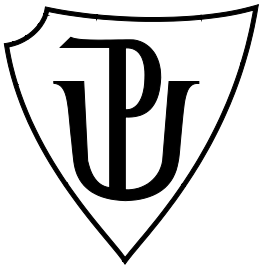
Palacký University Olomouc
Faculty of Science
Department of Cell Biology and Genetics

and

Institute of Experimental Botany of the Czech Academy of Sciences
Centre of Plant Structural and Functional Genomics



Genome analysis in selected representatives of *Poaceae*

Jana Zwyrtková

Ph.D. Thesis

Supervisor: Prof. Ing. Jaroslav Doležel, DrSc.

Co-supervisor: Mgr. Eva Hřibová, Ph.D.

Olomouc 2021

Declaration

I hereby declare that I have written the Ph.D. Thesis independently under the supervision of Prof. Ing. Jaroslav Doležel, DrSc. using the information sources listed in the References.

This work was supported by the ERDF project "Plants as a Tool for Sustainable Global Development" (No. CZ.02.1.01/0.0/0.0/16_019/0000827). The computing was supported and computational resources were supplied by the Ministry of Education, Youth and Sports of the Czech Republic under the Projects CESNET (Project No. LM2015042) and CERIT-Scientific Cloud (Project No. LM2015085) provided within the program Projects of Large Research, Development and Innovations Infrastructures.

In Olomouc ....

Acknowledgements

I would like to express my gratitude to my supervisor, the head of the laboratory, Prof. Ing. Jaroslav Doležel, DrSc., for the opportunity to work in the Centre of Plant Structural and Functional Genomics, for his professional guidance and invaluable ideas and help. Further, I would like to thank Mgr. Eva Hřibová, Ph.D. for her cooperation and guidance of the experimental part of my studies and also to all members of the Centre. Finally, I would like to thank my family and my partner Jan for supporting me during my studies.

Bibliographical identification

Author's first name and surname	Jana Zwyrtková
Title	Genome analysis in selected representatives of *Poaceae*
Type of thesis	Ph.D. Thesis
Department	Department of Cell Biology and Genetics, Palacký University Olomouc and Centre of Plant Structural and Functional Genomics, Institute of Experimental Botany of the Czech Academy of Sciences
Supervisor	Prof. Ing. Jaroslav Doležel, DrSc.
Co-supervisor	Mgr. Eva Hřibová, Ph.D.
The year of presentation	2021
Keywords	*Agropyron cristatum*, chromosome genomics, *Festuca* spp., Illumina sequencing, *Lolium* spp., repetitive DNA sequences, simple sequence repeats, wild wheat relatives.
Number of pages/appendices	88/IV
Language	English

Abstract

The Thesis aimed to improve the knowledge of genome organization in selected representatives of the family *Poaceae*. Their genomes were compared, both at the level of genera, species and cultivars. Important results were obtained using advanced molecular biology methods, including flow cytometry and next-generation sequencing.

In order to compare repetitively arranged DNA sequences and identify a novel centromeric element in the representatives of the *Festuca* and *Lolium* genera, various molecular and genomics approaches were employed. For comparative analyses, sequencing of their genomes by Illumina technology with a small depth of coverage and subsequent *in silico* analyses of the sequence data were used. The analyses showed that retrotransposons were the most abundant type of repetitive DNA sequences in the genomes of fescues and ryegrasses. The biggest difference in the representation

of repetitive sequences between genomes was observed for the Athila element, which belongs to the family of Ty3/gypsy retroelements, when the occurrence of this element in ryegrasses was up to five-fold more frequent than in the fescues. The Ty3/gypsy family includes the Cereba element, which typically localizes to (peri)centromeric chromosome regions, as shown for example in barley or wheat. In this work, the long terminal repeat (LTR) element Fesreba was identified, whose relationship with the Cereba element was confirmed by phylogenetic analysis and further confirmed by colocalization with the histone variant CENH3 in cytogenetic experiments. The location of the Cereba element in the centromeric and pericentromeric chromosome regions suggested a possible role of the Fesreba element in these chromosome domains in the studied grass species.

