi – of ‘ends and odds’. *Current Opinion in Microbiology* 20: 19–26.
- García S, Gálvez F, Gras A, Kovařík A, Garnatje T. 2014. Plant rDNA database: update and new features. *Database* 2014: bau063.
- García S, Kovařík A, Leitch AR, Garnatje T. 2017. Cytogenetic features of rRNA genes across land plants: analysis of the Plant rDNA database. *Plant Journal* 89: 1020–1030.
- Goldblatt P, Johnson DE. 1998. Index to plant chromosome numbers 1994–1995. *Monographs in Systematic Botany from the Missouri Botanical Garden* 69: 1–208.
- Goldblatt P, Johnson DE. 2000. Index to plant chromosome numbers 1996–1997. *Monographs in Systematic Botany from the Missouri Botanical Garden* 81: 1–188.
- Goldblatt P, Johnson DE. 2003. Index to plant chromosome numbers 1998–2000. *Monographs in Systematic Botany from the Missouri Botanical Garden* 94: 1–297.
- Goldblatt P, Johnson DE. 2006. Index to plant chromosome numbers 2001–2003. *Monographs in Systematic Botany from the Missouri Botanical Garden* 106: 1–242.
- Goldblatt P, Johnson DE. 2010. Index to plant chromosome numbers 2004–2006. In: Goldblatt P, Johnson ED, eds. *Regnum vegetabile, vol. 152.* Ruggell, Liechtenstein: ARG Gartner Verlag, 256.
- Gotoh K. 1924. Über die Chromosomenzahl von *Secale cereale* L. *The Botanical Magazine* 38: 135–152.
- Guo J, Liu H, He Y, Cui X, Du X, Zhu J. 2015. Origination of asexual plantlets in three species of Crassulaceae. *Protoplasma* 252: 591–603.
- Holmes DS, Bougourd SM. 1989. B-chromosome selection in *Allium schoenoprasum* L. Natural populations. *Heredity* 63: 83–87.
- Houben A. 2017. B chromosomes – a matter of chromosome drive. *Frontiers in Plant Science* 8: 210.
- Houben A, Banaei-Moghaddam A, Klemme S. 2013. Biology and evolution of B chromosomes. In: Greilhuber J, Doležel J, Wendel JF, eds. *Plant genome diversity, vol. 2.* Vienna, Austria: Springer, 149–165.
- Jang TS, Parker JS, Weiss-Schneeweiss H. 2016. Structural polymorphisms and distinct genomic composition suggest recurrent origin and ongoing evolution of B chromosomes in the *Prospero autumnale* complex (*Hyacinthaceae*). *New Phytologist* 210: 669–679.
- Jones N, Houben A. 2003. B chromosomes in plants: escapees from the A chromosome genome? *Trends in Plant Science* 8: 417–423.
- Jones RN. 1976. Genome organisation in higher plants. In: Pearson PL, Lewis KR, eds. *Chromosomes today, vol. 5.* Chichester, UK: John Wiley & Sons, 117–130.
- Jones RN, Díez M. 2004. The B chromosome database. *Cytogenetic and Genome Research* 106: 149–150.
- Jones RN, Rees H. 1982. *B chromosomes.* New York, NY, USA: Academic Press.
- Jones RN, Viegas W, Houben A. 2007. A century of B chromosomes in plants: so what? *Annals of Botany* 101: 767–775.
- Kantama L, Sharbel TF, Schranz ME, Mitchell-Olds T, de Vries S, de Jong H. 2007. Diploid apomicts of the *Boechera holboellii* complex display large-scale chromosome substitutions and aberrant chromosomes. *Proceedings of the National Academy of Sciences, USA* 104: 14026–14031.
- Kapraun DF. 1993. Karyology of marine green algae. *Phycologia* 32: 1–21.
- Leach CR, Houben A, Field B, Pistrick K, Demidov D, Timmis JN. 2005. Molecular evidence for transcription of genes on a B chromosome in *Crepis capillaris*. *Genetics* 171: 269–278.
- Levin DA, Palestis BG, Jones RN, Trivers R. 2005. Phyletic hot spots for B chromosomes in angiosperms. *Evolution* 59: 962–969.
- Lewis RJ. 1996. Chromosomes of the brown algae. *Phycologia* 35: 19–40.
- Makunin AI, Dementyeva PV, Graphodatsky AS, Volobouev VT, Kukekova AV, Trifonov VA. 2014. Genes on B chromosomes of vertebrates. *Molecular Cytogenetics* 7: 99.
- Mandáková T, Schranz ME, Sharbel TF, Jong H, Lysak MA. 2015. Karyotype evolution in apomictic *Boechera* and the origin of the aberrant chromosomes. *Plant Journal* 82: 785–793.
- Mongeon P, Paul-Hus A. 2016. The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics* 106: 213–228.
- Muravenko O, Selyakh I, Kononenko N, Stadnichuk I. 2001. Chromosome numbers and nuclear DNA contents in the red microalgae *Cyanidium caldarium* and three *Galdieria* species. *European Journal of Phycology* 36: 227–232.
- Palestis BG, Cabrero J, Trivers R, Camacho JPM. 2010. Prevalence of B chromosomes in Orthoptera is associated with shape and number of A chromosomes. *Genetica* 138: 1181–1189.
- Palestis BG, Trivers R, Burt A, Jones RN. 2004. The distribution of B chromosomes across species. *Cytogenetic and Genome Research* 106: 151–158.
- Rice A, Glick L, Abadi S, Einhorn M, Kopelman NM, Salman-Minkov A, Mayzel J, Chay O, Mayrose I. 2015. The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. *New Phytologist* 206: 19–26.
- Sharbel TF, Green DM, Houben A. 1998. B chromosome origin in the endemic New Zealand frog *Leiopelma hochstetteri* through sex chromosome evolution. *Genome* 41: 14–22.
- Smith-White S, Carter CR. 1970. The cytology of *Brachycome lineariloba*. *Chromosoma* 30: 129–153.
- Stukenbrock EH, Croll D. 2014. The evolving fungal genome. *Fungal Biology Reviews* 28: 1–12.

