Palacký University Olomouc

Faculty of Science

Department of Botany

and

Institute of Experimental Botany of the Czech Academy of Sciences

Centre of Plant Structural and Functional Genomics

Occurrence and evolution of B chromosomes in genus *Sorghum*

Ph.D. Thesis

Mgr. Martina Bednářová

Ph.D. Program of Biology – Botany

Supervisor: Mgr. Jan Bartoš, Ph.D.

Olomouc
2022

Acknowledgements

I would like to thank my supervisor Mgr. Jan Bartoš, Ph.D., for his professional guidance, advice and patience. I would also like to thank Mgr. Eva Hřibová, Ph.D., for her willingness and help with phylogenetic analysis. My thanks also belong to prof. Ing. Jaroslav Doležel, DrSc., for the opportunity to work in the Centre of Plant Structural and Functional Genomics. I am also grateful to all members of the laboratory for their help and friendly atmosphere.

Declaration

I hereby declare that I have written the Ph.D. thesis independently under the supervision of Mgr. Jan Bartoš, Ph.D., using the sources listed in references and with no conflict of interest.

....

This work was supported by the Czech Science Foundation (grant award no. 18-12338Y) and the European Regional Development Fund project ‘Plants as a tool for sustainable global development’ (CZ.02.1.01/0.0/0.0/16_019/0000827).

Bibliographical identification

Author's name:	Mgr. Martina Bednářová
Title:	Occurrence and evolution of B chromosomes in genus *Sorghum*
Type of Thesis:	Ph.D. thesis
Department:	Department of Botany
Supervisor:	Mgr. Jan Bartoš, Ph.D.
The year of presentation:	2022

Abstract:

This work is focused on B chromosomes (Bs) in the genus *Sorghum*. In this genus, Bs have been reported in five species. Since their first discovery in the first half of the 20th century, there has not been any significant advance in our knowledge of them. The available literature on this topic in *Sorghum* is limited to only a few cytogenetic studies and there are no molecular data whatsoever. The aim of this work was to expand our knowledge about Bs in *Sorghum* in the molecular field. The first goal was focused on the development of specific markers for B chromosome of *S. purpureosericeum*. For this purpose, haploid pollen nuclei from B-positive and B-negative plants were isolated using flow-cytometry. Extracted nuclear DNA was used for the preparation of sequencing libraries which were subjected to Illumina sequencing. Based on the raw data, a repeat analysis was performed using RepeatExplorer, which revealed nine putative B-specific clusters. These clusters were used to design primers whose B-specificity was verified in PCR. The second goal was to clarify phylogenetic relationships within the genus with respect to the occurrence of B chromosomes. Therefore, a phylogenetic reconstruction of 21 *Sorghum* species represented by 58 samples involving internal transcribed spacers (*ITS1* and *ITS2*) and two chloroplast intergenic spacers (*trnL-trnF* and *trnH-psbA*) as phylogenetic markers

was performed. For the reconstruction, BioNJ and PhyML method was applied. Based on the outcome of the analysis, the feasible scenarios of B chromosome origin in this genus were discussed.

Keywords: B chromosome, *Sorghum*, phylogenetic analysis, sequencing, evolution

Number of Pages/Appendices: 98/I

Language: English

Bibliografická identifikace

Jméno autora: Mgr. Martina Bednářová

Název práce: Výskyt a evoluce B chromozomů v rodu *Sorghum*

Typ práce: Dizertační práce

Katedra: Katedra botaniky

Školitel: Mgr. Jan Bartoš, Ph.D.

Rok obhajoby: 2022

Abstrakt:

Tato práce je zaměřena na B chromozomy v rodu *Sorghum* (čirok). Přestože přítomnost B chromozomů byla popsána u pěti druhů čiroku již v minulém století, od doby jejich objevu nedošlo k významnému posunu v jejich výzkumu. Dostupná literatura na toto téma je omezena pouze na několik cytogenetických studií a molekulární studie dosud publikovány nebyly. Cílem této práce bylo rozšířit naše poznatky o B chromozomech v rodu *Sorghum* především na molekulární úrovni. Prvním cílem byl vývoj markerů specifických pro B chromozom *S. purpureosericeum*. Za tímto účelem byla pomocí průtokové cytometrie izolována haploidní pylová jádra B-pozitivních a B-negativních rostlin *S. purpureosericeum*. Z jejich DNA byly připraveny sekvenační knihovny, které byly osekvenovány na platformě Illumina. Na základě získaných dat byla provedena analýza repetic pomocí nástroje RepeatExplorer. Tato analýza odhalila devět sekvenčních klastrů potenciálně specifických pro B chromozom, které byly dále použity pro navržení primerů, jejichž specifita byla ověřena pomocí PCR. Druhým cílem práce bylo objasnění fylogenetických vztahů mezi jednotlivými druhy čiroku, a to především s ohledem na výskyt B chromozomů. Do fylogenetické analýzy bylo zahrnuto celkem 21 druhů čiroku, které byly zastoupeny 58 vzorky. Jako fylogenetické markery byly použity vnitřní transkribované mezerníky (*ITS1* a *ITS2*) a dva chloroplastové mezigenové mezerníky (*trnL-trnF* a *trnH-psbA*). Na základě fylogenetického stromu, který byl

vytvořen pomocí metody BioNJ and PhyML, byly diskutovány možné scénáře vzniku B chromozomu v rodu *Sorghum*.

Klíčová slova: B chromozomy, *Sorghum*, fylogenetická analýza, sekvenování, evoluce

Počet stran/příloh: 98/I

Jazyk: Anglický

CONTENT

1 INTRODUCTION 10
2 LITERATURE OVERVIEW 12
2.1 B chromosomes...12
2.1.1 General features...12
2.1.2 Phenotypic effects of B chromosomes..13
2.1.3 Sequence composition of B chromosomes...................................14
2.1.4 Accumulation mechanisms of B chromosomes............................18
2.1.5 Mitotic and meiotic stability...21
2.1.6 Origin of B chromosomes..23
2.1.7 Evolutionary models of B chromosomes.......................................27
2.1.8 Importance and potential use of B chromosomes........................28
2.2 Genus *Sorghum* and B chromosomes..30
2.2.1 Characterization of the genus *Sorghum*.......................................30
2.2.2 B chromosomes in *Sorghum*...31
2.3. Methods of B chromosome study.. 33
2.3.1 Cytology and cytogenetics...33
2.3.2 Sequencing methods...36
2.4 Molecular phylogenetics..40
2.4.1 Molecular markers...42
2.4.2 Phylogenetic analysis...44
3 AIMS OF THE THESIS 46
4 RESULTS 47
4.1 The B chromosome of *S. purpureosericeum* reveals the first pieces of its sequence...47
4.2 B chromosomes in genus *Sorghum* (Poaceae)......................................61
5 GENERAL CONCLUSION 77
6 SOUHRN (Summary, in Czech)..79
7 REFERENCES 81
8 LIST OF ABBREVIATIONS 96
9 OTHER PUBLICATIONS 98

1 INTRODUCTION

B chromosomes (Bs) are extra chromosomes that occur in some individuals within populations of a particular species. They have been observed in all major groups of living organisms - plants, animals and fungi (D'Ambrosio *et al.* 2017). Unlike the basic set of chromosomes (A chromosomes, As), Bs are not essential for their host organism and are considered to be parasitic elements with their own evolutionary pathway (Baukeboom 1994). In addition, the behavior of Bs in their host organisms is in many ways unique, as they do not follow Mendelian inheritance and do not pair and recombine with A chromosomes during meiosis. The transmission rate of Bs is higher than 0.5, which is achieved due to the existence of their accumulation mechanisms developed during B's evolution. B chromosomes often show numerical variability between individuals, or even within an individual. Their presence does not have an effect on phenotype, therefore their detection relies on cytological observations (Jones 1995).

Although our knowledge of B chromosomes has advanced considerably in recent years, mainly due to the development of new sequencing technologies, there are still many species or groups of organisms with B chromosomes that have been so far largely neglected by the scientific community. Example of such group is the genus *Sorghum*. *Sorghum* belongs among monocot flowering plants in the grass family Poaceae. The genus covers about 23 annual and perennial species including an agriculturally important *Sorghum bicolor*. B chromosomes were reported in five *Sorghum* species: *S. purpureosericeum*, *S. nitidum*, *S. halepense*, *S. stipoideum*, and *S. bicolor* ssp. *verticilliflorum* (Janaki-Amal 1940, Raman and Krishnaswami 1960, Raman *et al.* 1964, Wu 1992, Huskins and Smith 1932). Typically, B chromosomes in this genus are absent in roots and most of the shoot tissues (Darlington and Thomas 1941). Further, B chromosomes in *Sorghum* are morphologically variable, and different morphotypes have been reported even within individual species (Raman and Krishnaswami 1960, Wu and Pi 1975). Due to the fact that B chromosomes have not yet been the target of greater research interest, our knowledge in this field is limited to a few cytogenetic studies. The presented work aims to broaden the knowledge about B chromosomes in the genus *Sorghum*, for the first time it presents

the results at the molecular level. Further, the work aims to contribute to the understanding of the evolution of the B chromosome within the genus.

2 LITERATURE OVERVIEW

2.1 B chromosomes

2.1.1 General features

B chromosomes are extra chromosomes that complement the basic set of chromosomes in some species. The first species, in which the extra chromosome was observed, was the hemipteran insect *Metapodius* (Wilson 1907, 1909). Randolph (1928) first used the name "B chromosome" to distinguish it from the basic set of "A chromosomes". So far, B chromosomes have been found in more than two thousand plant species, more than seven hundred animal species (including mammals), and several species of fungi (D'Ambrosio *et al.* 2017). Presumably, there are many other B containing species that has not been identified yet. The behavior, structure and evolutionary dynamics of B chromosomes is partially species-specific, however, they share several common features - they are dispensable, they do not pair and recombine with A chromosomes during meiosis, and have non-Mendelian mode of inheritance (Jones 1995). Due to a range of accumulation mechanisms, Bs are usually transmitted to the offspring at a frequency higher than 0.5 and therefore they are considered as parasitic elements (Baukeboom 1994). Within a given population, Bs may not be present in every individual and different individuals may carry a different number of Bs.

B chromosomes are usually smaller than As, but some Bs can be as large as As and even Bs larger than As have been observed. The centromere position often differs from the centromere position on A chromosomes (Jones and Houben 2003). In many species, morphological variants of Bs have been reported. For example, the chive *Allium schoenoprasum* carries up to 29 types of Bs, the plant *Brachycome dichromosomatica* carries two forms micro-B and a large B (Fig. 1) and two forms of B have been found also in the mottled grasshopper *Myrmeleotettix maculatus* (Bougourd and Parker 1979, Smith-White and Carter 1970, John and Hewitt 1965). For most species, a maximum of 3-4 Bs has been reported, however, in some species, Bs are very well tolerated even in large numbers. *Pachyphytum fittkaui* can harbor up to 50 Bs (Uhl and Moran 1973) and *Zea mays* up to 34 Bs (Jones and Reese 1982).

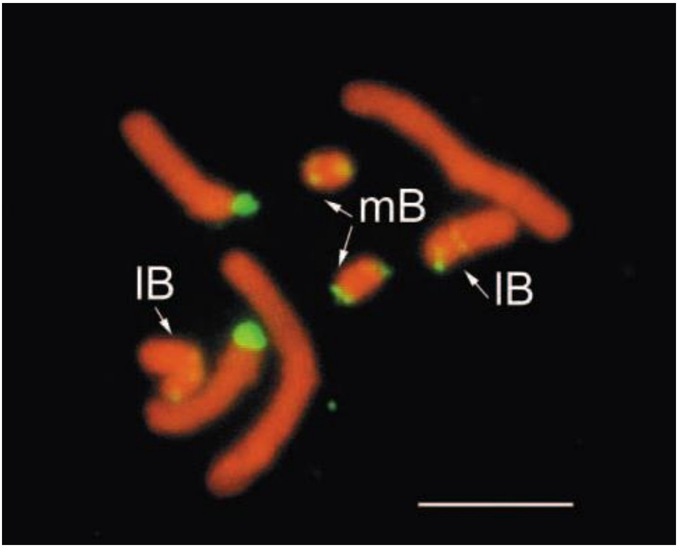

Fig. 1 An example of B chromosome polymorphism in *Brachycome dichromosomatica.* Yellow signal indicates 45S rDNA on the A, micro-B (mB), and large B chromosomes (lB). Scale bar = 5 mm (adopted from Marschner *et al.* 2007).

2.1.2 Phenotypic effects of B chromosomes

B chromosomes usually do not affect the phenotype, although there are some exceptions - for example leaf striping observed in maize with high number of Bs (Staub 1987). Presence of Bs is generally neither beneficial nor harmful, but at higher numbers Bs often reduce fertility and vigor due to their cumulative negative effect (Jones 1991). The presence of Bs increases the DNA content of the cell, which causes a burden on the carrier, especially with regard to nuclear physiology. Lengthened cell cycle, increased cell size and relatively reduced levels of nuclear proteins and RNA are the consequences of the presence of the Bs (Jones 2012). A higher number of Bs then results in impaired fertility and fitness, as demonstrated by the example of rye (Fig. 2). On the other hand, in some species Bs provide benefits such as resistance to rust in the oat (*Avena sativa*) (Dherawattana and Sadanaga 1973) or resistance to antibiotics of the plant pathogenic fungus *Nectria haematococca* (Han *et al.* 2001). In the frog *Leiopelma hochstetteri* Bs have been reported to affect sex determination (Green 1988).

Fig. 2 The harmful effects of B chromosomes on fitness in *Secale cereale*. Rye plants with 3Bs are phenotypically indiscernible from 0B plants (not shown), but 6Bs plants show a remarkable decrease in vigor and are completely sterile (Jones 2012).

2.1.3 Sequence composition of B chromosomes

Generally, B chromosomes are a mosaic of different sequences derived from the host genome. They contain large amounts of repetitive sequences, especially mobile elements and tandem repeats. A large number of repeats can be the underlying cause of strong heterochromatinization of many B chromosomes (Houben *et al.* 2013). B chromosomes also frequently contain gene sequences and sequences derived from plastids or mitochondria. Only a very few sequences were identified and confirmed as a B-specific (Marques *et al.* 2018).

The high proportion of repeats of different types and copy numbers may have accumulated in Bs due to uneven crossing over and reduced recombination (Camacho *et al.* 2000). Sometimes repeats can form a majority of the B chromosome mass, as in micro-B of *Brachycome dichromosomatica* (Houben *et al.* 2001) or PSR (paternal sex ratio) chromosome of the jewel wasp *Nasonia vitripennis* (McAllister and Werren 1997). In *Drosophila subsilvestris*, repeats form the entire B chromosome (Gutknecht *et al.* 1995). Rapid amplification of repeats is considered a possible

mechanism for stabilization and meiotic isolation of the emerging B chromosome. This process is documented on the B chromosome of *Plantago lagopus*, where a massive amplification of 5S rDNA was connected with the emergence of B chromosome (Dhar *et al.* 2002, Kumke *et al.* 2016).

The origin of the majority of repetitive sequences can be traced to A chromosomes, but repeats specific to B chromosomes have also been discovered. For example, the end of the long arm of the rye B chromosome contains specific high-copy repeats E3900 and D1100 (Sandery *et al.* 1990; Blunden *et al.* 1993). These repeats, which have probably evolved *de novo* on B chromosome, are tandem repeats created most likely from fragments of various sequences, including mobile elements (Langdon *et al.* 2000). The transcriptional activity of E3900 and D1100 in specific tissues and the production of non-coding RNA transcripts have been reported (Klemme *et al.* 2013, Carchilan *et al.* 2007, Carchilan *et al.* 2009). The composition of repeats on the B of rye is shown in Fig. 3. B-specific repeats were identified also on maize B chromosome. StarkB, derived mainly from retrotransposons, is located in the distal heterochromatin region and is transcriptionally active. Another B-specific repeat - ZmBs – is localized in the centromere region (Alfenito and Birchler 1993). It is interspersed between two other repeats at the centromere core and is associated with CENH3 (Jin *et al.* 2005). It has been proposed that ZmBs can be involved in nondisjunction (Han *et al.* 2007).

Ribosomal DNA on B chromosomes has been detected in many species, e.g. *Brachycome dichromosomatica*, *Crepis capillaris* or *Eyprepocnemis plorans*. In the latter two, the genes are even transcriptionally active (Donald *et al.* 1995, Leach *et al.* 2005, Ruiz-Estévez *et al.* 2012).

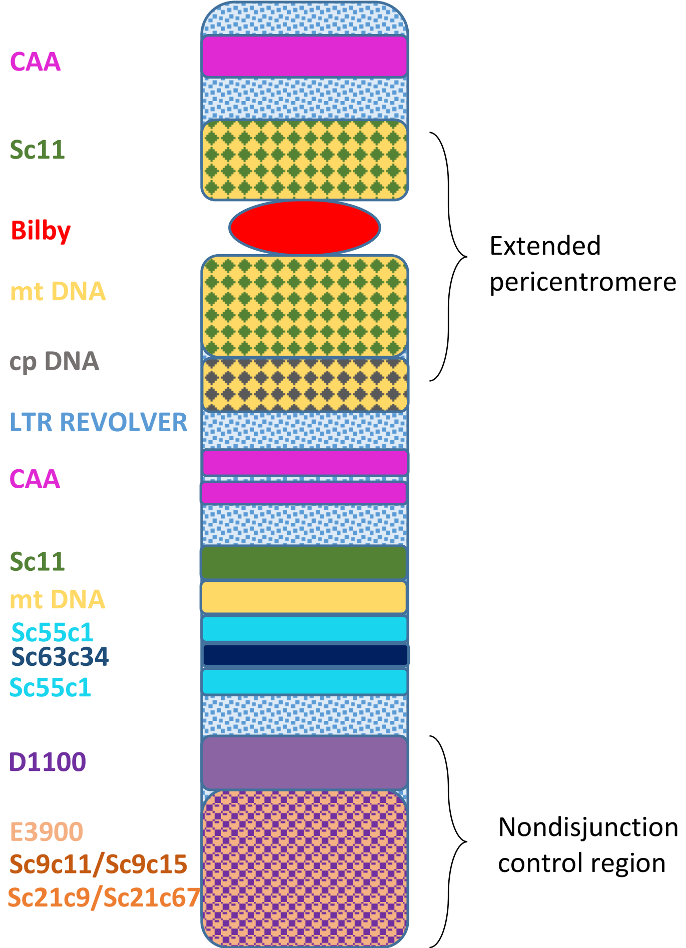

Fig. 3 Model for the distribution of rye B chromosome-enriched sequences. Rye B-specific repeats D1100 and E3900 are located in nondisjunction control region at the end of the long arm. cp DNA – chloroplast DNA, mt DNA – mitochondrial DNA (according to Marques *et al.* 2018).

B chromosomes typically contain transposable elements, which may be caused by reduced crossing over during meiosis, similarly to the case of sex chromosomes (Charlesworth 2008). In addition, unlike As, Bs are not the subject to selection pressure, thus transposons can accumulate indefinitely. One piece of evidence is the different proportion of transposons on B and As in some species. For example, rye A chromosomes are rich in the transcriptionally inactive ancient retroelement Sabrina, which is present in all Triticeae (Shirasu *et al.* 2000), whereas the B chromosome contains lower number of Sabrina copies and an expansion of the active Revolver element (Klemme *et al.* 2013). Another example is the spread of the NATE element on the PSR chromosome of *Nasonia vitripennis* (McAllister and Werren 1997).

It is thought that transposons may undergo ectopic recombination, which may be a successful tool for the transfer of sequences of various origins on Bs. Such event has been proposed by Franks *et al.* (1996) in *Brachycome dichromosomatica*, whose B chromosome contains chloroplast DNA. Due to the abundant presence of mobile elements on B chromosomes, it can be assumed that their insertions cause silencing of gene sequences and their subsequent pseudogenization. Increased transposon activity is thus another factor leading to structural variability and B chromosome evolution (Camacho *et al.* 2000).

The transfer of DNA sequences from mitochondria and plastid to the cell nucleus takes place continuously. Organellar DNA integrated into the nucleus has been found in almost all plant genomes studied (Ayliffe *et al.* 1998). Over time, this DNA is eroded and removed by mechanisms including mutations, deletions, and replacement with non-organellar DNA, as has been shown in rice and *Arabidopsis* (Matsuo *et al.* 2005, Noutsos *et al.* 2005). A role of nuclear mitochondrial DNA and nuclear plastid DNA in evolving new nuclear genes has been reported (Lloyd and Timmis 2011). Since Bs are not essential for their hosts, the accumulation of organelle-derived sequences does not have any detrimental effects here. It has been speculated that the mechanisms responsible for the removal of such sequences on As could be disrupted on Bs (Marques *et al.* 2018). Large amounts of mitochondrial DNA were found in rye B pericentromeric regions (Martis *et al.* 2012). Ruban *et al.* (2014) reported both plastid- and mitochondria-derived DNA accumulation on B chromosome of *Aegilops speltoides* and abundance of organellar DNA was revealed also on maize B chromosome (Blavet *et al.* 2021).

It has long been thought that Bs do not contain genes and are not transcriptionally active. However, more studies are now available that have revealed transcriptional activity and functional genes on B chromosomes. New analyses involving next generation sequencing have shown that some B chromosomes contain numerous protein-coding sequences. Transcriptional activity of B chromosomes was observed in *Drosophila albomicans*, *Nasonia vitripennis* or the Siberian roe deer (*Capreolus pygarus*) (Zhou *et al.* 2012, Akbari *et al.* 2013, Trifonov *et al.* 2013). In contrast, no or weak transcriptional activity was found in *Apodemus peninsulae* or *Myrmeleotettix maculatus* (Ishak *et al.* 1991, Fox *et al.* 1974). Gene sequences on Bs

are most likely derived from A chromosomes, but horizontal transfer was also demonstrated (Brelsfoard *et al.* 2014). The low abundance of functional genes on B chromosomes corresponds to their dispensability and thus the increased rate of mutations. The fate of genic sequences is erosion by mutations and following pseudogenization. But it is also very likely that if there are functional genes on the B chromosomes, they are downregulated, to preserve the gene dosage. In rye B, gene fragments representing copies of A chromosome genes were identified. The fragments showed increased polymorphism, proving ongoing pseudogenization, but some of them were transcribed. Furthermore, these sequences have been found to affect transcription on A chromosomes (Martis *et al.* 2012, Banaei-Moghaddam *et al.* 2013). Later, the functional Argonaut-like protein was identified on rye B by Ma *et al.* (2017). It is apparent that Bs and As may interact and there is mutual regulation between them. The mechanism of such regulation can involve RNA interference, as has been demonstrated in maize and *Nasonia vitripennis* (Huang *et al.* 2016, Li *et al.* 2017). Epigenetic mechanisms or B-affected spatial organization in the nucleus could also play a role in these relationship (Houben *et al.* 2014).

2.1.4 Accumulation mechanisms of B chromosomes

There is a wide variety of accumulation mechanisms by which B chromosomes ensure their transmission to offspring with the transmission rate higher than 0.5. This transmission advantage is referred to as “B chromosome drive”. In some cases, the accumulation mechanisms can have a drastic impact on their hosts. For example, in *Nasonia vitripennis*, PSR chromosome is transmitted to the progeny via male sperm. After fertilization of the egg, PSR induces elimination of the paternal genome and instead of females, haploid males bearing PSR are produced (Nur *et al.* 1988). However, not all B chromosomes show drive. For example, in *Prochilodus lineatus* and *Eyprepocnemis plorans*, no drive mechanism has been identified (Oliveira *et al.* 1997, López-León *et al.* 1992). B chromosome behavior in *E. plorans* has been described as "near neutral" (Camacho *et al.* 1997). It is believed that in some cases B chromosome may be maintained in the population due to its beneficial properties. *Allium schoenoprasum* could be such an example - the accumulation mechanism is

lacking in this species, but a positive effect of B chromosome on germination in dry conditions has been demonstrated (Plowman and Bougourd 1994, Bougourd and Plowman 1996).

Accumulation mechanisms of B chromosomes are species-specific and can be active before, during or after meiosis. They can either operate in both sexes, or be limited to only one. Whereas in animals, premeiotic and meiotic drive is the most common, in plants we typically observe postmeiotic drive during pollen mitosis, although there are some exceptions. The types of accumulation mechanisms in plant species are summarized in Tab. 1.

Tab. 1 Various mechanisms of B chromosome accumulation in plants (Jones 2018, amended).

Mechanism	Species
Nondisjunction at first pollen grain mitosis	*Aegilops speltoides, Alopecurus pratensis, Anthoxanthum aristatum, Brachycome lineariloba, Briza media, Dactylis glomerata, Deschampsia bottnica, Deschampsia caespitosa, Deschampsia wibeliana, Festuca arundinacea, Festuca pratensis, Haplopappus gracilis, Holcus lanatus, Phleum phleoides, Panicum maximum, Aegilops mutica, Secale cereale*
Nondisjunction at second pollen grain mitosis	*Zea mays*
Pollen grain mitosis of extra divisions	*Sorghum purpureosericeum*
Somatic nondisjunction in the developing inflorescences	*Crepis capillaris*
Female meiotic drive	*Lilium callosum, Phleum nodosum, Plantago serraria, Trillium grandiflorum*
Female meiotic drive and male meiotic drag	*Picea sitchensis, Hypochoeris maculata*
Male drive	*Haplopappus validus, Clarkia elegans, Iseilema laxum, Briza humilis*
Somatic nondisjunction coincident with flower initiation	*Crepis capillaris*
No apparent mechanism	*Allium schoenoprasum, Xanthisma texanum, Centauria scabiosa, Poa alpina, Ranunculus acris, Ranunculus ficaria*

Premeiotic drive involves nondisjunction of the B chromosome during mitoses that precede germ cell formation. Premeiotic drive was observed during spermatogonial mitosis in the testes of animals. For example, in the migratory locust *Locusta migratoria*, the B chromosome preferentially enters the male germ line through nondisjunction (Nur 1969, Kayano 1971) and the same way of drive was observed in the grasshopper *Neopodismopsis abdominalis* (Rothfels 1950). In plants, premeiotic drive has been described so far only in a single species, *Crepis capillaris*, where somatic nondisjunction of the B chromosome occurs during flower initiation – B chromosome is included in cell lines leading to inflorescence formation (Rutishauser and Röthlisberger 1966, Parker *et al.* 1989).

Meiotic drive was first observed in the plant *Lilium callosum*. During the division of the embryo sac, B chromosome moved to the micropylar side in 80% of the cases and subsequently was included in the egg cell (Kimura and Kayano 1961). Meiotic accumulation through females was also observed in the spruce *Picea sitchensis* (Kean *et al.* 1982), and also in some animals, for example in the grasshopper *Myrmeleotettix maculatus* (Hewitt 1976). Interestingly, different drive mechanism can even combine in one species, as in *Locusta migratoria*. Except of the premeiotic drive in males described above, it seems that there is also meiotic drive in females (Pardo *et al.* 1994).

Unlike premeiotic and meiotic drive, which are rare in plant kingdom, postmeiotic drive has been reported in a large number of plant species, including *Secale cereale* (rye) and *Zea mays* (maize). In these two species, the accumulation mechanisms have been investigated and described in detail.

In maize, three mechanisms are known to be responsible for maintaining of the B chromosome in the population: 1) nondisjunction during pollen mitosis (Roman 1947, Rusche *et al.* 1997), 2) preferential fertilization of the egg by a sperm cell bearing a B chromosome (Carlson 1969) and 3) prevention of the loss of B during meiosis when it is present as an unpaired univalent (Carlson and Roseman 1992). Nondisjunction of the B chromosome takes place during the second pollen mitosis, when the generative nucleus of the male gametophyte is divided to two sperm cells. The sister chromatids of the B chromosome do not split in the anaphase and both are drawn to one pole. As a result, only one of the newly formed sperm cells contains B.

The B-carrying sperm cell then preferentially fertilizes the egg, while the sperm cell without the B chromosome merges with the nucleus of the germ sac. The occurrence of nondisjunction depends on the factor localized at the end of the long arm of B chromosome and can act *in trans* (Roman 1947; Carlson 1978; Lamb *et al.* 2006). A locus controlling preferential fertilization is located in centromere proximal region of the B chromosome (Blavet *et al.* 2021), but is also influenced by factors located on A chromosomes (González-Sánchez *et al.* 2003).

In rye, nondisjunction occurs during the first pollen mitosis, when the vegetative and generative nuclei of the gametophyte are developed. During this division, the chromatids does not divide in anaphase and the entire B chromosome is in most cases included into the generative nucleus (Hasegawa 1934). In the second pollen mitosis, the chromatids of the B chromosome segregate normally, forming two sperm cells with an unreduced number of Bs. The factor controlling nondisjunction is located at the end of the long arm of the B chromosome and can act *in trans* (Endo 2008). Rye nondisjunction was documented in both sexes. Accumulation in females probably takes place during the first postmeiotic division in the embryo sac (Hakansson 1948).

2.1.5 Mitotic and meiotic stability

In addition to drive, B chromosome heredity is also affected by its behavior during mitotis and meiotic instability. In most species, regular B chromosome transmission was observed during mitosis and thus all cells carry the same number of Bs. However, there are also species in which mitotic instability has been described. If multiple morphotypes of Bs are present, each form can behave differently. For example, in *Brachycome dichromosomatica*, there are two types of Bs – while "large" B behaves regularly during mitosis, micro-B is subject to mitotic instability (Carter and Smith-White 1972). Mitotic instability can be manifested by nondisjunction of the B chromosome chromatids in the anaphase. Subsequently, B chromosome accumulates in one of the daughter cells, while the other remains without the B chromosome. As a result, different tissues or organs can have variable numbers of Bs. Mitotic instability sometimes represents a form of premeiotic drive, by which B

preferentially enters germline cells, as was shown in grasshoppers *Camnula pelucida* and *Locusta migratoria* (Nur 1969). Alternatively, B chromosome nondisjunction can cause lagging in anaphase. This can lead to cell populations where B is completely eliminated.

In plant species, the B chromosome is typically eliminated from the roots, while in aerial parts it is preserved. An example of such behavior is *Aegilops mutica*, *Ae. speltoides* or *Cymbopogon palmarosa* (Ohta 1995, Mendelson and Zohary 1972, Sreenath and Jagadishchandra 1988). In *Sorghum purpureosericeum*, B chromosome absence was reported not only in root, but also in stem and leaf tissues (Darlington and Thomas 1941). In animals, elimination has been observed, for example, in the ant *Leptothorax spinosior*, where the B chromosome was completely absent from somatic cells, was largely unstable in maternal germ cells, and stable occurrence was observed only in male germ cells (Imai 1974). The mechanism of elimination has been described in detail in *Ae. speltoides*. It is a strictly controlled process taking place during early embryonic development, when chromatid nondisjunction occurs in the proto-root cells. B laggards are excluded from daughter nuclei and the B chromosome is eliminated via micronucleation (Ruban *et al.* 2020).

The behavior of B chromosomes in meiosis is variable not only among species, but can be different also within one species. The only general rule is that there is no pairing between Bs and As. In some species Bs can pair and form chiasms with each other, whereas in others Bs do not form bivalents despite their apparent morphological similarity. These cases could be explained by structural polymorphism, high heterochromatin content, or insufficient length for successful pairing (Houben *et al.* 2013). In metaphase I, B chromosomes can form univalents, bivalents, or multivalents. The destiny of univalent is species- and condition-dependent. Although univalents are lost during the first meiosis or in the second anaphase in most species (as in *Ae. speltoides*, Mendelson and Zohary 1972), in some cases the univalents behave regularly and are maintained. For instance, in *Plantago serraria*, B univalent resides near one of the poles of the dividing spindle and is included in one of the daughter nuclei in telophase. During anaphase II it divides normally (Frost 1959). Regular univalent transmission is also maintained in *Myrmeleotettix maculatus*,

Allium cernuum, *Zea mays* and others (John and Hewitt 1965; Grun 1959; Carlson 1986).

2.1.6 Origin of B chromosomes

Based on the similarity of some sequences of B and A chromosomes, it has been hypothesized that B chromosomes are derived from autosomes. Indeed, many studies have shown that some Bs originated from autosomes of the host species, e.g. the B chromosome in *Zea mays* (Peacock *et al.* 1981, Alfenito and Birchler 1993, Stark *et al.* 1996) (Fig. 4), *Secale cereale* (Sandery *et al.* 1990, Blunden 1993, Houben *et al.* 1996) (Fig. 4), *Crepis capillaris* (Jamilena *et al.* 1994, 1995) or *Locusta migratoria* (Teruel *et al.* 2010).

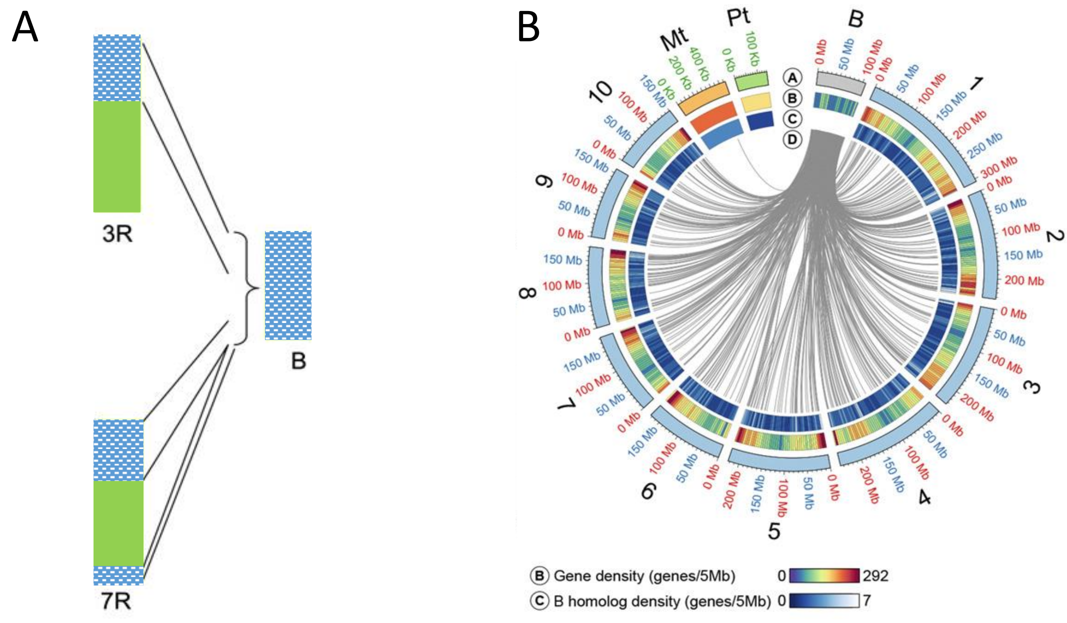

Fig. 4 Multichromosomal origin of B chromosome in rye (**A**) and maize (**B**). **A** – The B chromosome of rye shows extended similarity to regions of chromosomes 3R and 7R (modified from Martis *et al.* 2012). **B** – Comparison of the maize B chromosome and the B73 A chromosomal complement. Track A: chromosomes and organellar DNA; magnification of the mitochondrial and chloroplastic DNA by 200x and 500x, respectively. Track B: gene density in 5-Mb windows. Track C: density of paralogs of B chromosome genes in 5-Mb windows. Track D: link between B and A paralogs with mitochondrial gene links colored in orange (adopted from Blavet *et al.* 2021).

Sex chromosomes have also been suggested as the source of genetic material for B chromosome, as organisms are known to deal more easily with redundant sex chromosomes than with redundant autosomes (Hewitt 1974). The analysis of repetitive DNA sequences of the B chromosome of *Eyprepocnemis plorans* has revealed that it originated from the X chromosome (López-León *et al.* 1994). Similarly, based on the sequence analysis, the W sex chromosome of *Leiopelma hochstetteri* seems to be the ancestor of the B chromosome in this frog (Sharbel *et al.* 1998).

The B chromosome is not always derived from the chromosomes of the host genome, as interspecies origin is also possible. Sapre and Deshpande (1987) showed that the B chromosome can arise spontaneously during an interspecific cross. The hybridization of *Coix aquatica* and *C. gigantea* produced a univalent that showed B chromosome features. Through the phylogenetic analysis of the retrotransposon *NATE*, interspecies origin was also confirmed in the PSR chromosome of *N. vitripennis*. The B chromosome was found to be derived from a related genus *Trichomalopsis* (McAllister and Werren 1997). Further, the B chromosome of the frog *Hypsiboas albopunctatus* also has a possible interspecies origin (Gruber *et al.* 2014).

Hybridization-derived B chromosomes have the advantage of being immediately isolated from the A chromosomes of the host genome. Intraspecifically evolving B chromosomes, however, have to employ mechanism that prevents their recombination with source elements during meiosis. Such process, which is crucial for the initiation of independent B chromosome development, requires rapid structural rearrangement of the emerging B chromosome (Camacho *et al.* 2000).

How this process works is not yet known, but several hypotheses have been presented. The mechanism of rapid heterochromatinization and meiotic isolation of a *de novo* emerged DNA element has been observed and described in the grasshopper *E. plorans*. In this species, extra chromosomes in the germ line have been discovered, which are not transmitted to the next generation and which are created *de novo* from autosomes (Talavera *et al.* 1990). Once a recombination barrier is created between the B chromosome and its parental sequences, B chromosome follows its own evolutionary path and due to the absence of the selection pressure, the sequences in B chromosome can diverge rapidly. Accumulating mutations,

unequal crossing-over, and DNA polymerase slippage during replication are mechanisms which may contribute to its evolution.

B chromosomes can "capture" sequences of various origin, which can then spread within the B chromosome and thus contribute to its development. These sequences include mobile elements, which are abundant on B chromosomes. Mobile elements probably played an important role in the formation of Bs. Their activity may have led to the amplification of certain sequences, which could significantly accelerate divergence of the B chromosome from As (Cheng and Lin 2004; Lamb *et al.* 2007; Carchilan *et al.* 2009). The participation of transposons could also explain how various DNA sequences are transferred from the A chromosome to the B chromosome. The Bs of some species appear to be a mosaic of sequences derived from different A chromosomes as well as organellar DNA, as has been shown e.g. for maize or rye (Alfenito and Birchler 1993, Martis *et al.* 2012). In *B. dichromosomatica*, insertion of chloroplast DNA into the B-specific repeat is also attributed to transposon activity (Franks *et al.* 1996).

Ribosomal DNA is another element that could play a role in B chromosome origin. Ribosomal genes occur as clusters of repetitive units and form a secondary constriction in the nucleolus organizer region (NOR). The NOR region is believed to be prone to breakage, and if such a break occurs, the subsequent fragment could serve as the basis for the formation of the B chromosome. Houben *et al.* (1997) hypothesized that such event led to the emergence of Bs in *Brachycome dichromosomatica*, which has different types of B chromosomes with different rDNA representation. As a result of massive amplification of 5S rDNA, the B chromosome of *Plantago lagopus* have formed (Dhar *et al.* 2002, Kumke *et al.* 2016).

The origin of B chromosomes is different for different species. For ancient Bs, it is difficult to trace their origin. An example of such ancient B chromosome is the B of maize, wherein intraspecific origin is assumed (Stark *et al.* 1996), but a particular source element is unlikely to be traced because of the degeneration of the original sequences. As a study by Blavet *et al.* (2021) has shown, homologues of B-gene sequences were dispersed on all A chromosomes, but no region of synteny could be found. Based on the divergence time of A- and B-gene copies, the B chromosome could be up to 12 million years old. On the other hand, the B chromosome of rye

originated relatively recently. It has been estimated to be 1.1 to 1.3 million of years old (Martis *et al.* 2012). The genus *Secale* was arising at the same period, so it is possible that the B chromosome originated as a result of the chromosome rearrangement event. Houben *et al.* (2014) proposed a model of the rye B formation, which assumes initial whole genome or segmental duplication followed by translocations and insertions. The process required rapid accumulation of structural changes to prevent meiotic pairing with As and to provide the emergence of the drive mechanism. This was mediated through rapidly evolving repeats and the absence of selection pressure. As a result, B chromosome could begin to develop independently and the gene sequences were probably later suppressed and degenerated. The evolutionary model of rye B chromosome is shown in the Fig. 5. Unlike for maize Bs, the spread of transposon (and other sequences) can be relatively easily traced during the development of rye B chromosome (Klemme *et al.* 2013).

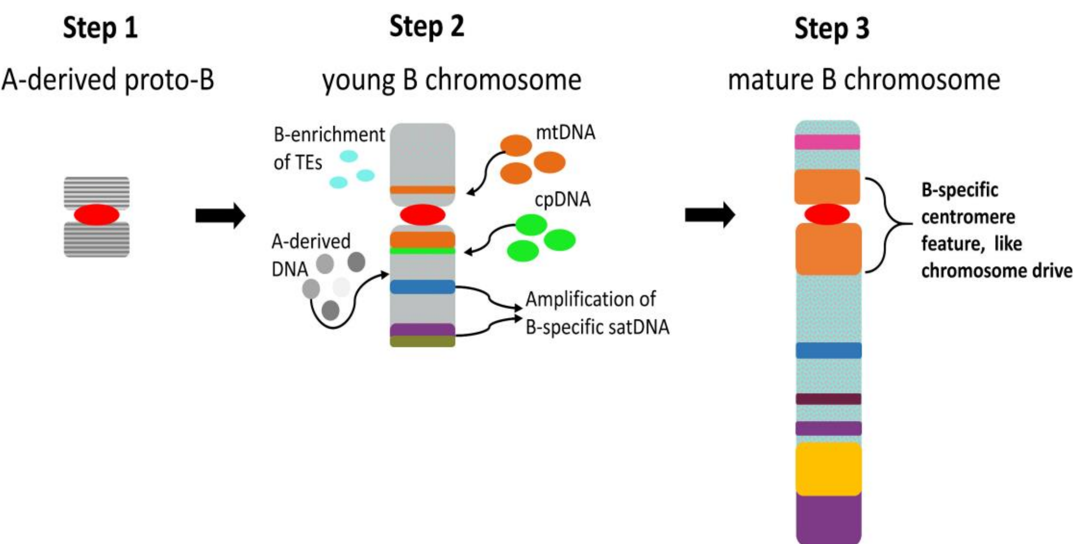

Fig. 5. Model for the repeat accumulation and evolution of plant B chromosomes (demonstrated on the evolution of rye B chromosome). (**1**) Proto-B chromosome derived as a result of multiple translocations and duplications of A chromosome fragments (harboring a functional centromere). (**2**) Gene erosion/silencing followed by a restriction of meiotic recombination triggers B-specific repeat accumulation. (**3**) Mature Bs could achieve a high degree of B-specific repeats, an efficient drive mechanism and a tolerable impact on the fitness of the host organism. mtDNA - mitochondrial DNA; cpDNA - chloroplast DNA; satDNA - satellite DNA (adopted from Marques *et al.* 2018).

2.1.7 Evolutionary models of B chromosomes

The ability of the B chromosome to persist in a population depends on two attributes: the extent to which it adversely affects the host's fitness and the extent of its transmission. Two evolutionary models for B chromosomes have been proposed: heterotic and parasitic (White 1973, Jones 1985). These models assume that the frequency of Bs in the population is in equilibrium established by two counteracting forces. According to the heterotic model, B chromosome does not have an accumulation mechanism, but provides a selection advantage to its host when present at low numbers, while at higher numbers the effect of Bs is detrimental (Camacho *et al.* 2000). In the parasitic model (Fig. 6), which appears to better describe the behavior of most B chromosomes, the balance is created by the rate of the B chromosome accumulation on one hand and its adverse effect on the host on the other. In the early stages of its existence, parasitic B chromosome has a detrimental effect on the host fitness, but this is outweighed by the high rate of accumulation. Its negative effect on plant fitness triggers counteracting processes in the host genome, which are directed to neutralization of the B chromosome. This can lead to a loss of accumulation mechanisms and B becomes a near-neutral. Such B chromosome is doomed to extinction, unless the accumulation mechanism is restored or unless it begins to provide some benefit to the host (Camacho *et al.* 2000).

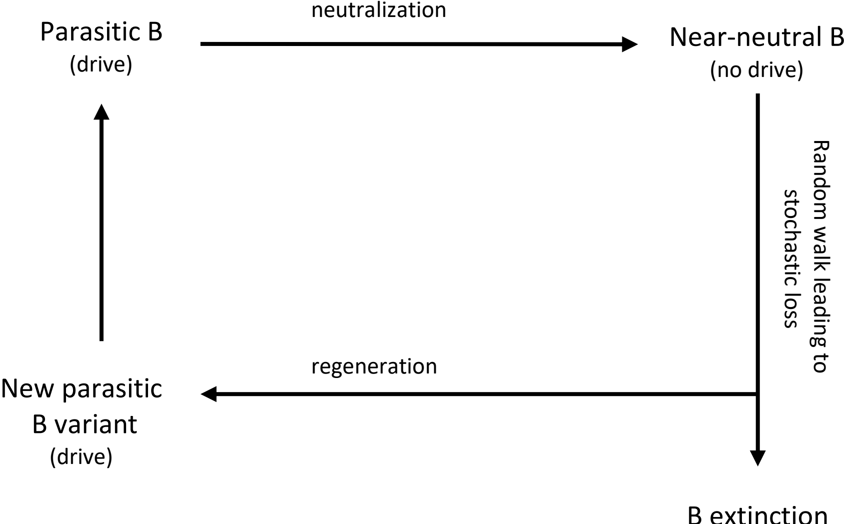

Fig. 6 Long term evolution of parasitic Bs (according to Camacho *et al.* 2000). The evolutionary cycle of the B chromosome begins with the parasitic stage, when B chromosome has a drive mechanism that allows it to spread in the population. At this stage, the B chromosome is harmful to its host. In the neutralization stage, the appropriate genes on the A chromosomes evolve, which suppress the B chromosome drive. At the same time, the ability of the A chromosomes to tolerate the effects of the B chromosome increases, and/or less harmful B variants arise. Near-neutral B chromosome has lost its accumulation mechanism and has only an insignificant effect on its carrier (at least at low numbers). The loss of the accumulation mechanism leads to a random walk towards stochastic loss. If the B chromosome does not restore its drive during a random walk, it will be eventually lost. If the drive is restored, the whole cycle begins again.

2.1.8 Importance and potential use of B chromosomes

B chromosomes are used to study centromere function and nondisjunction (Jin *et al.* 2005, Lamb *et al.* 2006, Han *et al.* 2007, Banaei-Moghaddam *et al.* 2012, Blavet *et al.* 2021). The practical use of B chromosomes is currently limited, but this may change in the future, especially in the field of plant artificial chromosomes.

a) Gene mapping.

So far, the maize B chromosome has found practical use in gene mapping. Already in the first half of the last century, B-A translocation lines were developed using X-rays, and these lines were used for the demonstration of nondisjunction of the B chromosome and preferential fertilisation in the second

pollen mitosis (Roman 1947). This way of the production of translocation lines later turned out to be useful for the mapping of genes in maize. The mapping method is based on the nondisjunction of the B^A chromosome in the second pollen mitosis. Homozygous recessive females are crossed with males carrying the tested dominant translocations. If a recessive gene of an unknown locus is crossed as a female against a dominant translocation series of males, the recessive phenotype can be unmasked by one of the translocation (when a sperm cell carrying B^A fertilizes the nucleus of the germ sac instead of the egg) and thus the locus can be mapped to the appropriate arm (Jones and Ruban 2019).

b) Plant artificial chromosomes.

B chromosomes could be used as a basis for the design of artificial plant chromosomes also called minichromosomes, through which the required genes or even entire metabolic pathways could be introduced into plants. This method could overcome the difficulties associated with introducing the transgene into the nucleus by conventional methods (Jones 2012). One of the approaches used for the production of minichromosomes is the telomere mediated truncation, which is based on the replacement of a part of the chromosomal arm by a desired insert (Yu *et al.* 2006). The advantage of B chromosome over As is that it is not essential and therefore its truncation does not have a detrimental effect on the plant. Another advantage is the absence of pairing with A chromosomes, so that multiple genes could be introduced in a single insert without disrupting the function of endogenous genes. In addition, it is a subject to accumulation mechanism via nondisjunction (Cody *et al.* 2015). The successful formation of a minichromosome from the maize B chromosome and its transmission to the next generation has already been achieved by Yu *et al.* (2007).

c) Chiasmata frequency and chromosome pairing.

Bs can influence the frequency and distribution of chiasmata in meiosis. This effect has been described in some plant and animal species, where the

presence of B could decrease or increase the frequency of chiasmata (Jones and Ruban 2019). Moreover, it has been found that B can modify the chromosome pairing in hybrid plants. Evans and Macefield (1972, 1973) discovered the diploidization effect of the B chromosome on tetraploid hybrids created by crossing *Lolium perenne* and *L. temulentum*. While hybrids without B chromosome showed both homologous and homoeologous pairing in pollen mother cells, in B-carrying hybrids only homologous pairing occurred. These features of B chromosomes could find use in crop improvement.

2.2 Genus *Sorghum* and B chromosomes

2.2.1 Characterization of the genus *Sorghum*

Sorghum is a genus of monocot flowering plants in the grass family Poaceae, subfamily Panicoideae and the tribe Andropogoneae. The genus encompasses about 23 annual and perennial species and a number of subspecies and races arising from hybridization. Based on their morphological traits, these species are classified into five subgenera: *Sorghum*, *Stiposorghum*, *Parasorghum*, *Chaetosorghum and Heterosorghum* (Garber, 1950).

Subgenus *Sorghum* includes cultivated sorghum (*Sorghum bicolor*) and its wild relatives with chromosome numbers 2n=2x=20 in diploids and 2n=4x=40 in tetraploids. Representatives of this subgenus originated in Africa and Asia (Lazarides *et al.* 1991). *S. bicolor* is one of the five most important cereal grains used for human consumption, as it being used as food for humans, feed for animals and also for ethanol production (Food and Agriculture Organization of the United Nations, 2021). The beginnings of its cultivation are located in Africa, but today it is grown worldwide, in addition to Africa mainly in America and Asia. *S. halepense*, commonly known as Johnson grass, is native to the Mediterranean region of Europe and Africa, and possibly to Asia Minor. It is grown for forage, similarly as *S. propinquum*, and was introduced to many areas including North America, Brazil, Argentina, Northern Australia and others (Lazarides *et al.* 1991, Warwick and Black 1983). However, at the

same time, *S. halepense* is also considered as one of ten most aggressive weeds in the world (Holm et at. 1977).

Subgenus *Stiposorghum*, encompasses 10 species occurring in Northern Australia. Their chromosome numbers range from 2n=2x=10, 20, 30 to 40. Subgenus *Parasorghum*, with the same range of chromosome numbers, includes seven species growing in Australia, Central America, Africa, and Asia (Lazarides *et al.* 1991). The last two subgenera, both represented by a single species, are *Chaetosorghum* and *Heterosorghum* with 2n=2x=40. *S. macrospermum*, as the representative of *Chaetosorghum*, occurs endemically in Australia's Northern Territory, and *S. laxiflorum*, the representative of *Heterosorghum*, is native to Northern Australia and Papua New Guinea (Lazarides *et al.* 1991). Wild sorghum species are considered a potential source of new genetic material for crop improvement such as resistance to biotic and abiotic stresses or increase of nutritional values (Kameswara *et al.* 2003, Abdelhalim *et al.* 2019).

2.2.2 B chromosomes in *Sorghum*

B chromosomes were reported in five sorghum species: *S. purpureosericeum*, *S. halepense*, *S. stipoideum*, *S. nitidum* and *S. bicolor* ssp. *verticilliflorum*. Typically, B chromosomes in this genus are stably present only in meristems leading to the development of inflorescences and are absent in roots and most of the shoot tissues. B chromosomes in sorghums are morphologically variable, and different morphotypes have been reported even within individual species. B chromosomes have been characterized by several cytological works during the last century, however, molecular data are not available yet.