With the aim to facilitate the use of *Agropyron cristatum*, a wild relative of wheat, in introgressive wheat breeding, highly polymorphic microsatellite (SSR) markers specific for individual chromosomes and chromosome arms of this species were developed. The first step in the experimental procedure was the sorting of chromosomes and chromosome arms by flow cytometry. This approach allowed targeted development of SSR markers. The chromosomal DNA was sequenced by Illumina technology and the data were analyzed using appropriate bioinformatics procedures. This led to *in silico* identification of 250 SSR markers that were verified using PCR. Out of these, 72 were specific for a particular chromosome, or chromosomal arm. The newly developed SSR markers will enable rapid identification of introduced segments of the *A. cristatum* genome into the wheat genome and subsequent monitoring of their transfer to future generations during the breeding process.

There is an increasing need to understand genome organization and function in important crops, many of them with large genomes. The last part of this work reviewed experimental approaches to facilitate the analysis of complex genomes. The focus was on flow cytometric chromosome sorting and on coupling this technology with the methods of molecular biology and genomics. The successful applications include targeted development of DNA markers, physical mapping, chromosome sequencing, and gene cloning.

In summary, the presented Thesis contributed to the improvement of the knowledge on the genome organization in important crops. The results obtained will facilitate the application of molecular and genomic techniques in breeding of new varieties of agricultural crops.

Bibliografická identifikace

Jméno a příjmení autora	Jana Zwyrtková
Název práce	Analýza genomů vybraných zástupců *Poaceae*
Typ práce	Disertační práce
Pracoviště	Katedra buněčné biologie a genetiky, Univerzita Palackého v Olomouci a Centrum strukturní a funkční genomiky rostlin, Ústav experimentální botaniky Akademie věd České republiky
Školitel	Prof. Ing. Jaroslav Doležel, DrSc.
Školitel specialista	Mgr. Eva Hřibová, Ph.D.
Rok obhajoby práce	2021
Klíčová slova	*Agropyron cristatum*, chromozómová genomika, divocí příbuzní pšenice, *Festuca* spp., Illumina sekvenování, jednoduché repetice, *Lolium* spp., repetitivní DNA sekvence.
Počet stran/příloh	88/IV
Jazyk	Anglický

Abstrakt

Cílem disertační práce bylo prohloubit znalost struktury genomů vybraných zástupců čeledi *Poaceae*. Jejich genomy byly porovnávány jak na úrovni rodů, tak na úrovni druhů a kultivarů. Pomocí metod molekulární biologie, včetně průtokové cytometrie a metod sekvenování nové generace, byly získány významné výsledky.

Různé molekulární a genomické přístupy byly použity za účelem srovnání zastoupení repetitivně uspořádaných sekvencí DNA a identifikace nového centromerického elementu u zástupců rodů kostřava (*Festuca*) a jílek (*Lolium*). Pro komparativní analýzy bylo využito sekvenování jejich genomů technologií Illumina s malou hloubkou pokrytí a následné *in silico* analýzy sekvenačních dat. Tyto analýzy ukázaly, že retrotranspozóny jsou nejčetnějším typem repetitivních DNA sekvencí v genomech kostřav a jílků. Největší rozdíl v zastoupení repetitivních sekvencí mezi genomy byl pozorován v případě elementu Athila, který patří do rodiny Ty3/gypsy

retroelementů, kdy výskyt tohoto elementu u jílků byl až pětkrát četnější než u kostřav. Do rodiny Ty3/gypsy patří Cereba element, který se typicky vyskytuje v (peri)centromerických oblastech, například u ječmene či pšenice. V této práci byla identifikována dlouhá terminální repetice (LTR) element Fesreba, jehož příbuznost s Cereba elementem byla potvrzena fylogenetickou analýzou a také kolokalizací s histonovou variantou CENH3 pomocí cytogenetických experimentů. Výskyt Cereba elementu byl potvrzen v centromerických a pericentromerických oblastech, což naznačuje jeho možnou roli v těchto chromozomálních doménách u studovaných druhů trav.