- Tegelström H, Rytman H. 1981. Chromosomes in birds (Aves): evolutionary implications of macro- and microchromosome numbers and lengths. *Hereditas* 94: 225–233.
- Trivers R, Burt A, Palestis BG. 2004. B chromosomes and genome size in flowering plants. *Genome* 47: 1–8.
- Valente GT, Nakajima RT, Fantinatti BEA, Marques DF, Almeida RO, Simões RP, Martins C. 2017. B chromosomes: from cytogenetics to systems biology. *Chromosoma* 126: 73–81.
- van Vugt JJFA, de Nooijer S, Stouthamer R, de Jong H. 2005. NOR activity and repeat sequences of the paternal sex ratio chromosome of the parasitoid wasp *Trichogramma kaykai*. *Chromosoma* 114: 410–419.
- Vujošević M, Blagojević J. 2004. B chromosomes in populations of mammals. *Cytogenetic and Genome Research* 106: 247–256.
- Werren JH, Stouthamer R. 2003. PSR (paternal sex ratio) chromosomes: the ultimate selfish genetic elements. *Genetica* 117: 85–101.
- Williams E, Barclay PC. 1968. The effects of B-chromosomes on vigour and fertility in *Dactylis* hybrids. *New Zealand Journal of Botany* 6: 405–416.
- Wilson EB. 1907. Note on the chromosome-groups of *Metapodius* and *Banasa*. *The Biological Bulletin* 12: 303–313.
- Wolff FD. 1971. Techniques for the vegetative propagation of maize (*Zea mays* L.). *Euphytica* 20: 524–526.
- 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 Genetics* 7: e1002203.

## Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

**Table S1** Results of the statistical analyses comparing plants and animals, on the one hand, and monocots and eudicots, on the other hand

**Table S2** Results of the statistical analyses at the family level

**Table S3** Results of the statistical analyses at the genus level

**Table S4** Results of the statistical analyses at the order level

**Table S5** Data used for the analysis performed in genera *Allium*, *Artemisia* and *Festuca*

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

**Key words:** B-chromosomes, cytogenetics, database, data-mining, evolution, karyology, karyotype.



## About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and **all the latest information** visit **www.newphytologist.com**