B chromosomes in *S. purpureosericeum* were first reported by Janaki-Amal, who observed one to six Bs in one cell (1940). Later, Darlington and Thomas observed three types of B chromosomes (long, medium and short), which did not pair with each other. B chromosome was an isochromosome and it often formed a ring during meiosis. All the observed types of chromosomes were heterochromatic. Meiotic behavior in 2B plants was almost regular, while in 1B plants the B chromosome passed undivided through the first meiotic division and split in the second division,

resulting in four microspores - two with B present and two with B absent. In B-carrying plants, there is an interesting phenomenon of extra divisions between the first and second pollen mitosis, which could indicate the existence of an accumulation mechanism that is unique across the plant kingdom (Darlington and Thomas 1941).

In *S. nitidum*, Bs of a size equal to As were observed by Raman and Krishnaswami (1960). Bs were found only in diploid plants and when two Bs were present, their behavior in meiosis was regular. B chromosomes in *S. nitidum* were further studied by Wu and Pi (1975) and Wu (1978). In their study, plants with one B chromosome were examined. B chromosome was completely heterochromatic and much shorter than any of As. It was an isochromosome showing inter-arm pairing during meiosis. During meiotic division, B chromosome behaved as a standard univalent (Wu and Pi 1975). The hypothesis of the B chromosome origin was postulated by the authors. They outlined the formation of an extra univalent after interspecies hybridization, which later succumbed to spontaneous fragmentation and thus the loss of one of the arms. They assume that the B chromosome could have arisen from the short arm of nucleolar organizing chromosome based on their similar length and heterochromatic nature (Wu and Pi 1975).

In *S. halepense*, four to six Bs were found in diploid plants. In both *S. nitidum* and *S. halepense*, tetraploids do not appear to host B chromosomes. Further, in *S. halepense*, three kinds of Bs were observed, two of them showed partial relationship. During meiosis in pollen mother cells, B-bivalents showed aberrant behavior as delayed disjunction, nondisjunction and elimination, and their number was instable. Fertility of B-carrying plants was significantly reduced (Raman *et al.* 1964, 1965 and 1976).

In *S. stipoideum*, B chromosome was found by Wu (1992). It was completely euchromatic and clearly shorter than any of the As. In meiosis, when two Bs were present, they formed bivalents and their behavior was nearly regular. When one B chromosome was present, inter-arm pairing was observable, which implies its isochromosome nature. In most of the cases, univalents divided precociously during anaphase I, eventually they moved to the pole undivided. Micronuclei were noticeable, which must represent eliminated Bs. In this species, mosaic occurrence

of B chromosome between spikes was reported. The variability in B chromosome number was noticed between spikelets (Wu 1992).

In *S. bicolor*, ssp. *verticilliflorum*, “an additional pair of fragments” was observed during male meiosis by Huskins and Smith (1932). They described it as much smaller than any of the As. This additional pair was attached to A-bivalent. Unfortunately, there is no other study describing the presence of the B chromosome in this species, therefore it is possible that the supposed supernumerary chromosomes were mere chromosomal fragments resulting form chromosome breakage.

2.3 Methods of B chromosome study

2.3.1 Cytology and cytogenetics

Cytology and cytogenetics play an indispensable role in B chromosome research. So far, in many species, cytological and cytogenetic studies are the only findings on their Bs. As mentioned above, the first B chromosome was discovered more than a century ago, when Wilson (1907) noticed an extra chromosome in the insect *Metapodius*. This moment ushered in the era of B chromosome research, which relied on cytological/cytogenetic techniques until molecular methods were developed. Currently, cytogenetic and molecular methods complement each other and this synergy enabled the development of other methods such as fluorescence *in situ* hybridization (FISH). Cytology and cytogenetics enable the detection of B chromosomes, study of mitotic and meiotic behavior of chromosomes in the presence of Bs and the behavior of B chromosome itself including the nondisjunction process.

Fluorescence *in situ* hybridization is a cytogenetic method that allows to visualize the position of specific DNA sequences on chromosomes or nuclei. The technique uses fluorescent probes, short labeled DNA or RNA fragments that hybridize to a complementary DNA sequence of a target (Levski and Singer 2003). The method has gradually achieved wide popularity and has been improved considerably. Today, FISH is used in sophisticated assays such as 3-D analysis of genome organization (Schwarz-Finsterle *et al.* 2005; Simonis *et al.* 2006), quantitative analysis

(Martens *et al.* 1998), expression analysis (Levski *et al.* 2002; Dirks and Raap, 1995), the study of chromosomal territories (Branco and Pombo 2006), high-resolution mapping of genes and chromosomal regions (Heiskanen *et al.* 1995) and many others.

In plants, repetitive DNA sequences and large-insert genomic DNA clones, have been most commonly used as a probe (Jiang and Gill 1994, 2006). Further, satellite repeats are also very popular target for the generation of FISH probes as they are highly abundant in centromeric and subtelomeric regions and are often used for chromosome identification (Jiang 2019). Alternatively, multicopy genes such as histone genes or rDNA are another frequently used source for probe design (Phillips and Reed 1996). More recently, synthetic oligo probes have become popular (Wallace *et al.* 1981, Cuadrado and Schwarzacher 1998; Schmidt and Heslop-Harrison 1996).

FISH is an essential method for the identification of B chromosomes in a karyotype. In combination with molecular methods, it helps to reveal the structure, function and evolution of B chromosomes. Knowledge of B-specific sequence is a key to the recognition of B chromosomes by FISH. These repeats have been used for the identification of B chromosomes in many plant and animal species, such as in rye (Blunden *et al.* 1993, Sandery *et al.* 1990), maize (Stark *et al.* 1996, Alfenito and Birchler 1993), *Festuca pratensis* (Ebrahimzadegan *et al.* 2019), *Brachycome dichromosomatica* (Leach *et al.* 1995), *Drosophila subsilvestris* (Gutknecht *et al.* 1995) or *Nasonia vitripennis* (Eickbush *et al.* 1992).

FISH has been an indispensable tool in many studies elucidating the composition and origin of Bs. An example of such study is the work of Lopéz-León *et al.* (1994). They focused on B2 type chromosome of *Eyprepocnemis plorans*. The study revealed a predominant content of tandem repeats and rDNA and suggested that B is derived from the X chromosome. FISH in combination with GISH (genomic *in situ* hybridization) was also used by Wilkes *et al.* (1995) for the characterization of the structure of the B chromosome in rye. They have determined the overall similarity between As and B, which indicated the intraspecific origin of the B. Vicari *et al.* (2011) used chromosome painting and FISH to analyze the composition of sequences on the A and B chromosomes in the fish *Astyanax scabripinnis.* The study supported the hypothesis of the intraspecific origin of the B and provided evidence of the

evolutionary enrichment of certain repetitive sequences on the B chromosome. Two naturally occurring variants of the maize B chromosome were also characterized by FISH in the work of Cheng *et al.* (2016). Their organization and origin from the standard B chromosome were determined.

FISH also significantly contributed to research of B chromosome elimination. Cabrero *et al.* (2017) investigated FISH elimination of B chromosome during spermiogenesis in grashoppers *Eumigus monticola* and *Eyprepocnemis plorans*. They observed the presence of microspermatides containing B chromosomes, which correlated with a significant decrease in the number of B-carying spermatids. They evaluated this process as a defense mechanism preventing the accumulation of the parasitic B chromosome. In *Ae. speltoides*, B chromosome elimination was studied on embryo tissue sections, where cells with B chromosome preserved/depleted were distinguished using B-specific high-copy repeats. It was found that the elimination of Bs in roots takes place in the early embryonic phase through the process of micronucleation (Ruban *et al.* 2020).

FISH with B-specific probes has been used for the elucidation of the effects of *r-X1* deletion on maize B chromosome. This deletion is an intercalary chromosome deficiency located on the long arm of chromosome 10 (Lin 1987). The ability of *r-X1* to induce B-chromosome nondisjunction during the first pollen mitosis was reported (Tseng *et al.* 2018). Recently, *r-X1* deletion was found to be able to induce terminal deletions of B chromosome in maize (Huang *et al.* 2021). Further, FISH enabled the localization of cDNA sequences of B-associated transcripts on Bs, like in rye (Carchilan *et al.* 2009), maize (Lin *et al.* 2014, Huang *et al.* 2016) or the cichlid fish *Astatotilapia latifasciata* (Ramos *et al.* 2017).

Employment of 3-D FISH, i.e. hybridization of DNA probes on three-dimensionally preserved cells, made it possible to study the functional organization within the nuclei carrying B chromosomes. Karamysheva *et al.* (2017) investigated the 3-D organization of fibroblasts and spermatocyte nuclei carrying B chromosomes in *Apodemus peninsulae*. Using laser scanning microscopy and 3-D FISH, colocalization of B chromosomes with constitutive heterochromatin of A chromosomes and non-random distribution of Bs in the spermatocyte nucleus were detected. Another work revealing spatial organization of B chromosome was

performed by Schemczssen-Graeff *et al.* (2020) on nuclei of the fish *Astyanax scabripinnis*.

2.3.2 Sequencing methods

The first widely used sequencing method was invented by Frederick Sanger in 1977 (Sanger *et al.* 1977). Later, Sanger sequencing was automated and sequencers were able to read about 1000 bp per reaction. The method was well applicable for smaller genomes, but sequencing of the large genomes remained a challenge. Sanger's method dominated until the beginning of the 21st century, when its era culminated in the successful sequencing of the human genome in the frame of the Human Genome Project (Collins *et al.* 2003). In the following twenty years, sequencing methods have made impressive progress and enabled a wide range of applications. The use of massively parallel analysis has reduced labor and significantly reduced costs (Fig. 7). Current sequencing methods are very fast and achieve high accuracy. These aspects made sequencing a standard and affordable method used in molecular biology.

Massively parallel sequencing (or next generation sequencing - NGS), entering the market in 2005, has made a breakthrough in sequencing technology. Sequencing has become available to the wider scientific community. While sequencing the human genome took over ten years and cost about $ 3 billion, after the advent of NGS, whole genome sequences began to be generated in a fraction of the time and cost. Currently, Illumina is the most used sequencing platform of the NGS. Sequencers of various parameters are available, such as Illumina MiSeq for small genome sequencing, or NovaSeq 6000, which is capable of generating up to 20 billion reads in about 2 days and is suited for whole-genome sequencing (Illumina.com).

However, NGS methods have also some limitations. Short read length and the need for amplification steps are aspects that can lead to sequencing or assembly errors. In order to overcome these limitations, new sequencing technologies such as Oxford Nanopore (Eisenstein 2012) or PacBio (Eid *et al.* 2009) have been developed more recently. These approaches use a single template molecule and generate reads of tens to hundreds of kilobases.

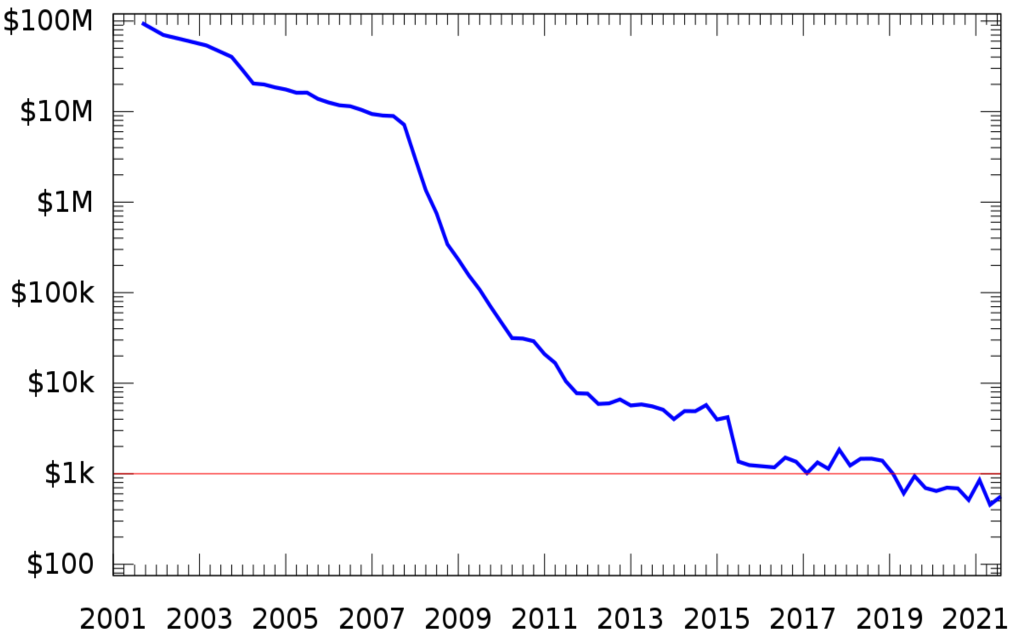

Fig. 7 The development of the sequencing costs of the human genome over the last 20 years (adopted from wikiwand.com)

Although sequencing technologies are constantly developing and more sophisticated techniques are emerging to cover new areas of research, B chromosome sequencing still presents certain challenges. Currently, two approaches are used to sequence B chromosomes (Ruban *et al.* 2017): the sequencing of isolated Bs and the comparative sequencing of the whole genome of individuals with B chromosome present and absent. The first approach was used for instance for sequencing of B chromosome of rye, the green anole lizard *Anolis carolinensis* or the fish *Astyanax paranae* (Martis *et al.* 2012, Kichigin *et al.* 2018, Silva *et al.* 2014), the latter was involved, for example, in sequencing of Bs in the grasshopers *Eumigus monticola* or *Eyprepocnemis plorans* (Ruiz-Ruano *et al.* 2016, Navarro-Domínguez *et al.* 2017). Alternatively, a combination of both approaches can be used, like in the sequencing of B chromosome of *Locusta migratoria* (Ruiz-Ruano *et al.* 2018).

In the first approach, the B chromosome must be separated from the A chromosomes by microdissection or sorting by flow-cytometry. This procedure generates a small amount of DNA, so it is necessary to amplify the DNA after isolation,

which mainly concerns microdissected chromosomes. The problem may be the presence of contaminating DNA derived from other chromosomes or PCR amplification bias (Ruban *et al.* 2017). The second approach is to sequence the genome of an individual with and without Bs. NGS then results in two datasets, which are compared and the sequences belonging to the B chromosome are identified. Three different methods can be used for comparative analysis: similarity-based clustering, coverage ratio analysis and *k*-mer frequency ratio analysis, the latter two requiring a reference sequence (Ruban *et al.* 2017). A combination of sequencing approaches in conjunction with other complementary methods has been shown to be a suitable solution for B chromosome sequencing. Long reads generated by PacBio and Oxford Nanopore, BioNano optical mapping or chromatin conformation capture (3C) libraries represent an improvement, through which it is possible to achieve an accurate assembly of B chromosomes (Ahmad and Martins, 2019). Such a combination of methods has been used to generate the reference sequence of the maize B chromosome (Blavet *et al.* 2021).

Numerous studies utilizing low coverage sequencing were published. Low coverage sequencing is not applicable to create a high-quality assembly, but it is sufficient for preliminary genomic characterization with reduced cost. It can serve as a basis for the analysis of B-enriched sequences like mobile elements or satellite repeats and for the extraction of molecular markers - in the case of B chromosomes, it is most often used to identify B-specific repeats. Such work was performed on *Festuca pratensis*, where B-bearing and B-lacking individuals were sequenced by Illumina and subsequently repeat analysis was conducted using RepeatExplorer. As a result, two B-specific repeats were revealed (Ebrahimzadegan *et al.* 2019). A similar approach has been used in analyzes of B-enriched sequences of *P. lagopus*, rye, the grasshoppers *Eumigus monticola* or *Abracris flavolineata* and others (Kumke *et al.* 2016, Martis *et al.* 2012, Ruiz-Ruano *et al.* 2017, Milani *et al.* 2021).

In order to clarify the origin and evolution of B chromosome in *Astatotilapia latifasciata,* sequencing study was performed, which compared genomes with and without Bs (Valente *et al.* 2014). “Blocks” of B chromosome sequences were identified by coverage ratio analysis and validated by sequencing of microdissected B chromosome. It has been shown that the B chromosome contains sequences

duplicated from almost all autosomes including intact genes and at least three of these genes were transcriptionally active.

Sequencing with sufficient depth and data processing with advanced bioinformatics tools lead to a high-quality assembly that allows to generate a reference sequence. Due to the nature of B chromosomes, this has long been difficult to achieve. However, rapid progress in technology and bioinformatics have already borne fruit. The first annotated B-chromosome reference sequence was reported by Blavet *et al.* (2021). By combining the methods of Illumina sequencing, BioNano optical mapping and Hi-C scaffolding, 125.9 Mb of B chromosome sequence of maize was assembled representing 89% of the predicted 141-Mb size of the B chromosome. More than 700 genes have been identified, out of which at least 88 were transcriptionally active. Another complex sequencing study was performed on PSR chromosome of *Nasonia vitripennis*. Illumina, Oxford Nanopore and PacBio sequencing platforms were used to elucidate the detailed composition and expression of PSR. 120 PSR-specific scaffolds were identified, which made 9.2 Mb of the sequence. The largest part of PSR was made up of satellite repeats, the rest was represented by transposable elements, coding sequences and telomeric repeats (Benetta *et al.* 2020).

An earlier hypothesis on the heterochromatic nature and transcriptional inactivity of Bs has been challenged. Recent works show that many B chromosomes exhibit transcriptional activity, including transcription of protein-coding sequences or regulatory RNAs (Valente *et al.* 2014, Navarro-Domínguez *et al.* 2017, Ma *et al.* 2017, Blavet *et al.* 2021, Trifonov *et al.* 2013, Benetta *et al.* 2020). Transcriptomic studies have focused on the search for B-located genes and their expression levels, in order to determine their relationship to the transmission and maintenance of Bs and also interactions between Bs and host genome. For this purpose, comparative analysis of transcriptomes generated from B-carrying and B-lacking individuals is involved. A number of studies focusing on this topic has been published in recent years. For instance, Navarro-Domínguez *et al.* (2017) found 10 protein-coding genes on B chromosome of *E. plorans*, with some of them being intact and up-regulated in B-carrying individuals. Later, in the same species, 188 transcripts showing altered expression in B-carrying individuals were identified. These transcripts belonged to the

categories of genes affecting microtubule movement, cell division and cell metabolism, gene silencing and protein stabilization (Navarro-Domínguez *et al.* 2019). In maize and rye, there is evidence of interaction between A and B chromosomes at transcriptional level. Huang *et al.* (2016) showed that in B-carrying maize plants, expression of 130 A-located genes involved in cell metabolism and nucleotide binding were altered. The rye transcriptome study revealed almost two thousand transcripts in generative tissues and more than one thousand transcripts in vegetative tissues which were derived from B chromosome. Moreover, a functional B-encoded Argonaute-like protein was identified in rye (Ma *et al.* 2017). Boudichevskaya *et al.* (2020) tried to shed light on the process of B chromosome elimination in embryos of *Ae. speltoides*. They have analyzed the transcriptome of the central meristematic region, where the elimination of B chromosomes occurs. Almost 14 600 differentially expressed and 341 B-specific transcript isoforms were identified during B-elimination. Benneta *et al.* (2020) performed transcriptome sequencing of PSR-carrying testes of *Nasonia vitripennis* using Nanopore sequencing of full-length complementary DNA. Besides the identification of 68 PSR-specific transcripts, PSR-expressed gene responsible for the elimination of sperm-derived chromatin in embryos, named *haploidizer*, was successfully identified. Some gene sequences have shown high similarities to other insect species, such as the closely related *Trichomalopsis sarcophagae* (Benneta *et al.* 2020).

2.4 Molecular phylogenetics

Molecular phylogenetics is a branch of phylogenetics that analyzes evolutionary relationships between organisms by means of molecular markers. Based on the degree of differences in the observed marker, with the involvement of statistical approaches, it is possible to estimate the divergence rate between the studied organisms and thus assess their evolutionary relationship. An output of the phylogenetic analysis is a phylogenetic tree - phylogram or cladogram (Fig. 8).

Generally, little is known about the evolutionary history and phylogeny of B chromosomes. So far, some progress in composing a mosaic of their origin has been

made only in the most intensively studied species. However, there are many other species, including the genus *Sorghum*, where Bs have only been recorded and some cytogenetic characteristics were carried out. Some B chromosomes are known to have originated a long time ago and appear to be relatively conserved. An example is the B chromosome of maize, whose age has been estimated to be up to about 12 million years (Blavet *et al.* 2021).

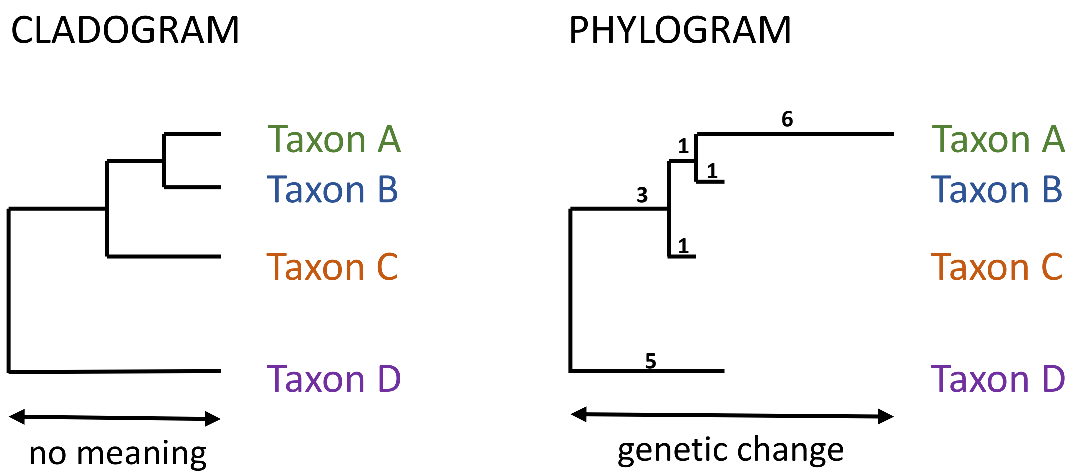

Fig. 8 Cladogram and phylogram. The result of molecular phylogenetic analysis is represented by a phylogenetic tree - phylogram or cladogram. In the cladogram, the length of the branches does not take into account the number of mutations that have taken place, so it has no meaning for this type of display. In a phylogram, the length of the branches corresponds to the number of character changes, and thus represents time since divergence (according to C. B. Stewart, NHGRI-lecture, 12/5/00).

On the other hand, there are species where the evolution of B chromosome is very dynamic, which is manifested by the emergence of new B chromosome variants in a relatively short periods of time, such as in *E. plorans* (López-León *et al.* 1993). The genus *Sorghum* is part of the Andropogoneae tribe, encompassing about 1200 species in 98 genera (Soreng *et al.* 2017). Species hosting B chromosomes are relatively abundant within this tribe. *Sorghum* comprises about 23 species, five of which carry the B chromosome. To elucidate the evolution of the B chromosome in a given taxonomic group, it is necessary to define the phylogenetic relationships of individual members in the group. Subsequently, B chromosomes have to be subjected to analysis involving molecular data, which would enable the

determination of their relationships and origin. In the genus *Sorghum*, no molecular data have been obtained yet.

2.4.1 Molecular markers

Phylogenetics used to rely on morphological markers, which, however, can be affected by environmental conditions. Therefore, with the progress in molecular biology, morphological markers were replaced by molecular markers, which reflect variability at the DNA level. Molecular phylogenetics uses three types of markers - hybridization-based markers, PCR-based markers and sequence-based markers.

Hybridization markers were the first molecular markers used. Their advantage is that they do not require the knowledge of the genetic sequence. Restriction fragment length polymorphism (RFLP) was the first molecular marker to be applied to plants (Beckmann and Soller *et al.* 1983). These markers are codominant, reliable and highly polymorphic.

Amplified fragment length polymorphism (AFLP) is based on selective PCR amplification of fragments resulting from the restriction digestion of DNA (Janssen *et al.* 1996). The method is relatively reproducible, there is no need for previous sequence knowledge. Simple sequence repeats (SSRs) are another type of PCR-based markers. SSRs are also known as microsatellite markers or short tandem repeats (STR). These are short tandemly repeating oligonucleotide motifs present in both non-coding and coding regions. These markers are codominant, highly reproducible, have a high degree of polymorphism and specificity and allow automated analysis. On the other hand, sequence knowledge is required for the design of SSR primers (Rakoczy-Trojanowska and Bolibok 2004).

With the massive development of sequencing technologies, the amount of genetic data usable for phylogenetic studies has increased significantly, which enabled the development of sequence-based phylogenetic markers. This type of markers includes single nucleotide polymorphisms (SNPs), defined as a variation at a single position in a DNA sequence among individuals. SNPs are codominant and the analysis is suitable for automation (Gupta *et al.* 2008). Beside SNPs, expressed sequence tags (ESTs) are used in phylogeny. ESTs are obtained by sequencing of several hundred nucleotides from the 5' or 3' ends of mRNA (Idrees and Irshad 2014).

One of the most widely used markers in phylogenetics is ribosomal RNA (rRNA). It is present in all cellular organisms and contains both conserved and variable domains. Ribosomal RNA is encoded by two separate genetic loci - 5S rDNA and 45S rDNA. These genes occur as tandem repeats in hundreds to thousands of copies. The units are highly uniform, which is a result of concerted evolution (Alvarez and Wendel 2003). While 5S rDNA locus encodes 5S RNA, 45S rDNA locus is formed by regions coding for 18S, 5.8S and 25S RNA, separated by internal transcribed spacers ITS1 and ITS2. ITS region in angiosperms is 500-700 bp long (Baldwin *et al.* 1995), so it is easily amplifiable by PCR. ITS have such a degree of variability that is suitable for phylogenetic reconstruction at the level of species, genus or family (Baldwin *et al.* 1995). The composition of 45S rDNA is shown in Fig. 9.

Other widely used markers are chloroplast and mitochondrial sequences. Many studies have been based on the analysis of chloroplast genes (e.g. *rbcL*, *ndhF* or *matK* Chase *et al.* 1993, Olmstead *et al.* 2000, Johnson and Soltis 1994) or chloroplast intergenic spacers e.g. *trnL-trnF* or *trnH-psbA* (Gielly and Taberlet 1994). From mitochondrial markers, cytochrome-b or cytochrome oxidase I/II genes are the most frequently used ones (Patwardhan *et al.* 2014).

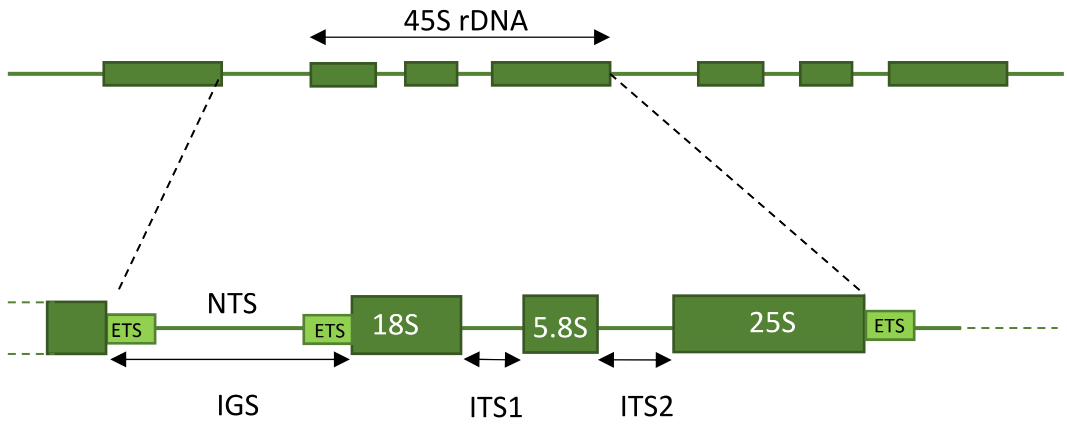

Fig. 9 45S rDNA locus organization in *Arabidopsis*. 45S rDNA is composed of 18S, 5.8S and 25S rRNA genes separated by internal transcribed spacers ITS1 and ITS2. 45S rDNA units are separated by intergenic spacers (IGS), which consist of two external transcribed spacers (ETS) and non-transcribed spacer (NTS) (according to Sáez-Vásquez and Delseny 2019).

2.4.2 Phylogenetic analysis

Phylogenetic analysis involves several consecutive steps (Patwardhan *et al.* 2014), schematically depicted in Fig. 10.

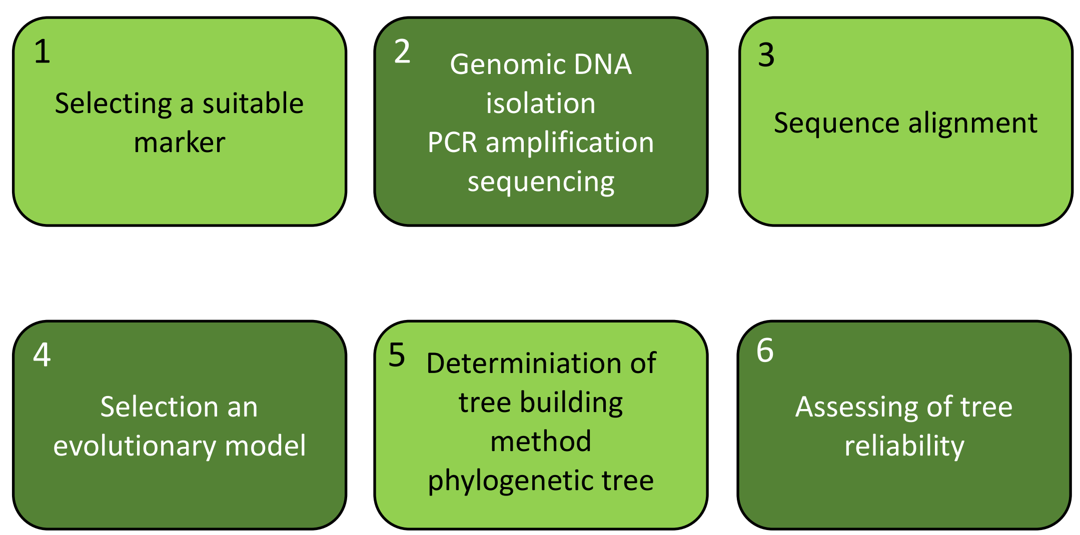

Fig. 10 Steps of phylogenetic analysis.

The selection of a marker depends on the evolutionary distance of the studied taxa. For the phylogenetic analysis of closely related taxa, DNA sequences are suitable due to their relative high variability. For evolutionarily distant groups of organisms, however, slowly evolving sequences are the most suitable. As in this case the use of nucleotide sequences can lead to difficulties in alignment due to their high divergence, protein sequences are employed more often for such type of analysis. (Ajawatanawong 2016). Phylogenetic reconstruction can be performed also based on multiple markers and phylogeny is then inferred from multiple concatenated sequences.

Alignment accuracy is critical for phylogenetic analysis, because it determines the positional correspondence in evolution (Xiong 2006). "Progressive sequence alignment", available in programs like MUSCLE, MAFFT or Clustal Omega, is frequently used (Edgar 2004, Katoh *et al.* 2002, Sievers *et al.* 2011). There are also other bioinformatics tools (such as GBLOCKS, SeqFIRE, or Rascal) that enable the

correction of alignment errors and removal of poorly aligned sections (Castresana 2000, Ajawatanawong 2012, Thompson *et al.* 2003).

To determine the evolutionary distance between sequences, it is necessary to choose a substitution model. The Jukes-Cantor model (Jukes and Cantor 1969) assumes that the substitution of purines is as likely as pyrimidines. This model is used for closely related sequences. The Kimura model (Kimura 1980, 1981) is more realistic, predicting that transitions occur more often than transversions.

There are two approaches to the construction of a phylogenetic tree: methods based on “distance” and methods based on “characters”. For distance-based methods, the first step is to calculate a matrix of pairwise distances between taxa from their sequence differences. It means that all sequence information is converted to single number per each pair of taxa, which is then analyzed using an algorithm for clustering (Ajawatanawong 2016). These methods include for example Unweighted Pair Group Method with Arithmetic Mean (UPGMA) (Sokal and Michener 1958) and Neighbor–joining (Saitou and Nei 1987). Advantage of this approach is its speed, which is achieved thanks to the reduction of all data into a simple distance matrix. Methods based on characters attempt to construct the phylogenetic tree directly, based on all the individual “sequence characters”. They take into account mutational events accumulated in the sequences and create more accurate trees (Patwardhan *et al.* 2014). These methods include Maximum parsimony, Maximum likelihood and Bayesian inference (Farris 1970, Fitch 1971, Felsenstein 1981, Rannala and Yang 1996).

Finally, the reliability of the tree must be verified using statistical methods. One of the most widely used methods is bootstrapping, which is based on the generation of pseudo-samples from the original data, which are then used for the construction of independent tree. The representation of nodes for individual bootstrap repetitions is evaluated. For sufficient robustness of the test, bootstrapping 500 – 1000 times should be performed. The bootstrap result is expressed as a bootstrap percentage or bootstrap p-value, which shows the degree of support of individual branches of the tree with respect to the input data (Soltis and Soltis 2003).

3 AIMS OF THE THESIS

The present work aims to contribute to the understanding of the evolution and function of B chromosomes in the genus *Sorghum*. Two main goals were:

I. Development of markers specific for the B chromosome of *Sorghum purpureosericeum*

II. Phylogenetic analysis of the genus *Sorghum* with respect to the hypothesis of the B chromosome origin in the genus

4 RESULTS

4.1 The B chromosome of *Sorghum purpureosericeum* reveals the first pieces of its sequence

Karafiátová M., Bednářová M., Said M., Čížková J., Holušová K., Blavet, N., & Bartoš J.

Journal of Experimental Botany 2021, 72(5): 1606-1616

doi: 10.1093/jxb/eraa548

IF: 5.9

Summary

To contribute to B chromosome research in *Sorghum*, specific markers for B chromosome of *Sorghum purpureosericeum* were developed. Employing flow cytometry, haploid pollen nuclei from B-positive and B-negative plants were collected. From extracted nuclear DNA sequencing libraries were prepared and subjected to Illumina sequencing on NovaSeq 6000. Based on raw data, repeat analysis using RepeatExplorer was performed. This analysis revealed nine putative B-specific clusters SpuCL115, SpuCL135, SpuCL144, SpuCL168, SpuCL169, SpuCL175, SpuCL189, SpuCL214 and SpuCL220. For each of these nine candidates, a set of primers was derived and their specificity was tested in PCR. Among primer pairs designed for short amplicons (up to 1300 bp), three were verified as a B-specific (Fig. 11). From the primer pairs designed for long amplicons (from 2 to 5 kbp), additional four were selected as B-specific (Fig. 12). As a result, seven specific markers were developed for reliable PCR detection of *S. purpureosericeum* B chromosome.

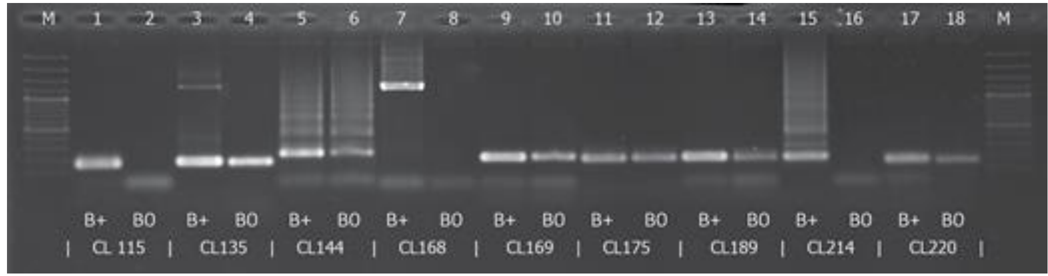

Fig. 11 Electrophoresis of short amplicons for nine putative B-specific clusters (SpuCL115, SpuCL135, SpuCL144, SpuCL168, SpuCL169, SpuCL175, SpuCL189, SpuCL214, and SpuCL220). For all clusters, primer specificity was tested in parallel on B+ DNA and B0 DNA. The presence of PCR products of SpuCL115, SpuCL168, and SpuCL214 exclusively in B+ samples indicates their B specificity. The remaining primer pairs showed non-specific amplifications. The M lane corresponds to the GeneRuler 100 bp Plus DNA Ladder (Karafiátová *et al.* 2021).

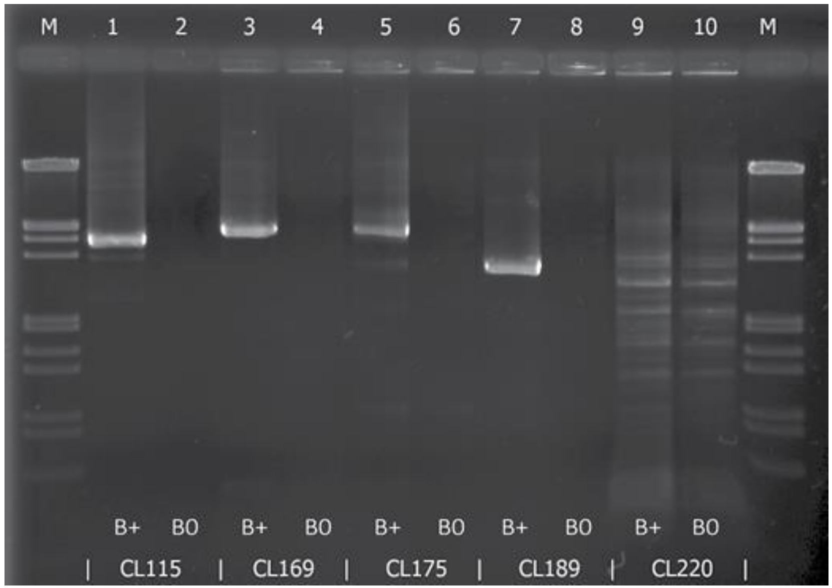

Fig. 12 Electrophoresis of long amplicons for five putative B-specific clusters (SpuCL115, SpuCL169, SpuCL175, SpuCL189, and SpuCL220). Primer specificity was tested in parallel on B+ DNA and B0 DNA. The presence of PCR products of SpuCL115, SpuCL169, SpuCL175, and SpuCL189 exclusively in B+ samples indicates their B specificity. The M lane corresponds to λ DNA digested with *Eco*RI and *Hin*dIII (Karafiátová *et al.* 2021).

Journal of Experimental Botany, Vol. 72, No. 5 pp. 1606–1616, 2021
doi:10.1093/jxb/eraa548 Advance Access Publication 20 November 2020
This paper is available online free of all access charges (see https://academic.oup.com/jxb/pages/openaccess for further details)

RESEARCH PAPER

The B chromosome of *Sorghum purpureosericeum* reveals the first pieces of its sequence

Miroslava Karafiátová*, Martina Bednářová*, Mahmoud Said, Jana Čížková, Kateřina Holušová, Nicolas Blavet and Jan Bartoš†

Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Hana for Biotechnological and Agricultural Research, Šlechtitelů 31, Olomouc, 779 00, Czech Republic

* These authors contributed equally to this work.
† Correspondence: bartos@ueb.cas.cz

Received 17 June 2020; Editorial decision 9 November 2020; Accepted 13 November 2020

Editor: Zoe Wilson, University of Nottingham, UK

Abstract

More than a century has passed since the B chromosomes were first discovered. Today we know much of their variability, morphology, and transmission to plant progeny. With the advent of modern technologies, B chromosome research has accelerated, and some of their persistent mysteries have since been uncovered. Building on this momentum, here we extend current knowledge of B chromosomes in *Sorghum purpureosericeum* to the sequence level. To do this, we estimated the B chromosome size at 421 Mb, sequenced DNA from flow-sorted haploid pollen nuclei of both B-positive (B+) and B-negative (B0) plants, and performed a repeat analysis on the Illumina raw sequence data. This analysis revealed nine putative B-specific clusters, which were then used to develop B chromosome-specific markers. Additionally, cluster SpuCL4 was identified and verified to be a centromeric repeat. We also uncovered two repetitive clusters (SpuCL168 and SpuCL115), which hybridized exclusively on the B chromosome under fluorescence *in situ* hybridization and can be considered as robust cytogenetic markers. Given that B chromosomes in *Sorghum* are rather unstable across all tissues, our findings could facilitate expedient identification of B+ plants and enable a wide range of studies to track this chromosome type *in situ*.

Keywords: B chromosomes, cytogenetics, flow cytometry, pollen nuclei, repeat analysis, *Sorghum purpureosericeum*

Introduction

B chromosomes are special genomic components with a fate of their own. They are dispensable parts of the genomes that escape Mendelian inheritance and mostly bring no advantage to the host but show strong negative effects on it when present in higher numbers. Nevertheless, there are a few species with fitness advantage of having a B chromosome reported, namely chives *Allium schoenoprasum* (Plowman and Bougourd, 1994) or fungal pathogens *Nectria haematococca* (Han *et al.*, 2001) and *Magnaporthe oryzae* (Peng *et al.*, 2019). The selfish nature of these supernumerary chromosomes secures their persistence in the population via the mechanism of non-disjunction, which goes directly against natural selection (Jones *et al.*, 2008). Much is now known about B chromosomes in general, yet we still have only vague knowledge of their significance to their carriers as well as of their emergence and evolution.

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Experimental Biology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

Despite the topic of supernumerary chromosomes becoming more conspicuous over the last three decades, its research scope across species remains largely uneven. Most of the assessed information still comes from a few favoured species of certain economic importance, making them preferred research targets. Being neither a staple food nor contributing extra nutrients in the human diet, wild sorghum plants have lingered on the periphery of science thus far and so our knowledge of these species is very poor. Nevertheless, knowing more about these wild sorghums may be beneficial given the rising demand for cultivated *Sorghum bicolor* in developing countries, particularly in tropical Africa, where *S. bicolor* is among the staples for rural people (Taylor, 2004; FAO, 1995). Wild sorghums represent a valuable tertiary genetic pool that is useful for breeding. Notably, perennial species are favoured by breeders aiming to produce and introduce perennial *S. bicolor* cultivars (Piper and Kulakow, 1994; Dweikat, 2005).

Though the exact number of species with reported supernumeraries is unknown (D'Ambrosio *et al.*, 2017), the total number is estimated at >2800 species widespread in all kingdoms. Of these, the vast majority have been reported from plants (Ahmad and Martins, 2019), with maize (Kuwada, 1925; Longley, 1927; Kato *et al.*, 2005; Carlson, 2007; Su *et al.*, 2018) and rye (Gotoh, 1924; Martis *et al.*, 2012; Banaei-Moghaddam *et al.*, 2019) receiving the most research attention.

The B chromosomes in the genus *Sorghum* remain almost entirely unexplored. Nearly all published reports focus exclusively on elementary karyological descriptions and B chromosome transmission via meiosis. Among the 22 *Sorghum* species, B chromosomes were reported in just five. First, they were discovered in *S. verticilliflorum* in 1934 (Huskins and Smith 1934) and later in *S. purpureosericeum* (Janaki-Ammal, 1939), *S. nitidum* (Raman and Krishnaswami, 1959), *S. halepense* (Raman *et al.*, 1964), and most recently in *S. stipoideum* (Wu, 1992). Apart from *S. verticilliflorum*, whose B chromosome was described as a fragment, the species' B chromosomes are of a standard size or are shorter when compared with chromosomes from the A complement. A maximum of six B chromosomes tolerated by the cell was recorded in *S. purpureosericeum* and *S. halepense* (Janaki-Ammal, 1940; Raman *et al.*, 1964).

Progress in acquiring more knowledge of accessory chromosomes in *Sorghum* spp. is challenging because of their high numerical instability. Generally, the transmission of B chromosomes during plant development and growth is regular in somatic cells. Nevertheless, there are exceptions to this rule. In *Crepis capillaris*, the numbers of B chromosomes in aerial parts differs from those in roots and claudicle leaves (Rutishauser and Röthlisberger, 1966). Somatic variation in B chromosome frequency was also found in both *Poa alpina* and *Agropyron cristatum*, for which B chromosomes are absent in adventitious roots yet preserved in primary roots (Müntzing and Nygren, 1955; Baenzinger 1962). Further, the elimination of B chromosomes from all roots is a specific phenomenon described in few plant species, namely *Aegilops* (Mochizuki, 1957; Mendelson and Zohary, 1972), *Haplopapus gracilis* (Östergen and Fröst, 1962), or *Poa timoleontis* (Nygren, 1957). Similarly, in *S. purpureosericeum*, the B chromosomes do not occur in roots, and are unstable in young shoots, ovaries, and tapetal cells. Their fixed distribution has been demonstrated only in fertile florets (Janaki-Ammal, 1940; Darlington *et al.*, 1941). Recently, the mechanism of tissue-specific B chromosome distribution was uncovered in *Aegilops speltoides* (Ruban *et al.*, 2020). This study reveals that B chromosome elimination from roots in goatgrass is a programmed and controlled process with onset only few days after pollination. It shows that elimination of the B chromosome is a consequence of non-disjunction of B chromatids in anaphase (Ruban *et al.*, 2020). Interestingly, similar numerical variability has never been observed either in rye or maize, whose distribution of B chromosomes is stable during their life cycle and the B chromosomes can be found in all common tissues (Jones and Rees, 1982).

Here, we report on the first attempt to gain a deeper insight into the genome of *Sorghum purpureosericeum* at the sequence level. Our study estimated B chromosome size and discovered a scattered shotgun sequence, which was acquired to derive B-specific markers. This is an essential prerequisite for broader studies to elucidate the mechanism(s) underpinning the biology and evolution of accessory chromosomes in wild sorghums.

Materials and methods

Plant material

The seeds of *S. purpureosericeum* (Hochst. ex A.Rich.) Schweinf. & Asch. (accession no. IS18947) with an unknown B chromosome status were supplied by the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT, India). Before sowing them, the hard seed coats were removed and seeds were soaked in water overnight. Seeds were germinated on Petri dishes in a thermal incubator under a 8 h light/16 h dark photoperiod at temperatures of 29 °C day/25 °C night. The ensuing seedlings were planted into soil mixed with sand (2:1) in 10 cm diameter pots and cultivated under the same conditions as for seeds.

Population screening

To distinguish those plants carrying B chromosomes, haploid nuclei from their pollen grains were isolated and analysed using flow cytometry. Briefly, a few fresh anthers from individual plants were collected and their pollen extracted in a LB01 buffer (Doležel *et al.*, 1989) by vortexing for 5 min at 10 000 rpm. Additionally, samples were shaken for another 12 min in a thermomixer (Eppendorf, Hamburg, Germany) at 2000 rpm and 25 °C to release any residual pollen grains from the anthers. Pollen grains were separated from open anthers by centrifugation at 850 *g* for 5 min, at 20 °C. The empty anthers were removed with tweezers, and pollen nuclei were released from tough pollen by adding glass beads (Sigma Aldrich, Cat. No. G8772) and vortexing this for 5 min at 10 000 rpm. The resulting suspension containing the haploid nuclei was then filtered through 20 μm mesh, stained with DAPI (2 μg ml^{-1}), and analysed using FACSAria SORP (BD Biosciences, San Jose, CA, USA).

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

The presence of the B chromosomes in plants analysed using flow cytometry was verified in meiosis for 10 B-negative (B0) individuals and all B-positive (B+) individuals. Immature anthers were collected and checked for their developmental stage. Anthers at meiotic metaphase I were fixed in 3:1 (ethanol:glacial acetic acid) for 7 d at 37 °C, and then stored in 70% ethanol at –20 °C. The number of bivalents was determined on simple squashed preparations (Masoudi-Nejad *et al.*, 2002) stained with DAPI under a fluorescent microscope Zeiss Axio Image Z2.0. For each plant, 50 meiocytes in metaphase I were scored to analyse the stability and number of B chromosomes present.

B chromosome size estimation

Nuclear genome size was measured using flow cytometry according to Doležel *et al.* (2007). With respect to B chromosome size estimation, we first estimated the genome size of B0 plants of *S. purpureosericeum* (IS18947). Samples for genome size estimation were prepared from young leaves and analysed three times on three different days. *Zea mays* cv. CE-777 with an estimated genome size of 5.43 pg/2C served as an internal reference standard. Genome size (2C value) was determined considering that 1 pg of DNA is equal to 0.978×10^9 bp (Doležel *et al.*, 2003).

Similarly, nuclei for estimation of B chromosome size were isolated from florets of 1B plants. The genome size of 1B plants was calculated from the ratio of peak positions for the two populations of nuclei (2C versus 2C+1B) based on the estimate of the genome size of B0 plants of *S. purpureosericeum* (IS18947). B chromosome size was determined by subtracting the genome size of *S. purpureosericeum* B0 plants from the genome size of 1B plants.

Isolation and sequencing of pollen nuclei

Haploid pollen nuclei from both B+ and B0 plants (accession no. IS18947) were isolated using flow cytometry. The samples were prepared and processed as described above. From each individual, 9000 of its 1C nuclei (corresponding to 20 ng of DNA) were flow-sorted into 40 µl of distilled water in a 0.5 ml tube; nuclear DNA was used separately for preparing the sequencing libraries. Sorted nuclei were treated with 1.8 µl of proteinase K (10 mg ml^{-1}), for 18 h at 50 °C, to perform the protein degradation. Proteinase K was deactivated at 85 °C for 15 min, after which the samples were frozen at –80 °C. After adding water up to 100 µl, DNA from nuclei was fragmented using a Bioruptor Plus (Diagenode, Denville, NJ, USA); this was done six times, for 30 s each, at its high setting. Fragmented DNA was purified with Ampure XP (Beckman Coulter, Brea, CA, USA). Libraries for sequencing were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (Ipswich, MA, USA) with the following modifications: (i) size selection was directed to obtain a larger final library size (~1000 bp) and (ii) PCR enrichment was carried out in 12 cycles. Finally, the prepared libraries were size selected within a range of 700–1000 bp using a Blue Pippin Instrument (Sage Science, Beverly, MA, USA). Libraries were sequenced on NovaSeq 6000, and 2×250 bp paired-end reads were eventually produced.

Repeat analysis

Analysis of repetitive sequences was implemented using the Galaxy-based server (Afgan *et al.*, 2018), supplemented with the RepeatExplorer2 (Novák *et al.*, 2013) and TAREAN (Novák *et al.*, 2017) tools (https://repeatexplorer-elixir.cerit-sc.cz). Sequence reads were pre-processed in this way: reads whose quality was not greater than 10 for >95% of bases were removed, as were any reads having ambiguous nucleotides (Ns); then, the first 20 nucleotides were removed and the reads were trimmed to the same length of 150 bp. Sequence data were down-sampled to 600 000 paired-end reads (2×150 bp) from each sample-B0 and B+. This corresponds to the equivalent of ~8% of the genome, assuming a genome size of 2.21 Gb/1C. Next, the reads of both samples were clustered with RepeatExplorer2, and putative tandem repeats were analysed with TAREAN. Clusters representing >0.01% of the entire dataset were investigated further, manually. Sequence reads used in this analysis, the resulting sequences of clusters, and the counts of B+/B0 reads in each cluster have all been submitted to the Dryad public repository (https://doi.org/10.5061/dryad.rxwdbrv5j; Karafiátová *et al.*, 2021).

PCR marker development

Nine putative B-specific clusters were revealed by the repeat analysis (Table 1). For each cluster, specific primers with short amplicons (up to 1.3 kbp, Supplementary Table S1) and long amplicons (in the range of 2–5 kbp, Supplementary Table S2) were designed using Primer3 software (Koressaar *et al.*, 2007; Untergasser *et al.*, 2012). Long amplicons were found to be missing for clusters SpuCL135, SpuCL144, CL 168, and SpuCL214 due to the limited length of their cluster sequence.