S cílem prohloubit znalosti o potenciálních zdrojích důležitých genů pro introgresní šlechtění pšenice byl studován její planý příbuzný druh *Agropyron cristatum*. V průběhu disertační práce byly vyvinuty vysoce polymorfní mikrosatelitní (SSR) markery specifické pro jednotlivé chromozómy a ramena chromozómů tohoto druhu. Výchozím experimentálním postupem bylo třídění chromozómů a chromozómových ramen pomocí průtokové cytometrie. Tento přístup umožnil zacílený vývoj SSR markerů. DNA chromozómů byla sekvenována technologií Illumina a data byla zpracována odpovídajícími bioinformatickými postupy. To umožnilo *in silico* identifikaci 250 SSR markerů, které byly ověřeny pomocí PCR. Z nich 72 bylo specifických pro určitý chromozóm či rameno chromozómu. Nově vyvinuté SSR markery umožní rychlou identifikaci vnesených segmentů genomu *A. cristatum* do genomu pšenice a sledování jejich přenosu do dalších generací v průběhu šlechtění.

S ohledem na rostoucí potřebu znalosti organizace a funkce genomu důležitých zemědělských plodin, které mají obvykle velké genomy, byly v přehledném článku shrnuty experimentální přístupy pro usnadnění jejich studia. Byla popsána kombinace metod třídění chromozómů pomocí průtokové cytometrie a metod molekulární biologie a genomiky. Úspěšné aplikace zahrnují cílený vývoj DNA markerů, fyzické mapování, sekvenování na chromozómové úrovni a klonování genů.

Ve svém souhrnu předkládaná disertační práce přispěla k prohloubení znalosti organizace genomů důležitých rostlin a vytvořila předpoklady pro snadnější aplikaci molekulárních a genomických technik ve šlechtění nových odrůd zemědělských plodin.

Content

1 Introduction

The growth of the world's population places great demands on producing enough food, including the efficient use of arable land. Representatives of the *Poaceae* family, including the important cereals such as wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), rye (*Secale cereale*), and oat (*Avena sativa*) play a significant role in human nutrition. This plant family also includes grass species, which are the main component of pastures and meadows. In addition to providing feed for livestock, they have an important anti-erosion role, help to retain water in the landscape, and perform other functions in the ecosystem. Many types of grasses are grown in decorative gardens, in sports fields, and have many more uses. This Thesis focuses on two groups of grasses, cereals, and fodder crops/amenity grasses.

Growing demands on the performance of new varieties of agricultural crops, including the resistance to biotic and abiotic stresses, are compromised by limited genetic variation in current cultivars, resulting in part from the loss of potentially important genes during the domestication and breeding. In contrast, wild relatives of crops, which are characterized by larger genetic variation, may serve as an important source of genes and their alleles for breeding. However, the use of this variability is hampered by limited knowledge of their nuclear genomes. Obtaining the missing information is thus necessary for a wider use in the so-called introgressive breeding, which has the potential to contribute to breeding of new varieties that will have higher and stable yields under environmentally sustainable conditions.

2 Aims of the Thesis

With the aim to improve the knowledge of genome organization in selected representatives of *Poaceae* and develop molecular resources to support the breeding of crops, this Thesis focuses on three activities:

I Comparative analyses of repetitive DNA elements and identification of centromeric element in representatives of fescues and ryegrasses

The first aim of the Thesis was to characterize repetitive DNA elements in the nuclear genomes in ten representatives of *Festuca* and *Lolium* and provide more insights into DNA sequences in their centromeres using molecular and cytogenetic approaches.

II Genome organization and molecular chromosome structure of *Agropyron cristatum*

The second aim was to characterize repetitive and genic DNA sequences in *Agroypron cristatum*, a wild relative of wheat (*Triticum aestivum*), develop chromosome- and chromosome-arm specific microsatellite (SSR) markers from *A. cristatum*, and perform a comparative genome analysis with the orthologous subgenomes of wheat.