Primer specificity was tested on B+ genomic DNA, for which B0 genomic DNA served as the control. DNA was extracted from whole, lyophilized spikelets of *S. purpureosericeum* plants (IS18947), by using the NucleoSpin® Plant II Kit (Macherey-Nagel, Düren, Germany) and following the manufacturer's recommendations. DNA concentration was measured using a Nanodrop® ND-1000 spectrophotometer (Saveen Werner, Malmo, Sweden).

Long amplicons were amplified with PrimeStar®GXL DNA polymerase (Takara Bio Inc., Shiga, Japan) on a C1000 Touch™ Thermal Cycler (BioRad, Hercules, CA, USA). Each PCR (20 µl) contained 50 ng of genomic DNA, 1× PrimeSTAR GXL buffer, 0.2 mM dNTPs, 0.25 µM of each primer, 0.5 U of PrimeSTAR GXL DNA polymerase, and distilled water. The products were amplified in 30 cycles, consisting

Table 1. Characteristics of B-specific/enriched repetitive clusters and centromeric satellites as revealed by the RepeatExplorer2/TAREN pipeline

Cluster	Consensus length (bp)	Genome proportion	Annotation	B-specific reads	PCR specificity	FISH signal
SpuCL4	137	1.00%	Centromeric repeat	58.6%	B0/B+	Centromeres
SpuCL115	N/A	0.20%	Class_II./hAT	99.4%	B+	B specific
SpuCL135	1 130	0.14%	Putative satellite	100%	B0/B+	
SpuCL144	366	0.11%	Putative satellite	99.9%	B0/B+	
SpuCL168	1 560	0.056%	Putative satellite	100%	B+	B specific
SpuCL169	6 860	0.055%	Putative satellite	98.2%	B+	B/As
SpuCL175	N/A	0.048%	Class_I./LINE	97.8%	B+	B/As
SpuCL189	3 500	0.028%	Putative satellite	94.7%	B+	B/As
SpuCL214	260	0.017%	Putative satellite	100%	B+	No signal
SpuCL220	N/A	0.015%	N/A	89.9%	No product	

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

of 98 °C/10 s, 60 °C/15 s, and 68 °C/270 s, and separated on a 0.8% agarose gel.

Short amplicons were amplified using *Taq* DNA polymerase with the Standard Taq Buffer (New England Biolabs, Ipswich, MA, USA). Each PCR contained 20 ng of genomic DNA, 1× Standard Taq Reaction Buffer, 0.2 mM dNTPs, 0.5 μM of each primer, 0.5 U of *Taq* DNA polymerase, and distilled water up to 20 μl. The reaction was conducted under these conditions: initial denaturation 98 °C/3 min; 35 cycles of 98 °C/30 s; 60 °C (62 °C)/30 s; 72 °C/60 s, followed by a final extension at 72 °C/5 min. The annealing temperature (T_a) was determined to be 60 °C for all primers, except for SpuCL144 and SpuCL214 (T_a=62 °C). The amplification products were visualized on a 1.5% agarose gel.

Microscopic slide preparation, probe labelling, and fluorescence in situ hybridization (FISH)

Meiotic chromosome spreads for FISH were prepared from fixed immature anthers of B+ plants and B0 plants serving as negative control. The anthers were first digested with 4% cellulase Onozuka R-10 (Yakult Honsa Co. Ltd, Minato-ku, Japan) and 1% Pectolyase Y-23 (MP Biochemicals, Santa Ana, CA, USA,) in a 1× KCl (0.15 M KCl, 5.4 mM EDTA, pH 4) buffer for 20 min at 37 °C. Macerated anthers were squashed and the slides prepared as described by Karafiátová *et al.* (2013).

All six clusters resulting in a B-specific product were tested as cytogenetic markers. Where possible, long amplicons were preferably used for probe preparation. For clusters SpuCL115, SpuCL169, SpuCL175, and SpuCL189, the long PCR products were labelled with aminoallyl-dUTP-5-FAM (fluorescein aminide; Jena Biosciences, Jena, Germany) or Texas Red d-UTP (Invitrogen, Camarillo, CA, USA), using the Nick translation Mix (Roche, Mannheim, Germany). In clusters SpuCL214 and SpuCL168, where no long amplicons could be designed, the probes were labelled using PCR. Additionally, cluster SpuCL4 was identified as a centromeric repeat. The centromeric probe and B-specific clusters SpuCL168 and SpuCL214 were labelled with fluorescein isothiocyanate (FITC)–dUTP (Roche, Mannheim, Germany) or Texas Red-dUTP (Invitrogen), using PCR. The 25 μl reaction consisted of 30 ng of genomic DNA, 1 μM of each forward and reverse primer (primer sequences for SpuCL214 and SpuCL168 are listed in Supplementary Table S1; primers for centromeric cluster SpuCL4 were as follows: F, AGTGGAAGCACGTTTCGGTA; R, ATCGGGTGCATCCAAAACTA), 1× *Taq* polymerase buffer, nucleotides (0.2 mM each of dATP, dCTP, dGTP; 0.1 mM each of dTTP and dUTP), and 0.5 U of *Taq* polymerase (New England Biolabs). The labelling was implemented under these conditions: initial denaturation, at 94 °C/60 s; 30 cycles of 94 °C/30 s, 60 °C/30 s, 72 °C/30 s; final extension, at 72 °C/10 min.

The slides with mounted chromosome preparations were fixed in 4% formaldehyde, as described in Said *et al.* (2018). Hybridization was performed following Karafiátová *et al.* (2013) and the signals were evaluated using Zeiss Axio Image Z2.0 equipped with a Cool Cube camera (Metasystems, Altussheim, Germany) and an appropriate set of filters.

Results and discussion

B chromosomes are consistently detectable in the 'germ line' only

Because the seed collections available in genebanks lack information about B chromosome status, here we screened anonymous accessions of *S. purpureosericeum*, with a preference for those originating in Sudan and India where the B chromosome(s) were previously reported (Janaki Ammal, 1939; Wu, 1984). Following the finding from old reports that B chromosomes were absent in roots, shoots, and leaves (Darlington *et al.*, 1941), we analysed the meiocytes and pollen nuclei in anthers; however, the cytological scoring of meiocytes is rather laborious and time consuming. To circumvent this limitation, using flow cytometry, we instead checked parts of whole plants for the presence of B chromosomes. Detection of B chromosomes in plant tissues using flow cytometry has been previously employed in goatgrass, where this approach turned out to be reliable to detect B chromosomes in leaves (Wu *et al.*, 2019) and even in immature embryos (Ruban *et al.*, 2020). In our material, the population of B+ nuclei formed a distinct, clearly resolved peak in samples prepared from anthers (Fig. 1) and the peduncle (Supplementary Fig. S1). Further, we found a residual population of B+ nuclei also in leaf meristems and the last node. Other tissues such as root, leaf, and stem were found to be lacking B chromosomes (Supplementary Fig. S1). Considering the results of our sampling, detection of the presence of B chromosomes relies on the proportion of B+/B0 nuclei in individual parts/tissues and thus on developmental stage. The results indicate that B chromosomes could be partially preserved in meristems. However, the amount of B+ nuclei is frequently below the detection limit of the approach. As such, we can state that B chromosomes are consistently detected in the cell line leading to the reproductive organs. Taking all recent and older findings into consideration, the hypothesis of B chromosome elimination accompanying cell differentiation could be reasonably proposed. This would be the first example of such B chromosome behaviour in plants. Nevertheless, the B chromosome elimination from somatic cells in adults was described earlier for the flatworm *Polycelis tenuis* (Melander, 1950). In a myricine ant, *Leptothorax spinosior*, the eradication of B chromosomes has even been extended and their distribution shown to be stable only in male germ lines (Imai, 1974).

In our study, a total of 120 plants were checked for their B chromosome status using flow cytometry. Of them, 19 individuals (16%) showed distinct additional peaks in the histogram, whereas in plants lacking the B chromosome in anthers two peaks appeared, representing haploid 1C pollen nuclei and 2C nuclei corresponding to residual cells from the tapetum (Fig. 1A). Specifically, the B+ individual population of haploid cells had split into two distinct subpopulations representing 1C and 1C+B nuclei (Fig. 1B). Occasionally, a population of 2C nuclei was also observed, most probably being nuclei from anther walls, where the B chromosome distribution is variable. The status of plants was further evaluated cytogenetically in metaphase I. In analysed plants showing no extra population in the histogram, five bivalents were observed in all meiocytes, confirming the absence of B chromosomes. In contrast, plants showing extra populations always had more than five bivalents. The B chromosomes formed univalents (if odd in number, Fig. 2A) or ring bivalents (Fig. 2B). Notably, rod bivalent formation was observed. A maximum of four B chromosomes were found in our plant material (Fig. 2C)

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

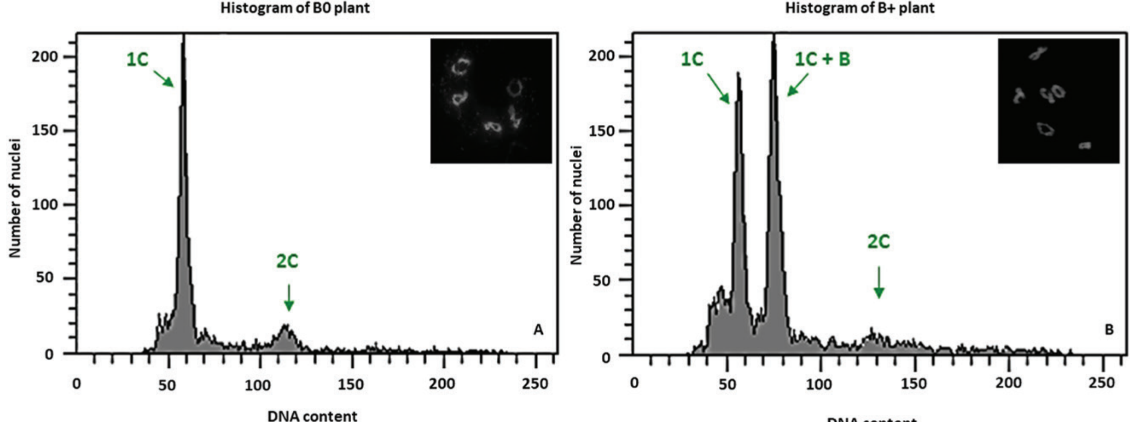

Fig. 1. Flow histograms of pollen nuclei isolated from B0 (A) and 2B (B) plants of *Sorghum purpureosericeum*. Insets show the number of bivalents in meiotic metaphase I of analysed plants. (A) Histogram of a B0 plant with two peaks representing haploid 1C pollen nuclei and 2C nuclei corresponding to residual cells from tapetum. (B) The population of haploid cells splits into two distinct subpopulations representing those harbouring 1C and 1C+B nuclei.

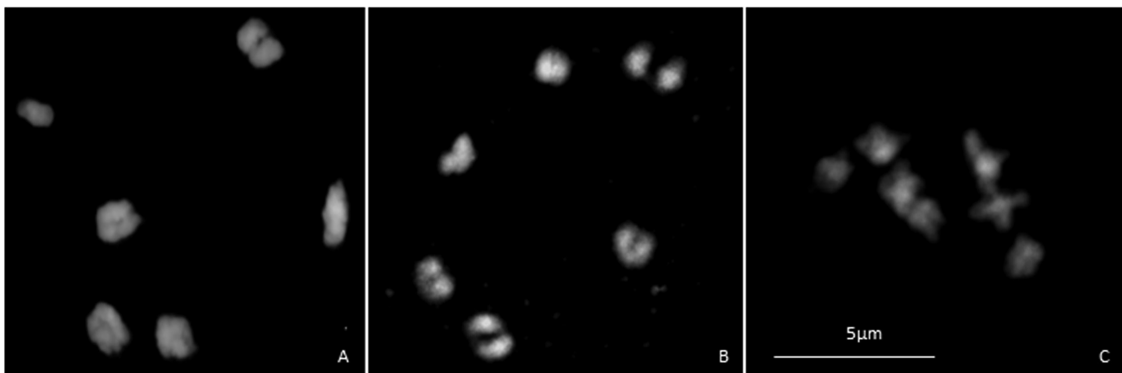

Fig. 2. Meiosis in B+ plants of *Sorghum purpureosericeum*. (A) B+ plant (2*n*=2*x*=10+B) forming five bivalents and one univalent in meiotic metaphase I; (B) B+ plant (2*n*=2*x*=10 + 2B) showing six ring bivalents; and (C) B+ plant (2*n*=2*x*=10 + 4B) with seven bivalents.

Further, the B chromosome was present in all meiocytes at metaphase I in B+ individuals, and the number of B chromosomes in all analysed cells was stable.

In addition to its numerical traits, structural polymorphism is a characteristic phenomenon of B chromosomes. More than one polymorphic form has been reported in >60 species (Jones and Rees, 1982). Although a polymorphism in size had been observed before in *S. purpureosericeum*, we have never seen such variability. All B chromosomes observed in our material were of the same size, being on a par with standard A chromosomes. Interestingly, the reports describing B polymorphism in *S. purpureosericeum* are not in agreement. Although Reddy *et al.* (1958) found two B chromosome types in his material: a short B chromosome (S) similar in size to A and an isochromosome (L) twice as large, Darlington *et al.* (1941) unveiled more additional types that probably arose from a morbid mitosis. Besides the sorghums, two polymorphic types of B chromosomes, metacentric and telocentric, were detected for instance in *Aegilops mutica* (Mochizuki, 1960). Further, there are six forms of Bs in rye (Jones and Rees, 1982), and up to 29 forms occur in *A. schoenoprasum* (Bougourd and Parker, 1979).

Accumulation of specific repetitive sequences in the B chromosomes of S. purpureosericeum

Isolated haploid nuclei of B+ and B0 plants were sequenced in this study. Our analysis of repetitive sequences was done using the RepeatExplorer2 pipeline (Novák *et al.* 2013). Out of 2.4 million reads, a total of 1 952 433 were analysed (975 698 from the B+ sample and 976 734 from the B0 sample). Among them, 66% were clustered, leaving 432 108 reads occurring as singlets. The analysis generated 231 clusters, each accounting for at least 0.01% of the original dataset. Comparing the count of reads in these clusters (between B+ and B0 samples) revealed that individual clusters consisted of 35.6–100% reads of the B+ sample. Among them, 194 clusters had a nearly equal representation (40–60%) of reads from the B+ sample as from the B0 sample, a ratio expected for clusters originating in the A chromosome complement (as it is present in both samples). In contrast, just nine clusters were significantly enriched for reads of the B+ sample, whose proportion was at least 80% (Fig. 3; Table 1). Sequences of these clusters were further checked for B specificity, with subsequent annotations

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

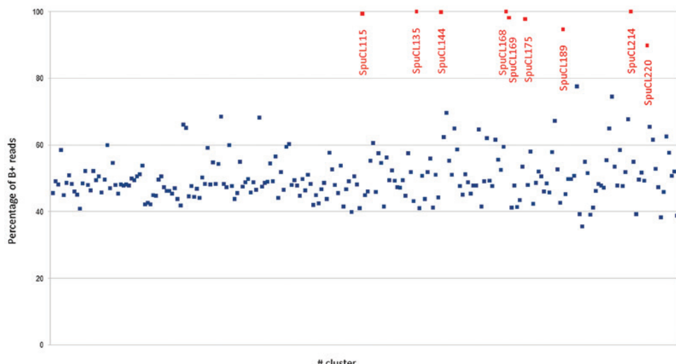

Fig. 3. Proportion of reads from the B+ sample in clusters generated and analysed using the RepeatExplorer pipeline. The clusters are organized along the *x*-axis based on their frequency in the genome (i.e. leftmost being most frequent). Values on the *y*-axis correspond to percentage of reads from the B+ sample. Clusters labelled in red are expected to be specific/enriched in the B+ sample. Note that the value should be ~50% for sequences (clusters) that originated in the A chromosome complement.

of cluster sequences based on a protein domain search within the pipeline. This revealed one B-specific DNA transposon/hAT (SpuCL115) and one long interspersed nuclear element (LINE; SpuCL175). Putative tandem repeats (satellites) were identified based on cluster layout topology, using the TAREAN tool (Novák *et al.*, 2017). For those clusters, k-mer frequencies in a set of oriented reads obtained from the graph were used to reconstruct the consensus sequence of the satellite monomer. Among putative B-enriched sequences, six qualified as being low-confidence satellites (SpuCL135, SpuCL144, SpuCL168, SpuCL169, SpuCL189, and SpuCL214; Supplementary Fig. S2). Similarly, B-specific repeats have been identified in the sequence of other plant B chromosomes (Alfenito and Bichler 1993; Martis *et al.*, 2012; Wu *et al.*, 2019; Ebrahimzadegan *et al.*, 2019). Moreover, our analysis of repetitive sequences revealed that one of the most abundant clusters (SpuCL4) is a promising candidate sequence for a centromeric satellite repeat.

B chromosome sequence reveals its specific landmarks

In some species, specific morphological features will let researchers distinguish the B chromosome in the karyotype based on simple cytological observations. Often, it is their size that is distinct, and this conspicuous trait can be gleaned in species with smaller B chromosomes, as in *Secale cereale* (Jones, 1991) and *A. cristatum* (Said *et al.*, 2019), or, conversely, those with B chromosomes larger than A chromosomes, such as in *Plantago serraria* (Frost, 1959) and *Rumex thyrsiflorum* (Zuk, 1969). Occasionally, the position of the centromere could serve as a landmark. The supernumeraries were described as telocentrics in *Allium pulchellum* (Tschermak-Woess and Schinmann, 1960) or *P. alpina* (Nygren, 1962), while in maize the B chromosomes are club shaped, missing any visible centromeric truncation (Randolph, 1941). In *S. purpureosericeum*, B chromosomes are indistinguishable from standard chromosomes, being of the same size and lacking any specific attribute. Sporadically, B chromosomes could be identified in meiotic metaphase I. Generally, they form ring bivalents when present in even numbers. Nevertheless, we observed that when a rod bivalent formation occurs, it is frequently composed of B chromosomes (Fig. 4B, C).

Unfortunately, all of these morphological features lose their significance and utility when it comes to the nuclei and observation of B chromosomes in interphase. To follow the supernumeraries through the cell cycle or at particular developmental stages of plants, more advanced markers must be employed. In animals, B-specific whole chromosome paints are very effective; in this way, the B chromosomes of the cichlid fish *Astatotilapia latifasciata* (Valente *et al.*, 2017) and of the characid fish *Astyanax scabripinnis* (Vicari *et al.*, 2011) were visualized. In plants, however, applying whole chromosome-specific probes is fraught with difficulty because of the extreme level of unspecific hybridization; hence, shorter probes specific for B chromosome segments are alternatively utilized in species, whereby the B chromosome sequence is acquired. To date, actually, this has been implemented in just a few species, namely rye (Martis *et al.*, 2012), maize (Alfenito and Bichler, 1993), goatgrass (Wu *et al.*, 2019), and meadow fescue (Ebrahimzadegan *et al.*, 2019)—in all of them, B-specific repeats were detected.

In our study, we broadened this not numerous group of species by adding to it *S. purpureosericeum*. The availability of B-specific PCR markers is crucial for further research of the *Sorghum* B chromosome, whether it is to facilitate the detection of B+ plants or to determine the number of B chromosomes within a plant, using either quantitative or digital droplet PCR. Here we provide reliable PCR-based markers with strong amplification. The set of primers was derived from nine putative B-specific clusters SpuCL115, SpuCL135, SpuCL144, SpuCL168, SpuCL169, SpuCL175, SpuCL189, SpuCL214, and SpuCL220 (Supplementary Tables S1, S2).

For the short amplicons, the amplification of only three primers (SpuCL115, SpuCL168, and SpuCL214) out of nine resulted in a product on the B+ template only. The others did not work specifically, in that they each showed amplification in both the B+ and B0 samples (Fig. 5).

Among the five primer pairs for the long amplicons, four were found to be B chromosome specific. Primers designed for clusters SpuCL115, SpuCL169, SpuCL175, and SpuCL189 provided products of an expected length from the B+ sample but no visible products for the B0 sample (Fig. 6). Primers derived from cluster SpuCL220 did not yield any B-specific products.

In aiming to track the B chromosome *in situ*, we developed strong and reliable cytogenetic markers specific for the B chromosome. Products of B-specific clusters SpuCL115, SpuCL168, SpuCL169, SpuCL175, SpuCL189, and SpuCL214 were labelled with Texas Red or FAM, and the probe specificity

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

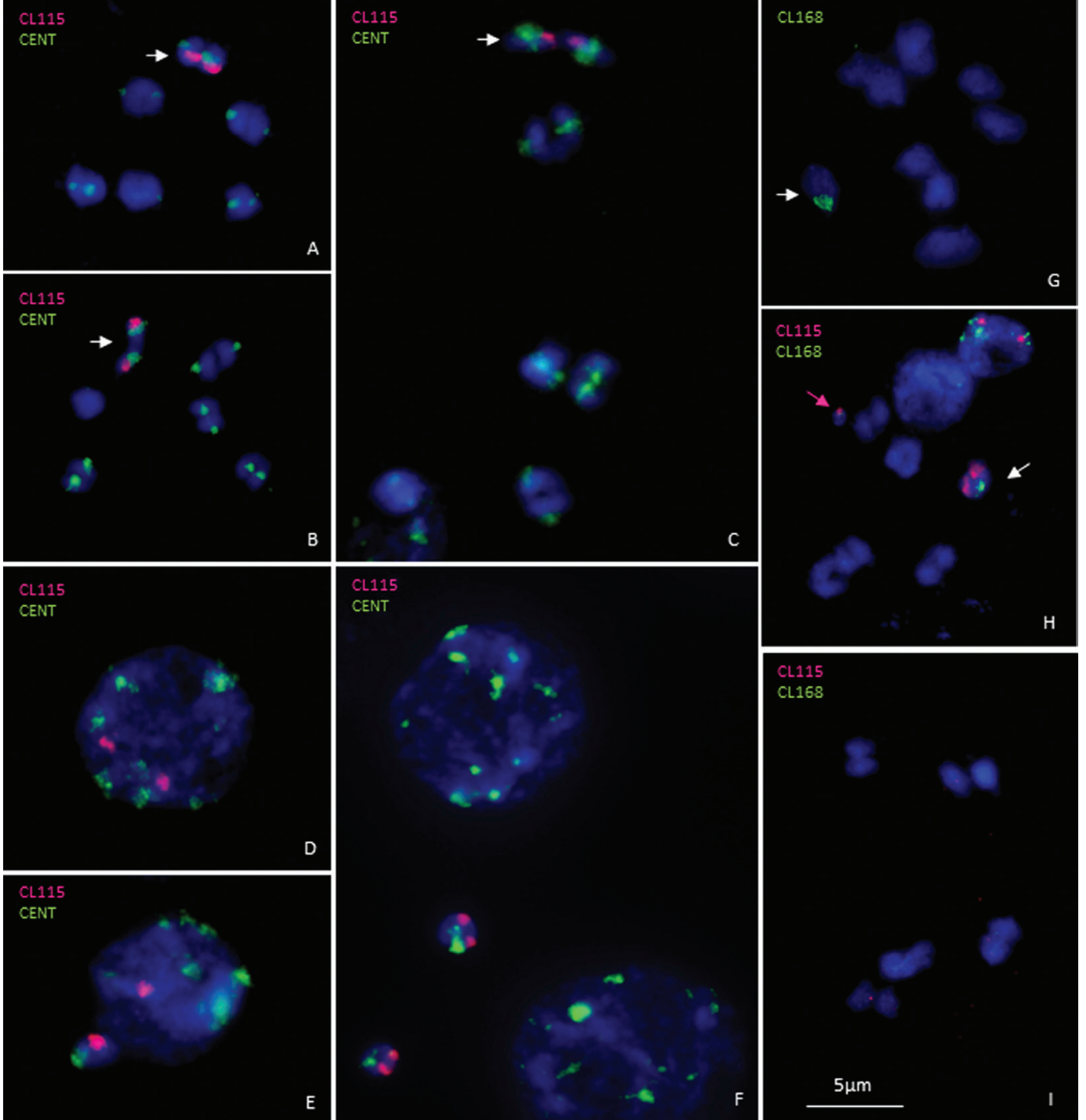

Fig. 4. FISH with the B-specific probes SpuCL115 and SpuCL168 and the centromeric probe SpuCL4. The SpuCL115 (red) and SpuCL4 (green) on 2B plants with B chromosomes forming ring bivalents (A) and rod bivalents with an opposite orientation (B, C) in meiotic metaphase I. Two distinct SpuCL115 (red) signals and a centromere (green) on the interphase nucleus (D). The SpuCL115 (red) on the nucleus with one B chromosome preserved and on the micronucleus with one B chromosomes eliminated (E); SpuCL115 signals on micronuclei with both B chromosomes, for which adjacent nuclei are free from signals (F). The B-specific cluster SpuCL168 (green) in meiotic metaphase I (G); co-localization of SpuCL115 (red) and SpuCL168 (green), with the purple arrow pointing to the micronucleus (H). Localization of B-specific probes on meiotic metaphase I of B0 plants showing no distinct signal (I). Chromosomes are counterstained with DAPI (blue) and B chromosome bivalents are indicated by arrows.

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

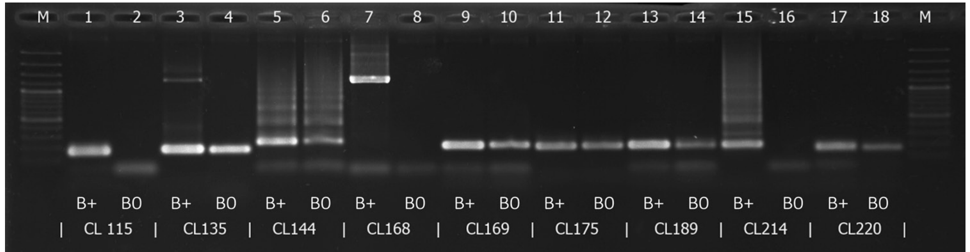

Fig. 5. Electrophoresis of short amplicons for nine putative B-specific clusters (SpuCL115, SpuCL135, SpuCL144, SpuCL168, SpuCL169, SpuCL175, SpuCL189, SpuCL214, and SpuCL220). For all clusters, primer specificity was tested in parallel on B+ DNA and B0 DNA. The presence of PCR products of SpuCL115, SpuCL168, and SpuCL214 exclusively in B+ samples indicates their B specificity. The remaining primer pairs showed non-specific amplifications. The M lane corresponds to the GeneRuler 100 bp Plus DNA Ladder.

was tested on the meiocytes of a B+ plant. According to these results, SpuCL214 provided no distinct signal on any chromosome. Abundant signals for the Spu169, SpuCL175, and SpuCL189 clusters were detected on B chromosomes, but the probe also hybridized to chromosomes of the A complement. Only clusters SpuCL115 and SpuCL168 were confirmed to be B chromosome specific (Table 1). Hybridization of the DNA transposon SpuCL115 and satellite SpuCL168 showed a single chromosome-specific signal, which was regularly detected on one bivalent in the *S. purpureosericeum* plant with 2B chromosomes (Fig. 4A–C, G). Importantly, both clusters hybridized to the same chromosome (Fig. 4H). To address whether the probes targeted the B chromosome, in parallel we performed the identical experiment but using B0 plants, for which a signal was never found in their preparations (Fig. 4I). Both probes hybridized to the distal parts of opposite chromosome arms and represent a segment of significant size. However, the signal from cluster SpuCL115 was larger, more robust, and compact when compared with that of cluster SpuCL168. These results agree with the repeat analysis findings from before; specifically, they suggested that the representation of SpuCL115 amounted to 0.2% (4.2 Mbp in 1C) of the genome, being only 0.056% (1.2 Mbp in 1C) for cluster SpuCL168 (Table 1). Still, both signals were distinguishable on the nuclei as well (Fig. 4D, E, H). Nevertheless, the respective number of the loci varied due to the process of B elimination. When the micronucleus was present, the signal from the B-specific repeat was consistently detected here; when we followed the signals in mother nuclei and micronuclei, the sum of their signals was stable (Fig. 4E, F).

Our analysis revealed two specific FISH markers. While the signal of the satellite SpuCL168 is characterized by a rather dispersed structure, that of DNA transposom SpuCL115 is compact and 'paints' nearly a whole single arm of the B chromosome and so it is robust enough for use in various applications. Similar blocks of repeats were found in *Festuca pratensis*, whereby, when using two satellite probes making up 0.33% and 0.015% of its genome, respectively, practically the whole B chromosome can be visualized (Ebrahimzadegan *et al.*, 2019). In contrast, no B-specific block of similar size was found, for example, in *Ae. speltoides*, where two B-specific satellites represented only 0.053% and 0.026% of the genome (Wu *et al.*, 2019) whose signals had a less dense, 'dotty' pattern.

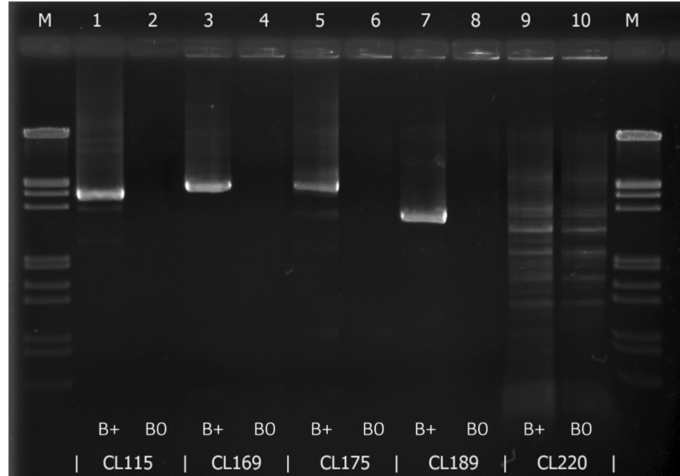

Fig. 6. Electrophoresis of long amplicons for five putative B-specific clusters (SpuCL115, SpuCL169, SpuCL175, SpuCL189, and SpuCL220). Primer specificity was tested in parallel on B+ DNA and B0 DNA. The presence of PCR products of SpuCL115, SpuCL169, SpuCL175, and SpuCL189 exclusively in B+ samples indicates their B specificity. The M lane corresponds to λ DNA digested with *Eco*RI and *Hind*III.

Additionally, cluster SpuCL4 proved to be a centromeric probe, able to hybridize to all centromeres. Nonetheless, its signal is not uniform. There are two chromosome pairs with a unique hybridization pattern: the signal of the B centromere is significantly larger, and the difference is so unambiguous that it could also serve as a marker for B chromosome identification (Fig. 4A–C). Next to the B chromosome, one chromosome pair from the A complement showed a remarkably weaker and smaller signal relative to the remaining bivalents.

An adverse effect of the B chromosomes on a host

The harm of a high dosage of B chromosomes on the fitness of their carriers is generally known, for which the onset of a negative effect on phenotype is anticipated just in higher dosage (Valente *et al.*, 2017). A maximum of 6Bs were observed

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

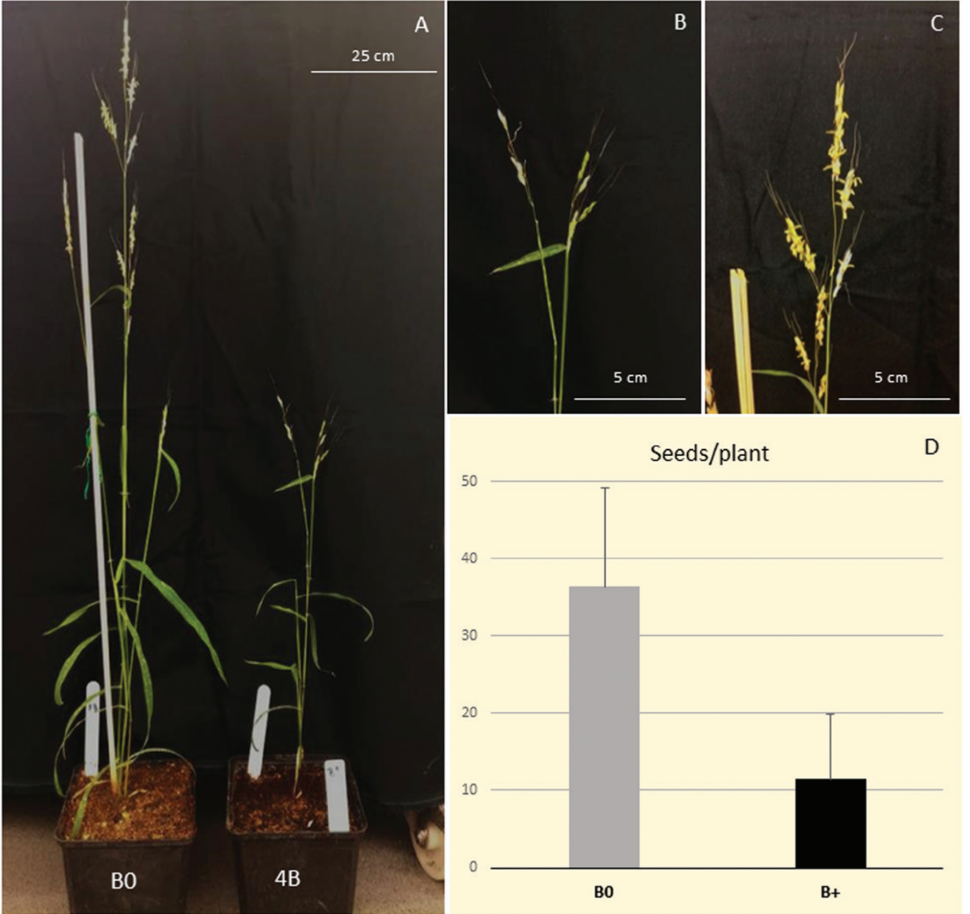

Fig. 7. Harmful effects of the B chromosome presence on phenotype. A 4B plant whose fitness was significantly decreased when compared with a B0 plant (A); spikes of a 4B plant characterized by morbid development and reduced fertility (B); normal inflorescence structure of a B0 plant (C); and comparison of seed production of 11 B0 and 2B plants (D).

in *S. purpureosericeum* (Janaki-Ammal, 1940). Nonetheless, the phenotypic effect of B chromosomes in *S. purpureosericeum* was apparent even under a disomic constitution, when plants with 2Bs were mostly smaller and their fertility was decreased, and the effect was even stronger in plants possessing 3Bs or 4Bs (Fig. 7A–C). Plants with 2Bs produced a significantly lower number of seeds when compared with B0 individuals (Fig. 7D; Supplementary Table S3). The seed production was evaluated for 11 B+ and 11 B0 plants. While B0 plants produced 36.4±12.8 seeds, B+ plants have significantly (Student's *t*-test, $P<0.001$) reduced fertility, producing only 11.5±8.4 seeds. In rye, wherein a maximum of 8Bs was reported, harbouring 3Bs in its genome is not harmful and the phenotype is identical to that of the B0 plant (Jones *et al.*, 2008). Nonetheless, our findings are in line with earlier work by Magoon *et al.* (1961), who reported a >95% sterile tiller in *S. purpureosericeum* plants having 3Bs.

We assumed that the onset of the adverse effects of B chromosomes on host plant vigour depends on the proportion of their DNA in the nucleus rather than the number of B chromosomes *per se*. According to earlier research on rye and maize, reductions in plant fitness are pronounced when the B chromosome mass exceeds 20% of DNA in the nucleus, a potential heuristic obtainable via calculations of previously published data. In maize with a genome size of 4.8 Gbp/2C (Schnable *et al.*, 2009) and a B chromosome size estimated at 140 Mbp (N. Blavet *et al.*, unpublished results), poor fitness and reduced fertility were observed in plants with 10 or more B chromosomes (Masonbrink et Birchler, 2012). In the 10Bs plant, B chromosomes constituted ~23% of DNA in the diploid nucleus. Similarly, in rye (15.8 Gbp/2C; Bartos *et al.*, 2008), where B chromosome size was reported to be 540 Mbp (Martis *et al.*, 2012), the presence of 6Bs (~20% of diploid DNA content) was shown to be detrimental (Jones *et al.*, 2008). We also observed a dramatic impact on plant fertility and condition when the genome carried two or more B chromosomes. We estimated the genome size of *S. purpureosericeum* (accession no. IS18947) to be 4.42 Gb/2C (Supplementary Fig. S3; Supplementary Table S4), which is similar to previous estimates (4.2 Gbp/2C; Price *et al.*, 2005). Together with B chromosome

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

size in this species being 421 Mbp (Supplementary Fig. S3; Supplementary Table S5), the accrued mass of 2Bs accounted for ~20% of extra DNA in the diploid nucleus. In conclusion then, the proportion of DNA content of 2Bs in relation to the genome size of *S. purpureosericeum* is comparable with that of 10Bs in maize or 6Bs in rye plants, and, as such, the substantial fitness impact of B chromosomes in *S. purpureosericeum* is not all that surprising.

Counterintuitively, this feature (i.e. the large size of *S. purpureosericeum* B chromosome) runs against all mechanisms the B chromosomes have evolved to themselves in hosting individuals. Any degree of sterility adversely affects the transmission of B chromosomes into the progeny and lessens their probability of being retained in the population. Taking this finding into due consideration, the reduction of B chromosome size would seem to be a meaningful evolutionary trend.

Supplementary data

The following supplementary data are available at *JXB* online.

Fig. S1. Detection of B+ nuclei in specific parts of a 2B plant using flow cytometry.

Fig. S2. Topology of the putative B-specific clusters

Fig. S3. Determination of genome size of a B0 plant of *Sorghum purpureosericeum* and the size of the B chromosome.

Table S1. List of primers designed for the short amplicons.

Table S2. List of primers designed for the long amplicons.

Table S3. Seed production of B0 and 2B plants.

Table S4. Genome size of B0 plants of *Sorghum purpureosericeum*

Table S5. Size of the B chromosome of *Sorghum purpureosericeum*.

Acknowledgements

We thank Helena Tvardíková and Jitka Weiserová for technical assistance. This work was supported by the Czech Science Foundation (grant award no. 18-12338Y) and the European Regional Development Fund project 'Plants as a tool for sustainable global development' (CZ.02.1.01/0.0/0.0/16_019/0000827). Computational resources were provided by the ELIXIR-CZ project (LM2015047), a component of the international ELIXIR infrastructure that is part of the project 'e-Infrastruktura CZ' (LM2018140) within the program Projects of Large Research, Development and Innovations Infrastructures supported by Ministry of Education, Youth and Sports. Further, our thanks go to the International Crop Research Institute for the Semi-Arid Tropics (ICRISAT) for providing valuable seed stocks.

Author contributions

JB designed the research and revised the manuscript, MK and MB contributed equally to conducting the experiments and manuscript writing, MS assisted with flow-cytometric screening and sorting of nuclei, JČ estimated genome and B chromosome size, KH performed DNA sequencing, JB and NB analysed sequence data.

Data availability

Sequence reads used in this work, the resulting sequences of clusters, and the counts of B+/B0 reads in each cluster have been submitted to the Dryad Digital Repository (https://doi.org/10.5061/dryad.rxwdbrv5j; Karafiátová *et al.*, 2021).

References

Afgan E, Baker D, Batut B, *et al.* 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Research **46**, W537–W544.

Ahmad SF, Martins C. 2019. The modern view of B chromosomes under the impact of high scale omics analyses. Cells **8**, 156.

Alfenito MR, Birchler JA. 1993. Molecular characterization of a maize B chromosome centric sequence. Genetics **135**, 589–597.

Baenziger H. 1962. Supernumerary chromosomes in diploid and tetraploid forms of crested wheatgrass. Canadian Journal of Botany **40**, 549–561.

Banaei-Moghaddam AM, Schubert V, Kumke K, *et al.* 2019. Nondisjunction in favor of a chromosome: the mechanism of rye B chromosome drive during pollen mitosis. The Plant Cell **24**, 4124–4134.

Bartos J, Paux E, Kofler R, *et al.* 2008. A first survey of the rye (*Secale cereale*) genome composition through BAC end sequencing of the short arm of chromosome 1R. BMC Plant Biology **8**, 95.

Bougourd S, Parker JS. 1979. The B chromosome system of *Allium schoenoprasum* II. Stability, inheritance and phenotypic effects. Chromosoma **75**, 36–383.

Carlson WR. 2007. Locating a site on the maize B chromosome that controls preferential fertilization. Genome **50**, 578–587.

D'Ambrosio U, Alonso-Lifante MP, Barros K, Kovařík A, Mas de Xaxars G, Garcia S. 2017. B-chrom: a database on B-chromosomes of plants, animals and fungi. New Phytologist **216**, 635–642.

Darlington CD, Thomas TRS, Thomas PT. 1941. Morbid mitosis and the activity of inert chromosomes in *Sorghum*. Proceedings of the Royal Society B: Biological Sciences **130**, 127–150.

Doležel J, Bartoš J, Voglmayr H, Greilhuber J. 2003. Nuclear DNA content and genome size of trout and human. Cytometry **51A**, 127–128.

Doležel J, Binarová P, Lucretti S. 1989. Analysis of nuclear DNA content in plant cells by flow cytometry. Biologia Plantarum **31**, 113–120.

Dolezel J, Greilhuber J, Suda J. 2007. Estimation of nuclear DNA content in plants using flow cytometry. Nature Protocols **2**, 2233–2244.

Dweikat I. 2005. A diploid, interspecific, fertile hybrid from cultivated sorghum, *Sorghum bicolor*, and the common Johnsongrass weed *Sorghum halepense*. Molecular Breeding **16**, 93–101.

Ebrahimzadegan R, Houben A, Mirzaghaderi G. 2019. Repetitive DNA landscape in essential A and supernumerary B chromosomes of *Festuca pratensis* Huds. Scientific Reports **9**, 19989.

FAO. 1995. Sorghum and millets in human nutrition. http://www.fao.org/3/t0818e/T0818E00.htm

Frost S. 1959. The cytological behaviour and mode of transmission of accessory chromosomes in *Plantago serraria*. Hereditas **45**, 191–210.

Gotoh K. 1924. Über die Chromosomenzahl von *Secale cereale* L. Botanical Magazine Tokyo **38**, 135–152.

Han Y, Liu X, Benny U, Kistler HC, VanEtten HD. 2001. Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. The Plant Journal **25**, 305–314.

Huskins CL, Smith SG. 1934. A cytological study of the genus *Sorghum* Pers. II. The meiotic chromosomes. Journal of Genetics **28**, 387–395.

Imai HT. 1974. B-chromosomes in the myrmicine ant, *Leptothorax spinosior*. Chromosoma **45**, 431–444.

Janaki-Ammal EK. 1939. Supernumerary chromosomes in para-*Sorghum*. Current Science **8**, 210–211.

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

Janaki-Ammal EK. 1940. Chromosome diminuition in a plant. Nature **146**, 839–840.

Jones RN. 1991. B-chromosome drive. The American Naturalist **137**, 430–442.

Jones RN, González-Sánchez M, González-García M, Vega JM, Puertas MJ. 2008. Chromosomes with a life of their own. Cytogenetic and Genome Research **120**, 265–280.

Jones RN, Rees H. 1982. B chromosomes. Academic Press Inc.

Karafiátová M, Bartoš J, Kopecký D, Ma L, Sato K, Houben A, Stein N, Doležel J. 2013. Mapping nonrecombining regions in barley using multicolor FISH. Chromosome Research **21**, 739–751.

Karafiátová M, Bednářová M, Said M, ČížkováJ, Holušová K, Blavet N, Bartoš J. 2021. Data from: The B chromosome of *Sorghum purpureosericeum* reveals the first pieces of its sequence. Dryad Digital Repository. (https://doi.org/10.5061/dryad.rxwdbrv5j

Kato A, Zheng YZ, Auger DL, Phelps-Durr T, Bauer MJ, Lamb JC, Birchler JA. 2005. Minichromosomes derived from the B chromosome of maize. Cytogenetic and Genome Research **109**, 156–165.

Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics **23**, 1289–1291.

Kuwada Y. 1925. On the number of chromosomes in maize. Botanical Magazine Tokyo **39**, 227–234.

Longley AE. 1927. Supernumerary chromosomes in Zea mays. Journal of Agricultural Research **35**, 769–784.

Magoon ML, Ramanna MS, Shambulingappa KGF. 1961. Desynapsis and spontaneous chromosome breakage in *Sorghum purpureosericeum*. Indian Journal of Genetics and Plant Breeding **21**, 87–97.

Martis MM, Klemme S, Banaei-Moghaddam AM, *et al.* 2012. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. Proceedings of the National Academy of Sciences, USA **109**, 13343–13346.

Masonbrink RE, Birchler JA. 2012. Accumulation of multiple copies of maize minichromosomes. Cytogenetic and Genome Research **137**, 50–59.

Masoudi-Nejad A, Nasuda S, McIntosh RA, Endo TR. 2002. Transfer of rye chromosome segments to wheat by a gametocidal system. Chromosome Research **10**, 349–357.

Melander Y. 1950. Accessory chromosomes in animals, especially in *Polycelis tenuis*. Hereditas **36**, 19–38.

Mendelson D, Zohary D. 1972. Behaviour and transmisson of supernumerary chromosomes in *Aegilops speltoides*. Heredity **29**, 329–339.

Mochizuki A. 1957. B chromosomes in *Aegilops mutica* Boiss. Wheat Information Service **5**, 9–11.

Mochizuki A. 1960. A note on the B chromosomes in natural populations of A*egilops mutica* Boiss. in central Turkey. Wheat Information Service **11**, 31.

Müntzing A, Nygren A. 1955. A new diploid variety of *Poa alpina* accessory chromosomes at meiosis. Hereditas **41**, 405–422.

Novák P, Ávila Robledillo L, Koblížková A, Vrbová I, Neumann P, Macas J. 2017. TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. Nucleic Acids Research **45**, e111.

Novák P, Neumann P, Pech J, Steinhaisl J, Macas J. 2013. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics **29**, 792–793.

Nygren A. 1957. *Poa timoleontis* Heldr., a new diploid species of the section Bolbophorum and Gr. with accessory chromosomes only in meiosis. LantbrHögsk. Annual Report Agricultural College of Sweden **23**: 489–495.

Nygren A. 1962. On *Poa badensis*, *Poa concina*, *Poa pumila*, *Poa xerophila* and the possible origin of *Poa alpina*. Annals of Academic Regional Science Uppsaliensis **6**, 1–29.

Ostergren G, Frost S. 1962. Elimination of accessory chromosomes from the roots in *Haplopappus gracilis*. Hereditas **48**, 363–366.

Peng Z, Oliveira-Garcia E, Lin G, *et al.* 2019. Effector gene reshuffling involves dispensable mini-chromosomes in the wheat blast fungus. PLoS Genetics **15**, e1008272.

Piper JK, Kulakow PA. 1994. Seed yield and biomass allocation in *Sorghum bicolor* and F1 and backcross generations of *S. bicolor*×*S. halepense* hybrids. Canadian Journal of Botany **72**, 468–474.

Plowman J, Bougourd S. 1994. Selectively advantageous effects of B chromosomes on germination behaviour in *Allium schoenoprasum* L. Heredity **72**, 587–593.

Price HJ, Dillon SL, Hodnett G, Rooney WL, Ross L, Johnston JS. 2005. Genome evolution in the genus *Sorghum* (Poaceae). Annals of Botany **95**, 219–227.

Raman VS, Krishnaswami D. 1959. Accessory chromosomes in *Sorghum nitidum* Pers. Journal of Indian Botanical Society **34**, 278–280.

Raman VS, Meenakshi K, Thangam MS, Sivagnanam L. 1964. The cytogenetical behaviour of B-chromosomes in *Sorghum halepense*. Madras Agricultural Journal **51**, 72–73

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

Reddy BVS. 1958. On differentiation of A and B chromosomes of *Sorghum purpureosericeum*, a Eu-sorghum. Journal of Indian Botanical Society **37**, 279–289.

Ruban A, Schmutzer T, Wu DD, *et al.* 2020. Supernumerary B chromosomes of *Aegilops speltoides* undergo precise elimination in roots early in embryo development. Nature Communications **11**, 2764.

Rutishauser A, Röthlisberger E. 1966. Boosting mechanism of B chromosomes in *Crepis capillaris*. Chromosomes Today **1**, 28–30.

Said M, Hřibová E, Danilova TV, Karafiátová M, Čížková J, Friebe B, Doležel J, Gill BS, Vrána J. 2018. The *Agropyron cristatum* karyotype, chromosome structure and cross-genome homoeology as revealed by fluorescence in situ hybridization with tandem repeats and wheat single-gene probes. Theoretical and Applied Genetics **131**, 2213–2227.

Said M, Kubaláková M, Karafiátová M, Molnár I, Doležel J, Vrána J. 2019. Dissecting the complex genome of crested wheatgrass by chromosome flow sorting. Plant Genome **12**, 180096.

Schnable PS, Ware D, Fulton RS, *et al.* 2009. The B73 maize genome: complexity, diversity, and dynamics. Science **326**, 1112–1115.

Su H, Liu Y, Liu Y, Birchler JA, Han F. 2018. The behavior of the maize B chromosome and centromere. Genes **9**, 476.

Taylor JRN. 2004. Grain production and consumption: Africa. In: Wrigley C, Corke H, Walker CE, eds. Encyclopedia of grain science. Elsevier, 70–78.

Tschermak-Woess E, Schinmann H. 1960. Strukturelle Hybridität und akzessorische Chromosomen bei *Allium pulchellum*. Österreichische Botanische Zeitschrift **107**, 212–227.

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

Valente GT, Nakajima RT, Fantinatti BE, Marques DF, Almeida RO, Simões RP, Martins C. 2017. B chromosomes: from cytogenetics to systems biology. Chromosoma **126**, 73–81.

Vicari MR, de Mello Pistune HF, Castro JP, de Almeida MC, Bertollo LA, Moreira-Filho O, Camacho JP, Artoni RF. 2011. New insights on the origin of B chromosomes in *Astyanax scabripinnis* obtained by chromosome painting and FISH. Genetica **139**, 1073–1081.

Wu D, Ruban A, Fuchs J, Macas J, Novák P, Vaio M, Zhou Y, Houben A. 2019. Nondisjunction and unequal spindle organization accompany the drive of *Aegilops speltoides* B chromosomes. New Phytologist **223**, 1340–1352.

Wu T. 1992. B-chromosomes in *Sorghum stipoideum*. Heredity **68**, 457–463.

Wu TP. 1984. B-chromosomes in *Sorghum purpureosericeum*. Proceedings of National Science Council, Republic of China **8**, 198–209.

Zuk J. 1969. The additional heterochromatic chromosome and its influence on sex chromosome pairing in *Rumex*. Heredity **24**, 69–74.

Downloaded from https://academic.oup.com/jxb/article/72/5/1606/5995917 by guest on 06 April 2022

4.2 B chromosomes in genus *Sorghum* (Poaceae)

Bednářová M., Karafiátová M., Hřibová E., & Bartoš J.

Plants 2021, 10(3): 505

doi: 10.3390/plants10030505

IF: 2.76

Summary

To shed more light on phylogenetic relationships in the genus with respect to B chromosome occurrence, phylogenetic reconstruction was performed. Twenty-one *Sorghum* species were included into analysis and represented by 58 samples. Internal transcribed spacers *ITS1* and *ITS2* and two chloroplast intergenic spacers *trnH-psbA* and *trnL-trnF* were selected as phylogenetic markers. Selected DNA regions were amplified by PCR and products of the amplification were sequenced on ABI 3730xl DNA analyzer. Multiple sequence alignment was conducted by MAFFT v7.029. Phylogenetic tree was reconstructed using phyML and BioNJ algorithm resulting in single tree with two strongly supported lineages (Fig. 13). B-carrying species were distributed independently in both major clades. Three evolutionary scenarios of the B chromosome formation in the genus were considered. B chromosome could have arise: A) in a single event in a common ancestor of all *Sorghum* species and then have been preserved in some lineages while disappeared in the others; B) in a single event in an ancestor of closely related species which have kept Bs up until today; or C) several times in independent events during *Sorghum* evolution which also include the eventuality of B chromosome formation at the level of individual species. Based on phylogenetic analysis, scenario 2), i.e. formation in the ancestor of a group of closely related species, was ruled out.

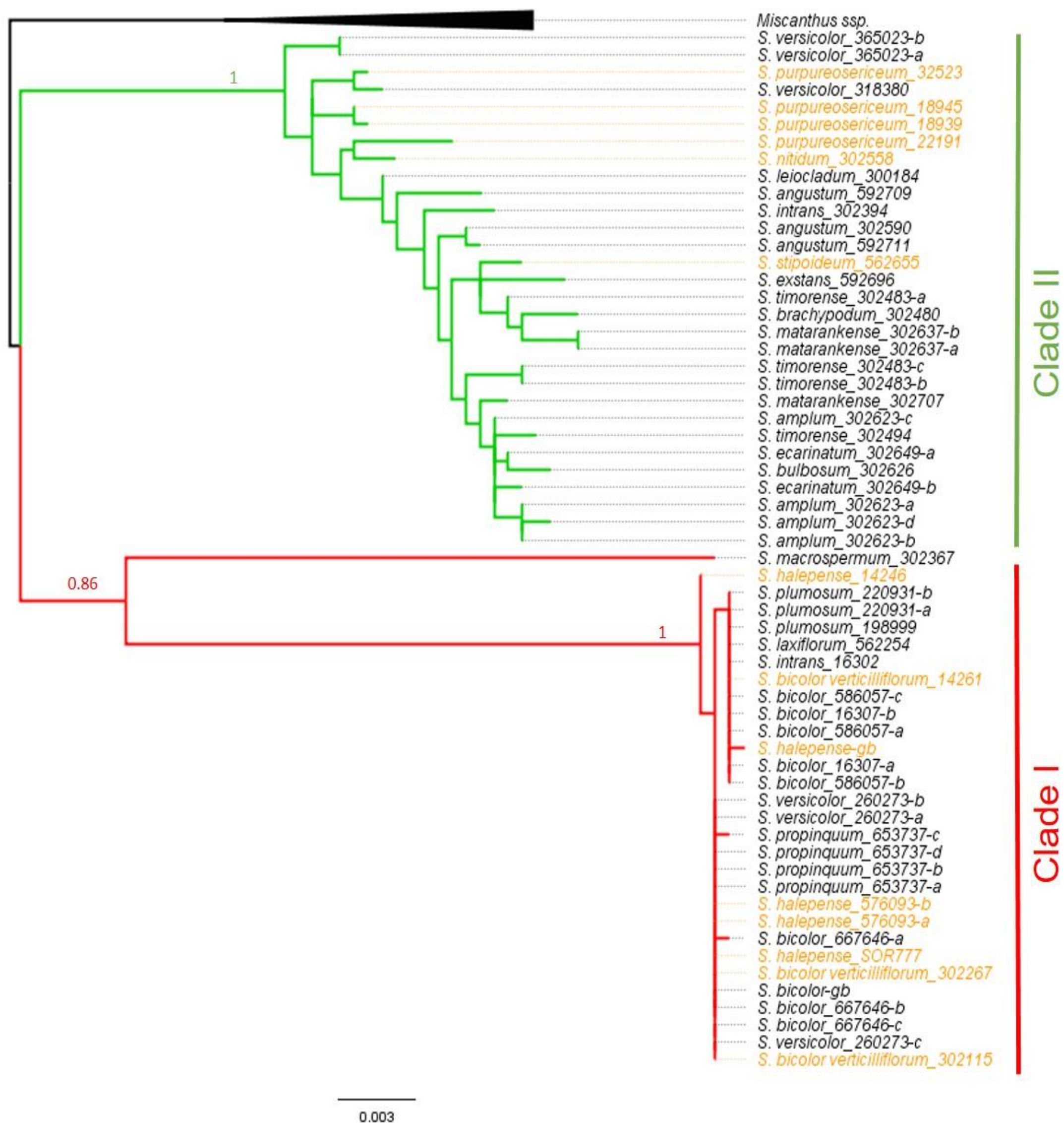

Fig. 13 Phylogenetic tree of 21 *Sorghum* species based on phyML analysis using concatenated nuclear and chloroplast sequences (*ITS1-ITS2-trnHpsbA-trnLtrnF*). Two main clades I and II with strong bootstrap support were resolved. B-carrying species are marked in orange (Bednářová *et al.* 2021).

Review

B Chromosomes in Genus *Sorghum* (Poaceae)

Martina Bednářová, Miroslava Karafiátová, Eva Hřibová and Jan Bartoš *

Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany of the Czech Academy of Sciences, Šlechtitelů 31, 779 00 Olomouc, Czech Republic; bednarova@ueb.cas.cz (M.B.); karafiatova@ueb.cas.cz (M.K.); hribova@ueb.cas.cz (E.H.)
* Correspondence: bartos@ueb.cas.cz

Abstract: B chromosomes (Bs) are supernumerary dispensable genomic elements that have been reported in several thousand eukaryotic species. Since their discovery, Bs have been subjected to countless studies aiming at the clarification of their origin, composition, and influence on the carriers. Despite these efforts, we still have very limited knowledge of the processes that led to the emergence of Bs, the mechanisms of their transmission, and the effects of Bs on the hosts. In the last decade, sophisticated molecular methods, including next-generation sequencing, have provided powerful tool to help answer some of these questions, but not many species have received much attention yet. In this review, we summarize the currently available information about Bs in the genus *Sorghum*, which has so far been on the periphery of scientific interest. We present an overview of the occurrence and characteristics of Bs in various *Sorghum* species, discuss the possible mechanisms involved in their maintenance and elimination, and outline hypotheses of the origin of Bs in this genus.

Keywords: B chromosomes; supernumerary chromosomes; *Sorghum*; chromosome elimination; phylogenesis; evolution

Citation: Bednářová, M.; Karafiátová, M.; Hřibová, E.; Bartoš, J. B Chromosomes in Genus *Sorghum* (Poaceae). *Plants* **2021**, *10*, 505. https://doi.org/10.3390/plants10030505

Academic Editor: João Loureiro

Received: 12 February 2021
Accepted: 4 March 2021
Published: 9 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

B chromosomes (Bs) are supernumerary, dispensable chromosomes that have been observed in all major groups of living organisms—animals, plants, and fungi [1]. The basic characteristics of Bs are their inability to pair with A chromosomes (As) during meiosis and their irregular mode of inheritance [2,3]. As the transmission rate of Bs is higher than 0.5, they can be viewed as parasitic elements with their own evolutionary pathway [4]. One of the common features of Bs is that they are present only in some individuals of a particular species, and this variability may also exist at the level of populations or even at the level of tissues of a single individual. Most Bs share the common basic features mentioned above, but they have also developed some species-specific attributes, which in some cases resulted in the emergence of unique systems of Bs [5].

Bs are known to be present in a number of plant species, however, they have been studied in more detail mainly in plants with agronomic importance (e.g., rye, maize) [6–13]. Information about Bs in other plant species is rather superficial. In angiosperms, Bs tend to be present in species that have relatively large genomes and a small number of chromosomes [14–16]. Bs were found in 8% monocots and 3% eudicots, and their distribution in various orders, families, and genera is not random [17]. Among monocotyledonous plants, the orders Commelinales and Liliales seem to be the "hotspots" of B chromosome occurrence [17].

Based on the dispensable nature of Bs and on their potentially detrimental effect on a host, it would be logical to expect that they will be gradually suppressed and subsequently eradicated from the population. However, Bs seem to successfully persist in the populations thanks to their specific accumulation mechanisms [18]. One of these mechanisms is nondisjunction, which has been relatively well described in rye and maize [12,19–22]. In maize, nondisjunction takes place during the second pollen mitosis, when two sperm cells are formed. Sister chromatids of B chromosome fail to disjoin at anaphase; both are

pulled to one pole and thus end up in one sperm cell. As a result, one of the sperm cells accumulates B chromosome at the expense of the other. The sperm cell with B chromosome then preferentially merges with the egg cell [20,23–25]. In rye, nondisjunction occurs during the first pollen mitosis, when vegetative and generative nuclei are formed: both chromatids of the B chromosome are in most cases included in the generative nucleus [26]. Nondisjunction of rye and maize Bs are both examples of post-meiotic drive. Besides this mechanism, pre-meiotic and meiotic drive have also been reported, but mainly in animals [27–30].

Generally, Bs are smaller than As, but Bs of similar size as As (“large” Bs) have also been reported [2,31,32]. In many species, different morphological variants of Bs have been observed within a single species, for example, in chives *Allium schoenoprasum*, grasshopper *Eyprepocnemis plorans*, or fish *Astyanax scabripinnis* [33–35]. Bs occur in different organisms at variable numbers, and the tolerable maximum depends on the particular species. When present in low numbers, Bs generally do not have any detrimental effect on the host, however, in higher numbers they can reduce the fitness of the carrier [2]. There have been a few reports suggesting some positive effects of the presence of Bs on their carrier [36–38], but in general, Bs do not provide any obvious benefits.

Due to the absence of selection pressure, Bs behave like a “genomic sponge” and accumulate sequences of various origins. As shown in rye and maize, Bs can accumulate organellar DNA, transposable elements, satellite sequences, ribosomal DNA, and other sequences from various As [39–41]. Bs can also contain genic sequences, but they are mostly not functional due to the pseudogenization [42–44]. All captured sequences can then diverge from their original homologues, and thus Bs can serve as a potential source of genetic variability. Although Bs have been considered transcriptionally inactive elements, recent studies indicate that at least some Bs are transcriptionally active and contain functional protein-coding genes [45–48]. Genes localized on Bs were shown to play a role in female sex determination in cichlid fish [49], in the processes related to cell division in *E. plorans* [48], and in the cell cycle and development in the red fox (*Vulpes vulpes*) and raccoon dog (*Nyctereutes procyonoides procyonoides*) [50]. These data suggest that Bs can carry genes controlling their specific behavior.

2. B Chromosomes in the Genus *Sorghum*

In the genus *Sorghum*, Bs have been reported in five species: *S. bicolor* ssp. *verticilliflorum* [51], *S. stipoideum* (Figure 1a) [52], *S. purpureosericeum* (Figure 1b) [53,54], *S. halepense* (Figure 1c) [55,56], and *S. nitidum* (Figure 1d) [31,57]. Despite the morphological variability of Bs described in sorghums, they share one common feature—they are well preserved in the cell lineages leading to the reproductive organs, but are absent in most somatic tissues. Several cytological studies on *Sorghum* Bs have been performed, and their morphology and behavior during meiosis have been relatively well documented in all species except *S. bicolor* ssp. *verticilliflorum*.

From all the Bs in genus *Sorghum*, the most detailed information is available about Bs from *S. purpureosericeum*. A maximum of six Bs in one cell was reported in this species [53], and the Bs described so far are not morphologically identical. Darlington and Thomas [54] described three types of heterochromatic Bs (long, medium, short), which did not pair with each other. Based on the published reports, the medium type of B chromosome with visible constriction seems to be the most common. The transmission of B chromosome(s) through meiosis in both 1B and 2B plants has been well documented, showing nearly regular behavior in 2B plants, resulting in four microspores with 1B chromosome. In 1B plants, however, the B chromosome passes undivided through the first meiotic division and divides in a second division, giving rise to two microspores with 1B and two without [58]. Meiosis of B-carrying plants was also previously studied by D’cruz and Deshmukh [59], who found precociously dividing B chromosome at metaphase I of male meiosis. An outline of the B chromosome behavior in the first pollen mitosis was proposed by Darlington and Thomas [54], who also suggested a unique manner of the accumulation of Bs in this species. They observed extra divisions (polymitosis) between the first and

second pollen mitosis, which they believed led to B chromosome multiplication [54]. The existence of the micronucleus containing Bs in resting cells has also been noticed, and the theory of B chromosome elimination via micronucleation has been previously proposed in earlier reports [52,54,58].

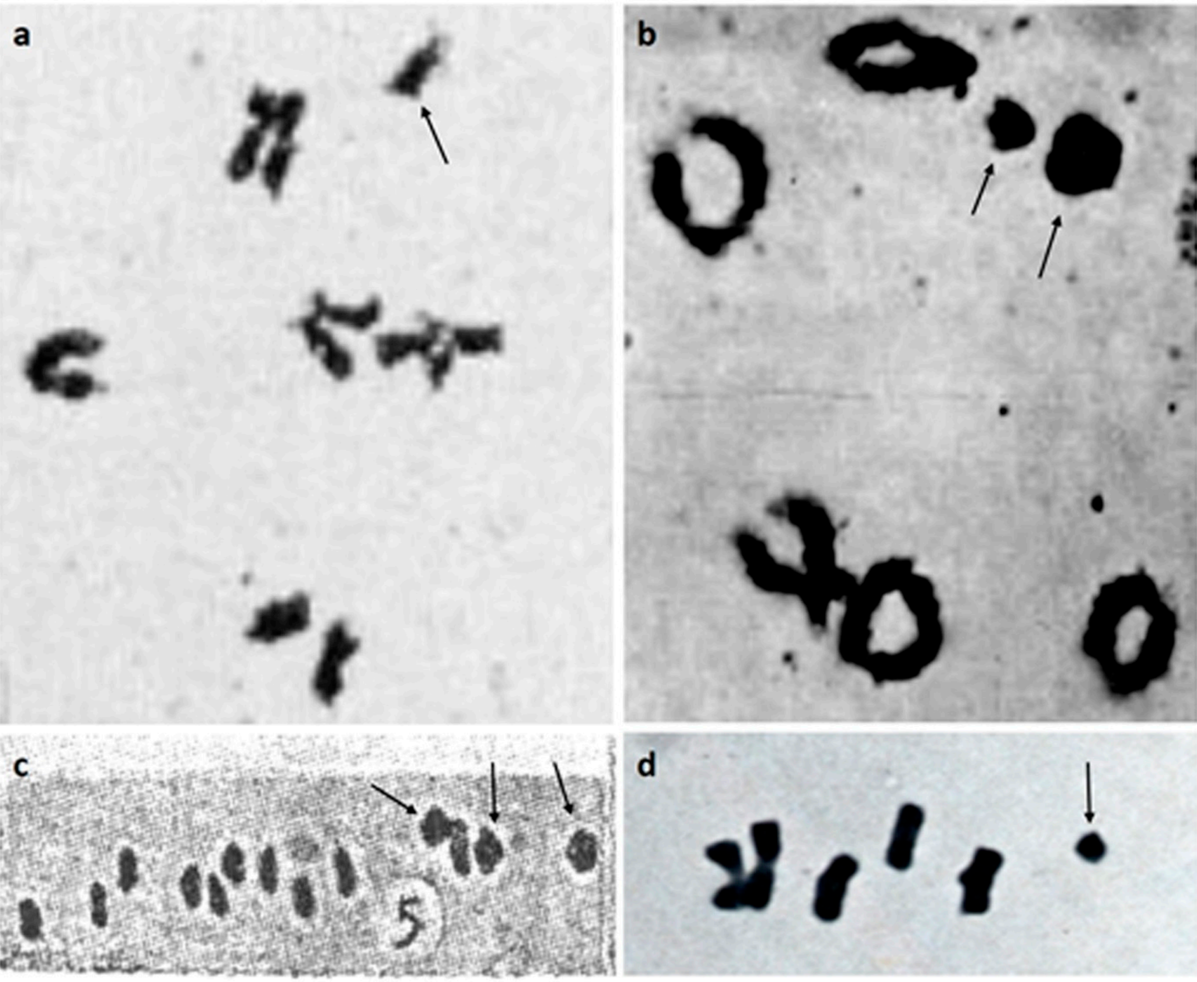

Figure 1. Chromosome pairing in meiotic metaphase I in pollen mother cells of *Sorghum* sp. Bs are marked by arrows. (**a**) Five A-bivalents and one B-univalent of *S. stipoideum* [52] (reprinted with permission from Springer Nature: Nature, Heredity, *B-chromosomes in Sorghum stipoideum, Wu*, Copyright 1992); (**b**) Five A-bivalents and two B-univalents in *S. purpureosericeum* [54] (copied from *Morbid mitosis and the activity of inert chromosomes in Sorghum*, Darlington and Thomas (1941) with the permission of the publisher); (**c**) Ten A-bivalents and three B-bivalents of *S. halepense* [56] (modified from *Paternal transmission of accessory chromosomes in a species of Eu-sorghum*, Raman et al. (1965)); (**d**) Five A-bivalents and one B-univalent of *S. nitidum* [57] (copied from *Accessory chromosome in Sorghum nitidum Pers.*, Wu and Pi (1975) with the permission of the publisher).

In *S. nitidum*, Raman and Krishnaswami [31] observed Bs in diploid plants (2n = 2x = 10), but not in tetraploids (2n = 4x = 20). The size of Bs was equal to the chromosomes of A complement. When two Bs were present, they paired regularly and behaved normally at meiosis. Wu and Pi [57], and later Wu [60], analyzed *S. nitidum* plants with one B chromosome (2n = 2x = 10 + 1B). They described B chromosome as an isochromosome, which folded back to pair with itself at the pachytene. The whole chromosome was heterochromatic with a terminal knob distal to the centromere, and its heterochromatic arm was separated from the knob by a constriction. This B chromosome was much shorter than any chromosome of the A-complement, which indicates that *S. nitidum* might contain more than one type of B chromosome.

The only study on the Bs in *S. stipoideum* was published by Wu [52]. The author found one type of B chromosome, which was distinctly shorter than any of the As and euchromatic

along its whole length. Its euchromatic nature is interesting in the context of the fact that the other Bs in the *Sorghum* genus are heterochromatic, and it implies transcriptional activity. However, the euchromatic nature of Bs is not striking, as, for instance, *Allium cernuum* and *Crepis pannonica* also carry partially or completely euchromatic Bs [61,62]. The B chromosome of *S. stipoideum* was described as an isochromosome, which exhibited inter-arm pairing when present only in one copy. At anaphase I of 1B plants, B chromosome divided precociously or moved undivided to one pole of the cell. During anaphase II, the majority of the cells had lagging B-chromatids and after division, micronuclei were observed, indicating B chromosome elimination. In 2B plants, meiotic behavior was nearly regular. However, numerical variability among and even within the spikelets was observed [52].

In *S. halepense*, four to six Bs have been observed, and their occurrence seems to be limited exclusively to diploids (2n = 2x = 20) [63]. The behavior of Bs during meiosis in pollen mother cells was aberrant; accessory bivalents exhibited delayed disjunction or nondisjunction leading to the subsequent elimination. Three types of Bs were reported, two of which showed partial homology [56,63].

The presence of B chromosome in *S. bicolor* ssp. *verticilliflorum* is rather questionable. Huskins and Smith [51] observed an additional pair of chromosome fragments during male meiosis in this species. These fragments were much smaller than As and were attached to a bivalent of As. They considered those fragments to be a pair of supernumerary chromosomes. However, as this study is the only existing work describing the presence of Bs in this species, it is questionable whether it was a real B chromosome or rather a mere chromosomal fragment. It has been documented that some chromosomes contain so-called "fragile sites" that are prone to breakage during cell division and are sensitive to replication stress [64]. In plants, these sites are associated with 45S rDNA. Fragile sites can lead to chromosomal rearrangements and affect genome organization [65–67]. In the case of *S. bicolor* ssp. *verticilliflorum*, fragile sites in some genotypes might lead to chromosomal fragments misinterpreted as Bs. Unfortunately, so far there has been no other research that could confirm or disprove the existence of Bs in this species.

3. Elimination and Maintenance of B Chromosomes

Although Bs are usually transmitted regularly in mitosis, sorghums belong to species where the transmission is irregular. Bs in genus *Sorghum* show a high level of numerical instability, which is frequently observed in somatic tissues. In all *Sorghum* species, B chromosome elimination or irregular transmission leads to a mosaic distribution of the Bs. In *S. purpureosericeum*, Janaki-Amal [53] reported B chromosome absence in roots. Darlington and Thomas [54] described B chromosome absence in root, stem, and leaf tissues in the same species. These findings correspond with the results of a recently published study [68], in which the authors analyzed parts of adult B-carrying plants in order to identify the tissues where the B chromosome is preserved. Except for the inflorescence, where Bs are stably present, the residual population of B-carrying nuclei was detected in leaf meristem, last node, and peduncle. Thus, in *S. purpureosericeum*, Bs probably persist only in the meristems from which generative organs are later established, and are likely to be eliminated from other vegetative tissues [68]. Similarly, B elimination from root tissue was noticed in *S. nitidum* [31] and *S. halepense* [56]. In microsporocytes and tapetal cells of *S. stipoideum*, Bs occurred mosaically, while root, stem, and leaf meristem cells were completely lacking Bs [52]. Recently, the process responsible for the elimination of B chromosome from the roots of *Aegilops speltoides* has been described [69]. The strictly controlled process is based on B chromatid nondisjunction in mitosis, lagging in anaphase, and the formation of a micronucleus, which is subsequently eliminated. Elimination mechanisms in *Sorghum* have not yet been investigated, but they might be similar to *Aegilops*.

Research on other B-containing species indicates that the existence of accumulating mechanisms is necessary to avoid the loss of Bs [70]. The Bs in genus *Sorghum* also undoubtedly had to evolve some accumulation mechanism(s) acting directly against natural

selection. However, our knowledge of these multiplicative mechanisms is only fragmentary. If we consider possible divisions where the nondisjunction can take place, the meiotic drive can be ruled out, as the division of pollen mother cells was proven to be regular [54,58]. As pre-meiotic drive is generally rare, attention should be focused mainly on pollen mitosis. A solitary study of Darlington and Thomas [54] is the only work dealing with the division following male meiosis. The authors did not find any irregularities in first pollen mitosis and suggested that nondisjunction occurs during second pollen division. Conclusions presented in this work were drawn from the statistical analysis of progenies of B-carrying plants, and the study lacks strong proof of this hypothesis. This approach, based on the analysis of the frequency and number of Bs in the offspring, can be replaced today by technologically advanced methods, which enable the visualization of the B chromosome in situ, directly on its way through both pollen mitoses. These modern approaches have already been used to elucidate the mechanism of nondisjunction in rye or *Aegilops* [12,71]. Markers recently developed for *S. purpureosericeum* [68] open up the possibility of also using these approaches in *Sorghum*.

The tissue-specific elimination of Bs complicates their detection in growing plants. In the sorghum model, the detection of Bs requires the cultivation of the plant up to the stage of inflorescence, when immature anthers are collected and the meiocytes are scored at metaphase I, when the presence of Bs can be determined. In species with a specific proportion of A and B chromosomes, an alternative approach based on flow cytometry can be used (Figure 2) [68]. Although the flow cytometry screening method also requires the inflorescence, it is less laborious and thus facilitates and speeds up the whole detection process. The protocol was originally established for *S. purpureosericeum* and worked well both for isolated haploid nuclei from pollen grains and for samples prepared from whole florets. However, this approach is not suitable for the detection of Bs in very young seedling/seeds, as it requires a relatively large amount of material and thus is destructive. The flow cytometry approach has previously been used to detect Bs in leaves and immature embryos of *Aegilops speltoides* [69,71]. The evaluation of B status in developing seeds would make the work significantly faster, however, unfortunately, this kind of approach based on PCR markers is not currently available.

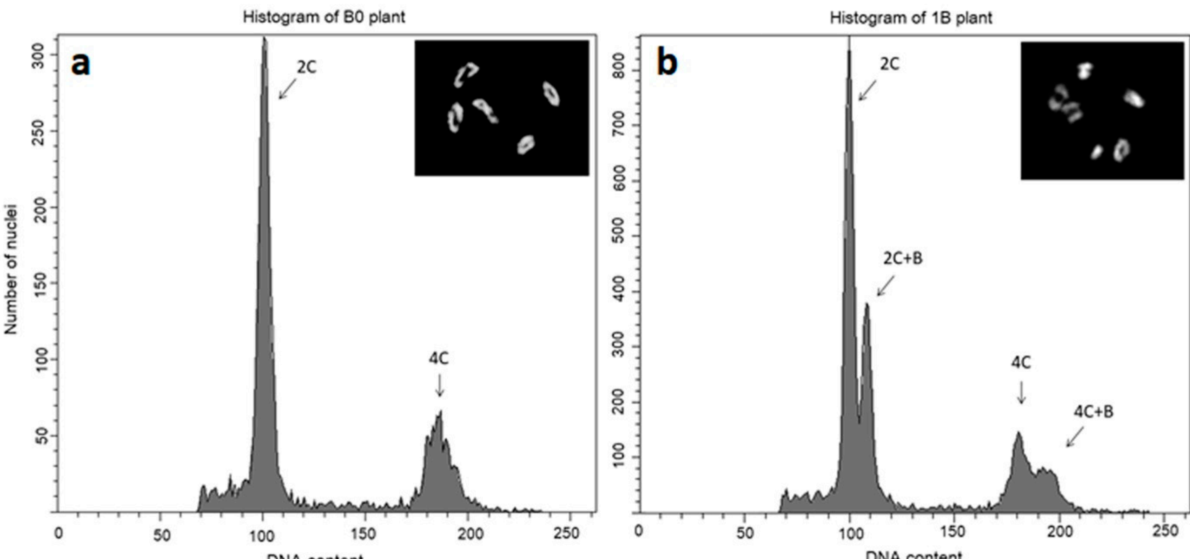

Figure 2. Flow cytometric analysis of the nuclei of *Sorghum purpureosericeum* isolated from spikelets. The cytological verification of B chromosome presence/absence in the analyzed plant is shown in the inset. (**a**) Histogram of B-negative plant showing two distinct peaks corresponding to 2C and 4C nuclei; (**b**) histogram of B-positive plant with a significant change in the flow karyotype. Even one copy of the B chromosome results in a clear separation of the populations of nuclei carrying B chromosome at both 2C and 4C ploidies.

Cytological techniques play an irreplaceable role in B chromosome research, but only molecular studies are able to provide us with information that is above the resolution of cytogenetics. Sequencing is a powerful tool that can bring us information about the genomic content, origin, evolution, and biological role of Bs. Next-generation sequencing, enabling large-scale analysis, provides an opportunity for a significant advance in B chromosome research. However, out of the five types of Bs reported in *Sorghum*, only the B chromosome in *S. purpureosericeum* has been subjected to molecular studies so far [68]. Sequence analysis has revealed several B-specific repeats in this chromosome, including DNA transposon/hAT and one LINE element. Based on the selected repetitive sequences, PCR and cytogenetic markers specific for B chromosome have been developed [68]. The accumulation of different types of repeats in Bs is common; these repeats are often strongly amplified and may even form a significant part of the B chromosome, like the PSR element in *Nasonia vitripennis* [72] or micro B of *Brachycome dichromosomatica* [73]. B-specific repeats have also been identified in other plant species, such as E3900 and D1100 in rye [10,13], ZmBs and StarkB in maize [7,74], and Bd49 in *B. dichromosomatica* [75].

4. Did B Chromosomes Emerge Several Times in the Genus *Sorghum*?

Sorghum is a genus of monocot flowering plants in the grass family Poaceae, subfamily Panicoideae, and the tribe Andropogoneae. The genus includes 23 annual and perennial species and a number of subspecies and races resulting from hybridization. Based on morphological traits, they are divided into five subgenera: *Sorghum*, *Parasorghum*, *Stiposorghum*, *Chaetosorghum*, and *Heterosorghum* (Figure 3) [76]. Subgenus *Sorghum* is represented by cultivated sorghum (*Sorghum bicolor* (L.) Moench) and its wild relatives. Representatives of this subgenus originated in Africa and Asia, and their chromosome numbers are 2n = 2x = 20 in diploids and 2n = 4x = 40 in tetraploids [77]. The subgenus *Parasorghum* includes seven species from Australia, Central America, Africa, and Asia [77], and their chromosome numbers vary from 2n = 2x = 10, 20, 30 to 40. *S. macrospermum* (2n = 2x = 40) is the only representative of the subgenus *Chaetosorghum* and can be found endemically in the Northern Territory of Australia. Subgenus *Heterosorghum* is represented by *S. laxiflorum*, growing in Northern Australia and Papua New Guinea (2n = 2x = 40). The last subgenus, *Stiposorghum*, includes 10 species occurring in Northern Australia with chromosome numbers ranging from 2n = 2x = 10, 20, 30 to 40 [77].

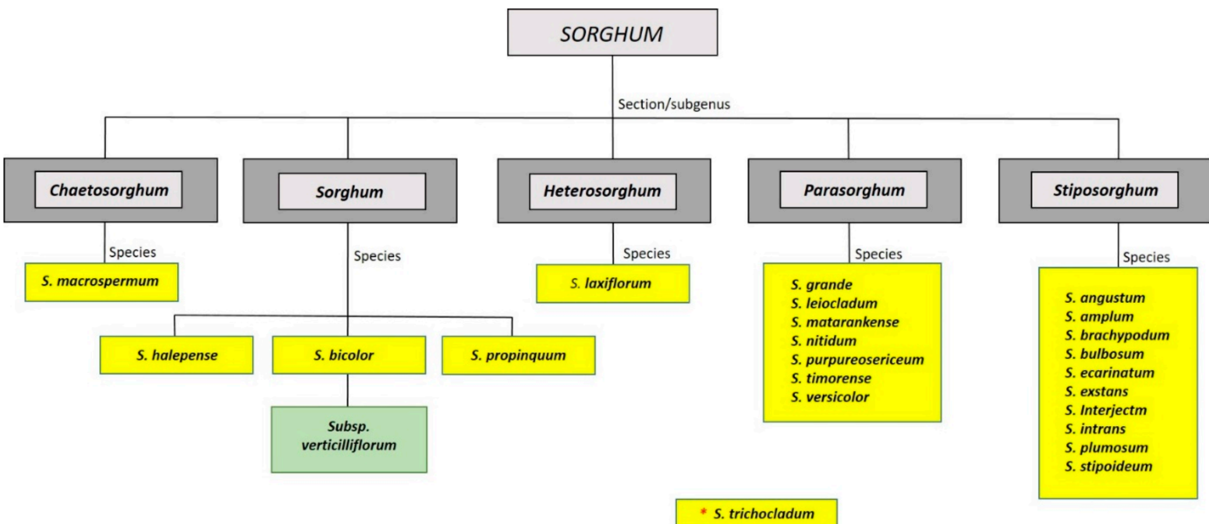

Figure 3. Classification of *Sorghum* (modified from Ananda et al. [78]) * *S. trichocladum* has not yet been assigned to any subgenus.

The genus *Sorghum* has been subjected to several phylogenetic analyses [79–82], most of which agree with this classification of species into the abovementioned sections, although relationships between some sister taxa are still under debate. For example, a

close relationship between *S. macrospermum* and *S. laxiflorum* led to proposals to merge the *Chaetosorghum* and *Heterosorghum* sections [80–82]. We have performed phylogenetic reconstruction of the genus with a focus on species possessing Bs. PhyML analysis of concatenated sequences of ITS1-ITS2, trnH-psbA, and trnL-trnF resulted in a phylogram in which two strongly supported major clades I and II were identified (Figure 4), which is in agreement with the phylogenetic analyses published previously.

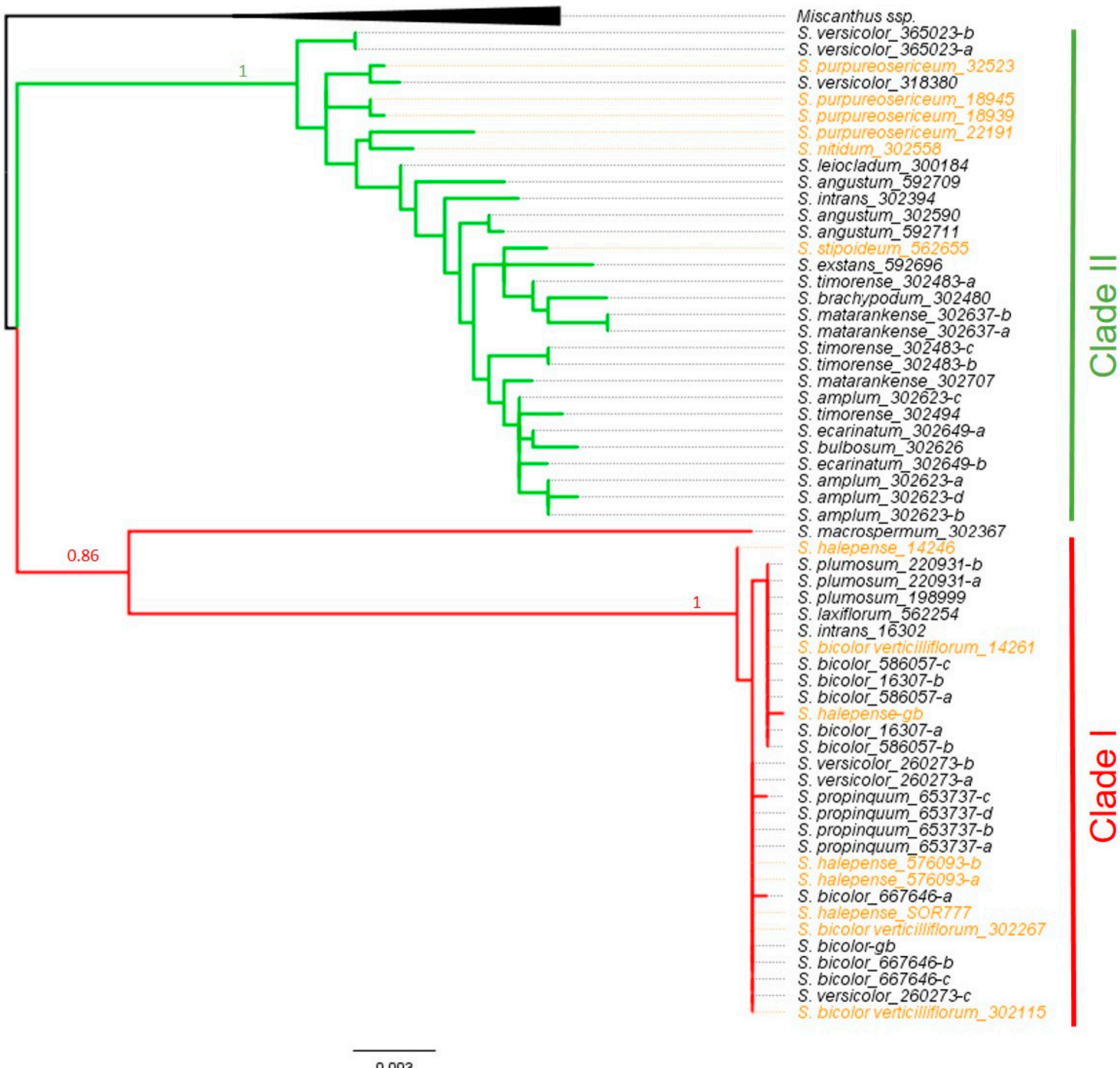

Figure 4. Phylogenetic tree of 21 *Sorghum* species based on phyML analysis using concatenated nuclear and chloroplastic sequences (ITS1-ITS2-trnHpsbA-trnLtrnF). With strong branch support, two main clades I and II are resolved. B-carrying species are marked in orange.

A widely accepted hypothesis of the formation of Bs assumes that they have autosomal origin. This is supported by the fact that sequences similar to those from As have often been found on Bs. Bs containing mosaically organized sequences derived from different As have been described, for example, in rye [39], maize [83], and *Brachycome dichromosomatica* [73]. In *Nasonia vitripennis*, B chromosome was formed from interspecies hybridization [84]. Additionally, Bs originating from sex chromosomes were described in grasshopper *Eyprepocnemis plorans* and frog *Leiopelma hochstetteri* [85,86]. Despite countless studies dealing with the possible origin and evolution of B chromosome(s) [7,39,84,87–90], these issues still remain unclear.

There are three possible evolutionary scenarios for B chromosome(s) in the genus *Sorghum* (Figure 5). B chromosome(s) may have formed (1) in a single event in a common ancestor of all *Sorghum* species and then have been preserved in some lines during evolution and disappeared in others (Figure 5a); (2) once in the ancestor of closely related species, all of which have kept Bs up until today (Figure 5b); or (3) several times during independent events in *Sorghum* evolution, which includes also the possibility that the B chromosome originated only at the level of individual species (Figure 5c).

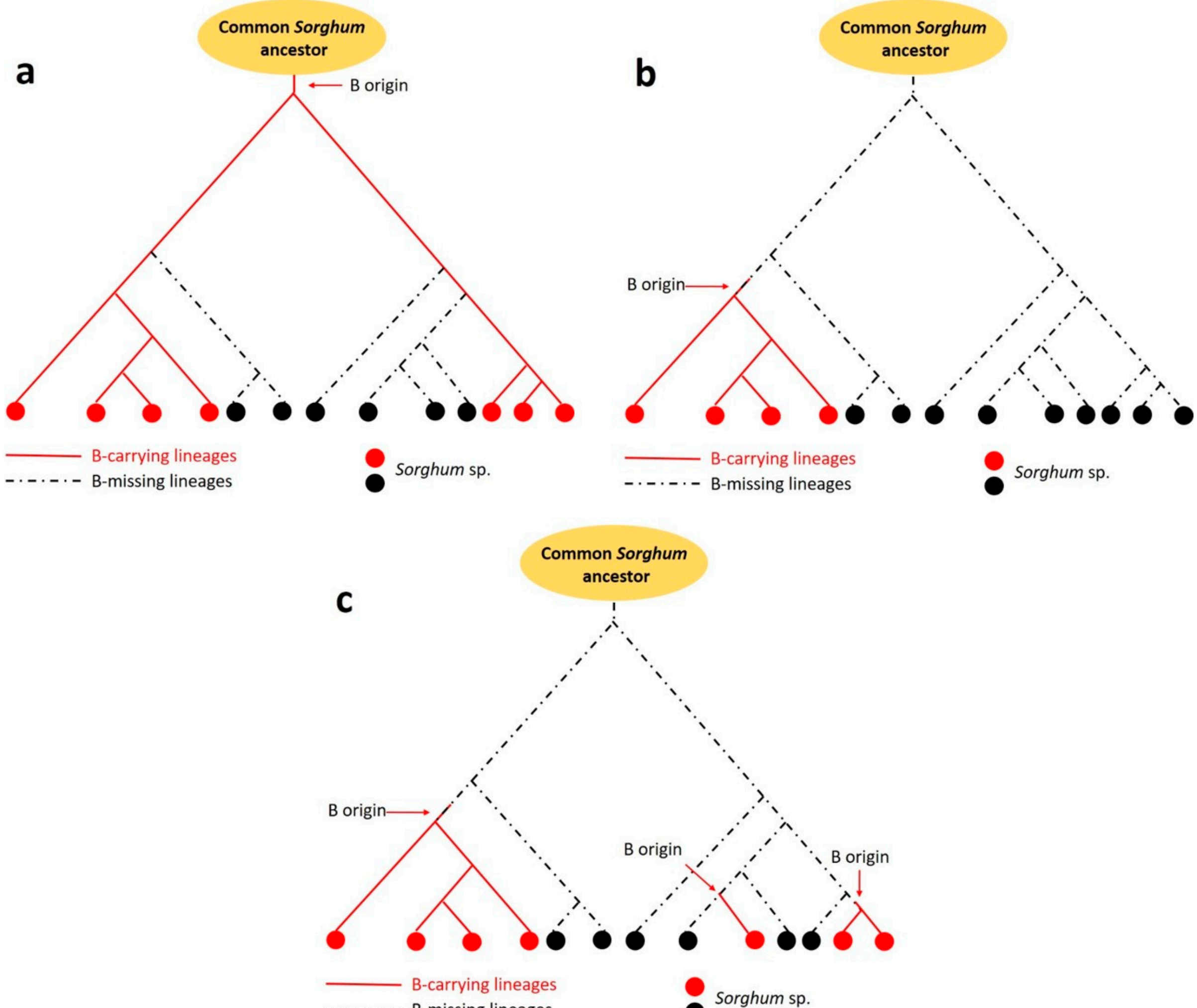

Figure 5. Schemes of the possible origin of B chromosome in the genus *Sorghum*. (**a**) Origin of B chromosome(s) in a common ancestor; (**b**) origin of B chromosome(s) in a group of closely related species; (**c**) multiple origins of B chromosome(s).

Considering the conclusions of the cytological studies [31,51–60,63], it is clear that there are several types of Bs within the genus, which differ in their morphology. One evolutionary scenario assumes the formation of the B chromosome in a single event and its persistence in closely related species. Given that Bs occur in both clades (Figure 4), we can exclude a close phylogenetic relationship between the species carrying Bs. However, the phylogenetic analysis cannot rule out the possibility that B chromosome originated only once during the early evolution of *Sorghum* or in the ancestor of the genus. This hypothesis could be supported or questioned based on the analysis of sequence similarity between the Bs from different species in the genus *Sorghum*, however, these data are not currently available. Recently, Wu et al. [71] identified B-specific tandem repeat shared by Bs in *Aegilops speltoides*, *Aegilops mutica*, and *Secale cereale*, however, they were not able to

conclude whether the chromosomes have a common origin or whether the shared repetitive sequence is a result of the exchange of genetic material among those species.

The hypothesis assuming multiple independent origins of Bs in the genus *Sorghum* is also feasible. This scenario of B chromosome origin appears to be supported by the nature of the B chromosome of *S. nitidum*—the appearance of B in this species might represent an example of the formation of the B chromosome at the species level, as has already been hypothesized by Wu and Pi [57] and Wu [60]. The B chromosome in *S. nitidum* is an isochromosome, and its arms strikingly resemble the short arm of the nucleolus-associated chromosome, which is also entirely heterochromatic and approximately similar in length. Since B chromosome does not pair with the nucleolus-associated chromosome, further structural and genic changes had to occur later, leading to the loss of homology and to the inability of B chromosome to pair with the short arm of the nucleolus-associated chromosome [60]. Such resemblance between Bs and As has not been found in any other *Sorghum* species. However, since Bs are expected to be prone to aberrations and the accumulation of mutations [11,91], a relatively rapid diversification and loss of ability to pair with the original A homologue can be assumed. It is possible that the B chromosome of *S. nitidum* is evolutionarily younger than other *Sorghum* Bs, and therefore a high level of similarity between A and B is still maintained in this species.

Another argument speaking for the independent origin of Bs in different *Sorghum* species is the possibility of intra- and inter-specific hybridization. It has been suggested that if hybridization occurs between separated, diverged subpopulations, various irregularities in meiosis may appear, which is a precondition for the B chromosome formation [60]. Namely, the subgenus *Sorghum* is a complex group that includes a number of closely related species, subspecies, and races that can interbreed freely, and some species (e.g., *S. halepense*) are assumed to have hybrid origin [92–94]. Frequent hybridization might represent conditions favorable for the formation of Bs. The origin of Bs through interspecific hybridization has been demonstrated in hybrid derivatives from spontaneous crossing between two *Coix* species [95] and has also been described in *Poecilia formosa* and *Nasonia vitripennis* [84,96,97].

5. Conclusions

B chromosomes (Bs) are unique genomic elements with a transmission rate higher than 0.5. Although our knowledge of Bs has advanced considerably since their discovery in the first half of the last century, many questions remain unanswered. The elusiveness of *Sorghum* Bs, resulting from their extensive elimination during early plant development, certainly contributed to the fact that they have not yet been subjected to any comprehensive research. In the last century, several authors have provided cytological characteristics of Bs in some *Sorghum* species, but since then the research in this field has hardly progressed. However, new technologies give us new opportunities to meet this challenge. The use of flow cytometric screening and sorting makes it easier to detect the presence of Bs and to obtain material for sequencing, which will be necessary for their thorough molecular characterization. So far, our knowledge of the mechanisms of accumulation or somatic elimination of Bs in *Sorghum* is only marginal. Analysis of the effects of *Sorghum* Bs on gene expression is another interesting topic that deserves thorough investigation. Research on Bs in *Sorghum* is still in its infancy, and there is a long way to go before we discover at least some of the mechanisms behind the unique behavior of these enigmatic elements in this genus.

Author Contributions: M.B., M.K., E.H. and J.B. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Czech Science Foundation (grant award no. 18-12338Y) and the European Regional Development Fund project "Plants as a tool for sustainable global development" (no. CZ.02.1.01/0.0/0.0/16_019/0000827).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported by the Czech Science Foundation (grant award no. 18-12338Y) and the European Regional Development Fund project "Plants as a tool for sustainable global development" (no. CZ.02.1.01/0.0/0.0/16_019/0000827). We thank the Australian Grains Genebank, the United States Department of Agriculture and International Crops Research Institute for the Semi-Arid Tropics for providing seed stocks, and the publisher Royal Society for permission to reprint Figure 1b. Further, we thank our colleague Iva Ilíková for language editing of the manuscript.

Conflicts of Interest: All authors declare no conflict of interest.

References

1. D'Ambrosio, U.; Alonso-Lifante, M.P.; Barros, K.; Kovařík, A.; Mas de Xaxars, G.; Garcia, S. B-chrom: A database on B-chromosomes of plants, animals and fungi. *New Phytol.* **2017**, *216*, 635–642. [CrossRef]
2. Jones, R.N. B chromosomes in plants. *New Phytol.* **1995**, *131*, 411–434. [CrossRef]
3. Camacho, J.P.M.; Sharbel, T.F.; Baukeboom, L.W. B-chromosome evolution. *Phil. Trans. R. Soc. Lond. B* **2000**, *355*, 163–178. [CrossRef] [PubMed]
4. Baukeboom, L.W. Bewildering Bs: An impression of the 1st B-chromosome conference. *Heredity* **1994**, *73*, 328–336. [CrossRef]
5. Werren, J.H.; Stouthamer, R. PSR (paternal sex ratio) chromosomes: The ultimate selfish genetic elements. *Genetica* **2003**, *117*, 85–101. [CrossRef] [PubMed]
6. Carlson, W.R. A procedure for localizing genetic factors controlling mitotic nondisjunction in the B chromosome of maize. *Chromosoma* **1973**, *42*, 127–136. [CrossRef]
7. Alfenito, M.R.; Birchler, J.A. Molecular characterization of a maize B chromosome centric sequence. *Genetics* **1993**, *135*, 589–597. [CrossRef] [PubMed]
8. Stark, E.A.; Connerton, I.; Bennet, S.T.; Barnes, S.R.; Parker, J.S.; Forster, J.W. Molecular analysis of the structure of the maize B-chromosome. *Chromosome Res.* **1996**, *4*, 15–23. [CrossRef] [PubMed]
9. Jin, W.; Lamb, J.C.; Vega, J.M.; Dawe, R.K.; Birchler, J.A.; Jiang, J. Molecular and functional dissection of the maize B chromosome centromere. *Plant Cell* **2005**, *17*, 1412–1423. [CrossRef] [PubMed]
10. Sandery, M.J.; Forster, J.W.; Blunden, R.; Jones, R.N. Identification of a family of repeated sequences on the rye B chromosome. *Genome* **1990**, *33*, 908–913. [CrossRef]
11. Klemme, S.; Banaei-Moghaddam, A.M.; Macas, J.; Wicker, T.; Novák, P.; Houben, A. High-copy sequences reveal distinct evolution of the rye B chromosome. *New Phytol.* **2013**, *199*, 550–558. [CrossRef] [PubMed]
12. Banaei-Moghaddam, A.M.; Schubert, V.; Kumke, K.; Wei, O.; Klemme, S.; Nagaki, K.; Macas, J.; González-Sánchez, M.; Heredia, V.; Gómez-Revilla, D.; et al. Nondisjunction in favor of a chromosome: The mechanism of rye B chromosome drive during pollen mitosis. *Plant Cell* **2012**, *24*, 4124–4134. [CrossRef]
13. Blunden, R.; Wilkes, T.J.; Forster, J.W.; Jimenez, M.M.; Sandery, M.J.; Karp, A.; Jones, R.N. Identification of the E3900 family, a second family of rye B chromosome specific repeated sequences. *Genome* **1993**, *36*, 706–711. [CrossRef] [PubMed]
14. Vinogradov, A.E. Mirrored genome size distributions in monocot and eudicot plants. *Acta Biotheor.* **2001**, *49*, 43–51. [CrossRef]
15. Trivers, R.; Burt, A.; Palestis, B.G. B chromosomes and genome size in flowering plants. *Genome* **2004**, *47*, 1–8. [CrossRef] [PubMed]
16. Palestis, B.G.; Trivers, R.; Burt, A.; Jones, R.N. The distribution of B chromosomes across species. *Cytogenet. Genome Res.* **2004**, *106*, 151–158. [CrossRef]
17. Levin, D.A.; Palestis, B.G.; Jones, R.N.; Trivers, R. Phyletic hotspots for B chromosomes in angiosperms. *Evolution* **2005**, *59*, 962–969. [CrossRef] [PubMed]
18. Jones, R.N. B-Chromosome Drive. *Am. Nat.* **1991**, *137*, 430–442. [CrossRef]
19. Müntzing, A. Cytological studies of extra fragment chromosomes in rye; the mechanism of non-disjunction at the pollen mitosis. *Hereditas* **1946**, *32*, 97–119. [CrossRef]
20. Roman, H. Mitotic nondisjunction in the case of interchanges involving the B-type chromosome in maize. *Genetics* **1947**, *32*, 391–409.
21. Carlson, W.R.; Chou, T.S. B chromosome nondisjunction in corn: Control by factors near the centromere. *Genetics* **1981**, *97*, 379–389.
22. Han, F.; Lamb, J.C.; Yu, W.; Gao, Z.; Birchler, J.A. Centromere function and nondisjunction are independent components of the maize B chromosome accumulation mechanism. *Plant Cell* **2007**, *19*, 524–533. [CrossRef]
23. Roman, H. Directed fertilization in maize. *Proc. Natl. Acad. Sci. USA* **1948**, *34*, 36–42. [CrossRef]
24. Carlson, W.R. Factors affecting preferential fertilization in maize. *Genetics* **1969**, *62*, 543–554. [CrossRef] [PubMed]
25. Rusche, M.L.; Mogensen, H.L.; Shi, L.; Keim, P.; Rougier, M.; Chaboud, A.; Dumas, C. B chromosome behavior in maize pollen as determined by a molecular probe. *Genetics* **1997**, *147*, 1915–1921. [CrossRef] [PubMed]
26. Hasegawa, N. A cytological study on 8-chromosome rye. *Cytologia* **1934**, *6*, 68–77. [CrossRef]
27. Nur, U. A mitotically unstable supernumerary chromosome with an accumulation mechanism in a grasshopper. *Chromosoma* **1963**, *14*, 407–422. [CrossRef] [PubMed]
28. Kayano, H. Accumulation of B chromosomes in the germ line of *Locusta migratoria*. *Heredity* **1971**, *27*, 119–123. [CrossRef]

29. Kayano, H. Cytogenetic studies in *Lilium callosum*. III. Preferential segregation of a supernumerary chromosome in EMCs. *Proc. Jpn. Acad.* **1957**, *33*, 553–558. [CrossRef]
30. Nur, U. A supernumerary chromosome with an accumulation mechanism in the lecanoid genetic system. *Chromosoma* **1962**, *13*, 249–271. [CrossRef]
31. Raman, V.S.; Krishnaswami, D. Accessory chromosomes in sorghum nitidum Pers. *J. Indian Bot. Soc.* **1960**, *39*, 278–280.
32. Zuk, J. The additional heterochromatic chromosome and its influence on sex chromosome pairing in *Rumex*. *Heredity* **1969**, *24*, 69–74. [CrossRef]
33. Holmes, D.S.; Bougourd, S.M. B-chromosome selection in *Allium schoenoprasum* II. Experimental populations. *Heredity* **1991**, *67*, 117–122. [CrossRef]
34. Bakkali, M.; Cabrero, J.; López-León, M. The B chromosome polymorphism of the grasshopper *Eyprepocnemis plorans* in North Africa. I. B variants and frequency. *Heredity* **1999**, *83*, 428–434. [CrossRef] [PubMed]
35. De Marco Ferro, D.A.; Moreira-Filho, O.; Bertollo, L.A.C. B chromosome polymorphism in the fish, *Astyanax scabripinnis*. *Genetica* **2003**, *119*, 147–153. [CrossRef]
36. Plowman, A.; Bougourd, S. Selectively advantageous effects of B chromosomes on germination behaviour in *Allium schoenoprasum* L. *Heredity* **1994**, *72*, 587–593. [CrossRef]
37. Dherawattana, A.; Sadanaga, K. Cytogenetics of a crown rust-resistant hexaploid oat with 42 + 2 fragment chromosomes1. *Crop Sci.* **1973**, *13*, 591–594. [CrossRef]
38. Han, Y.; Liu, X.; Benny, U.; Kistler, H.C.; VanEtten, H.D. Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. *Plant J.* **2001**, *25*, 305–314. [CrossRef]
39. Martis, M.M.; Klemme, S.; Banaei-Moghaddam, A.M.; Blattner, F.R.; Macas, J.; Schmutzer, T.; Scholz, U.; Gundlach, H.; Wicker, T.; Simkova, H.; et al. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13343–13346. [CrossRef]
40. Page, B.T.; Wanous, M.K.; Birchler, J.A. Characterization of a maize chromosome 4 centromeric sequence: Evidence for an evolutionary relationship with the B chromosome centromere. *Genetics* **2001**, *159*, 291–302.
41. Cheng, Y.M.; Lin, B.Y. Cloning and characterization of maize B chromosome sequences derived from microdissection. *Genetics* **2003**, *164*, 299–310. [PubMed]
42. Banaei-Moghaddam, A.M.; Meier, K.; Karimi-Ashtiyani, R.; Houben, A. Formation and expression of pseudogenes on the B chromosome of rye. *Plant Cell* **2013**, *25*, 2536–2544. [CrossRef] [PubMed]
43. Houben, A.; Banaei-Moghaddam, A.M.; Klemme, S.; Timmis, J.N. Evolution and biology of supernumerary B chromosomes. *Cell. Mol. Life Sci.* **2014**, *71*, 467–478. [CrossRef]
44. Makunin, A.I.; Kichigin, I.G.; Larkin, D.M.; O'Brien, P.C.; Ferguson-Smith, M.A.; Yang, F.; Proskuryakova, A.A.; Vorobieva, N.V.; Chernyaeva, E.N.; O'Brien, S.J.; et al. Contrasting origin of B chromosomes in two cervids (Siberian roe deer and grey brocket deer) unravelled by chromosome-specific DNA sequencing. *BMC Genom.* **2016**, *17*, 618. [CrossRef]
45. Miao, V.P.; Covert, S.F.; Van Etten, H.D. A fungal gene for antibiotic resistance on a dispensable ("B") chromosome. *Science* **1991**, *254*, 1773–1776. [CrossRef] [PubMed]
46. Valente, G.T.; Conte, M.A.; Fantinatti, B.E.A.; Cabral-de-Mello, D.C.; Carvalho, R.F.; Vicari, M.R.; Kocher, T.D.; Martins, C. Origin and evolution of B chromosomes in the cichlid fish *Astatotilapia latifasciata* based on integrated genomic analyses. *Mol. Biol. Evol.* **2014**, *31*, 2061–2072. [CrossRef] [PubMed]
47. Ma, W.; Gabriel, T.S.; Martis, M.M.; Gursinsky, T.; Schubert, V.; Vrána, J.; Doležel, J.; Grundlach, H.; Altschmied, L.; Scholz, U.; et al. Rye B chromosomes encode a functional Argonaute-like protein with in vitro slicer activities similar to its A chromosome paralog. *New Phytol.* **2017**, *213*, 916–928. [CrossRef] [PubMed]
48. Navarro-Domínguez, B.; Ruiz-Ruano, F.; Cabrero, J.; Corral, J.M.; López-León, M.D.; Sharbel, T.F.; Camacho, J.P.M. Protein-coding genes in B chromosomes of the grasshopper *Eyprepocnemis plorans*. *Sci. Rep.* **2017**, *7*, 45200.
49. Yoshida, K.; Terai, Y.; Mizoiri, S.; Aibara, M.; Nishihara, H.; Watanabe, M.; Kuroiwa, A.; Hirai, H.; Hirai, Y.; Matsuda, Y.; et al. B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. *PLoS Genet.* **2011**, *7*, e100220. [CrossRef]
50. Makunin, A.I.; Romanenko, S.A.; Beklemisheva, V.R.; Perelman, P.L.; Druzhkova, A.S.; Petrova, K.O.; Prokopov, D.Y.; Chernyaeva, E.N.; Johnson, J.L.; Kukekova, A.V.; et al. Sequencing of supernumerary chromosomes of red fox and raccoon dog confirms a non-random gene acquisition by B chromosomes. *Genes* **2018**, *9*, 405. [CrossRef]
51. Huskins, C.L.; Smith, S.G. A cytological study of the genus *Sorghum* PERS. II. The meiotic chromosomes. *J. Genet.* **1934**, *28*, 387–395. [CrossRef]
52. Wu, T. B-chromosomes in *Sorghum stipoideum*. *Heredity* **1992**, *68*, 457–463. [CrossRef]
53. Janaki-Ammal, E. Chromosome Diminution in a Plant. *Nature* **1940**, *146*, 839–840. [CrossRef]
54. Darlington, C.D.; Thomas, P.T. Morbid mitosis and the activity of inert chromosomes in sorghum. *Proc. R. Soc. Lond. B* **1941**, *130*, 127–150.
55. Raman, V.S.; Meenakshi, K.; Thangam, M.S.; Sivagnanam, L. The cytological behaviour of B-chromosomes in *Sorghum halepense*. *Abst. Madras Agric. J.* **1964**, *51*, 72–73.
56. Raman, V.S.; Thangam, M.S. Paternal transmission of accessory chromosomes in a species of *Eu-Sorghum*. *Sci. Cult.* **1965**, *31*, 150–151.
57. Wu, T.P.; Pi, C.P. Accessory chromosome in *Sorghum nitidum* Pers. *Taiwania* **1975**, *20*, 147–161.
58. Wu, T.P. B chromosomes in Sorghum purpureo-sericeum. *Proc. Natl. Sci. Counc. B* **1984**, *8*, 198–209.

59. D'cruz, R.; Deshmukh, J.N. Behaviour of B chromosomes in *Sorghum purpureosericeum*. *Poona Agric. Col. Mag.* **1960**, *51*, 30–34.
60. Wu, T.P. Pachytene morphology of *Sorghum nitidum* chromosome complement. *Cytologia* **1978**, *43*, 433–440. [CrossRef]
61. Friebe, B. Nucleolar activity of B-chromosomes in *Allium cernuum* (*Alliaceae*). *Plant Syst. Evol.* **1989**, *163*, 87–92. [CrossRef]
62. Fröst, S.; Östergren, G. *Crepis pannonica* and *Crepis conyzaefolia*—Two more species having accessory chromosomes. *Hereditas* **1959**, *45*, 211–214. [CrossRef]
63. Raman, V.S.; Meenakshi, K.; Thangam, M.S. Accessory chromosomes and their meiotic behaviour in hybrids of grain sorghum and Johnson grass. *Cytologia* **1976**, *41*, 193–200. [CrossRef]
64. Durkin, S.G.; Glover, T.W. Chromosome fragile sites. *Annu. Rev. Genet.* **2007**, *41*, 169–192. [CrossRef] [PubMed]
65. Huang, J.; Ma, L.; Yang, F.; Fei, S.; Li, L. 45S rDNA regions are chromosome fragile sites expressed as gaps in vitro on metaphase chromosomes of root-tip meristematic cells in Lolium spp. *PLoS ONE* **2008**, *3*, e2167. [CrossRef] [PubMed]
66. Bustamante, F.O.; Rocha, L.C.; Torres, G.A.; Davide, L.C.; Mittelmann, A.; Techio, V.H. Distribution of rDNA in diploid and polyploid *Lolium multiflorum* Lam. and fragile sites in 45S rDNA regions. *Crop Sci.* **2014**, *54*, 617–625. [CrossRef]
67. Lan, H.; Chen, C.L.; Miao, Y.; Yu, C.X.; Guo, W.W.; Xu, Q.; Deng, X.X. Fragile sites of 'Valencia' sweet orange (*Citrus sinensis*) chromosomes are related with active 45s rDNA. *PLoS ONE* **2016**, *11*, e0151512. [CrossRef]
68. Karafiátová, M.; Bednářová, M.; Said, M.; Čížková, J.; Holušová, K.; Blavet, N.; Bartoš, J. B chromosome of *Sorghum purpureosericeum* reveals the first pieces of its sequence. *J. Exp. Bot.* **2020**, *72*, 1606–1616. [CrossRef]
69. Ruban, A.; Schmutzer, T.; Wu, D.D.; Fuchs, J.; Boudichevskaia, A.; Rubtsova, M.; Pistrick, K.; Melzer, M.; Himmelbach, A.; Schubert, V.; et al. Supernumerary B chromosomes of *Aegilops speltoides* undergo precise elimination in roots early in embryo development. *Nat. Commun.* **2020**, *11*, 2764. [CrossRef]
70. Jones, R.N.; Gonzáles-Sanchez, M.; González-García, M.; Vega, J.M.; Puertas, M.J. Chromosome with their life of their own. *Cytogenet. Genome Res.* **2008**, *120*, 265–280. [CrossRef]
71. Wu, D.; Ruban, A.; Fuchs, J.; Macas, J.; Novák, P.; Vaio, M.; Zhou, Y.; Houben, A. Nondisjunction and unequal spindle organization accompany the drive of *Aegilops speltoides* B chromosomes. *New Phytol.* **2019**, *223*, 1340–1352. [CrossRef]
72. Eickbush, D.G.; Eickbush, T.H.; Werren, J.H. Molecular characterization of repetitive DNA sequences from a B chromosome. *Chromosoma* **1991**, *101*, 575–583. [CrossRef] [PubMed]
73. Houben, A.; Verlin, D.; Leach, C.R.; Timmis, J.N. The genomic complexity of micro B chromosomes of *Brachycome dichromosomatica*. *Chromosoma* **2001**, *110*, 451–459. [CrossRef] [PubMed]
74. Lamb, J.C.; Meyer, J.M.; Corcoran, B.; Kato, A.; Han, F.; Birchler, J.A. Distinct chromosomal distributions of highly repetitive sequences in maize. *Chromosome Res.* **2007**, *15*, 33–49. [CrossRef] [PubMed]
75. Franks, T.K.; Houben, A.; Leach, C.R.; Timmis, J.N. The molecular organisation of a B chromosome tandem repeat sequence from *Brachycome dichromosomatica*. *Chromosoma* **1996**, *105*, 223–230. [CrossRef] [PubMed]
76. Garber, E.D. Cytotaxonomic studies in the genus *Sorghum*. *Univ. Calif. Publ. Bot.* **1950**, *23*, 283–361.
77. Lazarides, M.; Hacker, J.B.; Andrew, M.H. Taxonomy, cytology and ecology of indigenous Australian sorghums (*Sorghum* Moench: Andropogoneae: Poaceae). *Aust. Syst. Bot.* **1991**, *4*, 591–635. [CrossRef]
78. Ananda, G.; Myrans, H.; Norton, S.L.; Gleadow, R.; Furtado, A.; Henry, R.J. Wild *Sorghum* as a promising resource for crop improvement. *Front. Plant Sci.* **2020**, *11*, 1108. [CrossRef]
79. Sun, Y.; Skinner, D.Z.; Liang, G.H.; Hulbert, S.H. Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theor. Appl. Genetics* **1994**, *89*, 26–32. [CrossRef]
80. Ng'uni, D.; Geleta, M.; Fatih, M.; Bryngelsson, T. Phylogenetic analysis of the genus *Sorghum* based on combined sequence data from cpDNA regions and ITS generate well-supported trees with two major lineages. *Ann. Bot.* **2010**, *105*, 471–480. [CrossRef] [PubMed]
81. Liu, Q.; Liu, H.; Wen, J.; Peterson, P.M. Infrageneric phylogeny and temporal divergence of *Sorghum* (Andropogoneae, Poaceae) based on low-copy nuclear and plastid sequences. *PLoS ONE* **2014**, *9*, e104933. [CrossRef]
82. Hawkins, J.S.; Ramachandran, D.; Henderson, A.; Freeman, J.; Carlise, M.; Harris, A.; Willison-Headley, Z. Phylogenetic reconstruction using four low-copy nuclear loci strongly supports a polyphyletic origin of the genus *Sorghum*. *Ann. Bot.* **2015**, *116*, 291–299. [CrossRef] [PubMed]
83. Peng, S.F.; Lin, Y.P.; Lin, B.Y. Characterization of AFLP sequences from regions of maize B chromosome defined by 12 B-10L translocations. *Genetics* **2005**, *169*, 375–388. [CrossRef]
84. McAllister, B.F.; Werren, J.H. Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. *Chromosoma* **1997**, *106*, 243–253. [CrossRef] [PubMed]
85. López-León, M.D.; Neves, N.; Schwarzacher, T.; Heslop-Harrison, J.S.; Hewitt, G.M.; Camacho, J.P.M. Possible origin of a B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Res.* **1994**, *2*, 87–92. [CrossRef]
86. Sharbel, T.F.; Green, D.M.; Houben, A. B-chromosome origin in the endemic New Zealand frog *Leiopelma hochstetteri* through sex chromosome devolution. *Genome* **1998**, *41*, 14–22. [CrossRef]
87. Donald, T.; Leach, C.; Clough, A.; Timmis, J.N. Ribosomal RNA genes and the B chromosome of *Brachycome dichromosomatica*. *Heredity* **1995**, *74*, 556–561. [CrossRef] [PubMed]
88. Jamilena, M.; Garrido-Ramos, M.; Rejón, M.R.; Rejón, C.R.; Parker, J.S. Characterization of repeated sequences from microdissected B chromosomes of *Crepis capillaris*. *Chromosoma* **1995**, *104*, 113–120. [CrossRef]

89. Dhar, M.K.; Friebe, B.; Koul, A.K.; Gill, B.S. Origin of an apparent B chromosome by mutation, chromosome fragmentation and specific DNA sequence amplification. *Chromosoma* **2002**, *111*, 332–340. [CrossRef]
90. Fantinatti, B.E.; Mazzuchelli, J.; Valente, G.T.; Cabral-de-Mello, D.C.; Martins, C. Genomic content and new insights on the origin of the B chromosome of the cichlid fish *Astatotilapia latifasciata*. *Genetica* **2011**, *139*, 1273–1282. [CrossRef] [PubMed]
91. Green, D.M. Muller's Ratchet and the evolution of supernumerary chromosomes. *Genome* **1990**, *33*, 818–824. [CrossRef]
92. Magoon, M.L.; Shambulingappa, K.G. Karyomorphology of *Sorghum propinquum* and its bearing on the origin of 40-chromosome sorghum. *Chromosoma* **1961**, *12*, 460–465. [CrossRef]
93. Barnaud, A.; Deu, M.; Garine, E.; Chantereau, J.; Bolteu, J.; Koïda, E.O.; McKey, D.; Joly, H.I. A weed–crop complex in sorghum: The dynamics of genetic diversity in a traditional farming system. *Am. J. Bot.* **2009**, *96*, 1869–1879. [CrossRef]
94. Sagnard, F.; Deu, M.; Dembélé, D.; Leblois, R.; Touré, L.; Diakité, M.; Calatayud, C.; Vaksmann, M.; Bouchet, S.; Mallé, Y.; et al. Genetic diversity, structure, gene flow and evolutionary relationships within the *Sorghum bicolor* wild–weedy–crop complex in a western African region. *Theor. Appl. Genet.* **2011**, *123*, 1231. [CrossRef] [PubMed]
95. Sapre, A.B.; Deshpande, D.S. Origin of B chromosomes in *Coix* L. through spontaneous interspecific hybridization. *J. Hered.* **1987**, *78*, 191–196. [CrossRef]
96. Schartl, M.; Nanda, I.; Schlupp, I.; Wilde, B.; Epplen, J.T.; Schmid, M.; Parzefall, J. Incorporation of subgenomic amounts of DNA as compensation for mutational load in a gynogenetic fish. *Nature* **1995**, *373*, 68–71. [CrossRef]
97. Perfectti, F.; Werren, J.H. The interspecific origin of B chromosomes: Experimental evidence. *Evolution* **2001**, *55*, 1069–1073. [CrossRef]

5 GENERAL CONCLUSION

B chromosomes are genomic elements that can spread in the population despite the fact that they do not carry any essential properties. To date, B chromosomes in the genus *Sorghum* have not been the subject of any extensive research, therefore our knowledge about them and their behavior is very limited and molecular data are completely missing. As was shown in previous studies, B chromosomes in *Sorghum* are eliminated from most tissues, which was also confirmed by our observations in *S. purpureosericeum*. One of the goals of this work was to develop markers specific for the B chromosome of *S. purpureosericeum*. Using data obtained by Illumina sequencing, repeat analysis was performed and candidates for B-specific repeats were identified. Based on them, a set of markers was designed and tested for their B-specifity. The verification confirmed that seven of them are indeed B-specific and can be used for the identification and study of B-carrying individuals of *S. purpureosericeum*. The developed markers will be valuable for research studies focused on the elimination process of Bs and will significantly simplify the identification of B-positive plants (or tissues) at an early developmental stage, bypassing the need of cytogenetic analysis of plants in the flowering phase.

In parallel, we focused on the evolutionary pathway of the B chromosome in the genus. In order to clarify the relationships among *Sorghum* species, a phylogenetic analysis of 21 species was performed, confirming the existence of two main developmental lineages, with B-carrying species being present in both of them. Three evolutionary scenarios for *Sorghum* Bs were considered: A) a single origin in a common ancestor of all *Sorghum* species; B) a single origin in the ancestor of closely related species, which have kept Bs up until today; and C) multiple independent origins during *Sorghum* evolution. Results of the phylogenetic reconstruction enabled us to rule out the possibility of the origin of B chromosome in an ancestor of a group of closely related species (scenario B). Our findings are just the first steps on the way to understand the evolution of the B chromosome in the genus *Sorghum*, and it is necessary to obtain much more information before a more comprehensive picture of B origin could be created.

B chromosomes are enigmatic elements that still rise many fundamental questions. Although they have basic common features, some of their properties can vary and can be unique for different species. The way how B chromosomes are formed, how they accumulate and transmit, how they affect the host - these are all aspects that can give us insight into fundamental biological processes and therefore, B chromosomes represent a great opportunity to better understand genome organization and evolution.

6 SOUHRN (Summary, in Czech)

B chromozomy jsou zvláštním typem chromozomů, které se mohou šířit v populaci i přesto, že nenesou žádné esenciální vlastnosti. B chromozomy rodu *Sorghum* dosud nebyly předmětem rozsáhlejšího výzkumu, naše znalosti o nich jsou velmi omezené a molekulární studie zcela chybí. Předchozí práce ukázaly, že B chromozomy v rodu *Sorghum* jsou z většiny tkání eliminovány, což bylo potvrzeno také v našich pozorováních u *S. purpureosericeum.* Jedním z cílů této práce bylo vyvinout markery specifické pro B chromozom *S. purpureosericeum*. Na základě dat získaných sekvenováním Illumina byla provedena analýza repetic, díky níž bylo identifikováno devět kandidátů na B-specifické repetice. Na jejich základě byly navrženy markery, z nichž sedm bylo po následném testování pomocí PCR ověřeno jako specifických pro B chromozom *S. purpureosericeum*. Vyvinuté markery budou cenné pro další výzkum zaměřený na proces eliminace B chromozomů a výrazně zjednoduší identifikaci B-pozitivních rostlin (nebo pletiv) v raném vývojovém stádiu, čímž bude možné vyhnout se nutnosti cytogenetické analýzy rostlin ve fázi květu.

Protože B chromozomy byly objeveny ve více druzích čiroku, využili jsme fylogenetickou analýzu k poodhalení evoluční cesty B chromozomu v rámci rodu. Za účelem objasnění vztahů mezi čirokovými druhy byla provedena fylogenetická analýza 21 druhů, která potvrdila existenci dvou hlavních vývojových linií, přičemž v obou byly přítomny druhy nesoucí B chromozomy. Byly uvažovány tři evoluční scénáře vzniku B chromozomu v rodu *Sorghum*: A) jednorázový vznik B chromozomu u společného předka všech čirokových druhů; B) jednorázový vznik B chromozomu u předka skupiny blízce příbuzných druhů, které si jej udržely až do současnosti; a C) nezávislý vznik B chromozomu vícekrát během evoluce čiroku. Výsledky fylogenetické rekonstrukce nám umožnily vyloučit možnost vzniku B chromozomu u předka skupiny blízce příbuzných druhů (scénář B). Naše poznatky jsou ovšem jen prvním krokem na cestě k pochopení evoluce B chromozomu v rodu *Sorghum* a je nutné získat mnohem více informací, než bude možné vytvořit komplexnější obraz o jejich původu.

B chromozomy jsou enigmatické elementy, které stále vyvolávají mnoho zásadních otázek. Přestože mají základní společné rysy, některé jejich vlastnosti mohou být u různých druhů organizmů odlišné nebo dokonce unikátní. Způsob, jakým

B chromozomy vznikají, jak fungují jejich akumulační mechanismy a jak ovlivňují hostitele – to vše jsou aspekty, které nám mohou poskytnout vhled do základních biologických procesů. B chromozomy tak představují příležitost pro lepší pochopení organizace a evoluce genomu.

7 REFERENCES

Abdelhalim, T. S., Kamal, N. M., & Hassan, A. B. (2019). Nutritional potential of wild sorghum: Grain quality of Sudanese wild sorghum genotypes (Sorghum bicolor L. Moench). *Food Science & Nutrition*, *7*(4), 1529-1539.

Ahmad, S. F., & Martins, C. (2019). The modern view of B chromosomes under the impact of high scale omics analyses. *Cells*, 8(2), 156.

Ajawatanawong, P. (2016). Molecular Phylogenetics: Concepts for a Newcomer. In: Nookaew, I. (eds) Network Biology. *Advances in Biochemical Engineering/Biotechnology*, vol 160. Springer, Cham.

Ajawatanawong, P., Atkinson, G. C., Watson-Haigh, N. S., MacKenzie, B., & Baldauf, S. L. (2012). SeqFIRE: a web application for automated extraction of indel regions and conserved blocks from protein multiple sequence alignments. *Nucleic acids research*, *40*(W1), W340-W347.

Akbari, O. S., Antoshechkin, I., Hay, B. A., & Ferree, P. M. (2013). Transcriptome profiling of Nasonia vitripennis testis reveals novel transcripts expressed from the selfish B chromosome, paternal sex ratio. *G3: Genes, Genomes, Genetics*, *3*(9), 1597-1605.

Alfenito, M. R., & Birchler, J. A. (1993). Molecular characterization of a maize B chromosome centric sequence. *Genetics*, *135*(2), 589-597.

Alfenito, M. R., & Birchler, J. A. (1993). Molecular characterization of a maize B chromosome centric sequence. *Genetics*, *135*(2), 589-597.

Álvarez, I. J. F. W., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular phylogenetics and evolution*, *29*(3), 417-434.

Ayliffe, M. A., Scott, N. S., & Timmis, J. N. (1998). Analysis of plastid DNA-like sequences within the nuclear genomes of higher plants. *Molecular Biology and Evolution*, *15*(6), 738-745.

Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., & Donoghue, M. J. (1995). The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri botanical garden*, *82*(2), 247-277.

Banaei-Moghaddam, A. M., Meier, K., Karimi-Ashtiyani, R., & Houben, A. (2013). Formation and expression of pseudogenes on the B chromosome of rye. *The Plant Cell*, *25*(7), 2536-2544.

Banaei-Moghaddam, A. M., Schubert, V., Kumke, K., Weiβ, O., Klemme, S., Nagaki, K., ... & Houben, A. (2012). Nondisjunction in favor of a chromosome: the mechanism of rye B chromosome drive during pollen mitosis. *The Plant Cell*, *24*(10), 4124-4134.

Beckmann, J. S., & Soller, M. (1983). Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. *Theoretical and Applied Genetics*, *67*(1), 35-43.

Bednářová, M., Karafiátová, M., Hřibová, E., & Bartoš, J. B. (2021). Chromosomes in Genus Sorghum (Poaceae). *Plants*, *10*, 505.

Beukeboom, L. W. (1994). Bewildering Bs: an impression of the 1st B-Chromosome Conference. *Heredity*, *73*(3), 328-335.

Blavet, N., Yang, H., Su, H., Solanský, P., Douglas, R. N., Karafiátová, M., ... & Bartoš, J. (2021). Sequence of the supernumerary B chromosome of maize provides insight into its drive mechanism and evolution. *Proceedings of the National Academy of Sciences*, *118*(23).

Blunden, R., Wilkes, T. J., Forster, J. W., Jimenez, M. M., Sandery, M. J., Karp, A., & Jones, R. N. (1993). Identification of the E3900 family, a second family of rye B chromosome specific repeated sequences. *Genome*, *36*(4), 706-711.

Boudichevskaia, A., Ruban, A., Thiel, J., Fiebig, A., & Houben, A. (2020). Tissue-Specific Transcriptome Analysis Reveals Candidate Transcripts Associated with the Process of Programmed B Chromosome Elimination in Aegilops speltoides. *International journal of molecular sciences*, *21*(20), 7596.

Bougourd, S. M., & Parker, J. S. (1979). The B-chromosome system of Allium schoenoprasum. *Chromosoma*, *75*(3), 369-383.

Bougourd, S. M., & Plowman, A. B. (1996). The inheritance of B chromosomes in Allium schoenoprasum L. *Chromosome Research*, *4*(2), 151-158.

Branco, M. R., & Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS biology*, *4*(5), e138.

Brelsfoard, C., Tsiamis, G., Falchetto, M., Gomulski, L. M., Telleria, E., Alam, U., ... & Aksoy, S. (2014). Presence of extensive Wolbachia symbiont insertions discovered in the genome of its host Glossina morsitans morsitans. *PLoS neglected tropical diseases*, *8*(4), e2728.

Cabrero, J., Martín-Peciña, M., Ruiz-Ruano, F. J., Gómez, R., & Camacho, J. P. M. (2017). Post-meiotic B chromosome expulsion, during spermiogenesis, in two grasshopper species. *Chromosoma*, *126*(5), 633-644.

Camacho, J. P. M., Sharbel, T. F., & Beukeboom, L. W. (2000). B-chromosome evolution. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1394), 163-178.

Camacho, J. P. M., Shaw, M. W., López-León, M. D., Pardo, M. C., & Cabrero, J. (1997). Population dynamics of a selfish B chromosome neutralized by the standard genome in the grasshopper Eyprepocnemis plorans. *The American Naturalist*, *149*(6), 1030-1050.

Carchilan, M., Delgado, M., Ribeiro, T., Costa-Nunes, P., Caperta, A., Morais-Cecílio, L., ... & Houben, A. (2007). Transcriptionally active heterochromatin in rye B chromosomes. *The Plant Cell*, *19*(6), 1738-1749.

Carchilan, M., Kumke, K., Mikolajewski, S., & Houben, A. (2009). Rye B chromosomes are weakly transcribed and might alter the transcriptional activity of A chromosome sequences. *Chromosoma*, *118*(5), 607-616.

Carlson, W. R. (1969). Factors affecting preferential fertilization in maize. *Genetics*, *62*(3), 543-554.

Carlson, W. R. (1978). The B chromosome of corn. *Annual review of genetics*, *12*(1), 5-5.

Carlson, W. R., & Phillips, R. L. (1986). The B chromosome of maize. *Critical Reviews in Plant Sciences*, *3*(3), 201-226.

Carlson, W. R., & Roseman, R. R. (1992). A new property of the maize B chromosome. *Genetics*, *131*(1), 211-223.

Carter, C. R., & Smith-White, S. (1972). The cytology of Brachycome lineariloba. *Chromosoma*, *39*(4), 361-379.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution*, *17*(4), 540-552.

Charlesworth, D. (2008). Plant sex chromosomes. In *Plant Genomes* (Vol. 4, pp. 83-94). Karger Publishers.

Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., ... & Albert, V. A. (1993). Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene rbcL. *Annals of the Missouri Botanical Garden*, *80(3)*, 528-580.

Cheng, Y. M., & Lin, B. Y. (2004). Molecular organization of large fragments in the maize B chromosome: indication of a novel repeat. *Genetics*, *166*(4), 1947-1961.

Cheng, Y. M., Feng, Y. R., Lin, Y. P., & Peng, S. F. (2016). Cytomolecular characterization and origin of de novo formed maize B chromosome variants. *Chromosome Research*, *24*(2), 183-195.

Cody, J. P., Swyers, N. C., McCaw, M. E., Graham, N. D., Zhao, C., & Birchler, J. A. (2015). Minichromosomes: vectors for crop improvement. *Agronomy*, *5*(3), 309-321.

Collins, F. S., Morgan, M., & Patrinos, A. (2003). The Human Genome Project: lessons from large-scale biology. *Science*, *300*(5617), 286-290.

Cuadrado, A., & Schwarzacher, T. (1998). The chromosomal organization of simple sequence repeats in wheat and rye genomes. *Chromosoma*, *107*(8), 587-594.

D'Ambrosio, U., Alonso-Lifante, M. P., Barros, K., Kovařík, A., de Xaxars, G. M., & Garcia, S. (2017). B-chrom: A database on B-chromosomes of plants, animals and fungi. *New Phytologist*, *216*(3), 635-642.

Dalla Benetta, E., Antoshechkin, I., Yang, T., Nguyen, H. Q. M., Ferree, P. M., & Akbari, O. S. (2020). Genome elimination mediated by gene expression from a selfish chromosome. *Science advances*, *6*(14), eaaz9808.

Darlington, C. D., & Thomas, P. T. (1941). Morbid mitosis and the activity of inert chromosomes in Sorghum. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, *130*(859), 127-150.

Dhar, M. K., Friebe, B., Koul, A. K., & Gill, B. S. (2002). Origin of an apparent B chromosome by mutation, chromosome fragmentation and specific DNA sequence amplification. *Chromosoma*, *111*(5), 332-340.

Dherawattana, A., & Sadanaga, K. (1973). Cytogenetics of a Crown Rust-Resistant Hexaploid Oat with 42+ 2 Fragment Chromosomes 1. *Crop Science*, *13*(6), 591-594.

Dirks, R. W., & Raap, A. K. (1995). Cell-cycle-dependent gene expression studied by two-colour fluorescent detection of a mRNA and histone mRNA. *Histochemistry and cell biology*, *104*(5), 391-395.

Donald, T. M., Houben, A., Leach, C. R., & Timmis, J. N. (1997). Ribosomal RNA genes specific to the B chromosomes in Brachycome dichromosomatica are not transcribed in leaf tissue. *Genome*, *40*(5), 674-681.

Ebrahimzadegan, R., Houben, A., & Mirzaghaderi, G. (2019). Repetitive DNA landscape in essential A and supernumerary B chromosomes of Festuca pratensis Huds. *Scientific reports*, *9*(1), 1-11.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, *32*(5), 1792-1797.

Eickbush, D. G., Eickbush, T. H., & Werren, J. H. (1992). Molecular characterization of repetitive DNA sequences from a B chromosome. *Chromosoma*, *101*(9), 575-583.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., ... & Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, *323*(5910), 133-138.

Eisenstein, M. (2012). Oxford Nanopore announcement sets sequencing sector abuzz. *Nature biotechnology*, *30*(4), 295-296.

Endo, T. R., Nasuda, S., Jones, N., Dou, Q., Akahori, A., Wakimoto, M., ... & Tsujimoto, H. (2008). Dissection of rye B chromosomes, and nondisjunction properties of the dissected segments in a common wheat background. *Genes & genetic systems*, *83*(1), 23-30.

Evans, G. M., & Macefield, A. J. (1973). The effect of B chromosomes on homoeologous pairing in species hybrids. *Chromosoma*, *41*(1), 63-73.

Evans, G. M., & Macefield, J. (1972). Suppression of homoeologous pairing by B chromosomes in a Lolium species hybrid. *Nature New Biology*, *236*(65), 110-111.

Farris, J. S. (1970). Methods for computing Wagner trees. *Systematic Biology*, *19*(1), 83-92.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of molecular evolution*, *17*(6), 368-376.

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Biology*, *20*(4), 406-416.

Fox, D. P., Hewitt, G. M., & Hall, D. J. (1974). DNA replication and RNA transcription of euchromatic and heterochromatic chromosome regions during grasshopper meiosis. *Chromosoma*, *45*(1), 43-62.

Franks, T. K., Houben, A., Leach, C. R., & Timmis, J. N. (1996). The molecular organisation of a B chromosome tandem repeat sequence from Brachycome dichromosomatica. *Chromosoma*, *105*(4), 223-230.

Fröst, S. (1959). The cytological behavior and mode of transmission of accessory chromosomes in Plantago serraria. *Hereditas*, *45*(2-3), 191-210.

Garber, E. D. (1950). Cytotaxonomic studies in the genus Sorghum. *University of California Publications in Botany*, *23*, 283-362.

Gielly, L., & Taberlet, P. (1994). The use of chloroplast DNA to resolve plant phylogenies: noncoding versus rbcL sequences. *Molecular biology and evolution*, *11*(5), 769-777.

González-Sánchez, M., González-González, E., Molina, F., Chiavarino, A. M., Rosato, M., & Puertas, M. J. (2003). One gene determines maize B chromosome accumulation by preferential fertilisation; another gene (s) determines their meiotic loss. *Heredity*, *90*(2), 122-129.

Green, D. M. (1988). Cytogenetics of the endemic New Zealand frog, Leiopelma hochstetteri: extraordinary supernumerary chromosome variation and a unique sex-chromosome system. *Chromosoma*, *97*(1), 55-70.

Gruber, S. L., Diniz, D., Sobrinho-Scudeler, P. E., Foresti, F., Haddad, C. F. B., & Kasahara, S. (2014). Possible interspecific origin of the B chromosome of Hypsiboas albopunctatus (Spix, 1824) (Anura, Hylidae), revealed by microdissection, chromosome painting, and reverse hybridisation. *Comparative Cytogenetics*, *8*(3), 185.

Grun, P. (1959). Variability of accessory chromosomes in native populations of Allium cernuum. *American journal of Botany*, *46*(3), 218-224.

Gupta, P. K., Rustgi, S., & Mir, R. R. (2008). Array-based high-throughput DNA markers for crop improvement. *Heredity*, *101*(1), 5-18.

Gutknecht, J., Sperlich, D., & Bachmann, L. (1995). A species specific satellite DNA family of Drosophila subsilvestris appearing predominantly in B chromosomes. *Chromosoma*, *103*(8), 539-544.

Hakansson, A. (1948). Behaviour of accessory rye chromosomes in the embryo-sac. *Hereditas*, *34*(1-2), 35-59.

Han, F., Lamb, J. C., Yu, W., Gao, Z., & Birchler, J. A. (2007). Centromere function and nondisjunction are independent components of the maize B chromosome accumulation mechanism. *The Plant Cell*, *19*(2), 524-533.

Han, Y., Liu, X., Benny, U., Kistler, H. C., & VanEtten, H. D. (2001). Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen Nectria haematococca. *The Plant Journal*, *25*(3), 305-314.

Hasegawa, N. (1934). A cytological study on 8-chromosome rye. *Cytologia*, *6*(1), 68-77.

Heiskanen, M., Hellsten, E., Kallioniemi, O. P., Mäkelä, T. P., Alitalo, K., Peltonen, L., & Palotie, A. (1995). Visual mapping by fiber-FISH. *Genomics*, *30*(1), 31-36.

Hewitt, G. M. (1974). The integration of supernumerary chromosomes into the orthopteran genome. In *Cold Spring Harbor Symposia on Quantitative Biology* (Vol. 38, pp. 183-194). Cold Spring Harbor Laboratory Press.

Hewitt, G. M. (1976). Meiotic drive for B-chromosomes in the primary oocytes of Myrmekotettix maculatus (Orthoptera: Acrididae). *Chromosoma*, *56*(4), 381-391.

Holm, L. G., Plucknett, D. L., Pancho, J. V., & Herberger, J. P. (1977). *The world's worst weeds. Distribution and biology*. University press of Hawaii.

Houben, A., Banaei-Moghaddam, A. M., & Klemme, S. (2013). Biology and evolution of B chromosomes. In *Plant Genome Diversity Volume 2* (pp. 149-165). Springer, Vienna.

Houben, A., Banaei-Moghaddam, A. M., Klemme, S., & Timmis, J. N. (2014). Evolution and biology of supernumerary B chromosomes. *Cellular and Molecular Life Sciences*, *71*(3), 467-478.

Houben, A., Kynast, R. G., Heim, U., Hermann, H., Jones, R. N., & Forster, J. W. (1996). Molecular cytogenetic characterisation of the terminal heterochromatic segment of the B-chromosome of rye (Secale cereale). *Chromosoma*, *105*(2), 97-103.

Houben, A., Leach, C. R., Verlin, D., Rofe, R., & Timmis, J. N. (1997). A repetitive DNA sequence common to the different B chromosomes of the genus Brachycome. *Chromosoma*, *106*(8), 513-519.

Houben, A., Verlin, D., Leach, C. R., & Timmis, J. N. (2001). The genomic complexity of micro B chromosomes of Brachycome dichromosomatica. *Chromosoma*, *110*(7), 451-459.

Huang, W., Du, Y., Zhao, X., & Jin, W. (2016). B chromosome contains active genes and impacts the transcription of A chromosomes in maize (Zea mays L.). *BMC Plant Biology*, *16*(1), 1-14.

Huang, Y. H., Lin, T. C., Chiou, W. Y., & Cheng, Y. M. (2021). The r-X1 deletion induces terminal deficiencies in the maize B chromosome. *Chromosome Research*, *29*(3), 351-360.

Huskins, C. L., & Smith, S. G. (1932). A cytological study of the genus Sorghum Pers. *Journal of Genetics*, *25*(2), 241-249.

Idrees, M., & Irshad M. (2014). Molecular markers in plants for analysis of genetic diversity: a review. *European academic research*, *2*(1), 1513-1540.

Imai, H. T. (1974). B-chromosomes in the myrmicine ant, Leptothorax spinosior. *Chromosoma*, *45*(4), 431-444.

Ishak, B., Jaafar, H., Maetz, J. L., & Rumpler, Y. (1991). Absence of transcriptional activity of the B-chromosomes of Apodemus peninsulae during pachytene. *Chromosoma*, *100*(4), 278-281.

Jamilena, M., Garrido-Ramos, M., Rejon, M. R., Rejon, C. R., & Parker, J. S. (1995). Characterisation of repeated sequences from microdissected B chromosomes of Crepis capillaris. *Chromosoma*, *104*(2), 113-120.

Jamilena, M., Ruiz Rejón, C., & Ruiz Rejón, M. (1994). A molecular analysis of the origin of the Crepis capillaris B chromosome. *Journal of Cell Science*, *107*(3), 703-708.

Janaki-Ammal, E. K. (1940). Chromosome diminution in a plant. *Nature*, *146*(3713), 839-840.

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., ... & Kersters, K. (1996). Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology*, *142*(7), 1881-1893.

Jiang, J. (2019). Fluorescence in situ hybridization in plants: recent developments and future applications. *Chromosome Research*, *27*(3), 153-165.

Jiang, J., & Gill, B. S. (1994). Nonisotopic in situ hybridization and plant genome mapping: the first 10 years. *Genome*, *37*(5), 717-725.

Jiang, J., & Gill, B. S. (2006). Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. *Genome*, *49*(9), 1057-1068.

Jin, W., Lamb, J. C., Vega, J. M., Dawe, R. K., Birchler, J. A., & Jiang, J. (2005). Molecular and functional dissection of the maize B chromosome centromere. *The Plant Cell*, *17*(5), 1412-1423.

John, B., & Hewitt, G. M. (1965). The B-chromosome system of Myrmeleotettix maculatus (Thunb.). *Chromosoma*, *16*(5), 548-578.

Johnson, L. A., & Soltis, D. E. (1994). matK DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic botany*,*19*, 143-156.

Jones, N. (2012). B chromosomes in plants. *Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology*, *146*(3), 727-737.

Jones, N., & Houben, A. (2003). B chromosomes in plants: escapees from the A chromosome genome? *Trends in plant science*, *8*(9), 417-423.

Jones, N., & Ruban, A. (2019). Are B chromosomes useful for crop improvement? *Plants, People, Planet*, *1*(2), 84-92.

Jones, R. N. (1985). Are B chromosomes selfish? In *The evolution of genome size*, pp. 397-425. London, Wiley.

Jones, R. N. (1991). B-chromosome drive. *The American Naturalist*, *137*(3), 430-442.

Jones, R. N. (1995). B chromosomes in plants. *New Phytologist*, *131*(4), 411-434.

Jones, R. N. (2018). Transmission and drive involving parasitic B chromosomes. *Genes*, *9*(8), 388.

Jones, R. N., & Rees, H. (1982). B Chromosomes. Academic Press, London, New York.

Jukes TH, Cantor CR (1969). Evolution of Protein Molecules, pp. 21–132. New York, Academic Press.

Kameswara Rao, N., Reddy, L. J., & Bramel, P. J. (2003). Potential of wild species for genetic enhancement of some semi-arid food crops. *Genetic Resources and Crop Evolution*, *50*(7), 707-721.

Karafiátová, M., Bednářová, M., Said, M., Čížková, J., Holušová, K., Blavet, N., & Bartoš, J. (2021). The B chromosome of Sorghum purpureosericeum reveals the first pieces of its sequence. *Journal of experimental botany*, *72*(5), 1606-1616.

Karamysheva, T. V., Torgasheva, A. A., Yefremov, Y. R., Bogomolov, A. G., Liehr, T., Borodin, P. M., & Rubtsov, N. B. (2017). Spatial organization of fibroblast and spermatocyte nuclei with different B-chromosome content in Korean field mouse, Apodemus peninsulae (Rodentia, Muridae). *Genome*, *60*(10), 815-824.

Katoh, K., Misawa, K., Kuma, K. I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research*, *30*(14), 3059-3066.

Kayano, H. (1971). Accumulation of B chromosomes in the germ line of Locusta migratoria. *Heredity*, *27*(1), 119-123.

Kean, V. M., Fox, D. P., & Faulkner, R. (1982). The Accumulation mechanism of the supernumerary (B-) chromosome in Picea sitchensis (Bong.) Carr. and the effect of this chromosome on male and female flowering. *Silvae genetica*, *31*(4), 126-131.

Kichigin, I. G., Lisachov, A. P., Giovannotti, M., Makunin, A. I., Kabilov, M. R., O'Brien, P., ... & Trifonov, V. A. (2019). First report on B chromosome content in a reptilian species: The case of Anolis carolinensis. *Molecular Genetics and Genomics*, *294*(1), 13-21.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution*, *16*(2), 111-120.

Kimura, M. (1981). Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences*, *78*(1), 454-458.

Kimura, M., & Kayano, H. (1961). The maintenance of supernumerary chromosomes in wild populations of Lilium callosum by preferential segregation. *Genetics*, *46*(12), 1699-1712.

Klemme, S., Banaei-Moghaddam, A. M., Macas, J., Wicker, T., Novák, P., & Houben, A. (2013). High-copy sequences reveal distinct evolution of the rye B chromosome. *New Phytologist*, *199*(2), 550-558.

Kumke, K., Macas, J., Fuchs, J., Altschmied, L., Kour, J., Dhar, M. K., & Houben, A. (2016). Plantago lagopus B chromosome is enriched in 5S rDNA-derived satellite DNA. *Cytogenetic and genome research*, *148*(1), 68-73.

Lamb, J. C., Han, F., Auger, D. L., & Birchler, J. A. (2006). A trans-acting factor required for non-disjunction of the B chromosome is located distal to the TB-4Lb breakpoint on the B chromosome. *Maize Genetics Cooperation Newsletter*, *80*, 4.

Lamb, J. C., Meyer, J. M., Corcoran, B., Kato, A., Han, F., & Birchler, J. A. (2007). Distinct chromosomal distributions of highly repetitive sequences in maize. *Chromosome Research*, *15*(1), 33-49.

Lamb, J. C., Riddle, N. C., Cheng, Y. M., Theuri, J., & Birchler, J. A. (2007). Localization and transcription of a retrotransposon-derived element on the maize B chromosome. *Chromosome Research*, *15*(3), 383-398.

Langdon, T., Seago, C., Jones, R. N., Ougham, H., Thomas, H., Forster, J. W., & Jenkins, G. (2000). De novo evolution of satellite DNA on the rye B chromosome. *Genetics*, *154*(2), 869-884.

Lazarides, M., Hacker, J. B., & Andrew, M. H. (1991). Taxonomy, cytology and ecology of indigenous Australian sorghums (Sorghum Moench: Andropogoneae: Poaceae). *Australian Systematic Botany*, *4*(4), 591-635.

Leach, C. R., Donald, T. M., Franks, T. K., Spiniello, S. S., Hanrahan, C. F., & Timmis, J. N. (1995). Organisation and origin of a B chromosome centromeric sequence from Brachycome dichromosomatica. *Chromosoma*, *103*(10), 708-714.

Leach, C. R., Houben, A., Field, B., Pistrick, K., Demidov, D., & Timmis, J. N. (2005). Molecular evidence for transcription of genes on a B chromosome in Crepis capillaris. *Genetics*, *171*(1), 269-278.

Levsky, J. M., & Singer, R. H. (2003). Fluorescence in situ hybridization: past, present and future. *Journal of cell science*, *116*(14), 2833-2838.

Levsky, J. M., Shenoy, S. M., Pezo, R. C., & Singer, R. H. (2002). Single-cell gene expression profiling. *Science*, *297*(5582), 836-840.

Li, Y., Jing, X. A., Aldrich, J. C., Clifford, C., Chen, J., Akbari, O. S., & Ferree, P. M. (2017). Unique sequence organization and small RNA expression of a “selfish” B chromosome. *Chromosoma*, *126*(6), 753-768.

Lin, B. Y. (1987). Cytological evidence of terminal deficiencies produced by the r-X1 deletion in maize. *Genome*, *29*(5), 718-721.

Lin, H. Z., Lin, W. D., Lin, C. Y., Peng, S. F., & Cheng, Y. M. (2014). Characterization of maize B-chromosome-related transcripts isolated via cDNA-AFLP. *Chromosoma*, *123*(6), 597-607.

Lloyd, A. H., & Timmis, J. N. (2011). The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome. *Molecular biology and evolution*, *28*(7), 2019-2028.

López-León, M. D., Cabrero, J., Pardo, M. C., Viseras, E., Camacho, J. P. M., & Santos, J. L. (1993). Generating high variability of B chromosomes in Eyprepocnemis plorans (grasshopper). *Heredity*, *71*(4), 352-362.

López-León, M. D., Neves, N., Schwarzacher, T., Hewitt, G. M., & Camacho, J. P. M. (1994). Possible origin of a B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Research*, *2*(2), 87-92.

López-León, M. D., Pardo, M. C., Cabrero, J., & Camacho, J. P. M. (1992). Random mating and absence of sexual selection for B chromosomes in two natural populations of the grasshopper Eyprepocnemis plorans. *Heredity*, *69*(6), 558-561.

Ma, W., Gabriel, T. S., Martis, M. M., Gursinsky, T., Schubert, V., Vrána, J., ... & Houben, A. (2017). Rye B chromosomes encode a functional Argonaute-like protein with in vitro slicer activities similar to its A chromosome paralog. *New Phytologist*, *213*(2), 916-928.

Makunin, A. I., Romanenko, S. A., Beklemisheva, V. R., Perelman, P. L., Druzhkova, A. S., Petrova, K. O., ... & Trifonov, V. A. (2018). Sequencing of supernumerary chromosomes of red fox and raccoon dog confirms a non-random gene acquisition by B chromosomes. *Genes*, *9*(8), 405.

Marques, A., Klemme, S., & Houben, A. (2018). Evolution of plant B chromosome enriched sequences. *Genes*, *9*(10), 515.

Marschner, S., Meister, A., Blattner, F. R., & Houben, A. (2007). Evolution and function of B chromosome 45S rDNA sequences in Brachycome dichromosomatica. *Genome*, *50*(7), 638-644.

Martens, U. M., Zijlmans, J. M. J. M., Poon, S. S., Dragowska, W., Yui, J., Chavez, E. A., ... & Lansdorp, P. M. (1998). Short telomeres on human chromosome 17p. *Nature genetics*, *18*(1), 76-80.

Martis, M. M., Klemme, S., Banaei-Moghaddam, A. M., Blattner, F. R., Macas, J., Schmutzer, T., ... & Houben, A. (2012). Selfish supernumerary chromosome reveals its origin as a mosaic

of host genome and organellar sequences. *Proceedings of the National Academy of Sciences*, *109*(33), 13343-13346.

Matsuo, M., Ito, Y., Yamauchi, R., & Obokata, J. (2005). The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast–nuclear DNA flux. *The Plant Cell*, *17*(3), 665-675.

McAllister, B. F., & Werren, J. H. (1997). Hybrid origin of a B chromosome (PSR) in the parasitic wasp Nasonia vitripennis. *Chromosoma*, *106*(4), 243-253.

Mendelson, D., & Zohary, D. (1972). Behaviour and transmission of supernumerary chromosomes in Aegilops speltoides. *Heredity*, *29*(3), 329-339.

Milani, D., Ruiz-Ruano, F. J., Camacho, J. P. M., & Cabral-de-Mello, D. C. (2021). Out of patterns, the euchromatic B chromosome of the grasshopper Abracris flavolineata is not enriched in high-copy repeats. *Heredity*, *127*(5), 475-483.

Mochizuki, A. (1957). B chromosomes in Aegilops mutica Boiss. *Wheat Inf. Serv.*, *5*, 9-11.

Navarro-Domínguez, B., Martín-Peciña, M., Ruiz-Ruano, F. J., Cabrero, J., Corral, J. M., López-León, M. D., ... & Camacho, J. P. M. (2019). Gene expression changes elicited by a parasitic B chromosome in the grasshopper Eyprepocnemis plorans are consistent with its phenotypic effects. *Chromosoma*, *128*(1), 53-67.

Navarro-Domínguez, B., Ruiz-Ruano, F. J., Cabrero, J., Corral, J. M., López-León, M. D., Sharbel, T. F., & Camacho, J. P. M. (2017). Protein-coding genes in B chromosomes of the grasshopper Eyprepocnemis plorans. *Scientific Reports*, *7*(1), 1-12.

Noutsos, C., Richly, E., & Leister, D. (2005). Generation and evolutionary fate of insertions of organelle DNA in the nuclear genomes of flowering plants. *Genome Research*, *15*(5), 616-628.

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

Nur, U., Werren, J. H., Eickbush, D. G., Burke, W. D., & Eickbush, T. H. (1988). A" selfish" B chromosome that enhances its transmission by eliminating the paternal genome. *Science*, *240*(4851), 512-514.

Ohta, S. (1995). Distinct numerical variation of B-chromosomes among different tissues in Aegilops mutica Boiss. *The Japanese Journal of Genetics*, *70*(1), 93-101.

Oliveira, C., Maria Rodrigues Saboya, S., Foresti, F., Augusto Senhorini, J., & Bernardino, G. (1997). Increased B chromosome frequency and absence of drive in the fish Prochilodus lineatus. *Heredity*, *79*(5), 473-476.

Olmstead, R. G., Kim, K. J., Jansen, R. K., & Wagstaff, S. J. (2000). The phylogeny of the Asteridae sensu lato based on chloroplast ndhF gene sequences. *Molecular Phylogenetics and Evolution*, *16*(1), 96-112.

Pardo, M. C., López-León, M. D., Cabrero, J., & Camacho, J. P. M. (1994). Transmission analysis of mitotically unstable B chromosomes in Locusta migratoria. *Genome*, *37*(6), 1027-1034.

Parker, J. S., Jones, G. H., Edgar, L., & Whitehouse, C. (1989). The population cytogenetics of Crepis capillaris. II. The stability and inheritance of B-chromosomes. *Heredity*, *63*(1), 19-27.

Patwardhan, A., Ray, S., & Roy, A. (2014). Molecular markers in phylogenetic studies-a review. *Journal of Phylogenetics & Evolutionary Biology*, *2014*.

Peacock, W. J., Dennis, E. S., Rhoades, M. M., & Pryor, A. J. (1981). Highly repeated DNA sequence limited to knob heterochromatin in maize. *Proceedings of the National Academy of Sciences*, *78*(7), 4490-4494.

Phillips, R. B., & Reed, K. M. (1996). Application of fluorescence in situ hybridization (FISH) techniques to fish genetics: a review. *Aquaculture*, *140*(3), 197-216.

Plowman, A. B., & Bougourd, S. M. (1994). Selectively advantageous effects of B chromosomes on germination behaviour in Allium schoenoprasum L. *Heredity*, *72*(6), 587-593.

Rakoczy-Trojanowska, M., & Bolibok, H. (2004). Characteristics and a comparison of three classes of microsatellite-based markers and their application in plants. *Cellular and Molecular Biology Letters*, *9*(2), 221-238.

Raman V. S., Meenakshi, K. & Thangam M. S. (1976). Accessory chromosomes and their meiotic behaviour in hybrids of grain sorghum and Johnson grass. *Cytologia*, *41*(2), 193-200.

Raman V.S., Meenakshi K., Thangam M.S., & Sivagnanam L. (1964). The cytological behaviour of B-chromosomes in *Sorghum halepense*. *Abst. Madras Agricultural Journal*, *51*, 72–73.

Raman, V. S., & Krishnaswami, D. (1960). Accessory chromosomes in Sorghum nitidum Pers. *J. Indian Bot. Soc.*, *39*, 278-280.

Raman, V.S.; & Thangam, M.S. (1965). Paternal transmission of accessory chromosomes in a species of Eu-Sorghum. *Sci. Cult.*, *31*, 150–151.

Ramos, É., Cardoso, A. L., Brown, J., Marques, D. F., Fantinatti, B. E., Cabral-de-Mello, D. C., ... & Martins, C. (2017). The repetitive DNA element BncDNA, enriched in the B chromosome of the cichlid fish Astatotilapia latifasciata, transcribes a potentially noncoding RNA. *Chromosoma*, *126*(2), 313-323.

Randolph L. F. (1928). Types of supernumerary chromosomes in maize. (Abtract) *Anat. Rec.*, *41*, 102.

Rannala, B., & Yang, Z. (1996). Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of molecular evolution*, *43*(3), 304-311.

Roman, H. (1947). Mitotic nondisjunction in the case of interchanges involving the B-type chromosome in maize. *Genetics*, *32*(4), 391.

Rothfels, K. H. (1950). Chromosome complement, polyploidy and supernumeraries in Neopodismopsis abdominalis (Acrididae). *Journal of Morphology*, *87*(2), 287-315.

Ruban, A., Fuchs, J., Marques, A., Schubert, V., Soloviev, A., Raskina, O., ... & Houben, A. (2014). B chromosomes of Aegilops speltoides are enriched in organelle genome-derived sequences. *PLoS One*, *9*(2), e90214.

Ruban, A., Schmutzer, T., Scholz, U., & Houben, A. (2017). How next-generation sequencing has aided our understanding of the sequence composition and origin of B chromosomes. *Genes*, *8*(11), 294.

Ruban, A., Schmutzer, T., Wu, D. D., Fuchs, J., Boudichevskaia, A., Rubtsova, M., ... & Houben, A. (2020). Supernumerary B chromosomes of Aegilops speltoides undergo precise elimination in roots early in embryo development. *Nature communications*, *11*(1), 1-12.

Ruiz-Estevez, M., Ma Dolores López-León, Cabrero, J., & Camacho, J. P. M. (2012). B-chromosome ribosomal DNA is functional in the grasshopper Eyprepocnemis plorans. *PLoS One*, *7*(5), e36600.

Ruiz-Ruano, F. J., Cabrero, J., López-León, M. D., & Camacho, J. P. M. (2017). Satellite DNA content illuminates the ancestry of a supernumerary (B) chromosome. *Chromosoma*, *126*(4), 487-500.

Ruiz-Ruano, F. J., Cabrero, J., López-León, M. D., Sánchez, A., & Camacho, J. P. M. (2018). Quantitative sequence characterization for repetitive DNA content in the supernumerary chromosome of the migratory locust. *Chromosoma*, *127*(1), 45-57.

Ruiz-Ruano, F. J., López-León, M. D., Cabrero, J., & Camacho, J. P. M. (2016). High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Scientific reports*, *6*(1), 1-14.

Rusche, M. L., Mogensen, H. L., Shi, L., Keim, P., Rougier, M., Chaboud, A., & Dumas, C. (1997). B chromosome behavior in maize pollen as determined by a molecular probe. *Genetics*, *147*(4), 1915-1921.

Rutishauser, A. (1966). Boosting mechanism of B chromosomes in Crepis capillaris. *Chromosomes today*, *1*, 28-30.

Sáez-Vásquez, J., & Delseny, M. (2019). Ribosome biogenesis in plants: from functional 45S ribosomal DNA organization to ribosome assembly factors. *The Plant Cell*, *31*(9), 1945-1967.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution*, *4*(4), 406-425.

Sandery, M. J., Forster, J. W., Blunden, R., & Jones, R. N. (1990). Identification of a family of repeated sequences on the rye B chromosome. *Genome*, *33*(6), 908-913.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences*, *74*(12), 5463-5467.

Sapre, A. B., & Deshpande, D. S. (1987). Origin of B chromosomes in Coix L. through spontaneous interspecific hybridization. *Journal of Heredity*, *78*(3), 191-196.

Schemczssen-Graeff, Z., Barbosa, P., Castro, J. P., Silva, M. D., Almeida, M. C. D., Moreira-Filho, O., & Artoni, R. F. (2020). Dynamics of replication and nuclear localization of the B Chromosome in kidney tissue cells in Astyanax scabripinnis (Teleostei: Characidae). *Zebrafish*, *17*(2), 147-152.

Schmidt, T., & Heslop-Harrison, J. S. (1996). The physical and genomic organization of microsatellites in sugar beet. *Proceedings of the National Academy of Sciences*, *93*(16), 8761-8765.

Schwarz-Finsterle, J., Stein, S., Großmann, C., Schmitt, E., Schneider, H., Trakhtenbrot, L., ... & Hausmann, M. (2005). COMBO-FISH for focussed fluorescence labelling of gene domains: 3D-analysis of the genome architecture of abl and bcr in human blood cells. *Cell biology international*, *29*(12), 1038-1046.

Sharbel, T. F., Green, D. M., & Houben, A. (1998). B-chromosome origin in the endemic New Zealand frog Leiopelma hochstetteri through sex chromosome devolution. *Genome*, *41*(1), 14-22.

Shirasu, K., Schulman, A. H., Lahaye, T., & Schulze-Lefert, P. (2000). A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. *Genome Research*, *10*(7), 908-915.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, *7*(1), 539.

Silva, D. M. D. A., Pansonato-Alves, J. C., Utsunomia, R., Araya-Jaime, C., Ruiz-Ruano, F. J., Daniel, S. N., ... & Foresti, F. (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*(4), e94896.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., De Wit, E., ... & De Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nature genetics*, *38*(11), 1348-1354.

Smith-White, S., & Carter, C. R. (1970). The cytology of Brachycome lineariloba. *Chromosoma*, *30*(2), 129-153.

Sokal, R. R. Michener C.D. (1958). A statistical method for evaluating systematic relationships. *Univ. Kansas, Sci. Bull.*, *38*, 1409-1438.

Soltis, P. S., & Soltis, D. E. (2003). Applying the bootstrap in phylogeny reconstruction. *Statistical Science*, *18*(2), 256-267.

Soreng, R. J., Peterson, P. M., Romaschenko, K., Davidse, G., Teisher, J. K., Clark, L. G., ... & Zuloaga, F. O. (2017). A worldwide phylogenetic classification of the Poaceae (Gramineae) II: An update and a comparison of two 2015 classifications. *Journal of Systematics and evolution*, *55*(4), 259-290.

Sreenath, H. L., & Jagadishchandra, K. S. (1988). In vivo and in vitro instability of B chromosomes in palmarosa grass (Cymbopogon martinii var. motia). *Genome*, *30*(6), 966-973.

Stark, E. A., Connerton, I., Bennett, S. T., Barnes, S. R., Parker, J. S., & Forster, J. W. (1996). Molecular analysis of the structure of the maize B-chromosome. *Chromosome Research*, *4*(1), 15-23.

Staub, R. W. (1987). Leaf striping correlated with the presence of B chromosomes in maize. *Journal of Heredity*, *78*(2), 71-74.

Talavera, M., Lopez-Leon, M. D., Cabrero, J., & Camacho, J. P. M. (1990). Male germ line polysomy in the grasshopper Chorthippus binotatus: extra chromosomes are not transmitted. *Genome*, *33*(3), 384-388.

Teruel, M., Cabrero, J., Perfectti, F., & Camacho, J. P. M. (2010). B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma*, *119*(2), 217-225.

Thompson, J. D., Thierry, J. C., & Poch, O. (2003). RASCAL: rapid scanning and correction of multiple sequence alignments. *Bioinformatics*, *19*(9), 1155-1161.

Trifonov, V. A., Dementyeva, P. V., Larkin, D. M., O'Brien, P., Perelman, P. L., Yang, F., ... & Graphodatsky, A. S. (2013). Transcription of a protein-coding gene on B chromosomes of the Siberian roe deer (Capreolus pygargus). *BMC biology*, *11*(1), 1-11.
Trifonov, V. A., Dementyeva, P. V., Larkin, D. M., O'Brien, P., Perelman, P. L., Yang, F., ... & Graphodatsky, A. S. (2013). Transcription of a protein-coding gene on B chromosomes of the Siberian roe deer (Capreolus pygargus). *BMC biology*, *11*(1), 1-11.
Tseng, S. H., Peng, S. F., & Cheng, Y. M. (2018). Analysis of B chromosome nondisjunction induced by the r-X1 deficiency in maize. *Chromosome Research*, *26*(3), 153-162.

Uhl, C. H., & Moran, R. (1973). The chromosomes of Pachyphytum (Crassulaceae). *American Journal of Botany*, *60*(7), 648-656.

Valente, G. T., Conte, M. A., Fantinatti, B. E., Cabral-de-Mello, D. C., Carvalho, R. F., Vicari, M. R., ... & Martins, C. (2014). Origin and evolution of B chromosomes in the cichlid fish Astatotilapia latifasciata based on integrated genomic analyses. *Molecular biology and evolution*, *31*(8), 2061-2072.

Vicari, M. R., de Mello Pistune, H. F., Castro, J. P., de Almeida, M. C., Bertollo, L. A. C., Moreira-Filho, O., ... & Artoni, R. F. (2011). New insights on the origin of B chromosomes in Astyanax scabripinnis obtained by chromosome painting and FISH. *Genetica*, *139*(8), 1073-1081.

Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E. H., & Itakura, K. (1981). The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit β-globin DNA. *Nucleic acids research*, *9*(4), 879-894.

Warwick, S. I., & Black, L. D. (1983). The biology of Canadian weeds.: 61. Sorghum halepense (L.) Pers. *Canadian Journal of Plant Science*, *63*(4), 997-1014.

White, M. J. D. (1973). Animal cytology and evolution. 3rd ed. Cambridge, University Press, London.

Wilkes, T. M., Francki, M. G., Langridge, P., Karp, A., Jones, R. N., & Forster, J. W. (1995). Analysis of rye B-chromosome structure using fluorescencein situ hybridization (FISH). *Chromosome Research*, *3*(8), 466-472.

Wilson, E. B. (1907). Note on the chromosome-groups of Metapodius and Banasa. *The Biological Bulletin*, *12*(5), 303-313.

Wilson, E. B. (1909). Studies on chromosomes. V. The chromosomes of Metapodius. A contribution to the hypothesis of the genetic continuity of chromosomes. *Journal of Experimental Zoology*, *6*(2), 147-205.

Wu, T. P. (1978). Pachytene morphology of Sorghum nitidum chromosome complement. *Cytologia*, *43*(2), 433-440.

Wu, T. P. (1992). B-Chromsomes in Sorghum stipoideum. *Heredity*, *68*(5), 457-463.

Wu, T. P., & Pi, C. P. (1975). Accessory chromosome in Sorghum nitidum Pers. *Taiwania*, *20*(2), 147-161.

Xiong, J. (2006). *Essential bioinformatics*. Cambridge University Press.

Yu, W., Han, F., Gao, Z., Vega, J. M., & Birchler, J. A. (2007). Construction and behavior of engineered minichromosomes in maize. *Proceedings of the National Academy of Sciences*, *104*(21), 8924-8929.

Yu, W., Lamb, J. C., Han, F., & Birchler, J. A. (2006). Telomere-mediated chromosomal truncation in maize. *Proceedings of the National Academy of Sciences*, *103*(46), 17331-17336.

Zhou, Q., Zhu, H. M., Huang, Q. F., Zhao, L., Zhang, G. J., Roy, S. W., ... & Bachtrog, D. (2012). Deciphering neo-sex and B chromosome evolution by the draft genome of Drosophila albomicans. *BMC genomics*, *13*(1), 1-12.

8 LIST OF ABBREVIATIONS

3-D	three-dimensional
A(s)	A chromosome(s)
AFLP	amplified fragment length polymorphism
B(s)	B chromosome(s)
bp	base pairs
cDNA	complementary DNA
cpDNA	chloroplast DNA
DNA	deoxyribonucleic acid
EST	expressed sequence tag
ETS	external transcribed spacer
FISH	fluorescence *in situ* hybridization
GISH	genomic *in situ* hybridization
IGS	intergenic spacer
ITS	internal transcribed spacer
mb	mega base pairs
mRNA	mediator RNA
mtDNA	mitochondrial DNA
n	number of chromosomes in a haploid cell
NGS	next generation sequencing
NOR	nucleolus organizer region
NTS	non-transcribed spacer
PCR	polymerase chain reaction
PSR chromosome	paternal sex ratio chromosome
rDNA	ribosomal DNA
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
satDNA	satellite DNA

SNP	single nucleotide polymorphism
SSR	simple sequence repeat
STR	short tandem repeat
UPGMA	Unweighted pair group method with arithmetic mean

9 OTHER PUBLICATIONS

Flow cytometry in plant research. In Applied plant cell biology

Vrána J., Cápal P., Bednářová M. & Doležel J.

Editors: Peter Nick, Zdeněk Opatrný, pp 395-430

Springer-Verlag Berlin Heidelberg 2014

(Appendix I)

Flow Cytometry in Plant Research: A Success Story

Jan Vrána, Petr Cápal, Martina Bednářová, and Jaroslav Doležel

Abstract Flow cytometry is a powerful technique with numerous applications in biomedical research, including immunology, haematology, oncology and other fields. It has also found important applications in plant science where it accompanied scientists for almost four decades. Without its invaluable outcomes, some areas of plant research would not be in the position where they are now. This chapter focuses on exploitation of this state-of-the-art technology for studying plants at cellular and subcellular level, first providing a general overview and then focusing on nuclei and nuclear DNA content – by far the most frequent and most important application of flow cytometry in plant science. We review applications of the method from the early days to recent advances and discuss its applied aspects.

1 Introduction

Since its inception and, in particular, its commercialization in the late 1960s and the early 1970s, flow cytometry (FCM) played important roles in many areas of the life sciences. The field where this technique had the deepest impact is without any doubt biomedical research. However, it did not take plant biologists a long time before they realized the potential of the new technology. The main advantages, which attracted plant science and biotechnology, were (a) high throughput (hundreds or even thousands of particles can be analysed per second) providing statistically relevant data; (b) high accuracy and resolution; (c) negligible destructiveness (only small tissue samples are required); and (d) low operating costs per sample (once the initial instrument purchase costs have been invested). One publicly acclaimed disadvantage of FCM is the lack of visual information when compared

J. Vrána • P. Cápal • M. Bednářová • J. Doležel (✉)
Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Šlechtitelů 31, 783 71 Olomouc, Czech Republic
e-mail: dolezel@ueb.cas.cz

P. Nick and Z. Opatrný (eds.), *Applied Plant Cell Biology*, Plant Cell Monographs 22,
DOI 10.1007/978-3-642-41787-0_13, © Springer-Verlag Berlin Heidelberg 2014

to laser scanning cytometers and image analysers. This is true for all but one type of flow cytometer (Basiji et al. 2007). Nevertheless, the speed of analysis together with the growing army of fluorescent dyes, fluorescent proteins, fluorescently labelled antibodies and other reagents makes flow cytometry a formidable force capable of analysing over 15 parameters simultaneously in each particle (Chattopadhyay et al. 2008; Perfetto et al. 2004).

Flow cytometry requires samples in a form of aqueous suspension of single particles and this imposes some limits on its use in plant biology. Plant cells are tied together by a complex extracellular matrix in various tissues and organs and it may be difficult to isolate single cells. The cells have rigid cell walls, which may be autofluorescent, bind fluorescent probes nonspecifically and hamper staining of intracellular components. Due to irregular shape, plant cells disturb laminar flow in the narrow liquid stream compromising the precision of analysis. Moreover, many plant cell types may be larger than the diameter of the orifices in the flow chambers (typically 100 μm). One way how to prepare plant cells with regular shape suitable for flow cytometry is to remove their cell walls by digestion with hydrolytic enzymes and obtain protoplasts. Various projects relied on the analysis of protoplasts using flow cytometry (Galbraith 2007). Unfortunately, protoplast preparation is time-consuming, and as the tissues are exposed to various compounds and digestion represents a stress factor, protoplasts cannot be considered a good surrogate for intact cells. As with intact cells, protoplast analysis is complicated by their size; protoplast fragility and laborious preparation pose further constraints. Last but not least, protoplasts can be prepared only from some species and limited types of tissues.

Given the difficulties to analyse intact cells and/or protoplasts, it comes as no surprise that the most frequent application of flow cytometry in plant science has been the analysis of subcellular organelles – mitochondria, plastids, nuclei and chromosomes. Among the organelles, nuclei have been the most studied. This may be due to the fact that, unlike intact cells and protoplasts, they are small and can be relatively easily analysed. The nuclei harbour the majority of cellular DNA, and their analysis provides data on ploidy, genome size and cell cycle, to name just a few. The analysis of nuclei by flow cytometry and the application of their analysis are the main topics of this chapter.

Probably, the first paper ever on FCM in plants was published in 1973 (Heller 1973), but as it seems from today's perspective, plant biology was not ready for the revolution as the technique did not gain much attention and was largely overlooked, perhaps also because the paper was written in German. No papers on this topic were published for the long 10 years after this report. All changed in 1983, when David Galbraith and his colleagues at the University of Nebraska published a breakthrough paper, which brought FCM to a wider audience thanks to a simple protocol to prepare suspensions of intact nuclei suitable for flow cytometric analysis (Galbraith et al. 1983). This paved the way for this technology, and the number of plant flow cytometric papers has been steadily increasing.

Currently, flow cytometry has been used both in basic and applied plant research, as well as industrial applications, including breeding. It has become the

number one tool for quick, reliable and reproducible analysis of nuclear genome size and ploidy level. It has quickly found its way into taxonomy where it has been used to study speciation processes and population dynamics. The gender of plants and even the gender of the determining pollen grains can be determined using this method. This chapter discusses the most important applications of flow cytometry in plant biology with a special emphasis on the estimation of nuclear DNA content.

2 Principles of Flow Cytometry

Flow cytometry is an analytical and preparative technique offering high throughput and precision. Its unique feature is the ability to analyse multiple optical properties of single particles at a rate of several hundreds or thousands per second. Although flow cytometers are sophisticated instruments utilizing fluidics, lasers, optics and electronics, the basic principles are simple (Fig. 1). By virtue of a fluidic system, individual particles of the sample are brought to flow in a single file in a core of a narrow stream of liquid and pass individually through a beam of light (a typical light source is laser). Optical signals arising from the interaction between the particles and light (light scatter and fluorescence) are then steered by elements of the optical system to spectrally separated detectors (photomultiplier tubes or photodiodes), where optical pulses are transformed into electrical pulses which are then processed by electronic processing. The results are displayed as monoparametric frequency distribution histograms, biparametric (2D) dot plots or multiparametric dot plots.

A special group of flow cytometers, called sorters, can physically isolate particles of interest for further analyses, and they can do it simultaneously with the analysis and at high speed. There are two basic sorting systems: fluid switching and droplet, with the latter being the most common and efficient sorting type (Fig. 1). In this system, the fluid stream emerging from the flow chamber is subdivided into small droplets (the place where this occurs is called break off point). When a particle of interest is detected, the instrument waits until the particle reaches the break-off point and electrically charges the droplet containing the particle. Further downstream, the charged droplet is electrostatically deflected and collected in a tube.

Flow cytometers were originally developed to analyse blood cells but later were found suitable for the analysis of cell organelles (e.g. nuclei, mitochondria, plastids, chromosomes) and even particles as small as bacteria and viruses. Although it may be surprising, a plethora of physical and physiological parameters and processes could be inferred (e.g. size, internal complexity, total DNA content, pH and viability, to name just few) from the analysis of optical parameters of particles (light scattered at different angles, total fluorescence, fluoresce pulse profile).

Apart from the number of scientific publications, the popularity of flow cytometry can be judged from the increasing number of companies, which produce flow cytometers and sorters and the number of models on the market. The

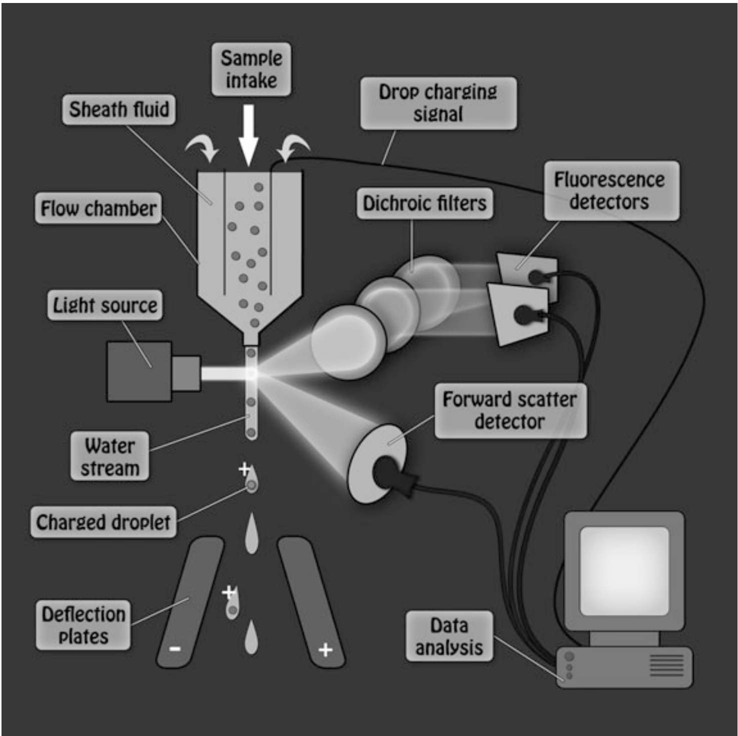

Fig. 1 Schematic view of a flow cytometer and sorter with one excitation light source, one detector of forward light scatter and two detectors of fluorescence signals. This type of instrument can sort two different populations of particles simultaneously

instruments differ in the number of functions, technical design and price tag. The range of flow cytometers is broad, starting with basic small footprint analysers capable of measuring just one or a few parameters, continuing with a strong middle class of cell analysers usually equipped with automatic sample loaders for quick analyses in multi-user facilities and finally ending with high-end multiparametric analysers and sorters which are often customized for particular needs. The reader is referred to the recent paper of Picot et al. (2012), which provides a qualified overview. The progress that flow cytometry has made since its beginnings (and a lot of useful information) can be found in Shapiro's excellent "Practical flow cytometry" book series (Shapiro 1985, 1988, 1995, 2003).

3 Plants and Flow Cytometry: An Uneasy but Fruitful Marriage

With only a few exceptions represented by the pollen grains of seed plants and other microspores of some nonvascular plants such as algae, plants are constituted of complex solid multicellular tissues, where individual cells are tied together in complex 3D matrix. Therefore, analysing intact plant cells using flow cytometry is not straightforward and found its applications in only a few areas of research. One of them is aquatic science, as most of phytoplankton exists as single cells. Phytoplankton analysis using flow cytometry is used mostly in marine ecology and biodiversity studies (reviewed in Dubelaar et al. 2007). The analysis in situ often requires customized (semi)automated instrumentation (Dubelaar et al. 1999). Another application of flow cytometry with plant cells is metabolic engineering. Plants themselves work as bioreactors; they synthesize a wide range of secondary metabolites, which may be used in pharmaceutical and food industry. In order to achieve higher yields of these compounds, cell cultures with standardized growth conditions have to be used (see also the chapter by Opatrný et al., this volume). Flow cytometry can be applied to analyse a broad range of plant products and culture properties, provided a desired trait correlates with fluorescence or light-scattering parameters that can be measured by FCM. The technique can be used for counting cells in culture, estimating their viability, growth potential and, most importantly, for detection of the desired metabolite accumulation (Gaurav et al. 2010). Moreover, flow sorting permits selection of subpopulations from cell cultures to establish highly producing cell lines (Schulze and Paulz 1998, Gaurav et al. 2010). Flow cytometry has also been used in plant pathology for detection, characterization and quantification of pathogens (reviewed by D'Hondt et al. 2011). The pathogens (viruses, bacteria, fungi and oomycetes) have been analysed for their presence and abundance, genome size and their biological status (viability, gene expression).

Cells bereft of cell walls are called protoplasts. Their advantage over whole cells is a round shape and hence regular passage in a laminar flow, but other disadvantages limit the usefulness of protoplasts for flow cytometry. The first is connected to lengthy preparation as plant tissues must be treated with cell wall-digesting enzymes. Thus, protoplasts are generally not suitable to study dynamic processes. Protoplasts may be large (up to 150 μm), depending on the tissue and species (Fox and Galbraith 1990), and therefore not compatible with most of commercial flow cytometers where nozzle diameters are typically in the range of 70–120 μm. The use of larger nozzles puts restrictions on the speed of analysis and sorting (Galbraith 1994). Without cell wall, plant protoplasts are fragile, which may limit their use and they cannot be prepared from all species and tissue types (Galbraith 2007). Since nuclei are located near plasma membrane and not in the centre of protoplasts, the protoplasts are generally not suited for estimation of DNA content as the nuclei are irregularly positioned against the focus of excitation light, resulting in a variability of signals (Galbraith 1990; Ulrich et al. 1988). Despite the difficulties, there have

been interesting applications of flow cytometry involving plant protoplasts. The most important uses of protoplasts include the analysis of gene expression (Birnbaum et al. 2003) and sorting of heterokaryons after protoplast fusion for production of somatic hybrid plants (Liu et al. 1995). Other applications of plant protoplast analysis include estimation of protein contents (Naill and Roberts 2005), measurement of physiological parameters such as pH (Giglioli-Guivarc'h et al. 1996) and analysis of apoptosis (Yao et al. 2004; Watanabe et al. 2002).

Apart from whole cells, also cell organelles can be analysed by flow cytometry. While nuclei are the most studied organelles using flow cytometry, flow cytometry was used to some extent to study chloroplasts and mitochondria. Analysis of isolated intact chloroplasts discriminated two populations of chloroplasts based on different chlorophyll fluorescence intensities (Kausch and Bruce 1994). Pfündel and Meister (1996) succeeded in discriminating chloroplast thylakoids obtained from mesophyll and bundle sheath cells of maize based on their differences in fluorescence spectra and sorted them for further analysis. In order to investigate the process of DNA depletion in chloroplasts during leaf development, Rowan et al. (2007) developed a flow cytometric method for detection of DNA in chloroplasts. Unfortunately, only a few studies have been undertaken on isolated plant mitochondria using flow cytometry. Petit et al. (1986) measured the binding of concanavalin A to mitochondria of potato, and Petit (1992) successfully monitored changes in membrane potential using Rhodamine 123. The only application of flow cytometry in plant research which can rival nuclear analysis and sorting in its importance and abundance is flow cytometric analysis and sorting of plant metaphase chromosomes. Flow-sorted plant chromosomes found many applications in cytogenetics and genomics. The recent article by Doležel et al. (2012) provides an extensive overview of this field.

4 Plant Nuclei in Focus

Nuclei are important cellular organelles where the majority of hereditary information is stored and important processes related to genome replication, repair and response to various stimuli and transcription into a range of RNA species take place. These processes impact the destiny of a cell, tissue and the whole organism as well as its progeny. Details of these processes and their significance are far from understood and flow cytometry has been revealing and certainly will continue revealing many of the enigmas of Mother Nature.

But before flow cytometry of cell nuclei could become a useful tool in plant biology, many problems and obstacles have to be solved and new approaches and protocols developed. The main difficulties were and to some extent still are associated with preparation of samples accessible for this method. The obstacles are numerous due to plant cell structure and physiology mentioned above – rigid cell wall, presence of secondary metabolites in cytosol, auto-fluorescence from

photosynthetic pigments and others, which will be dealt with in following paragraphs.

But all these difficulties are worth to be overcome as flow cytometry is able to give important results. The two most frequent applications of FCM on plant nuclei are the measurement of genome size in absolute units and the estimation of ploidy level. Other applications include cell cycle analysis, tissue-specific gene expression analysis and nuclear sorting.

5 Principles and Methodology

5.1 Sample Preparation

Flow cytometry requires samples in a form of aqueous suspensions of particles. Plants are not very cooperative in this regard as their cells have rigid walls and it may not be easy to release intact nuclei. The oldest method is based on enzymatic treatment of plant tissues in order to obtain wall-free cells – protoplasts – from which the nuclei are released into the isolation buffer by osmotic lysis (Heller 1973). This method is laborious and time-consuming and therefore impractical for wider use. However, as the nuclei are released gently, this method yields histograms of DNA content with excellent resolution and minimum of debris background (Ulrich and Ulrich 1991). The fact that protoplasts isolation takes time makes this method unsuitable for the analysis of cell cycle kinetics.

To date, the most frequently used method for preparation of nuclear suspensions is the ingeniously simple method developed by Galbraith et al. (1983). The method relies on homogenization of fresh tissue using a razor blade in a small volume of isolation buffer into which the cellular contents are released. Filtering the crude homogenate through a nylon mesh removes large tissue debris to avoid blockage of sample tubing and nozzle. The main advantages of this method are the speed and the need of only small amounts of practically any plant tissue (leaves, roots, stems, flowers, etc.). Silva et al. (2010) went so far that they developed a protocol for successful isolation of intact nuclei from only a single root meristem. The only drawback of the chopping method is the difficulty to automatize sample preparation for high-throughput applications. Therefore, Roberts (2007) adapted the so-called bead-beating method, routinely used for releasing DNA from cells for genomic studies (Haymes et al. 2004; Harmon et al. 2006), where small amount of plant tissue and isolation media are shaken together with zirconia/silica beads (2.5 mm in diameter) and intact nuclei are released. Several samples could be prepared at the same time using this method. The authors concluded that this method was suitable not only for fresh material but also for dried herbarium specimens and pollen. Cousin et al. (2009) took advantage of this sample preparation method, and using a cytometer equipped with an automatic sample loader, they created a high-throughput pipeline where one operator could analyse several hundred samples

per working day. One of few exceptions, where razor chopping was not very successful, is pollen grains, and therefore, Kron and Husband (2012) developed a method for isolating pollen nuclei by inducing bursting of pollen through a nylon mesh filter.

5.2 Nuclear Isolation Buffers

Chemical composition of the nuclear isolation buffers is important and a correct choice of a buffer may be the key to successful flow cytometric analysis. A number of isolation buffers has been published, differing in chemical composition, but they all should facilitate the release of nuclei free of cytoplasm and, in sufficient quantities, maintain the integrity of isolated nuclei, protect DNA against nucleases and facilitate stoichiometric DNA staining (Doležel 1991). Generally, the isolation buffers contain substances that stabilize nuclear chromatin such as magnesium ions (Galbraith et al. 1983; Arumuganathan and Earle 1991; Pfosser et al. 1995) or spermine (Doležel et al. 1989), chelating agents which serve as nuclease cofactors such as EDTA (Doležel et al. 1989; Marie and Brown 1993) or sodium citrate (Galbraith et al. 1983; Marie and Brown 1993) and inorganic salts (KCl, NaCl) to adjust appropriate ionic strength. Detergents (e.g. Triton X-100 or Tween 20) are included to facilitate nuclear release from the cytoplasm and to prevent nuclei from aggregation. Doležel and Bartoš (2005) and Greilhuber et al. (2007) list popular isolation buffers and discuss their advantages and limitations. Loureiro et al. (2006a) conducted a series of tests for the four most popular buffers. Although some buffers performed better than others, the authors concluded that there is no universal buffer suitable for every species and tissue due to diversity of plant tissues in structure and chemical composition, and thus, suitable buffers must be tested individually for a given material. Based on these findings, Loureiro et al. (2007a) developed two improved buffers. One performs well in species relatively free of cytosolic compounds (general purpose buffer (GPB)), while the other is more suitable for more problematic tissues (woody plant buffer (WPB)). They claim that WPB works better in problematic tissues/species than other nuclear isolation buffers and therefore is recommended as the first choice.

5.3 Parameters Analysed on Plant Nuclei

Fluorescence has been the most common optical parameter in flow cytometric analysis of plant nuclei. Although light scatter is helpful in detection and determination of single nuclei populations and also in detection of effects of cytosolic compounds on estimation of DNA amounts (Loureiro et al. 2006a, b), their use in flow cytometric analysis of plants is rather scarce. Fluorescence is a physical phenomenon caused by absorption of light energy by a molecule (fluorochrome)

and subsequent emission of light with longer wavelength. The fluorochromes can be either innate to the examined particle (so-called auto-fluorescence) or they are added artificially (this is the usual case). The choice of fluorochromes is very rich these days and new dyes are added to the list continuously. One can choose dyes for a wide range of structural or physiological features of the nuclei, including nucleic acid dyes (e.g. DAPI, propidium iodide, acridine orange), protein dyes, which can be used as fluorescent "tags" (e.g. FITC, PE, APC) and fluorescent proteins to monitor gene expression (e.g. GFP, YFP, DsRed). The most frequently used fluorochromes to analyse plant nuclei are nucleic acid dyes.

5.4 Nuclear Staining

The most common dyes used to stain plant nuclear DNA are DAPI and PI (Doležel et al. 1992a). DAPI (4′,6-diamidino-2-phenylindole) is a minor groove-binding molecule with selectivity for AT bases. It is used in non-saturating concentrations, typically 2–4 μg/ml (Pfosser 1989; Ulrich et al. 1988). DNA-bound DAPI is best excited by near UV light (358 nm excitation maximum), and therefore, it has become popular among plant researchers who preferred cheaper arc lamp-based instruments. DAPI binding to DNA is mostly independent of chromatin structure and the peaks on histograms of nuclei fluorescence are characterized by low variability. As DAPI is base selective and therefore not suitable for absolute DNA measurements (Doležel et al. 1992a), it is used predominantly for estimation of relative amounts to determine ploidy levels. PI (propidium iodide) is an intercalating dye and has no apparent base specificity. PI is used in saturating concentrations, typically 50–100 μg/ml (Taylor and Milthorpe 1980). Its most frequent use is the measurement of absolute DNA amounts, but as it also binds to double-stranded RNA, samples must be treated with RNAse. PI provides a broad range of excitation wavelengths (325–568 nm), but the most common wavelengths for PI excitation in commercial instruments are blue (488 nm) and green (532 nm) laser lines.

The progress in chemistry offers new fluorochromes, which may replace the traditionally used PI. For instance, Clarindo and Carvalho (2011) used SYBR Green 1 for measurement of genome size in two *Coffea* species and compared it with results of PI-stained nuclei. The results for genome size obtained either with SYBR Green 1 or PI were statistically identical. Together with the observation that the peaks of G0/G1 nuclei stained by SYBR Green 1 had lower coefficients of variation compared to those stained with PI and the fact that SYBR Green 1 is less mutagenic than propidium iodide, these findings indicate that SYBR Green 1 can be used for flow cytometric experiments with plants. Another advantage is that the excitation maximum of SYBR Green 1, compared to PI, is closer to 488 nm, the most common laser line in most commercial analysers, and therefore, the excitation of fluorescence is more efficient.

Fluorescent antibiotics such as chromomycin A3, mithramycin and olivomycin represent a further class of DNA dyes which bind preferentially to the GC-rich

regions of DNA. Their use in plant nuclear research is limited mainly to base composition studies (Meister and Barow 2007), as described later in this chapter. Acridine orange (AO) has been used even less frequently (Bergounioux et al. 1988). The dye binds to DNA and dsRNA and yields a green fluorescence upon binding to DNA, while the RNA-bound dye produces red fluorescence. This property of AO is used in human and animal cells to estimate simultaneously DNA and RNA content and to discriminate between dividing and quiescent cells (Darzynkiewicz 1994).

5.5 Secondary Metabolites

Secondary metabolites present in cytosol can affect the accessibility of the dye to the DNA. Their effects have been recognized first by micro-spectrophotometry (Greilhuber 1986, 1988) and only later in plant FCM analyses (Noirot et al. 2000, 2002, 2003, 2005; Price et al. 2000). Interference of cytosolic compounds with staining of DNA was originally observed as fluorescence inhibition but can also be manifest as an increase in fluorescence (Noirot et al. 2003; Loureiro et al. 2006b). Loureiro et al. (2006b) showed that cytosolic compounds can change light-scattering properties of nuclei (increase in side scatter signal intensity and decrease in resolution of histograms) and termed this phenomenon “tannic acid effect”. Changes in light scatter provide an opportunity to detect the interference of cytosol. To eliminate unfavourable effects of cytosolic compounds in studies on genome size, a good practice has been to carefully choose the most suitable isolation buffers, inhibitors of phenolic compounds such as mercaptoethanol and PVP (Price et al. 2000; Noirot et al. 2003, Yokoya et al. 2000) and to select tissues with lower content (or preferably even the absence) of phenolic compounds (Loureiro et al. 2007a).

5.6 Materials for Nuclear Isolation

Almost any fresh plant tissue is suitable to prepare suspensions of intact nuclei for flow cytometry using the mechanical chopping method. This, however, holds true only if the tissues are healthy as degradation of DNA results in low resolution of DNA content histograms. A problem may occur in field conditions where rapid dispatch of fresh material to the laboratory may not be possible. The difficulty may be overcome either by using a flow cytometer *on site* or by shipping preserved plant tissues to the laboratory. As the development of instrumentation continues, there are several small portable cytometers available on the market. However, their use in exotic and remote localities may be hampered by shortage of continuous supply of power and consumables. Plant tissues can be preserved in several ways. For short-term (several days) storage/transport, tissues can be bagged with moistened paper and kept at low temperatures (ca. 4°C). To ensure preservation over longer periods

of chemical fixation of the material (Doležel et al. 1992b; Suda and Trávníček 2006), dried samples (silica gel-dried samples or traditional herbarium vouchers; Suda et Trávníček 2006) or glycerol-preserved nuclei (Hopping 1993, Kolář et al. 2012) have been employed. Alternatively, dormant seeds may be used (Matzk et al. 2000; Śliwinska et al. 2005).

Chemical fixation is routinely used to store human and animal cells, but for plants it has turned out not to be practical, as the release of nuclei from cells fixed in some fixatives such as 3:1 (ethanol to acetic acid) is difficult or impossible. Importantly, fluorescence of some DNA dyes is altered by the fixation, and fixed materials are not suitable for some applications. For example, formaldehyde-fixed samples are not amenable to estimation of DNA content in absolute units, as the fixation interferes with propidium iodide staining (Becker and Mikel 1990; Overton and McCoy 1994). However, for some applications (e.g. ploidy level estimation and cell cycle analysis), combination of DAPI staining and formaldehyde-fixed nuclei is feasible (Jarret et al. 1995, Sgorbati et al. 1986). Flow cytometric estimation of ploidy levels was carried out successfully with plant tissues that have been rapidly frozen (Dart et al. 2004; Nsabimana and Van Staden 2006; Halverson et al. 2008; Cires et al. 2009). Several authors (Baranyi et al. 1996; Matzk et al. 2000; Śliwinska et al. 2005, 2009; Jedrzejcyzk and Śliwinska 2010) investigated the possibilities to estimate DNA content from dry seeds, which can be easily transported from a collection site to the laboratory. The use of seeds may be advantageous especially for species that contain cytosolic compounds in leaf cells, affecting quality of DNA content histograms (Śliwinska et al. 2005, Matzk 2007). However, this approach bears some limitations. The seeds must be collected during seed maturation season, and there are additional difficulties linked with seed germination ex situ and taxonomic complexity (Śliwinska et al. 2005, Kolář et al. 2012). Śliwinska et al. (2009) conclude that it is possible to use seeds for flow cytometric measurements of nuclear DNA content, but detailed understanding of seed biology is needed to interpret the results correctly. However, when the seeds are used up for flow cytometric analyses, no further analysis, such as karyological or molecular investigations, is possible. As an alternative approach seeds can be collected and germinated to analyse the growing plants in the laboratory (Suda et al. 2005).

Suda and Trávníček (2006) analysed nuclei isolated from herbarium and silica gel-dried material (up to 2 years old) in a set of plant groups. They concluded that as the quality of DNA histograms might be compromised (shifts in fluorescence intensity compared to that of fresh samples and decrease of uniformity of fluorescence, resulting in higher coefficients of variation (CV)), this method should be generally avoided for estimation of DNA content in absolute units. For example, Šmarda (2006) reported up to 10 % difference in estimates of DNA content between fresh and dry tissues in *Festuca*. On the other hand, dehydrated vouchers of mosses were suitable even for genome size estimation in absolute units (Voglmayr 2000). In contrast to genome size estimation, herbarium and silica gel-dried materials proved suitable for large-scale ploidy screening (Šmarda et al. 2005, Šmarda and Stančík 2006, Schönswetter et al. 2007, Suda and Trávníček 2006, Whittemore and Olsen 2011). But the idea of using silica gel-dried material for genome size studies

is still pursued, and Bainard et al. (2011a) discussed criteria and conditions under which this method might be used more broadly. Kolář et al. (2012) revisited a method to preserve suspensions of isolated nuclei in glycerol published by Chiatante et al. (1990) and Hopping (1993). This method is suitable even for genome size estimation and Kolář et al. (2012) state that this method, although requiring sample preparation in situ, is probably the most reliable way to preserve plant material for all flow cytometric applications. The same authors predict that the shipping of cooled fresh material for short-term and rapid silica-gel drying for long-term field trips, respectively, will still remain the methods of choice for the near future.

5.7 DNA Flow Cytometry

To correctly interpret the results of flow cytometric analysis of DNA content, one needs to consider its changes during cell cycle. Cycling cells undergo cyclical changes in nuclear DNA content, which can be described using C-values. In G_0/G_1 phase, the nuclei of somatic cells have an identical DNA amount – 2C (i.e. two copies of nuclear genome) – while nuclei in G_2/M phases have twice as much DNA described as 4C. During DNA synthesis (S) phase, cellular DNA content varies between 2C and 4C. As typical result of FCM analysis of DNA content histograms for nuclei in G1 and G2 phase yield narrow peaks, while the nuclei in S show a distribution ranging from 2C to 4C levels (Fig. 2). The resolution of flow cytometric measurements is influenced by many factors as discussed above. Quality of peaks is characterized by a coefficient of variation (CV), which is defined as standard deviation of the peak divided by the mean value of the peak. The lower the CV, the more accurate is the measurement. Other factors influencing reproducibility of histograms are the presence of debris and aggregates. Debris can be removed from histograms by combination of electronical thresholding and good gating strategies. Doublets and other aggregates are usually gated out from populations of single nuclei using 2-D (biparametric) histograms where one parameter is total fluorescence (DNA content) and the other is the width of fluorescence pulse (aggregates are usually larger/longer than singlets).

6 Biological Aspects: Analysis of Nuclei

6.1 Genome Size

Nuclear genome size (DNA amount representing one copy of nuclear genome) is an inherent and unique attribute of every eukaryote and its knowledge is critical for many areas of research, including taxonomy, ecology and evolutionary biology.

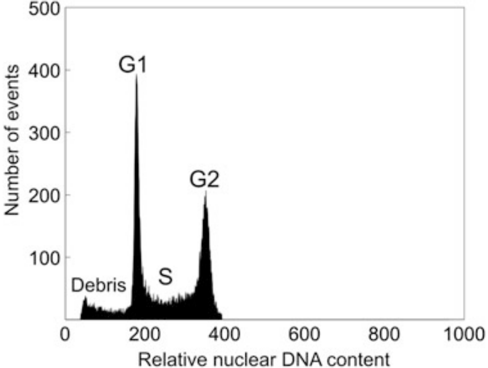

Fig. 2 Distribution of nuclear DNA content in a population of cycling cells. Histogram of relative DNA content obtained after the analysis of DAPI-stained nuclei isolated from actively growing root tips of bread wheat. As cells progress from G_1 to G_2 phase of cell cycle, their DNA content doubles. This situation reflected on the histogram by two peaks representing G_1 and G_2 nuclei with 2C and 4D DNA amounts, respectively. The distribution between the two peaks represents nuclei in S phase with DNA amounts ranging between 2C and 4C. The distribution to the left of G_1 peak represents cellular and nuclear debris

Data on genome size are essential for genomic studies as sequencing programmes need this information to estimate costs and manage the complexity of individual projects (Cardoso et al. 2012). Yet, despite all efforts, it was estimated in 2007 that only about 1.8 % of angiosperm plants have been analysed for genome size (Leitch and Bennett 2007). The most recent report lists DNA content estimates for 6287 angiosperm species (Bennett and Leitch 2011). A high-throughput method is thus needed to estimate genome size in a majority of plants, and flow cytometry has been a major player in this area.

Historically, the first method to estimate nuclear DNA content was chemical extraction (Schmidt and Thannhauser 1945), but it was laborious and time-consuming, therefore impractical for wider use. A further disadvantage was that it provided average values for the population and the results were compromised by the occurrence of cells in S and G2 phases of cell cycle. The most popular method before the advent of FCM was Feulgen microdensitometry (Typas and Heale 1980; Voglmayr and Greilhuber 1998). This method estimates DNA amount by measuring absorption of monochromatic light by the stained nucleus and comparing it to known standards. A modern version of densitometry is the so-called image cytometry (ICM), relying on CCD camera and image analysis software (Hardie et al. 2002). Although very reliable, the main disadvantages of microdensitometry are the rather limited speed of analysis and a need for mitotically active tissues (Greilhuber 2008). In contrast, FCM offers both speed and reliability, does not require dividing cells and provides greater statistical accuracy. These may be the reasons why this method is today by far the most popular (Bennett and Leitch 2005; Leitch and Bennett 2007; Greilhuber et al. 2007). The trend is clearly visible in the compilations of plant C-values made by the Bennett group; while in 2005 the

percentages of first estimates made by FCM were some 58 % (Bennett and Leitch 2005), in the most recent release of C-value data compilation, they reached almost 85 % (Bennett and Leitch 2011). Several studies (e.g. Doležel et al. 1998; Vilhar et al. 2001) verified that results obtained by both FCM and Feulgen densitometry are reliable and comparable and can be used in parallel.

DNA amounts of eukaryotic organisms are expressed as C-values (C for constant); this term was first used by Swift (1950) based on a concept first proposed by Boivin et al. (1948) that within any tissue of an organism, the actual amount of DNA per somatic nucleus is constant. The DNA amount in the unreplicated gametic nuclear chromosome complement (n) is 1C, while DNA amount of a diploid (somatic) nucleus is 2C. Ambiguities in terminology may occur in polyploid organisms, and therefore, Greilhuber et al. (2005) proposed a unified terminology, where 1C value refers to half of somatic DNA content (2C), while 1Cx is 2C value divided by the ploidy level of the respective organism. 1C and 1Cx values correspond to the holoploid and monoploid genome sizes, respectively. These two values are the same only in diploids. A more complete version of the terminology on genome size which deals with various cytogenetic conditions, life cycle segments and nuclear phases was proposed by Greilhuber and Doležel (2009).

In order to report on absolute DNA amounts, C-values are quoted either in picograms (pg) or base pairs (bp), with the conversion formula being 1 pg DNA = 978 Mbp (Doležel et al. 2003). The first compilation of published data on plant DNA C-values was published by Bennet and Smith (1976), and since then, new C-values have been reviewed on a regular basis (Bennett et al. 1982, 2000; Bennett and Leitch 1995, 1997, 2005, 2011; Zonneveld et al. 2005). From 1997 onwards, the C-values database is available online (http://data.kew.org/cvalues/). But, as some published data for the same species could vary significantly, Bennet and Smith (1976) stressed already in the first release of their database and others agreed later (Temsch and Greilhuber 2000, 2001; Doležel et al. 1998) that some results should be critically re-evaluated as not all data were obtained in experiments carried out in the right way. Problems with standardization, wrong choice of fluorochromes and ignorance of the effects of secondary metabolites can lead to misleading results.

Estimation of DNA content in absolute units is based on comparison of relative DNA amounts of the unknown sample with a sample with known C-value (reference standard):

$$\begin{aligned}&\text{2C value of unknown sample} =\\&[\mathrm{G_0/G_1}\ \text{peak mean of unknown sample}/\mathrm{G_0/G_1}\ \text{peak mean of reference standard}]\\&\times\ \text{2C value of reference standard (pg).}\end{aligned}$$

There are two principal ways of standardization: external and internal. External standards are measured separately from the sample, while the internal standard is processed together with the sample. In order to avoid errors due to variation in sample preparation and staining, internal standardization has been recommended (Doležel et al. 1992a, Greilhuber 2008). Doležel et al. (1998) compared results of

genome size estimations in four different laboratories and concluded that flow cytometry is a reliable method, when a set of reference standards and their calibration are adopted. The issue of standardization has been dealt with in more detail by several authors (e.g. Doležel et al. 1992a; Johnston et al. 1999; Bennet et al. 2003; Doležel and Bartoš 2005; Greilhuber et al. 2007; Suda and Leitch 2010; Praça-Fontes et al. 2011).

In general, a reference standard must have well-defined genome size, preferably not too distant from the examined sample (Bennett et al. 2003; Doležel et al. 1992a; Doležel and Bartoš 2005) but not too close to avoid that the peaks of G_1 nuclei of sample and standard would overlap (Greilhuber et al. 2007), and the standard must be biologically similar, i.e. plant standards should be used for plant samples (Suda and Leitch 2010), it must be cytologically stable, it must contain low (or no) level of secondary metabolites and it must be generally available. As 1C DNA amounts in higher plants range from 0.065 pg in *Genlisea margaretae* (Greilhuber et al. 2006) to 152.2 pg in *Paris japonica* (Pellicer et al. 2010), which corresponds to a difference of approximately 2,400-fold, it is impossible to use a single standard for all species. Doležel et al. (1998, 2007) give a list of recommended standards with 1C genome sizes ranging from 1.1 to 34.89 pg DNA. These standards were calibrated using human leukocytes considering 2C values of 7 pg DNA (Doležel et al. 1998), a value recommended by Tiersch et al. (1989) as reference value for estimation of genome size in animals. The standards are freely available from the Doležel lab (Doležel et al. 2007). Ideally, 1C values of reference standards should be known exactly. Unfortunately, due to difficulties in assembling repetitive parts of genomes, the exact genome size is not known for humans and also not for plants that would be otherwise suitable as reference standards (Doležel and Greilhuber 2010).

To conclude, some of the most important critical preconditions for successful and reliable estimation of DNA content in absolute units are (a) precisely aligned instrument, (b) awareness of possible interference of secondary metabolites with the staining, (c) use of intercalating DNA stains and RNAse treatment of nuclei before staining, (d) proper standardization procedure and suitable reference standards and (e) other factors, such as buffer composition and dye concentration (Bainard et al. 2010). Moreover, it has been recommended that measurements have to be performed with at least three different plants from the same species or population and that the analyses are repeated at least three times, ideally on different days (Lysák et al. 1999).

6.2 Ploidy

Probably the most common use of FCM in plants has been the estimation of nuclear DNA content in relative units to determine ploidy levels. The number of chromosome sets is especially important especially for plant taxonomy and plant breeding programmes (Doležel 1997; Eeckhaut et al. 2005; Ochatt 2008). The most reliable

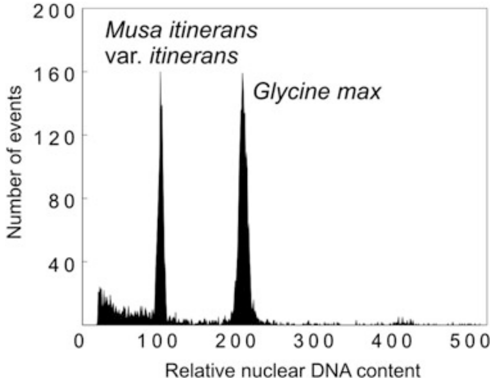

Fig. 3 Estimation of nuclear DNA amount in absolute units. Histogram of relative DNA content was obtained after simultaneous analysis of propidium iodide-stained nuclei of a banana species *Musa itinerans* var. *itinerans* and soybean (*Glycine max*). The soybean with 2C = 2.50 pg DNA was used as internal reference standard to estimate DNA amount of the banana species according to the formula:

2C value of *Musa* = (2.5 × G_1 peak mean of *Musa*)/(G_1 peak mean of *Glycine*)

The estimated 2C amount of *Musa* was 1.217 pg DNA. The mass of DNA in pg can be converted to the number of base pairs considering 1 pg DNA = 0.978 × 10^9 bp (Doležel et al. 2003)

method to determine ploidy levels is chromosome counting. However, this is laborious and time-consuming and should be done by experienced cytologists (especially in species with numerous and small chromosomes). Moreover, this approach requires actively dividing cells. There are indirect techniques for ploidy estimation including the estimation of leaf stomatal density and size (van Duren et al. 1996), determination of chloroplast number in guard cells and pollen diameter (Mishra 1997), but none of them is reliable enough for routine use. Thus, flow cytometry has become the method of choice in ploidy level estimation (Fig. 3).

The first report on estimation of ploidy levels using FCM was published by de Laat et al. (1987) and has been followed by an ever-increasing number of reports. As the amount of data on ploidy levels obtained by flow cytometry increases, Doležel (1991) and later Suda et al. (2006) appeal for consistent terminology first proposed by the Committee on Nomenclature, Society of Analytical Cytology (now International Society for Advancement of Cytometry; Hiddemann et al. 1984), distinguishing between the results obtained by conventional chromosome counting using microscopy (ploidy) and data obtained by flow cytometric analysis (DNA ploidy). As the main reason for this terminological distinction, Suda et al. (2006) list and discuss several cases in which discrepancies between the results from karyology and flow cytometry could arise.

A prerequisite for reliable estimation of ploidy levels using flow cytometry is a correlation between chromosome number and nuclear DNA amount. The results must be calibrated against a sample with known number of chromosomes as standard (Suda et al. 2006), which must be an individual from the same species. Unlike in the estimation of genome size, it is generally accepted to use external

standards in ploidy screening. However, as internal standardization eliminates potential errors due to variation in sample preparation and instrumental "drifts", internal standards may be preferred, especially in cases where aneuploidy is expected. Samples for ploidy screening are often stained by DAPI instead of PI or ethidium bromide (EB), because its binding to DNA is less affected by chromatin structure, and DNA peaks show lower variation. Moreover, DAPI can be used in lower concentrations, does not bind to RNA and can be excited by mercury lamps, which were common in older models of benchtop instruments popular among botanists.

Apart from the estimation of the number of complete chromosome sets, under certain conditions FCM may be used to detect minor changes in nuclear DNA amounts due to gain or loss of a few or even only a single chromosome (aneuploidy). Bashir et al. (1993) and Pfosser et al. (1995) analysed genome sizes of several wheat-rye addition lines and confirmed that flow cytometry was able to detect differences in DNA content as small as 1.8 %, corresponding to the presence of single telocentric rye chromosomes on the background of an entire wheat chromosome set. Roux et al. (2003) used flow cytometry for rapid detection of aneuploidy in triploid *Musa*, and the results were confirmed by conventional chromosome counting. Flow cytometry was also successful in detection of monosomy and trisomy in hops (Šesek et al. 2000).

In many plant species, differentiated tissues comprise cells with DNA contents higher than 2C (i.e. 4C, 8C). This phenomenon is called endopolyploidy and is usually caused by endoreduplication, which involves repeated rounds of DNA synthesis without intervening mitoses, leading to chromosomes with 4, 8, 16 and more chromatids. These cells do not divide, and flow cytometry is probably the best method to identify them and establish the frequency of individual levels of endopolyploidy. Endopolyploidy is common in angiosperms (Barow 2006; Barow and Jovtchev 2007) and mosses (Bainard and Newmaster 2010) but is rare in gymnosperms and ferns (Barow and Jovtchev 2007; Barow and Meister 2003) and even absent in liverworts (Bainard and Newmaster 2010). Although the biological significance of endoreduplication is poorly understood, it was proposed that it plays an important role in plant cell and tissue growth and differentiation (reviewed by Chevalier et al. 2011). In flow cytometric experiments, endopolyploidy is manifest in DNA content histograms as series of distinctive peaks, each corresponding to nuclei with different level of endoreduplication (Fig. 4). The degree of endopolyploidy (mean value of endopolyploidization) in different samples (organism, organ or tissue level) has been quantified as mean C-level and cycle value (Barow and Jovtchev 2007; Barow and Meister 2003).

6.3 Base Composition

Base composition (ratio of AT to GC base pairs) in nuclear DNA differs between species, may have a biological relevance and can be used as additional parameter in

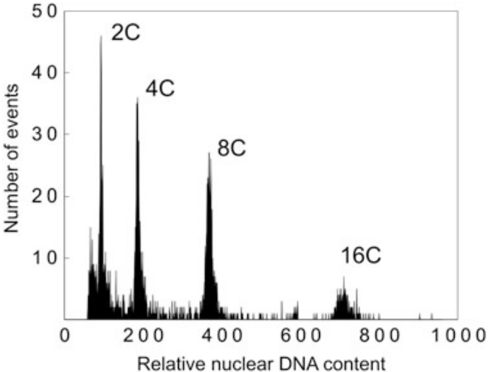

Fig. 4 Analysis of endoreduplication. Histogram of relative DNA amount was obtained after the analysis of DAPI-stained nuclei isolated from 14-day-old plants of *Arabidopsis thaliana*. The histogram comprises three peaks – one representing G_1 phase cells with 2C DNA content, one representing G_2 phase cells with 4C DNA amount, while the third peak represents nuclei with 8C DNA amounts which resulted from one round of endoreduplication and fourth peak stands for nuclei that experienced two rounds of endoreduplication and thus have 16C DNA content

taxonomic studies. DNA base content can be estimated using physicochemical methods such as chromatography, density centrifugation, DNA temperature melting analysis and UV absorbance and, more recently, by complete genome sequencing (Meister and Barow 2007; Šmarda et al. 2011). Nevertheless, flow cytometry offers comparable results for a fraction of time, cost and from the smallest amount of material as compared to other methods.

Flow cytometric measurements of base composition exploit DNA dyes with different affinity for DNA bases: intercalators with no base preference (typically propidium iodide) are compared to AT-specific dyes (e.g. DAPI, Hoechst dyes) or fluorescent antibiotics with GC preference (e.g. chromomycin A3, mithramycin A and olivomycin). Sample and standard are processed together and their fluorescence intensity profiles are measured independently with both the intercalating dye and the base-specific dye (either AT or GC specific). AT or GC content (depending on the base-specific dye used) is then calculated using the formula of Godelle et al. (1993). However, the fluorescence intensity found for the base-specific dyes does not correlate with base composition in a linear manner, and dye binding to DNA is influenced by the binding length (minimum number of consecutive complementary bases necessary to bind one dye molecule) and by biased base distribution over the length of DNA molecule (Godelle et al. 1993; Barow and Meister 2002). Therefore, all estimations of AT (or GC content) by flow cytometry should be taken as approximates. Barow and Meister (2002) conducted measurements of genome sizes and base contents for 54 plant species and concluded that there is no general correlation between genome size and AT/GC ratio in higher plants. The latest compilation of base composition values (Meister and Barow 2007) included 215 species. Since then, new estimates were published (Favoreto et al. 2012; Šmarda et al. 2008, 2011) and to date, there are data for about 300 species. These

numbers contrast with the number of data on genome size or DNA ploidy estimates. The main reasons for the scarcity of data may be higher laboriousness of the method (the need to consecutively measure the samples with two different dyes) and the need for internal standard with known C-value and AT/GC ratio, as well as uncertainties for some input parameters (binding length of the dye, nonrandomness of base distribution). Several studies tried to find some correlation between base composition and other biological parameters. For example, Vinogradov (1994) found a high correlation between GC content and genome size in angiosperms. However, other studies did not confirm these results (Cerbah et al. 2001; Ricroch et al. 2005), or even found opposite correlations (Barow and Meister 2002).

6.4 Cell Cycle

Basic information about the distribution of cells over the various phases of cell cycle can be obtained from a single parametric analysis of DNA content. However, there are several drawbacks to this method: cycling cannot be distinguished from quiescent cells, mathematical algorithms must be applied to resolve all major phases and lastly, this method is not amenable for kinetic studies and not suited for disturbed populations. Monoparametric flow cytometric cell cycle analyses were employed, for example, in studies on cell cycle activity and microtubule organization in seeds (Fujikura et al. 1999; Pawlowski et al. 2004), or in roots (Binarová et al. 1993), or to evaluate the effects of various drugs on the cell cycle (Binarová and Doležel 1993; Binarová et al. 1998a, b).

More refined approaches involve simultaneous analysis of DNA content and several key processes or molecules involved in the cell cycle (e.g. DNA synthesis, cyclin-dependent kinases). There are several methods which rely on detection of DNA synthesis, such as incorporation of the thymidine analogue BrdU (bromodeoxyuridine) into newly synthesized DNA and its detection either using fluorescently labelled antibodies against BrdU, or quenching fluorescence of Hoechst-DNA complex. In plants, this approach has been utilized by several groups (Glab et al. 1994; Lucretti et al. 1999; Sgorbati et al. 1991; Yanpaisan et al. 1998) but was not adopted widely, probably due to complicated sample preparation. For a long time, the BrdU-based assay was considered gold standard in flow cytometric cell cycle studies in humans and animals, but recently a new assay (Buck et al. 2008) based on incorporation of the thymidine analogue EdU (5-ethynyl-2′-deoxyuridine) and its detection using click chemistry started to replace it (Diermeier-Daucher et al. 2009; Sun et al. 2012). Kotogány et al. (2010) were the first to verify the feasibility of this new method in plants. They compared the EdU-based assay with traditional BrdU-based assay using cultured cells and root meristems of several plants and concluded that the EdU assay was superior, considerably faster, simpler, did not require digestion of the cell wall nor denaturation of the DNA and can be used in different plant systems.

6.5 Gene Expression

A further attractive but still underutilized application of flow cytometry is the analysis of tissue-specific transcription. Flow-sorted nuclei can be used as a source of transcripts for gene expression analysis in particular tissue types. This approach, developed by Macas et al. (1998), assumes that the transcript levels within the karyoplasm reflect the state of gene expression more accurately than total RNA samples whose majority is comprised of cytoplasmic RNA. For this purpose, intact nuclei are released by a quick tissue homogenization on ice (Galbraith et al. 1983) which preserves transcriptional state in the moment of homogenization. The method utilizes transgenic plants expressing fusion protein, comprising GFP marker and nuclear localization sequence under control of tissue-specific promoter (Zhang et al. 2005). Tissue-specific nuclei can then be sorted based on the GFP signal (Zhang et al. 2008), and RNA is extracted followed by gene expression analysis on microarrays or by next-generation sequencing (Zhang et al. 2008; Macas et al. 1998). This approach overcomes drawbacks of whole-cell sorting, namely, the risk of changing gene expression during protoplast isolation. Flow cytometry was also used to study transcriptional activity in nuclei with different levels of endoreduplication and showed positive correlation of increased transcription and endoreduplication level (Bourdon et al. 2012). This work demonstrated for the first time in plants that endoreduplication correlates with elevated transcriptional activity.

7 Examples for the Application of DNA Flow Cytometry in Plant Research

Information gained on plant nuclei using flow cytometry has been exploited in several fields of plant biology and applied research and industry, including plant taxonomy (e.g. Suda et al. 2007a, Jersáková et al. 2013), ecology (Leitch and Leitch 2012; Herben et al. 2011), evolutionary and population biology (reviewed by Kron et al. 2007), breeding (reviewed by Ochatt 2008; Ochatt et al. 2011) and cell biology (e.g. Binarová et al. 1993, 2000, Petrovská et al. 2012).

7.1 Taxonomy

Until recently, taxonomy has relied on classical morphological approaches. This has changed and molecular methods are being increasingly used. Flow cytometry complements these molecular methods as an invaluable tool and has been contributing significantly to taxonomic research. Genome size and ploidy estimations are important in this regard. Although ploidy studies can be conducted based on

karyological methods, the main advantage of FCM is a possibility to analyse many individuals in a short time and using almost any tissue as sample.

Flow cytometry is ideal to study variation in ploidy levels within the same species (intraspecific cytotype diversity) as it facilities large-scale sampling (cf. Suda et al. 2007b). Ploidy variation can be assessed among the analysed populations, including the occurrence of aneuploidy, and absolute DNA contents can be estimated. For example, analysis of 59 populations of *Cardamine* species in Japan and South Korea revealed large cytotype diversity (Marhold et al. 2010). The patterns of cytotype distribution differed between species, and while some were strictly uniform, containing only one cytotype, multiple cytotypes were present in populations of other species. Thus, *Cardamine yezoensis* was found to comprise six cytotypes ranging from common hexaploids to rare dodecaploids, while *C. amaraeiformis* comprised only one tetraploid cytotype. Similar studies were carried out in *Allium oleraceum* (Šafářová et al. 2011), *Centaurea phrygia* (Koutecký et al. 2012a) and *Odontites vernus* (Koutecký et al. 2012b) where spatial distribution of different cytotypes was observed.

Apart from studying differences in genome copy number, flow cytometry has been used in taxonomic and ecological research on homoploid plants (reviewed in Loureiro et al. 2010). Differences in genome sizes obtained by flow cytometry were useful for distinguishing between subspecies with the same ploidy level, for example, in genus *Festuca* (Loureiro et al. 2007b), *Equisetum* (Obermayer et al. 2002) and *Taraxacum* (Záveský et al. 2005). Flow cytometry was also successful in detection of homoploid hybrids as documented for the genera *Amaranthus* (Jeschke et al. 2003) and *Hieracium* (Morgan-Richards et al. 2004).

7.2 Evolution of Plant Genomes

Evolution of genome size in plants is dynamic and has been accompanied by increases and decreases within lineages (e.g. Leitch et al. 2005). As a consequence, there is at least 2,400-fold variation in genome size among angiosperms. Flow cytometry has become an important tool to study this variation and the underlying mechanisms. According to Bennett and Leitch (2005), the most important components needed to understand the evolution of plant genome sizes include distributional patterns of variation within and among taxa, historical trends that generated current patterns, mechanisms of genome size changes and phenotypic consequences influencing both taxonomic and geographical distribution for the variation of genome size. While the amplification of transposable elements and polyploidization events are the main forces behind increasing genome size, mechanisms of genome reduction are still poorly understood. Data on genome size in different plant groups together with improved phylogenetic knowledge allow us to understand genome size diversity in a phylogenetic context (Leitch et al. 1998). The first group of plants, where genome size values were superimposed onto phylogenetic trees, were the angiosperms (Leitch et al. 1998; Soltis et al. 2003). This allowed to

reconstruct ancestral genome size of angiosperms, which was assumed to be small (1C $\leq$ 1.4 pg). Similar studies followed the suite taking advantage of the ever-increasing data on genome sizes, including other groups of land plants (Leitch et al. 2005; Leitch and Leitch 2013).

7.3 Ecology and Plant Population Biology

Genome size is known to correlate with cell size and length of the cell cycle which in turn affect other physiological and phenotypic traits of plants (Knight and Beaulieu 2008). Therefore, ecological preferences of a species might be predicted based on their genome size. Leitch and Bennett (2007) and more recently Greilhuber and Leitch (2013) discuss interesting implications of genome size on phenotypic traits of plants. For example, Herben et al. (2012) examined variation in genome sizes, several plant traits and regional abundance in more than 400 herbaceous species of the Central European flora and found a weak but significant correlation between genome size and the characteristics of regional abundance for a given species. Similarly, Beaulieu et al. (2007) used data from more than 1,000 species and confirmed correlations between genome size and seed mass. After a comparison of genome sizes of weed and non-weed species, Bennett et al. (1998) concluded that weed species are characterized by small genome sizes. Suda et al. (2010) used genome size as a reliable marker of invasiveness in knotweed species and found that naturalized and invasive plants harbour also significantly smaller genomes than their non-invading relatives. Several studies tried to correlate genome size with climate (temperature, precipitation and length of growing season) connected to altitude and latitude, but the results were not always consistent (Knight et al. 2005). Temsch et al. (2010) and Vidic et al. (2009) studied relationships between heavy-metal soil pollution and genome size of surviving plant species. Both studies concluded that all surviving plants had on average smaller genome sizes.

7.4 Plant Sex and Reproduction

DNA flow cytometry has been used to determine the gender of plants already in early stages of growth and to characterize the mode of reproduction in plants, both of which can influence ecological behaviour. Flow cytometry can be used to distinguish the two genders of dioecious plant species when gender is determined by heteromorphic sex chromosomes. Using high-resolution flow cytometry, Doležel and Göhde (1995) analysed nuclear DNA contents for the two dioecious species *Melandrium album* and *Melandrium rubrum* and were able to discriminate between male and female nuclei based on small differences in DNA amounts. It is possible to reveal this difference very early in the development, and one can even

tell whether the pollen nuclei are male- or female-determining (Stehlik et al. 2007). FCM can thus serve in plant sex determination as an alternative to classical approaches – sex-specific molecular markers or cytology (Stehlik and Barrett 2005).

The use of flow cytometry for testing the mode of reproduction was termed flow cytometric seed screen (FCSS) and to a large extent substituted formerly used approaches (Matzk et al. 2000). Sporophytic or gametophytic mutants in sexual species can be evaluated by flow cytometry to distinguish between purely sexual, obligatory apomictic and facultative apomictic species. Each pathway of seed formation is characterized by different combination of DNA amounts in embryo and endosperm (Matzk et al. 2000), which can be estimated using FCM. Sexually raised seeds show ratios of embryo to endosperm DNA of 2C:3C, while apomictic seeds show 2C:4C. In case of facultative apomixis, more complicated histograms comprising several peaks representing different C-values are observed. Diploid nuclei isolated from leaf tissue are usually used as a standard for these measurements. The method remains popular and has been improved by mathematically estimating the male and female genomic contributions to the embryo and endosperm independent of the mode of gametophyte formation and ploidy of parental plants (Dobeš et al. 2013).

7.5 Cell Biology

One of the popular applications of DNA flow cytometry in cell biology is the analysis of cell cycle activity and dynamics in heterogeneous plant cell suspensions (Yanpaisan et al. 1998; Lee et al. 2004) and the cellular response to different stresses or cell cycle inhibitors. Glab et al. (1994) analysed effects of olomoucine, an inhibitor of cdc2/cdk2 kinases activity, on plant cells and found that this compound blocked cell cycle transitions at G_1 to S and G_2 to M phases. Other studies investigated the effects of oxidative stress and drought on plants and used flow cytometry to determine the cell cycle phases which were the most affected (Reichheld et al. 1999; Bagniewska-Zadworna 2008). Flow cytometric analysis was also the critical tool to verify the function of plant Aurora kinases that participate in the switch from meristematic cell proliferation to differentiation and endoreduplication (Petrovská et al. 2012). Similarly, the function of nitrilase 1 plant homologues in the regulation of cell cycle exit was studied using flow cytometry (Doskočilová et al. 2013). Combination of flow cytometry with immunofluorescence helped to understand the role of several important cytoskeletal proteins in cell division (Pawlowski et al. 2004; Binarová et al. 1993, 2000; Fujikura et al. 1999).

Also the analysis of cell and tissue differentiation in connection with endopolyploidy was advanced by flow cytometry. In order to get insight into the incidence of endopolyploid cells, Barrow and Meister (2003) probed for differences in ploidy status at organ and tissue levels. Zhang et al. (2005) used *Arabidopsis thaliana* roots as a model and confirmed, using simultaneous analysis of cell-specific transgenic

reporter lines and DNA content measurements, that different cell types differ in nuclear ploidy. The biological significance of endopolyploidy is not clear, although it has been linked to cell expansion (Cookson et al. 2006, Jovtchev et al. 2006, Gendreau et al. 1998, Melaragno et al. 1993), metabolic activity (Larkins et al. 2001, Vilhar et al. 2002), fruit size (Sugimoto-Shirasu et al. 2003) and the response abiotic stress (Ceccarelli et al. 2006). Cheniclet et al. (2005) studied variability of tomato fruit size and demonstrated a correlation of increased ploidy levels with the size of pericarp cells. Subsequently, Bourdon et al. (2011) sorted nuclei from tomato pericarp tissues and used them as template for BAC-FISH. Using this method, they were able to establish a ploidy map of the tomato fruit pericarp in intact tissues. The extent of endoreduplication was also used by Rewers and Śliwinska (2012) as a marker for seed developmental stages in five species of Fabaceae. Bainard et al. (2011b) analysed 37 species from 16 angiosperms families and found that mycorrhizal symbiosis positively stimulates endoreduplication.

7.6 Plant Breeding

FCM has been used extensively in plant breeding. According to Eeckhaut et al. (2005), the most important applications of flow cytometry in this field are the identification and characterization of parent plants suitable for breeding programmes, screening of the offspring and determination of ploidy levels after haploidization and/or polyploidization. To characterize of parent plants, there are many germplasm collections from which suitable candidates for breeding can be chosen. Screening for ploidy levels and genome sizes of different accessions using flow cytometry is a useful tool, and this procedure has been used for many species, e.g. bananas (Nsabimana and van Staden 2006; Pilay et al. 2006; Doleželová et al. 2005), agave (Palomino et al. 2003), grass pea (Ochatt et al. 2001) and water yam (Egesi et al. 2002). In some breeding programmes it might be of agronomic importance to screen for genotypes of specific gender (e.g. in asparagus; Ozaki et al. 1998). In plants where gender is determined by heteromorphic sex chromosomes differing in either size or base pair content, flow cytometry can discriminate between male and female plants based on their slightly different DNA content or AT/GC ratios, as successfully shown for White campion (Doležel and Göhde 1995) and date palm (Siljak-Yakovlev et al. 1996).

DNA flow cytometry has also been useful to characterize the progeny from interspecific hybridizations. The method facilitated identification of interspecific hybrids based on intermediate DNA contents in onion (Keller et al. 1996) and coffee (Barre et al. 1998). Triploidy is often connected with low fertility and seedlessness, which is a much sought character in a number of crops. In hops, seedless triploids are sought after for their better brewing qualities (Beatson et al. 2003). In citruses, triploid plants are produced after diploid x tetraploid hybridization (Aleza et al. 2012). In these cases, flow cytometry is an ideal tool to verify ploidy as it can quickly and efficiently screen large numbers of plants.

Doubled haploid (dihaploid) plants are useful in breeding to shorten breeding cycles. As a common practice dihaploids are generated by polyploidizing haploids produced in vitro from immature anthers. However, plants may regenerate not only from immature pollen but also from diploid somatic anther tissues. Flow cytometry is an effective tool to select haploid regenerants, as shown, for example, in chickpea (Grewal et al. 2009). Alternatively microspores can be cultured in vitro to produce haploids, and also in this case flow cytometry is invaluable to confirm the haploid status of regenerated plants (Weber et al. 2005). Some breeding programmes require development of autopolyploids, which are bigger and more robust compared to their diploid parents and have bigger flowers and fruits. Autopolyploids are also needed to produce triploids after crossing with diploids. Polyploidization is typically induced by exposing multicellular explants in vitro to mitotic spindle poisons. As some cells in the explant are not cycling and hence are not polyploidized, regenerated plants may be diploid, mixoploid (chimaeric) or tetraploid. Identification of solid tetraploids during early stages is best achieved using DNA flow cytometry (Awoleye et al. 1994, van Duren et al. 1996). The avoidance of chimaeras may be a problem, and Roux et al. (2001) used flow cytometry to follow dissociation of chimaeras after repeated cycles of in vitro propagation.

Rapid propagation of new and superior genotypes is often needed and micropropagation in vitro is often the method of choice (see also the chapter by Opatrný, this volume). Micropropagation has been applied without proper genetic verification in many commercial operations to supply planting material to producers and farmers. However, depending on the source of the material (e.g. shoot tip, embryo culture, callus culture), the culture may be prone to genetic instability, also called somaclonal variation (Neelakandan and Wang 2012), leading to high variability in culture-derived individuals. Deviations among the regenerants and the frequency of these aberrations have been assessed by various methods, and DNA flow cytometry enabled identification of plants differing in ploidy (e.g. Zhao et al. 2012). For example, genome stability of six medicinal plants that were propagated in vitro for a year was assessed by Sliwinska and Thiem (2007), and true-to-typeness of zygotic embryos, somatic embryos and somatic embryogenesis-derived plantlets of *Pinus* was verified to exclude major changes in ploidy level in this economically significant species (Marum et al. 2009). Screening for culture-induced variation is of enormous significance in forest trees and woody plants, as they have long life cycles (Rani and Raina 2000; Loureiro et al. 2005). When hundreds of accessions have to be screened, the speed of FCM analysis becomes the major advantage. It took some time to screen ploidy in the world's largest banana germplasm collection located in Leuven (Belgium), comprising more than a thousand accessions maintained in vitro under slow growth conditions (Doleželová et al. 2005), but no other method could achieve this goal so efficiently.

Acknowledgements We thank our colleagues Beáta Petrovská and Jana Čížková for sharing unpublished results and useful suggestions. This work was supported by the Czech Science Foundation (awards P501/10/1740, P501/10/1778 and P501/12/G090) and by the Ministry of

Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01.

References

Aleza P, Juárez J, Hernández M, Ollitrault P, Navarro L (2012) Implementation of extensive citrus triploid breeding programs based on 4x × 2x sexual hybridisations. Tree Genet Genome 8:1293–1306

Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. Plant Mol Biol Rep 9:208–218

Awoleye F, van Duren M, Doležel J, Novák FJ (1994) Nuclear DNA content and in vitro induced somatic polyploidization cassava (*Manihot esculenta* Crantz) breeding. Euphytica 76:195–202

Bagniewska-Zadworna A (2008) The root microtubule cytoskeleton and cell cycle analysis through desiccation of *Brassica napus* seedlings. Protoplasma 233:177–185

Bainard JD, Newmaster SG (2010) Endopolyploidy in bryophytes: widespread in mosses and absent in liverworts. J Bot 2010, Article ID 316356

Bainard JD, Fazekas AJ, Newmaster SG (2010) Methodology significantly affects genome size estimates: quantitative evidence using bryophytes. Cytometry 77A:725–732

Bainard JD, Husband BC, Aldwin SJ, Fazekas AJ, Gregory TR, Newmaster SG, Kron P (2011a) The effects of rapid desiccation on estimates of plant genome size. Chromosome Res 19:825–842

Bainard LD, Bainard JD, Newmaster SG, Klironomos JN (2011b) Mycorrhizal symbiosis stimulates endoreduplication in angiosperms. Plant Cell Environ 34:1577–1585

Baranyi M, Greilhuber J, Swiecicki WK (1996) Genome size in wild *Pisum* species. Theor Appl Genet 93:717–721

Barow M (2006) Endopolyploidy in seed plants. Bioassays 28:271–281

Barow M, Jovtchev G (2007) Endopolyploidy in plants and its analysis by flow cytometry. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 349–372

Barow M, Meister A (2002) Lack of correlation between AT frequency and genome size in higher plants and the effect of nonrandomness of base sequences on dye binding. Cytometry 47A:1–7

Barow M, Meister A (2003) Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. Plant Cell Environ 26:571–584

Barre P, Layssac M, D'Hont A, Louarn J, Charrier A, Hamon S, Noirot M (1998) Relationship between parental chromosomic contribution and nuclear DNA content in the coffee interspecific hybrid *C. pseudozanguebariae* × *C. liberica* var 'dewevrei'. Theor Appl Genet 96:301–305

Bashir A, Auger JA, Rayburn AL (1993) Flow cytometric DNA analysis of wheat-rye addition lines. Cytometry 14A:843–847

Basiji DA, Ortyn WE, Liang L, Venkatachalam V, Morrissey P (2007) Cellular image analysis and imaging by flow cytometry. Clin Lab Med 27:653–670

Beatson RA, Ferguson AR, Weir IE, Graham LT, Ansell KA, Ding H (2003) Flow cytometric identification of sexually derived polyploids in hop (*Humulus lupulus* L.) and their use in hop breeding. Euphytica 134:189–194

Beaulieu JM, Moles AT, Leitch IJ, Bennett MD, Dickie JB, Knight CA (2007) Correlated evolution of genome size and seed mass. New Phytol 173:422–437

Becker RL, Mikel UV (1990) Interrelation of formalin fixation, chromatin compactness and DNA values as measured by flow cytometry and image cytometry. Anal Quant Cytol 12:333–341

Bennet MD, Leitch IJ (2011) Nuclear DNA amounts in angiosperms: targets, trends and tomorrow. Ann Bot 107:467–590

Bennett MD, Leitch IJ (1995) Nuclear DNA amounts in angiosperms. Ann Bot 76:113–176

Bennett MD, Leitch IJ (1997) Nuclear DNA amounts in angiosperms – 583 new estimates. Ann Bot 80:169–196

Bennett MD, Leitch IJ (2005) Nuclear DNA amounts in angiosperms – progress, problems and prospects. Ann Bot 95:45–90

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Philos Trans R Soc Lond Ser B Biol Sci 274:227–274

Bennett MD, Smith JB, Smith RIL (1982) DNA amounts of angiosperms from the Antarctic and South Georgia. Environ Exp Bot 22:307–318

Bennett MD, Leitch IJ, Hanson L (1998) DNA amounts in two samples of angiosperm weeds. Ann Bot 82:121–134

Bennett MD, Bhandol P, Leitch IJ (2000) Nuclear DNA amounts in angiosperms and their modern uses – 807 new estimates. Ann Bot 86:859–909

Bennett MD, Leitch IJ, Price HJ, Johnston JP (2003) Comparisons with *Caenorhabditis* (100 Mb) and Drosophila (175 Mb) using flow cytometry show genome size in Arabidopsis to be 157 Mb and thus 25 % larger than the Arabidopsis genome initiative estimate of 125 Mb. Ann Bot 91:547–557

Bergounioux C, Perennes C, Brown SC, Gadal P (1988) Nuclear RNA quantification in protoplast cell cycle phases. Cytometry 9:84–87

Binarová P, Doležel J (1993) Effect of anti-microtubular drug amiprophos-methyl on somatic embryogenesis and DNA ploidy levels in alfalfa and carrot cell suspension cultures. Biol Plant 35:329–339

Binarová P, Číhalíková J, Doležel J (1993) Localization of MPM–2 recognized phosphoproteins and tubulin during cell cycle progression in synchronized *Vicia faba* root meristem cells. Cell Biol Int 17:847–856

Binarová P, Doležel J, Dráber P, Heberle-Bors E, Strnad M, Bogre L (1998a) Treatment of *Vicia faba* root tip cells with specific inhibitors to cyclin–dependent kinases leads to abnormal spindle formation. Plant J 16:697–707

Binarová P, Hause B, Doležel J, Dráber P (1998b) Association of γ–tubulin with kinetochore/centromeric region of plant chromosomes. Plant J 14:751–757

Binarová P, Cenklová V, Hause B, Kubátová E, Lysák M, Doležel J, Bögre L, Dráber P (2000) Nuclear γ-tubulin during acentriolar plant mitosis. Plant Cell 12:433–442

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN (2003) A gene expression map of the *Arabidopsis* root. Science 302:1956–1960

Boivin A, Vendrely R, Vendrely C (1948) L'acide désoxyribonucléique du noyau cellulaire dépositaire des caractéres héréditaires; arguments d'ordre analytique. C R Acad Sci 226:1061–1063

Bourdon M, Coriton O, Pirrello J, Cheniclet C, Brown SC, Poujol C, Chevalier C, Renaudin JP, Frangne N (2011) In planta quantification of endoreduplication using fluorescent in situ hybridization (FISH). Plant J 66:1089–1099

Bourdon M, Pirrello J, Cheniclet C, Coriton O, Bourge M, Brown S, Moïse A, Peypelut M, Rouyère V, Renaudin JP, Chevalier C, Frangne N (2012) Evidence for karyoplasmic homeostasis during endoreduplication and a ploidy-dependent increase in gene transcription during tomato fruit growth. Development 139:3817–3826

Buck SB, Bradford J, Gee KR, Agnew BJ, Clarke ST, Salic A (2008) Detection of S–phase cell cycle progression using 5–ethynyl–2′–deoxyuridine incorporation with click chemistry, an alternative to using 5–bromo–2′–deoxyuridine antibodies. Biotechniques 44:927–929

Cardoso DC, Carvalho CR, Cristiano MP, Soares FAF, Tavares MG (2012) Estimation of nuclear genome size of the genus *Mycetophylax emery*, 1913: evidence of no whole–genome duplication in Neoattini. C R Biol 335:619–624

Ceccarelli M, Sanantonio E, Marmottini F, Amzallag GN, Cionini PG (2006) Chromosome endoreduplication as a factor of salt adaptation in *Sorghum bicolour*. Protoplasma 227:113–118

Cerbah M, Mortreau E, Brown S, Siljak-Yakovlev S, Bertrand H, Lambert C (2001) Genome size variation and species relationships in the genus *Hydrangea*. Theor Appl Genet 103:45–51

Chattopadhyay PK, Hogerkorp CM, Roederer M (2008) A chromatic explosion: the development and future of multiparameter flow cytometry. Immunology 125:441–449

Cheniclet C, Rong WY, Causse M, Frangne N, Bolling L, Carde J-P, Renaudin J-P (2005) Cell expansion and endoreduplication show a large genetic variability in pericarp and contribute strongly to tomato fruit growth. Plant Physiol 139:1984–1994

Chevalier C, Nafati M, Mathieu-Rivet E, Bourdon M, Frangne N, Cheniclet C, Renaudin JP, Gévaudant F, Hernould M (2011) Elucidating the functional role of endoreduplication in tomato fruit development. Ann Bot 107:1159–1169

Chiatante D, Brusa P, Levi M, Sgorbati S, Sparvoli E (1990) A simple protocol to purify fresh nuclei from milligram amounts of meristematic pea root tissue for biochemical and flow cytometry applications. Physiol Plant 78:501–506

Cires E, Cuesta C, Peredo EL, Revilla MA, Prieto JAF (2009) Genome size variation and morphological differentiation within *Ranunculus parnassifolius* group (Ranunculaceae) from calcareous screes in the Northwest of Spain. Plant Syst Evol 281:193–208

Clarindo WR, Carvalho CR (2011) Flow cytometric analysis using SYBR Green I for genome size estimation in coffee. Acta Histochem 113:221–225

Cookson SJ, Radziejwoski A, Granier C (2006) Cell and leaf size plasticity in *Arabidopsis*: what is the role of endoreduplication? Plant Cell Environ 29:1273–1283

Cousin A, Heel K, Cowling WA, Nelson MN (2009) An efficient high-throughput flow cytometric method for estimating DNA ploidy level in plants. Cytometry 75A:1015–1019

D'Hondt L, Höfte M, Van Bockstaele E, Leus L (2011) Applications of flow cytometry in plant pathology for genome size determination, detection and physiological status. Mol Plant Pathol 12:815–828

Dart S, Kron P, Mable BK (2004) Characterizing polyploidy in *Arabidopsis lyrata* using chromosome counts and flow cytometry. Can J Bot 82:185–197

Darzynkiewicz Z (1994) Simultaneous analysis of cellular RNA and DNA content. Methods Cell Biol 41:401–420

de Laat AMM, Göhde W, Vogelzakg MJDC (1987) Determination of ploidy of single plants and plant populations by flow cytometry. Plant Breed 99:303–307

Diermeier-Daucher S, Clarke ST, Hill D, Vollmann-Zweren A, Bradford JA, Brockhoff G (2009) Cell type specific applicability of 5–ethynyl–2′–deoxyuridine (EdU) for dynamic proliferation assessment in flow cytometry. Cytometry 75A:536–546

Dobeš C, Luckl A, Hulber K, Paule J (2013) Prospects and limits of the flow cytometric seed screen – insights from *Potentilla* sensu lato (Potentilleae, Rosaceae). New Phytol 198:605–616

Doležel J (1991) Flow cytometric analysis of nuclear DNA content in higher plants. Phytochem Anal 2:143–154

Doležel J (1997) Application of flow cytometry for the study of plant genomes. J Appl Genet 38:285–302

Doležel J, Bartoš J (2005) Plant DNA flow cytometry and estimation of nuclear genome size. Ann Bot 95:99–110

Doležel J, Göhde W (1995) Sex determination in dioecious plants *Melandrium album and M. rubrum* using high–resolution flow cytometry. Cytometry 19:103–106

Doležel J, Greilhuber J (2010) Nuclear genome size: are we getting closer? Cytometry 77A:635–642

Doležel J, Binarová P, Lucretti S (1989) Analysis of nuclear DNA content in plant cells by flow cytometry. Biol Plant 31:113–120

Doležel J, Sgorbati S, Lucretti S (1992a) Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. Physiol Plantarum 85:625–631

Doležel J, Číhalíková J, Lucretti S (1992b) A high–yield procedure for isolation of metaphase chromosomes from root tips of *Vicia faba* L. Planta 188:93–98

Doležel J, Greilhuber J, Lucretti S, Meister A, Lysák MA, Nardi L, Obermayer R (1998) Plant genome size estimation by flow cytometry: inter–laboratory comparison. Ann Bot 82:17–26
Doležel J, Bartoš J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. Cytometry 51A:127–128
Doležel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc 2:2233–2244
Doležel J, Vrána J, Šafář J, Bartoš J, Kubaláková M, Šimková H (2012) Chromosomes in the flow to simplify genome analysis. Funct Integr Genomics 12:397–416
Doleželová M, Doležel J, Van den Houwe I, Roux N, Swennen R (2005) Focus on the *Musa* collection: ploidy levels revealed. InfoMusa 14:34–36
Doskočilová A, Kohoutová L, Volc J, Kourová H, Benada O, Chumová J, Plíhal O, Petrovská B, Halada P, Bögre L, Binarová P (2013) Nitrilase1 regulates the exit from proliferation, genome stability and plant development. New Phytol 198:685–698
Dubelaar GBJ, Gerritzen PL, Beeker AER, Jonker RR, Tangen K (1999) Design and first results of CytoBuoy: a wireless flow cytometer for in situ analysis of marine and fresh waters. Cytometry 37:247–254
Dubelaar GBJ, Casotti R, Tarran GA, Biegala IC (2007) Phytoplankton and their analysis by flow cytometry. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 287–322
Eeckhaut T, Leus L, van Huylenbroeck J (2005) Exploitation of flow cytometry for plant breeding. Acta Phys Plant 27:743–750
Egesi CN, Pillay M, Asiedu R, Egunjobi JK (2002) Ploidy analysis in water yam, *Dioscorea alata* L., germplasm. Euphytica 128:225–230
Favoreto FC, Carvalho CR, Lima ABP, Ferreira A, Clarindo WR (2012) Genome size and base composition of Bromeliaceae species assessed by flow cytometry. Plant Syst Evol 298:1185–1193
Fox MH, Galbraith DW (1990) The application of flow cytometry and sorting to higher plant systems. In: Melamed MR, Lindmo T, Mendelsohn ML (eds) Flow cytometry and sorting, 2nd edn. Wiley-Liss, New York, pp 633–650
Fujikura Y, Doležel J, Číhalíková J, Bögre L, Heberle-Bors E, Hirt H, Binarová P (1999) Vicia faba germination: synchronized cell growth and localization of nucleolin and α–tubulin. Seed Sci Res 9:297–304
Galbraith DW (1990) Flow cytometric analysis of plant genomes. Methods Cell Biol 33:549–562
Galbraith DW (1994) Flow cytometry and sorting of plant protoplasts and cells. Methods Cell Biol 42B:539–561
Galbraith DW (2007) Protoplast analysis using flow cytometry. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 231–250
Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 220:1049–1051
Gaurav V, Kolewe ME, Roberts SC (2010) Flow cytometric methods to investigate culture heterogeneities for plant metabolic engineering. In: Germano A (ed) Plant secondary metabolism engineering. Methods in molecular biology. Humana Press, New York, pp 243–262, Springer protocols
Gendreau E, Hofte H, Grandjean O, Brown S, Traas J (1998) Phytochrome controls the number of endoreduplication cycles in the *Arabidopsis thaliana* hypocotyls. Plant J 13:221–230
Giglioli-Guivarc'h N, Pierre JN, Vidal J, Brown S (1996) Flow cytometric analysis of cytosolic pH of mesophyll cell protoplasts from the crabgrass *Digitaria sanguinalis*. Cytometry 23:241–249
Glab N, Labidi B, Qin LX, Trehin C, Bergounioux C, Meijer L (1994) Olomoucine, an inhibitor of the cdc2/cdk2 kinases activity, blocks plant cells at the G1 to S and G2 to M cell cycle transitions. FEBS Lett 353:207–211

Godelle B, Cartier D, Marie D, Brown SC, Siljak-Yakovlev S (1993) Heterochromatin study demonstrating the non–linearity of fluorometry useful for calculating genomic base composition. Cytometry 14A:618–626

Greilhuber J (1986) Severely distorted Feulgen–DNA amounts in *Pinus* (Coniferophytina) after nonadditive fixations as a result of meristematic self–tanning with vacuole contents. Can J Genet Cytol 28:409–415

Greilhuber J (1988) Self–tanning: a new and important source of stoichiometric error in cytophotometric determination of nuclear DNA content in plants. Plant Syst Evol 158:87–96

Greilhuber J (2008) Cytochemistry and C-values: the less-well-known world of nuclear DNA amounts. Ann Bot 101:791–804

Greilhuber J, Doležel J (2009) 2C or not 2C: a closer look at cell nuclei and their DNA content. Chromosoma 118:391–400

Greilhuber J, Leitch IJ (2013) Genome size and the phenotype. In: Leitch IJ, Greilhuber J, Doležel J, Wendel JF (eds) Plant genome diversity. Vol. 2: Physical structure, behaviour and evolution of plant genomes. Springer, Wien, pp 323–344

Greilhuber J, Doležel J, Lysák MA, Bennett MD (2005) The origin, evolution and proposed stabilization of the terms 'genome size' and 'C–value' to describe nuclear DNA contents. Ann Bot 95:255–260

Greilhuber J, Bosch T, Muller K, Worberg A, Porembski S, Barthlott W (2006) Smallest angiosperm genomes found in lentibulariaceae, with chromosomes of bacterial size. Plant Biol 8:770–777

Greilhuber J, Temsch EM, Loureiro JCM (2007) Nuclear DNA content measurement. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 423–438

Grewal RK, Lulsdorf M, Croser J, Ochatt S, Vandenberg A, Warkentin TD (2009) Doubled-haploid production in chickpea (*Cicer arietinum* L.): role of stress treatments. Plant Cell Rep 28:1289–1299

Halverson K, Heard SB, Nason JD, Stireman JO (2008) Differential attack on diploid, tetraploid, and hexaploid *Solidago altissima* L. by five insect gallmakers. Oecologia 154:755–761

Hardie DC, Gegory TR, Hebert PDN (2002) From pixels to picograms – A beginners' guide to genome quantification by Feulgen image analysis densitometry. J Histochem Cytochem 50:735–749

Harmon AF, Zarlenga DS, Hildreth MB (2006) Improved methods for isolating DNA from *Ostertagia ostertagi* eggs in cattle feces. Vet Parasitol 135:297–302

Haymes KM, Ibrahim IA, Mischke S, Scott DL, Saunders JA (2004) Rapid isolation of DNA from chocolate and date palm tree crops. J Agric Food Chem 52:5456–5462

Heller FO (1973) DNS–Bestimmung an Keimwurzeln von *Vicia faba* L. mit Hilfe der Impulscytophotometrie. Bericht der Deutschen Botanischen Gesellschaft 86:437–441

Herben T, Suda J, Klimešová J, Mihulka S, Říha P, Šímová I (2012) Ecological effects of cell–level processes: genome size, functional traits and regional abundance of herbaceous plant species. Ann Bot 110:1357–1367

Hiddemann W, Schumann J, Andreef M, Barlogie B, Herman CJ, Leif RC, Mayall BH, Murphy RF, Sandberg AA (1984) Convention on nomenclature for DNA cytometry. Cancer Genet Cytogenet 13:181–183

Hopping ME (1993) Preparation and preservation of nuclei from plant–tissues for quantitative DNA analysis by flow cytometry. N Z J Bot 31:391–401

Jarret RL, Oziasakins P, Phatak S, Nadimpalli R, Duncan R, Hiliard S (1995) DNA contents in *Paspalum* spp. determined by flow cytometry. Genet Resour Crop Evol 42:237–242

Jedrzejczyk I, Sliwinska E (2010) Leaves and seeds as materials for flow cytometric estimation of the genome size of 11 Rosaceae woody species containing DNA-staining inhibitors. J Bot 2010, Article ID 930895

Jersáková J, Trávníček P, Kubátová B, Krejčíková J, Urfus T, Liu ZJ, Lamb A, Ponert J, Schulte K, Čurn V, Vrána J, Leitch IJ, Suda J (2013) Genome size variation in Orchidaceae subfamily Apostasioideae: filling the phylogenetic gap. Bot J Linn Soc 172:95–105

Jeschke MR, Tranel PJ, Rayburn AL (2003) DNA content analysis of smooth pigweed (*Amaranthus hybridus*) and tall waterhemp (*A–tuberculatus*): implications for hybrid detection. Weed Sci 51:1–3

Johnston JS, Bennett MD, Rayburn AL, Galbraith DW, Price HJ (1999) Reference standards for determination of DNA content of plant nuclei. Am J Bot 86:609–613

Jovtchev G, Schubert V, Meister A, Schubert I (2006) Nuclear DNA content and nuclear and cell volume are positively correlated in angiosperms. Cytogenet Genome Res 114:77–82

Kausch AP, Bruce BD (1994) Isolation and immobilization of various plastid subtypes by magnetic immunoabsorption. Plant J 6:767–779

Keller ERJ, Schubert I, Fuchs J, Meister A (1996) Interspecific crosses of onion with distant *Allium* species and characterization of the presumed hybrids by means of flow cytometry, karyotype analysis and genomic in situ hybridization. Theor Appl Genet 92:417–424

Knight CA, Beaulieu JM (2008) Genome size scaling through phenotype space. Ann Bot 101:759–766

Knight CA, Molinari NA, Petrov DA (2005) The large genome constraint hypothesis: evolution, ecology and phenotype. Ann Bot 95:177–190

Kolář F, Lučanová M, Těšitel J, Loureiro J, Suda J (2012) Glycerol–treated nuclear suspensions – an efficient preservation method for flow cytometric analysis of plant samples. Chromosome Res 20:303–315

Kotogány E, Dudits D, Horváth GV, Ayaydin F (2010) A rapid and robust assay for detection of S-phase cell cycle progression in plant cells and tissues by using ethynyl deoxyuridine. Plant Methods 6:5–19

Koutecký P, Štěpánek J, Baďurová T (2012a) Differentiation between diploid and tetraploid *Centaurea phrygia*: mating barriers, morphology and geographic distribution. Preslia 84:1–32

Koutecký P, Tuleu G, Baďurová T, Kosnar J, Stech M, Tesitel J (2012b) Distribution of cytotypes and seasonal variation in the *Odontites vernus* group in central Europe. Preslia 84:887–904

Kron P, Husband BC (2012) Using flow cytometry to estimate pollen DNA content: improved methodology and applications. Ann Bot 110:1067–1078

Kron P, Suda J, Husband BC (2007) Applications of flow cytometry to evolutionary and population biology. Annu Rev Ecol Evol Syst 38:847–876

Larkins BA, Dilkes BP, Dante RA, Coelho CM, Woo Y, Liu Y (2001) Investigating the hows and whys of DNA endoreduplication. J Exp Bot 52:183–192

Lee TJ, Shultz RW, Hanley-Bowdoin L, Thompson WF (2004) Establishment of rapidly proliferating rice cell suspension culture and its characterization by fluorescence-activated cell sorting analysis. Plant Mol Biol Rep 22:259–267

Leitch IJ, Bennett MD (2007) Genome size and its uses: the impact of flow cytometry. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 153–176

Leitch AR, Leitch IJ (2012) Ecological and genetic factors linked to contrasting genome dynamics in seed plants. New Phytol 194:629–646

Leitch AR, Leitch IJ (2013) Genome size diversity and evolution in land plants. In: Leitch IJ, Greilhuber J, Doležel J, Wendel JF (eds) Plant genome diversity. Vol. 2: Physical structure, behaviour and evolution of plant genomes. Springer, Wien, pp 307–322

Leitch IJ, Chase MW, Bennett MD (1998) Phylogenetic analysis of DNA C–values provides evidence for a small ancestral genome size in flowering plants. Ann Bot 82:85–94

Leitch IJ, Soltis DE, Soltis PS, Bennett MD (2005) Evolution of DNA amounts across land plants (Embryophyta). Ann Bot 95:207–217

Liu JH, Dixelius C, Eriksson I, Glimelius K (1995) *Brassica napus* (+) *B. tournefortii*, a somatic hybrid containing traits of agronomic importance for rapeseed breeding. Plant Sci 109:75–86

Loureiro J, Pinto G, Lopes T, Doležel J, Santos C (2005) Assessment of ploidy stability of the somatic embryogenesis process in *Quercus suber* L. using flow cytometry. Planta 221:815–822

Loureiro J, Rodriguez E, Doležel J, Santos C (2006a) Comparison of four nuclear isolation buffers for plant DNA flow cytometry. Ann Bot 98:679–689

Loureiro J, Rodriguez E, Doležel J, Santos C (2006b) Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. Ann Bot 98:515–527

Loureiro J, Rodriguez E, Doležel J, Santos C (2007a) Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. Ann Bot 100:875–888

Loureiro J, Kopecký D, Castro S, Santos C, Silveira P (2007b) Flow cytometric and cytogenetic analyses of Iberian Peninsula *Festuca* spp. Plant Syst Evol 269:89–105

Loureiro J, Trávníček P, Rauchová J, Urfus T, Vít P, Štech M, Castro S, Suda J (2010) The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. Preslia 82:3–21

Lucretti S, Nardi L, Nisini PT, Moretti F, Gualberti G, Doležel J (1999) Bivariate flow cytometry DNA/BrdUrd analysis of plant cell cycle. Methods Cell Sci 21:155–166

Lysák MA, Doleželová M, Horry JP, Swennen R, Doležel J (1999) Flow cytometric analysis of nuclear DNA content in *Musa*. Theor Appl Genet 98:1344–1350

Macas J, Lambert GM, Doležel D, Galbraith DW (1998) Nuclear expressed sequence tag (NEST) analysis: a novel means to study transcription through amplification of nuclear RNA. Cytometry 33:460–468

Marhold K, Kudoh H, Pak JH, Watanabe K, Španiel S, Lihová J (2010) Cytotype diversity and genome size variation in eastern Asian polyploid *Cardamine* (Brassicaceae) species. Ann Bot 105:249–264

Marie D, Brown SC (1993) A cytometric exercise in plant DNA histograms, with 2C values for 70 species. Biol Cell 78:41–51

Marum L, Rocheta M, Maroco J, Oliveira MM, Muguel C (2009) Analysis of genetic stability at SSR loci during somatic embryogenesis in maritime pine (*Pinus pinaster*). Plant Cell Rep 28:673–682

Matzk F (2007) Reproduction mode screening. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 131–152

Matzk F, Meister A, Schubert I (2000) An efficient screen for reproductive pathways using mature seeds of monocots and dicots. Plant J 21:97–108

Meister A, Barow M (2007) DNA base composition of plant genomes. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells. analysis of genes, chromosomes, and genomes. Viley-VCH, Weinheim, pp 177–215

Melaragno JE, Mehrotra B, Coleman AW (1993) Relationship between endopolyploidy and cell size in epidermal tissue of *Arabidopsis*. Plant Cell 5:1661–1668

Mishra MK (1997) Stomatal characteristics at different ploidy levels in *Coffea* L. Ann Bot 80:689–692

Morgan-Richards M, Trewick SA, Chapman HM, Krahulcova A (2004) Interspecific hybridization among *Hieracium* species in New Zealand: evidence from flow cytometry. Heredity 93:34–42

Naill MC, Roberts SC (2005) Flow cytometric analysis of protein content in *Taxus* protoplasts and single cells as compared to aggregated suspension cultures. Plant Cell Rep 23:528–533

Neelakandan AK, Wang K (2012) Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications. Plant Cell Rep 31:597–620

Noirot M, Barre P, Louarn J, Duperray C, Hamon S (2000) Nucleus-cytosol interactions – A source of stoichiometric error in flow cytometric estimation of nuclear DNA content in plants. Ann Bot 86:309–316

Noirot M, Barre P, Louarn J, Duperray C, Hamon S (2002) Consequences of stoichiometric error on nuclear DNA content evaluation in *Coffea liberica* var. *dewevrei* using DAPI and propidium iodide. Ann Bot 89:385–389

Noirot M, Barre P, Duperray C, Louarn J, Hamon S (2003) Effects of caffeine and chlorogenic acid on propidium iodide accessibility to DNA: consequences on genome size evaluation in coffee tree. Ann Bot 92:259–264

Noirot M, Barre P, Duperray C, Hamon S, De Kochko A (2005) Investigation on the causes of stoichiometric error in genome size estimation using heat experiments: consequences on data interpretation. Ann Bot 95:111–118

Nsabimana A, van Staden J (2006) Ploidy investigation of bananas (*Musa* spp.) from the National Banana Germplasm Collection at Rubona–Rwanda by flow cytometry. S Afr J Bot 72:302–305

Obermayer R, Leitch IJ, Hanson L, Bennett MD (2002) Nuclear DNA C – values in 30 species double the familial representation in Pteridophytes. Ann Bot 90:209–217

Ochatt SJ (2008) Flow cytometry in plant breeding. Cytometry 73A:581–598

Ochatt SJ, Durieu P, Jacas L, Pontécaille C (2001) Protoplast, cell and tissue cultures for the biotechnological breeding of grass pea (*Lathyrus sativus* L.). Lathyrus Lathyrism Newsl 2:35–38

Ochatt SJ, Patat-Ochatt EM, Moessner A (2011) Ploidy level determination within the context of in vitro breeding. Plant Cell Tissue Organ Cult 104:329–341

Overton WR, McCoy JP (1994) Reversing the effect of formalin on the binding of propidium iodide to DNA. Cytometry 16:351–356

Ozaki Y, Kumiko N, Michikazu H, Kenji U, Hiroshi O (1998) Application of flow cytometry for rapid determination of ploidy levels in Asparagus (*Asparagus officinalis* L.). J Fac Agric Kyushu Univ 43:83–88

Palomino G, Doležel J, Mendez I, Rubluo A (2003) Nuclear genome size analysis of *Agave tequilana*. Caryologia 56:37–46

Pawlowski TA, Bergervoet JHW, Bino RJ, Groot SPC (2004) Cell cycle activity and β-Tubulin accumulation during dormancy breaking of *Acer platanoides* L. seeds. Biol Plant 48:211–218

Pellicer J, Fay MF, Leitch IJ (2010) The largest eukaryotic genome of them all? Bot J Linn Soc 164:10–15

Perfetto SP, Chattopadhyay PK, Roederer M (2004) Seventeen–colour flow cytometry: unravelling the immune system. Nat Rev Immunol 4:648–655

Petit PX (1992) Flow Cytometric Analysis of rhodamine 123 fluorescence during modulation of the membrane potential in plant mitochondria. Plant Physiol 98:279–286

Petit PX, Diolez P, Muller P, Brown SC (1986) Binding of concanavalin A to the outer membrane of potato tuber mitochondria detected by flow cytometry. Febs Lett 196:65–70

Petrovská B, Cenklová V, Pochylová Z, Kourová H, Doskočilová A, Plíhal O, Binarová L, Binarová P (2012) Plant Aurora kinases play a role in maintenance of primary meristems and control of endoreduplication. New Phytol 193:590–604

Pfosser M (1989) Improved method for critical comparison of cell cycle data of asynchronously dividing and synchronized cell cultures of *Nicotiana tabacum*. J Plant Physiol 134:741–745

Pfosser M, Heberle-Bors E, Amon A, Lelley T (1995) Evaluation of sensitivity of flow cytometry in detecting aneuploidy in wheat using disomic and ditelosomic wheat–rye addition lines. Cytometry 21A:387–393

Pfündel E, Meister A (1996) Flow cytometry of mesophyll and bundle sheath chloroplast thylakoids of maize (*Zea mays* L.). Cytometry 23A:97–105

Picot J, Guerin CL, Le Van KC, Boulanger CM (2012) Flow cytometry: retrospective, fundamentals and recent instrumentation. Cytotechnology 64:109–130

Pillay M, Ogundiwin E, Tenkouano A, Doležel J (2006) Ploidy and genome composition of Musa germplasm at the International Institute of Tropical Agriculture (IITA). Afr J Biotechnol 5:1224–1232

Praça-Fontes MM, Carvalho CR, Clarindo WR, Cruz CD (2011) Revisiting the DNA C-values of the genome size-standards used in plant flow cytometry to choose the "best primary standards". Plant Cell Rep 30:1183–1191

Price HJ, Hodnett G, Johnston JS (2000) Sunflower (*Helianthus annuus*) leaves contain compounds that reduce nuclear propidium iodide fluorescence. Ann Bot 86:929–934

Rani V, Raina SN (2000) Genetic fidelity of organized meristem–derived micropropagated plants: a critical reappraisal. In Vitro Cell Dev Biol Plant 36:319–330

Reichheld JP, Vernoux T, Lardon F, van Montagu M, Inzé D (1999) Specific checkpoints regulate plant cell cycle progression in response to oxidative stress. Plant J 17:647–656

Rewers M, Sliwinska E (2012) Endoreduplication Intensity as a marker of seed developmental stage in the Fabaceae. Cytometry 81A:1067–1075

Ricroch A, Yockteng R, Brown SC, Nadot S (2005) Evolution of genome size across some cultivated Allium species. Genome 48:511–520

Roberts AV (2007) The use of bead beating to prepare suspensions of nuclei for flow cytometry from fresh leaves, herbarium leaves, petals and pollen. Cytometry 71A:1039–1044

Roux N, Doležel J, Swennen R, Zapata-Arias FJ (2001) Effectiveness of three micropropagation techniques to dissociate cytochimeras in *Musa* spp. Plant Cell Tissue Organ Cult 66:189–197

Roux N, Toloza A, Radecki Z, Zapata-Arias FJ, Doležel J (2003) Rapid detection of aneuploidy in *Musa* using flow cytometry. Plant Cell Rep 21:483–490

Rowan BA, Oldenburg DJ, Bendich AJ (2007) A high–throughput method for detection of DNA in chloroplasts using flow cytometry. Plant Methods 3:5

Šafářová L, Duchoslav M, Jandová M, Krahulec F (2011) *Allium oleraceum* in Slovakia: cytotype distribution and ecology. Preslia 83:513–527

Schmidt G, Thannhauser SJ (1945) A method for the determination of deoxyribonucleic acid, ribonucleic acid and phosphoproteins in animal tissues. J Biol Chem 161:83–89

Schönswetter P, Suda J, Popp M, Weis-Schneeweiss H, Brochmann C (2007) Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. Mol Phylogen Evol 42:92–103

Schulze D, Pauls KP (1998) Flow cytometric characterization of embryogenic and gametophytic development in *Brassica napus* microspore cultures. Plant Cell Physiol 39:226–234

Šesek P, Šuštar-Vozlič J, Bohanec B (2000) Determination of aneuploids in hop (*Humulus lupulus* L.) using flow cytometry. Pflugers Arch 439:16–18

Sgorbati S, Levi M, Sparvoli E, Trezzi F, Lucchini G (1986) Cytometry and flow cytometry of 4′,6-diamidino-2-phenylindole (DAPI)-stained suspensions of nuclei released from fresh and fixed tissues of plants. Physiol Plant 68:471–476

Sgorbati S, Sparvoli E, Levi M, Galli MG, Citterio S, Chiatante D (1991) Cell cycle kinetic analysis with flow cytometry in pea root meristem synchronized with aphidicolin. Physiol Plant 81:507–512

Shapiro HM (1985) Practical flow cytometry. Wiley, New York

Shapiro HM (1988) Practical flow cytometry, 2nd edn. Wiley-Liss, New York

Shapiro HM (1995) Practical flow cytometry, 3rd edn. Wiley-Liss, New York

Shapiro HM (2003) Practical flow cytometry, 4th edn. Wiley, Hoboken

Siljak-Yakovlev S, Benmale S, Cerbah M, Coba de la Pena T, Bounaga N, Brown SC, Sarr A (1996) Chromosomal sex determination and heterochromatin structure in date palm. Sex Plant Reprod 9:127–132

Silva TCR, Abreu IS, Carvalho CR (2010) Improved and reproducible flow cytometry methodology for nuclei isolation from single root meristem. J Bot 2010, Article ID 320609

Śliwinska E, Thiem B (2007) Genome size stability in six medicinal plant species propagated in vitro. Biol Plant 51:556–558

Śliwinska E, Zielinska E, Jedrzejcyk I (2005) Are seeds suitable for flow cytometric estimation of plant genome size? Cytometry 64A:72–79

Śliwinska E, Bassel GW, Bewley D (2009) Germination of *Arabidopsis thaliana* seeds is not completed as a result of elongation of the radicle but of the adjacent transition zone and lower hypocotyl. J Exp Bot 60:3578–3594

Šmarda P (2006) DNA ploidy levels and intraspecific DNA content variability in Romanian fescues (*Festuca*, Poaceae) measured in fresh and herbarium material. Folia Geobot 41:417–432

Šmarda P, Stančík D (2006) Ploidy level variability in South American fescues (*Festuca* L., Poaceae): use of flow cytometry in up to 5 1/2–year–old caryopses and herbarium specimens. Plant Biol 8:73–80

Šmarda P, Müller J, Vrána J, Kočí K (2005) Ploidy level variability of some Central European fescues (*Festuca* subg. Festuca, Poaceae). Biologia (Bratislava) 60:25–36

Šmarda P, Bureš P, Horová L, Foggi B, Rossi G (2008) Genome size and GC content evolution of *Festuca*: ancestral expansion and subsequent reduction. Ann Bot 101:421–433

Šmarda P, Bureš P, Šmerda J, Horová L (2011) Measurements of genomic GC content in plant genomes with flow cytometry: a test for reliability. New Phytol 193:513–521

Soltis DE, Soltis PS, Bennett MD, Leitch IJ (2003) Evolution of genome size in the angiosperms. Am J Bot 90:1596–1603

Stehlík I, Barrett SCH (2005) Mechanisms governing sex-ratio variation in dioecious *Rumex nivalis*. Evolution 59:814–825

Stehlík I, Kron P, Barrett SCH, Husband BC (2007) Sexing pollen reveals female bias in a dioecious plant. New Phytol 175:185–194

Suda J, Leitch IJ (2010) The quest for suitable reference standards in genome size research. Cytometry 77A:717–720

Suda J, Trávníček P (2006) Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry: new prospects for plant research. Cytometry 69A:273–280

Suda J, Kyncl T, Jarolímová V (2005) Genome size variation in Macaronesian angiosperms: forty percent of the Canarian endemic flora completed. Plant Syst Evol 252:215–238

Suda J, Krahulcová A, Trávníček P, Krahulec F (2006) Ploidy level versus DNA ploidy level: an appeal for consistent terminology. Taxon 55:447–450

Suda J, Kron P, Husband BC, Trávníček P (2007a) Flow cytometry and ploidy: applications in plant systematics, ecology and evolutionary biology. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 103–130

Suda J, Krahulcová A, Trávníček P, Rosenbaumová R, Peckert T, Krahulec F (2007b) Genome size variation and species relationships in *Hieracium* subgen. *Pilosella* (Asteraceae) as inferred by flow cytometry. Ann Bot 100:1323–1335

Suda J, Trávníček P, Mandák B, Berchová-Bímová K (2010) Genome size as a marker for identifying the invasive alien taxa in *Fallopia* section *Reynoutria*. Preslia 82:97–106

Sugimoto-Shirasu K, Roberts K (2003) "Big it up": endoreduplication and cell-size control in plants. Curr Opin Plant Biol 6:544–553

Sun YL, Sun Y, Lin GG, Zhang R, Zhang K, Xie JH, Wang LN, Li JM (2012) Multicolor flow cytometry analysis of the proliferations of T-lymphocyte subsets in vitro by EdU incorporation. Cytometry 81A:901–909

Swift H (1950) The constancy of desoxyribose nucleic acid in plant nuclei. Proc Natl Acad Sci U S A 36:643–654

Taylor IW, Milthorpe BK (1980) An evaluation of DNA fluorochromes, staining techniques, and analysis for flow cytometry. I. Unperturbed cell populations. J Histochem Cytochem 28:1224–1232

Temsch EM, Greilhuber J (2000) Genome size variation in *Arachis hypogae*a and *A. monticola* re-evaluated. Genome 43:449–451

Temsch EM, Greilhuber J (2001) Genome size in *Arachis duranensis*: a critical study. Genome 44:826–830

Temsch EM, Temsch W, Ehrendorfer–Schratt L, Greilhuber J (2010) Heavy metal pollution, selection, and genome size: the species of the Žerjav study revisited with flow cytometry. J Bot 2010, Article ID 596542

Tiersch TR, Chandler RW, Wachtel SS, Elias S (1989) Reference standards for flow cytometry and application in comparative studies of nuclear DNA content. Cytometry 10A:706–710

Typas MA, Heale JB (1980) DNA content of germinating sores, individual hyphal cells and resting structure cells of *Verticillium* spp. measured by microdensitometry. Microbiology 121:231–242

Ulrich I, Ulrich W (1991) High-resolution flow cytometry of nuclear DNA in higher plants. Protoplasma 165:212–215

Ulrich I, Fritz B, Ulrich W (1988) Application of DNA fluorochromes for flow cytometric DNA analysis of plant protoplasts. Plant Sci 55:151–158

van Duren M, Morpurgo R, Doležel J, Afza R (1996) induction and verification of autotetraploids in diploid banana (*Musa acuminata*) by in vitro techniques. Euphytica 88:25–34

Vidic T, Greilhuber J, Vilhar B, Dermastia M (2009) Selective significance of genome size in a plant community with heavy metal pollution. Ecol Appl 19:1515–1521

Vilhar B, Greilhuber J, Koce JD, Temsch EM, Dermastia M (2001) Plant genome size measurement with DNA image cytometry. Ann Bot 87:719–728

Vilhar B, Kladnik A, Blejec A, Chourey PS, Dermastia M (2002) Cytometrical evidence that the loss of seed weight in the miniature seed mutant of maize is associated with reduced mitotic activity in the developing endosperm. Plant Physiol 129:23–30

Vinogradov AE (1994) Measurement by flow cytometry of genomic AT/GC ratio and genome size. Cytometry 16A:34–40

Voglmayr H (2000) Nuclear DNA amounts in mosses (Musci). Ann Bot 85:531–546

Voglmayr H, Greilhuber J (1998) Genome size determination in Peronosporales (Oomycota) by Feulgen image analysis. Fungal Genet Biol 25:181–195

Watanabe M, Setoguchi D, Uehara K, Ohtsuka W, Watanabe Y (2002) Apoptosis–like cell death of *Brassica napus* leaf protoplasts. New Phytol 156:417–426

Weber S, Unker F, Friedt W (2005) Improved doubled haploid production protocol for *Brassica napus* using microspore colchicine treatment in vitro and ploidy determination by flow cytometry. Plant Breed 124:511–513

Whittemore AT, Olsen RT (2011) *Ulmus americana* (Ulmaceae) is a polyploid complex. Am J Bot 98:754–760

Yanpaisan W, King NJC, Doran PM (1998) Analysis of cell cycle activity and population dynamics in heterogeneous plant cell suspensions using flow cytometry. Biotechnol Bioeng 58:515–528

Yao N, Eisfelder BJ, Marvin J, Greenberg JT (2004) The mitochondrion – an organelle commonly involved in programmed cell death in *Arabidopsis thaliana*. Plant J 40:596–610

Yokoya K, Roberts AV, Mottley J, Lewis R, Brandham PE (2000) Nuclear DNA amounts in roses. Ann Bot 85:557–561

Záveský L, Jarolímová V, Štěpánek J (2005) Nuclear DNA content variation within the genus *Taraxacum* (Asteraceae). Folia Geobotanica 40:91–104

Zhang CQ, Gong FC, Lambert GM, Galbraith DW (2005) Cell type-specific characterization of nuclear DNA contents within complex tissues and organs. Plant Methods 1:7

Zhang CQ, Barthelson RA, Lambert GM, Galbraith DW (2008) Characterization of cell-specific gene expression through fluorescence–activated sorting of nuclei. Plant Physiol 147:30–40

Zhao J, Cui J, Liu J, Liao F, Henny RJ, Chen J (2012) Direct somatic embryogenesis from leaf and petiole explants of *Spathiphyllum* 'Supreme' and analysis of regenerants using flow cytometry. Plant Cell Tissue Organ Cult 110:239–249

Zonneveld BJM, Leitch IJ, Bennett MD (2005) First nuclear DNA amounts in more than 300 angiosperms. Ann Bot 96:229–244

Palacký University Olomouc

Faculty of Science

Department of Botany

and

Institute of Experimental Botany AS CR

Centre of Plant Structural and Functional Genomics

Mgr. Martina Bednářová

Occurrence and evolution of B chromosomes in genus *Sorghum*

P1527 Biology - Botany

Summary of Ph.D. Thesis

Olomouc

2022

Ph.D. thesis was carried out at the Department of Botany, Faculty of Science, Palacký University Olomouc, in years 2012–2022.

Candidate: **Mgr. Martina Bednářová**

Supervisor: **Mgr. Jan Bartoš, Ph.D.**

Reviewers: **prof. Ing. Vladislav Čurn, Ph.D.**
University of South Bohemia, České Budějovice, Czech Republic

RNDr. Jiří Široký, CSc.
Institute of Biophysics of the Czech Academy of Sciences, Brno, Czech Republic

Mgr. Karol Krak, Ph.D.
Czech University of Life Sciencies, Prague, Czech Republic

The evaluation of the Ph.D. thesis was written by…………………………..............................

…………………………………………………………………………………………………..........................

The summary of the Ph.D. thesis was sent for distribution on…………………………...............

The oral defence will take place on……………………in front of the commission for the Ph.D. study of the program Botany in…………………………………………………………………………….

…………………………………………………………………………………………………..........................

The Ph.D. thesis is available in the Library of the biological departments of the Faculty of Science of Palacký University Olomouc, Šlechtitelů 11, Olomouc-Holice.

Prof. Ing. Aleš Lebeda, DrSc.
Chairman of the Commission for the Ph.D.
Thesis of the study program Botany
Department of Botany, Faculty of Science
Palacký University Olomouc

Content

1 Introduction 4
2 Aims of the thesis 6
3 Materials and methods 7
4 Summary of results...9
5 Summary 12
6 References.. 13
7 List of author’s publications 16
8 Shrnutí (Summary, in Czech) 17

1 Introduction

B chromosomes (Bs) are extra chromosomes that occur in some individuals within populations of a particular species. They have been observed in all major groups of living organisms - plants, animals and fungi (D'Ambrosio et al. 2017). Unlike the basic set of chromosomes (A chromosomes, As), Bs are not essential for their host organism and are considered to be parasitic elements with their own evolutionary pathway (Baukeboom 1994). The behavior of Bs in their host organisms is in many ways unique, as they do not follow Mendelian inheritance and do not pair and recombine with A chromosomes during meiosis. Presence of Bs is generally neither beneficial nor harmful, but at higher numbers Bs often reduce fertility and vigor due to their cumulative negative effect (Jones 1991). B chromosomes often show numerical variability between individuals, or even within an individual. Their presence does not have an effect on phenotype, therefore their detection relies on cytological observations. B chromosomes are usually smaller than As, but some Bs can be as large as As and even Bs larger than As have been observed (Jones 1995). Generally, B chromosomes are a mosaic of different sequences derived from the host genome. They contain large amounts of repetitive sequences, especially mobile elements and tandem repeats (Houben et al. 2013). B chromosomes also may contain gene sequences and sequences derived from plastids or mitochondria. Only a very few sequences were identified and confirmed as a B-specific (Marques et al. 2018).

B chromosomes possess a wide variety of accumulation mechanisms by which they ensure their transmission to offspring with the transmission rate higher than 0.5. This transmission advantage is referred to as "B chromosome drive" (Jones 1991). Accumulation mechanisms of B chromosomes are species-specific and can be active before, during or after meiosis. Whereas in animals, premeiotic and meiotic drive is the most common, in plants we typically observe postmeiotic drive during pollen mitosis, although there are some exceptions (Jones 1991). The origin of B chromosomes has long time seemed mysterious. However, with ongoing research at molecular level, various scenarios of their possible origin have begun to be revealed. Based on the similarity of some sequences of B and A chromosomes, it has been found that some B chromosomes are derived from autosomes, e.g. the B chromosome in *Zea mays* (Peacock *et al.* 1981, Alfenito and Birchler 1993, Stark *et al.* 1996), *Secale cereale* (Sandery *et al.* 1990, Blunden 1993, Houben *et al.* 1996) or *Locusta migratoria* (Teruel *et al.* 2010). Sex chromosomes have also been suggested as the source of genetic material for B chromosome, as is the case of the B chromosome of the grasshopper *Eyprepocnemis plorans* (López-León *et al.* 1994). The B chromosome is not always derived from the chromosomes of the host genome - interspecies origin is also possible, as was confirmed in the PSR chromosome of *N. vitripennis* (McAllister and Werren 1997).

Little is known about the evolutionary history and phylogeny of B chromosomes. So far, some progress in composing a mosaic of their origin has been made only in the most intensively studied species. Some B chromosomes have probably originated a long time ago and appear to be relatively conserved. An example is the B chromosome of maize, whose age has been estimated to be up to about 12 million years (Blavet *et al.* 2021). On the other hand, there are species where the evolution of B chromosome is very dynamic, which is manifested by the emergence of new B chromosome variants in a relatively short periods of time, such as in *Eyprepocenmis plorans* (López-León *et al.* 1993).

Genus *Sorghum* belongs among monocot flowering plants in the grass family Poaceae, the tribe Andropogoneae. Andropogoneae encompasses about 1200 species in 98 genera with numerous species bearing Bs (Soreng *et al.* 2017). *Sorghum* covers about 23 annual and perennial species including an agriculturally important *Sorghum bicolor*. B chromosomes were reported in five *Sorghum* species: *S. purpureosericeum*, *S. nitidum*, *S. halepense*, *S. stipoideum*, and *S. bicolor* ssp. *verticilliflorum* (Janaki-Amal 1940, Raman and Krishnaswami 1960, Raman *et al.* 1964, Wu 1992, Huskins and Smith 1932). Typically, B chromosomes in this genus are absent in roots and most of the shoot tissues (Darlington and Thomas 1941). Further, B chromosomes in sorghums are morphologically variable, and different morphotypes have been reported even within individual species (Raman and Krishnaswami 1960, Wu and Pi 1975). Due to the fact that B chromosomes have not yet been the target of greater research interest, our knowledge about *Sorghum* Bs is limited to a few cytogenetic studies. Molecular data are still lacking.

The presented work aims to broaden the knowledge about B chromosomes in the genus *Sorghum* at the molecular level and to contribute to the understanding of the evolution of the B chromosome within the genus.

2 Aims of the thesis

The present work aims to contribute to the understanding of the evolution and function of B chromosomes in the genus *Sorghum*. Two main goals were:

I. Development of markers specific for the B chromosome of *Sorghum purpureosericeum*

II. Phylogenetic analysis of the genus *Sorghum* with respect to the hypothesis of the B chromosome origin in the genus *Sorghum*

3 Materials and methods

Plant material

Sorghum seeds were obtained from Australian Grains Genebank (Australia), United States Department of Agriculture (United States) and International Crops Research Institute for the Semi-Arid Tropics (India). Seeds were germinated on Petri dishes in a thermal incubator under 8 h light/16 h dark photoperiod at temperatures of 29 °C day/25 °C night. The ensuing seedlings were planted into soil mixed with sand (2:1) in 10 cm diameter pots and cultivated under the same conditions as for seeds. Illumina sequencing and PCR marker development was performed on *S. purpureosericeum*. The phylogenetic analysis, was performed involving 21 *Sorghum* species represented by 58 samples.

Isolation and sequencing of pollen nuclei

Fresh anthers from individual plants were collected and their pollen extracted in a LB01 buffer by vortexing and centrifugation. Haploid pollen nuclei from both B+ and B0 plants were isolated using flow cytometry. From each individual, 9000 of its 1C nuclei (corresponding to 20 ng of DNA) were flow-sorted and nuclear DNA was used separately for preparing the sequencing libraries. Libraries for sequencing were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (Ipswich, MA, USA) and sequenced on NovaSeq 6000, and 2×250 bp paired-end reads were eventually produced.

Repeat analysis and PCR marker development

Analysis of repetitive sequences was implemented using the Galaxybased server (Afgan et al., 2018), supplemented with the RepeatExplorer2 (Novák et al., 2013) and TAREAN (Novák et al., 2017) tools. Nine putative B-specific clusters were revealed by the repeat analysis. For each cluster, specific primers with short amplicons (up to 1.3 kbp) and long amplicons (in the range of 2–5 kbp) were designed. Primer specificity was tested on B+ genomic DNA, for which B0 genomic DNA served as the control. DNA was extracted from whole, lyophilized spikelets of *S. purpureosericeum* plants, by using the NucleoSpin® Plant II Kit (Macherey-Nagel, Düren, Germany). Long amplicons were amplified with PrimeStar®GXL DNA polymerase (Takara Bio Inc., Shiga, Japan) on a C1000 Touch™ Thermal Cycler (BioRad, Hercules, CA, USA) and separated on a 0.8% agarose gel. Short amplicons were amplified using Taq DNA polymerase with the Standard Taq Buffer (New England Biolabs, Ipswich, MA, USA). The amplification products were visualized on a 1.5% agarose gel.

DNA isolation, amplification and sequencing of ITS, trnL-trnF, trnH-psbA regions

Genomic DNA was extracted from lyophilised young leaf tissue of 21 *Sorghum* species using NucleoSpin® Plant II Kit (Macherey-Nagel, Düren, Germany). Nuclear ribosomal *ITS1-5.8S-ITS2* sequence region and two chloroplastic intergenic spacers *trnL-trnF* and *psbA-trnH* were amplified from each DNA sample using specific primers common for all species involved in analysis. PCR was performed using Jumpstart™ REDTaq® ReadyMix™ Reaction Mix (Merck KGaA, Kenilworth, USA). Alternatively, amplification was carried out using Phire Plant Direct PCR Master Mix (Thermo Fisher Scientific, Waltham, USA) for accessions poorly growing or not germinating. Products of all amplification reactions were separated and visualised on 1.2 % agarose gel.

Amplified DNA was purified using ExoSAP-IT® (USB, Cleveland, USA). The sequencing was accomplished using the BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, USA). All regions were sequenced using reverse and forward primer on ABI 3730xl DNA analyzer (Applied Biosystems).

Sequence analysis and construction of phylogenetic tree

Nucleotide sequence were edited using Staden Package 2.0.0b11 (Staden et al. 2000). The boundaries of sequences for alignment were determined by comparison to known *S. bicolor ITS* sequence and to annotated chloroplast sequence of *S. bicolor* in the case of *trnL-trnF* and *psbA-trnH* sequence regions. Multiple sequence alignment of each region was done using MAFFT v7.029 (--globallpair–maxiterate 1000) (Katoh & Toh 2008) and graphically shown in SeaView v5.0.2 (Gouy et al. 2010). GC content and sequence identity to consensus was calculated using BioEdit v.7.2.6.1 (Hall 1999) for individual sequences of *ITS*, *trnL-trnF* and *psbA-trnH* regions.

To create phylogenetic trees, three datasets consisting of concatenated *ITS* sequences (*ITS1-ITS2*); concatenated plastid sequences (*trnL-trnF* + *psbA-trnH*) and concatenated both nuclear and chloroplastic datasets (*ITS1-ITS2* + *trnL-trnF* + *psbA-trnH*) were used with *Miscanthus* sp. as an outgroup. Phylograms were created with BioNJ (Gascuel 1997) and PhyML 3.0 (Guindon et al. 2010) implemented in SeaView v5.0.2 (Gouy et al. 2010). Approximate likelihood ratio tests (Anisimova & Gascuel 2006) was performed to assess branch support of PhyML trees and bootstraping was conducted to support BioNJ clustering. Final phylogenetic trees were depicted with FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

4 Summary of results

To contribute to B chromosome research in *Sorghum*, specific markers for B chromosome of *Sorghum purpureosericeum* were developed. Based on raw data from Illumina sequencing, repeat analysis using RepeatExplorer was performed. This analysis revealed nine putative B-specific clusters SpuCL115, SpuCL135, SpuCL144, SpuCL168, SpuCL169, SpuCL175, SpuCL189, SpuCL214 and SpuCL220. These clusters were used to design primers and their specificity was tested in PCR. Among primer pairs designed for short amplicons (up to 1300 bp), three were verified as a B-specific (Fig. 1). From the primer pairs designed for long amplicons (from 2 to 5 kbp), additional four were selected as B-specific (Fig. 2). As a result, seven B-specific markers were developed for reliable PCR detection of *S. purpureosericeum* B chromosome.

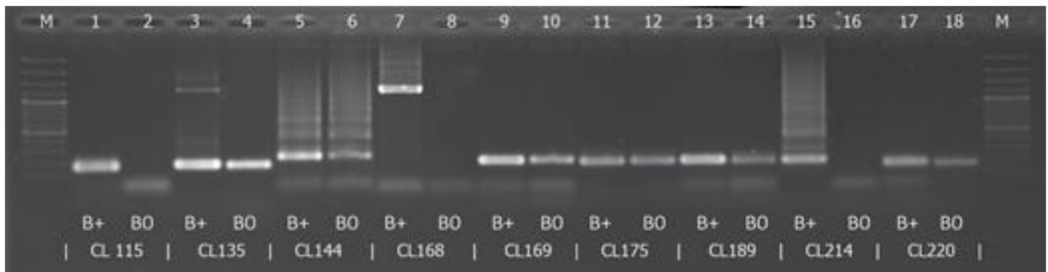

Fig. 1 Electrophoresis of short amplicons for nine putative B-specific clusters (SpuCL115, SpuCL135, SpuCL144, SpuCL168, SpuCL169, SpuCL175, SpuCL189, SpuCL214, and SpuCL220). For all clusters, primer specificity was tested in parallel on B+ DNA and B0 DNA. The presence of PCR products of SpuCL115, SpuCL168, and SpuCL214 exclusively in B+ samples indicates their B specificity. The remaining primer pairs showed non-specific amplifications. The M lane corresponds to the GeneRuler 100 bp Plus DNA Ladder (Karafiátová *et al.* 2021).

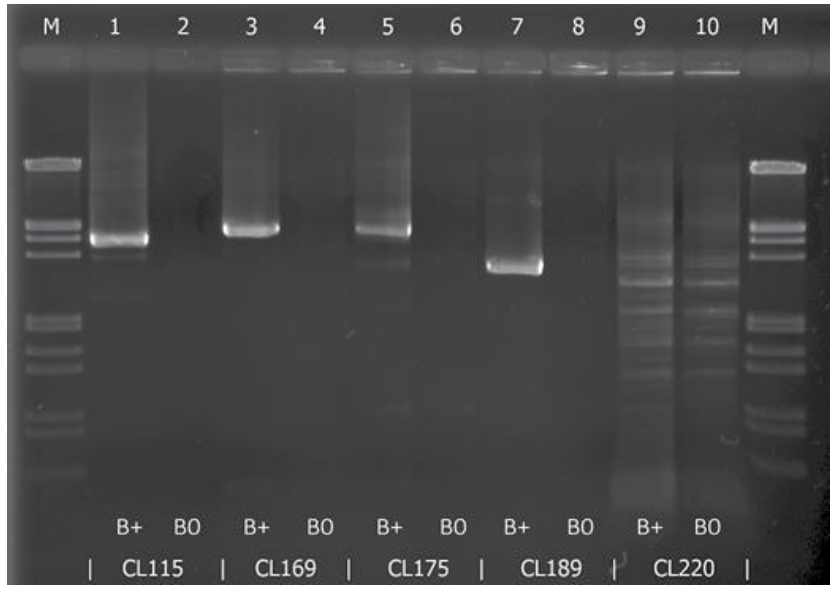

Fig. 2 Electrophoresis of long amplicons for five putative B-specific clusters (SpuCL115, SpuCL169, SpuCL175, SpuCL189, and SpuCL220). Primer specificity was tested in parallel on B+ DNA and B0 DNA. The presence of PCR products of SpuCL115, SpuCL169, SpuCL175, and SpuCL189 exclusively in B+ samples indicates their B specificity. The M lane corresponds to λ DNA digested with *Eco*RI and *Hin*dIII (Karafiátová *et al.* 2021).

To shed more light on phylogenetic relationships in the genus with respect to B chromosome occurrence, phylogenetic reconstruction was performed. Twenty-one *Sorghum* species were included into analysis and represented by 58 samples. As a phylogenetic markers three noncoding DNA regions were selected: nuclear spacer *ITS1 - ITS2*, and two chloroplastic spacers *trnH-psbA* and *trnL-trnF*. Phylogenetic tree was reconstructed using phyML and BioNJ algorithm resulting in single tree with two strongly supported lineages (Fig. 3). B-carrying species were distributed independently in both major clades. Three evolutionary scenarios of the B chromosome formation in the genus were considered. B chromosome could have arise: 1) in a single event in a common ancestor of all *Sorghum* species and then have been preserved in some lineages while disappeared in the others; 2) in a single event in an ancestor of closely related species which have kept Bs up until today; or 3) several times in independent events during *Sorghum* evolution which also include the eventuality of B chromosome

formation at the level of individual species. Based on phylogenetic analysis, scenario 2), i.e. formation in the ancestor of a group of closely related species, was ruled out.

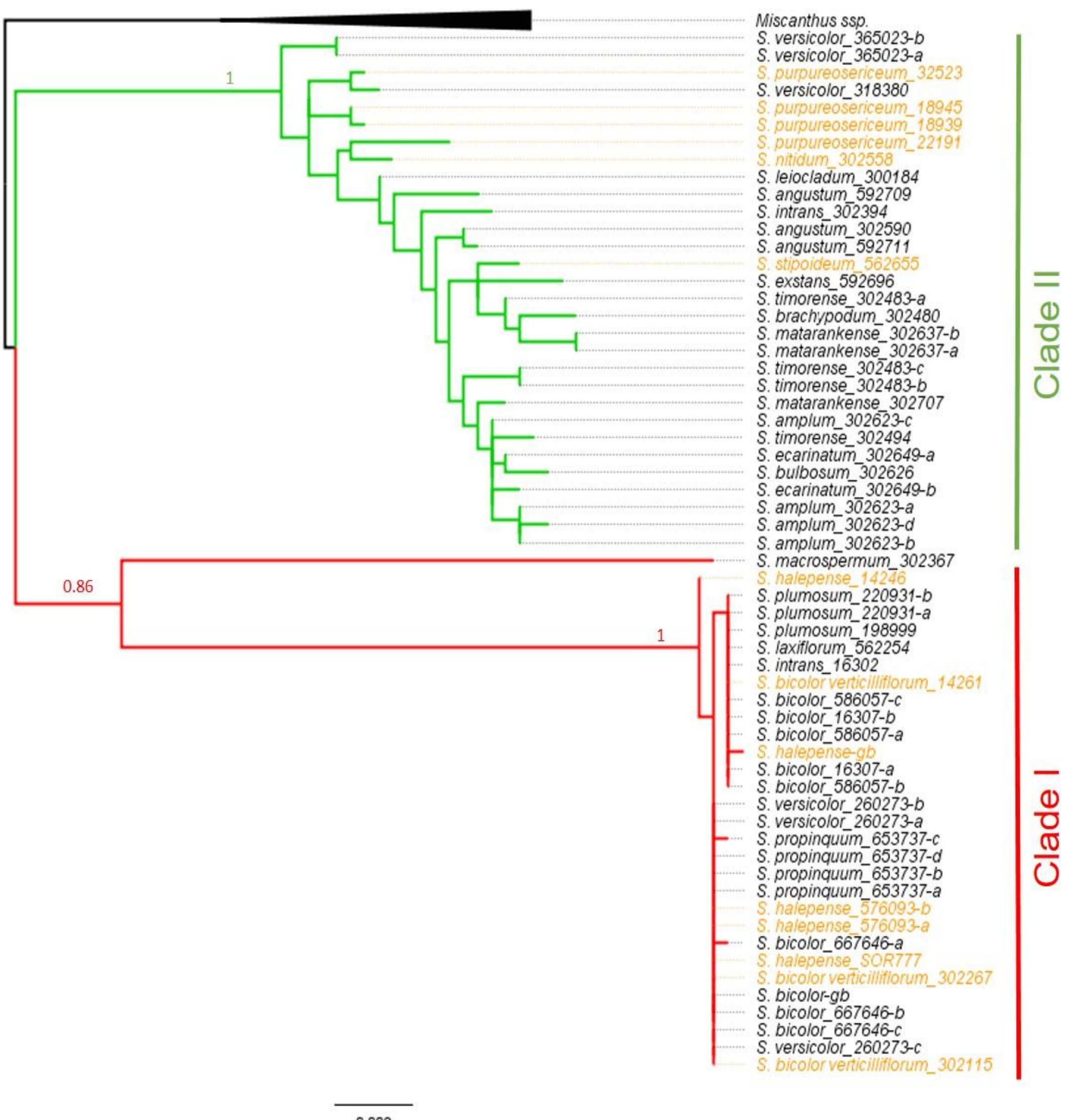

Fig. 3 Phylogenetic tree of 21 *Sorghum* species based on phyML analysis using concatenated nuclear and chloroplast sequences (*ITS1-ITS2-trnHpsbA-trnLtrnF*). Two main clades I and II with strong bootstrap support were resolved. B-carrying species are marked in orange (Bednářová *et al.* 2021).

5 Summary

B chromosomes are genomic elements that can spread in the population despite the fact that they do not carry any essential properties. To date, B chromosomes in the genus *Sorghum* have not been the subject of any extensive research, therefore our knowledge about them and their behavior is very limited and molecular data are completely missing. As was shown in previous studies, B chromosomes in *Sorghum* are eliminated from most tissues, which was also confirmed by our observations in *S. purpureosericeum*. One of the goals of this work was to develop markers specific for the B chromosome of *S. purpureosericeum*. Using data obtained by Illumina sequencing, repeat analysis was performed and candidates for B-specific repeats were identified. Based on them, a set of markers was designed and tested for their B-specifity in PCR. The verification confirmed that seven of them are indeed B-specific and can be used for the identification and study of B-carrying individuals of *S. purpureosericeum*. The developed markers will be valuable for research studies focused on the elimination process of Bs and will significantly simplify the identification of B-positive plants (or tissues) at an early developmental stage, bypassing the need of cytogenetic analysis of plants in the flowering phase.

In parallel, we focused on the evolutionary pathway of the B chromosome in the genus. In order to clarify the relationships among *Sorghum* species, a phylogenetic analysis of 21 species was performed, confirming the existence of two main developmental lineages, with B-carrying species being present in both of them. Three evolutionary scenarios for *Sorghum* Bs were considered: A) a single origin in a common ancestor of all *Sorghum* species and then have been preserved in some lineages while disappeared in the others; B) a single origin in the ancestor of closely related species, which have kept Bs up until today; and C) multiple independent origins during *Sorghum* evolution. Results of the phylogenetic reconstruction enabled us to rule out the possibility of the origin of B chromosome in an ancestor of a group of closely related species (scenario B).

6 References

Alfenito, M. R., & Birchler, J. A. (1993). Molecular characterization of a maize B chromosome centric sequence. *Genetics*, *135*(2), 589-597.

Beukeboom, L. W. (1994). Bewildering Bs: an impression of the 1st B-Chromosome Conference. *Heredity*, *73*(3), 328-335.

Blavet, N., Yang, H., Su, H., Solanský, P., Douglas, R. N., Karafiátová, M., ... & Bartoš, J. (2021). Sequence of the supernumerary B chromosome of maize provides insight into its drive mechanism and evolution. *Proceedings of the National Academy of Sciences*, *118*(23).

Blunden, R., Wilkes, T. J., Forster, J. W., Jimenez, M. M., Sandery, M. J., Karp, A., & Jones, R. N. (1993). Identification of the E3900 family, a second family of rye B chromosome specific repeated sequences. *Genome*, *36*(4), 706-711.

D'Ambrosio, U., Alonso-Lifante, M. P., Barros, K., Kovařík, A., de Xaxars, G. M., & Garcia, S. (2017). B-chrom: A database on B-chromosomes of plants, animals and fungi. *New Phytologist*, *216*(3), 635-642.

Darlington, C. D., & Thomas, P. T. (1941). Morbid mitosis and the activity of inert chromosomes in Sorghum. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, *130*(859), 127-150.

Houben, A., Banaei-Moghaddam, A. M., & Klemme, S. (2013). Biology and evolution of B chromosomes. In *Plant Genome Diversity Volume 2* (pp. 149-165). Springer, Vienna.

Houben, A., Kynast, R. G., Heim, U., Hermann, H., Jones, R. N., & Forster, J. W. (1996). Molecular cytogenetic characterisation of the terminal heterochromatic segment of the B-chromosome of rye (Secale cereale). *Chromosoma*, *105*(2), 97-103.

Huskins, C. L., & Smith, S. G. (1932). A cytological study of the genus Sorghum Pers. *Journal of Genetics*, *25*(2), 241-249.

Janaki-Ammal, E. K. (1940). Chromosome diminution in a plant. *Nature*, *146*(3713), 839-840.

Jones, R. N. (1991). B-chromosome drive. *The American Naturalist*, *137*(3), 430-442.

Jones, R. N. (1995). B chromosomes in plants. *New Phytologist*, *131*(4), 411-434.

López-León, M. D., Cabrero, J., Pardo, M. C., Viseras, E., Camacho, J. P. M., & Santos, J. L. (1993). Generating high variability of B chromosomes in Eyprepocnemis plorans (grasshopper). *Heredity*, *71*(4), 352-362.

López-León, M. D., Neves, N., Schwarzacher, T., Hewitt, G. M., & Camacho, J. P. M. (1994). Possible origin of a B chromosome deduced from its DNA composition using double FISH technique. *Chromosome Research*, *2*(2), 87-92.

Marques, A., Klemme, S., & Houben, A. (2018). Evolution of plant B chromosome enriched sequences. *Genes*, *9*(10), 515.

McAllister, B. F., & Werren, J. H. (1997). Hybrid origin of a B chromosome (PSR) in the parasitic wasp Nasonia vitripennis. *Chromosoma*, *106*(4), 243-253.

Peacock, W. J., Dennis, E. S., Rhoades, M. M., & Pryor, A. J. (1981). Highly repeated DNA sequence limited to knob heterochromatin in maize. *Proceedings of the National Academy of Sciences*, *78*(7), 4490-4494.

Raman V.S., Meenakshi K., Thangam M.S., & Sivagnanam L. (1964). The cytological behaviour of B-chromosomes in *Sorghum halepense*. *Abst. Madras Agricultural Journal*, *51*, 72–73.

Raman, V. S., & Krishnaswami, D. (1960). Accessory chromosomes in Sorghum nitidum Pers. *J. Indian Bot. Soc.*, *39*, 278-280.

Sandery, M. J., Forster, J. W., Blunden, R., & Jones, R. N. (1990). Identification of a family of repeated sequences on the rye B chromosome. *Genome*, *33*(6), 908-913.

Soreng, R. J., Peterson, P. M., Romaschenko, K., Davidse, G., Teisher, J. K., Clark, L. G., ... & Zuloaga, F. O. (2017). A worldwide phylogenetic classification of the Poaceae (Gramineae) II: An update and a comparison of two 2015 classifications. *Journal of Systematics and evolution*, *55*(4), 259-290.

Stark, E. A., Connerton, I., Bennett, S. T., Barnes, S. R., Parker, J. S., & Forster, J. W. (1996). Molecular analysis of the structure of the maize B-chromosome. *Chromosome Research*, *4*(1), 15-23.

Teruel, M., Cabrero, J., Perfectti, F., & Camacho, J. P. M. (2010). B chromosome ancestry revealed by histone genes in the migratory locust. *Chromosoma*, *119*(2), 217-225.

Wu, T. P. (1992). B-Chromsomes in Sorghum stipoideum. *Heredity*, *68*(5), 457-463.

Wu, T. P., & Pi, C. P. (1975). Accessory chromosome in Sorghum nitidum Pers. *Taiwania*, *20*(2), 147-161.

Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Čech, M., ... & Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic acids research*, *46*(W1), W537-W544.

Novák, P., Ávila Robledillo, L., Koblížková, A., Vrbová, I., Neumann, P., & Macas, J. (2017). TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic acids research*, *45*(12), e111-e111.

Staden, R., Beal, K. F., & Bonfield, J. K. (2000). The staden package, 1998. In Bioinformatics methods and protocols (pp. 115-130). Humana Press, Totowa, NJ.

Katoh, K., & Toh, H. (2008). Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC bioinformatics*, *9*(1), 1-13.

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular biology and evolution*, *27*(2), 221-224.

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In Nucleic Acids Symp. Ser. (Vol. 41, pp. 95-98).

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*, *59*(3), 307-321.

Anisimova, M., & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic biology*, *55*(4), 539-552.

Novák P., Neumann P., Pech J., Steinhaisl J. & Macas J. (2013). RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics*, *29*, 792–793.

7 List of author’s publications

Karafiátová M., Bednářová M., Said M., Čížková J., Holušová K., Blavet N., Bartoš J. (2021). The B chromosome of *Sorghum purpureosericeum* reveals the first pieces of its sequence, *Journal of Experimental Botany*, *72* (5),1606-1616, https://doi.org/10.1093/jxb/eraa548

Bednářová, M., Karafiátová, M., Hřibová, E., Bartoš, J. (2021). B Chromosomes in genus *Sorghum* (Poaceae). *Plants*, *10*, 505. https://doi.org/10.3390/plants10030505

Vrána J., Cápal P., Bednářová M., Doležel J. (2014). Flow Cytometry in Plant Research: A Success Story. In: Nick, P., Opatrny, Z. (eds) Applied Plant Cell Biology. Plant Cell Monographs, vol 22. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-41787-0_13

8 Souhrn (Summary, in Czech)

B chromozomy jsou zvláštním typem chromozomů, které se mohou šířit v populaci i přesto, že nenesou žádné esenciální vlastnosti. B chromozomy rodu *Sorghum* dosud nebyly předmětem rozsáhlejšího výzkumu, naše znalosti o nich jsou velmi omezené a molekulární studie zcela chybí. Předchozí práce ukázaly, že B chromozomy v rodu *Sorghum* jsou z většiny tkání eliminovány, což bylo potvrzeno také v našich pozorováních u *S. purpureosericeum.* Jedním z cílů této práce bylo vyvinout markery specifické pro B chromozom *S. purpureosericeum*. Na základě dat získaných sekvenováním Illumina byla provedena analýza repetic, díky níž bylo identifikováno devět kandidátů na B-specifické repetice. Na jejich základě byly navrženy markery, z nichž sedm bylo po následném testování pomocí PCR ověřeno jako specifických pro B chromozom *S. purpureosericeum*. Vyvinuté markery budou cenné pro další výzkum zaměřený na proces eliminace B chromozomů a výrazně zjednoduší identifikaci B-pozitivních rostlin (nebo pletiv) v raném vývojovém stádiu, čímž bude možné vyhnout se nutnosti cytogenetické analýzy rostlin ve fázi květu.

Protože B chromozomy byly objeveny ve více druzích čiroku, využili jsme fylogenetickou analýzu k poodhalení evoluční cesty B chromozomu v rámci rodu. Za účelem objasnění vztahů mezi čirokovými druhy byla provedena fylogenetická analýza 21 druhů, která potvrdila existenci dvou hlavních vývojových linií, přičemž v obou byly přítomny druhy nesoucí B chromozomy. Byly uvažovány tři evoluční scénáře vzniku B chromozomu v rodu *Sorghum*: A) jednorázový vznik B chromozomu u společného předka všech čirokových druhů, přičemž u některých linií se zachoval do současnosti, zatímco z jiných vymizel; B) jednorázový vznik B chromozomu u předka skupiny blízce příbuzných druhů, které si jej udržely až do současnosti; a C) nezávislý vznik B chromozomu vícekrát během evoluce čiroku. Výsledky fylogenetické rekonstrukce nám umožnily vyloučit možnost vzniku B chromozomu u předka skupiny blízce příbuzných druhů (scénář B).