III Comparison of chromosome genomic approaches

The third aim was to review the past and current methods for genome complexity reduction with a focus on flow cytometric chromosome sorting and the use of this method in plant molecular biology and genomics.

3 Literature overview

3.1 Family *Poaceae* and agricultural crops

Poaceae is a large family of flowering monocotyledonous plants. These wind-pollinated species with the widespread occurrence of polyploidy are very successful in colonizing various habitats and high species richness (Linder et al., 2018). It includes cereals, grasses, and bamboos, therefore it is the most economically important plant family providing either staple food for the human population and feed for animals.

Representatives of the family *Poaceae* differ in chromosome number, ploidy level, and genome size. Whole-genome duplication and diploidization are typical processes accompanying genome evolution and speciation of eukaryotes. These events caused an increase in genome size also in polyploid wheat (Akhunov et al., 2013). Paleopolyploidy in plant genomes emphasized the role of whole-genome duplication in the evolutionary success of flowering plants (Paterson et al., 2003; Soltis et al., 2008; Van De Peer et al., 2009).

The well-known representative of cereals and agricultural crops, bread wheat (*Triticum aestivum*, 2n = 6x = 42) is the most produced cereal because it is rich in protein, carbohydrates, and minerals (D'Appolonia and Rayas-Duarte, 1994; Hussain et al., 2010). China is the world's leading wheat producer, ahead of India, the Russian Federation, and the United States of America (Figure 1, http://www.fao.org/faostat/en/#rankings/countries_by_commodity). Since the eighties, the wheat area and yields in the European Union have remained almost stable. While the human population increases, cultivated areas cannot be increased due to gradual loss of arable land and climatic changes. That is the motive why it is needed to employ new genomic approaches to improve wheat to secure the production needed by future generations. Classical breeding approaches are unable to fast produce cultivars with improved yield (http://www.wheatgenome.org).

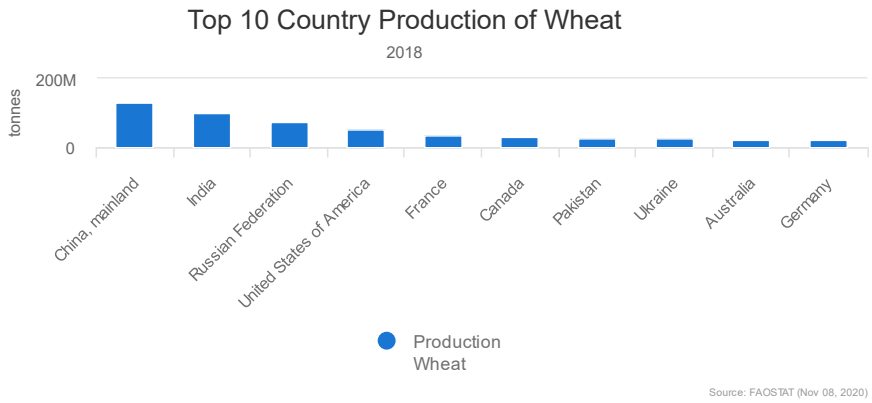


Figure 1: Comparison of the biggest wheat country production (http://www.fao.org/faostat/en/#rankings/countries_by_commodity).

The size of the bread wheat genome is 17 Gbp (Bennett and Smith, 1976). Its large and complex genome originated by two hybridization and allopolyploidization events. *Triticum urartu* (A subgenome of wheat) hybridized with an until now unknown relative *Aegilops speltoides* (B subgenome of wheat) to give rise to tetraploid *Triticum turgidum* (genome formula AABB) that subsequently hybridized with *Aegilops tauschii* (D subgenome of wheat) to give rise to hexaploid *Triticum aestivum* (genome formula AABBDD; Figure 2; Feldman et al., 1997; International Wheat Genome Sequencing Consortium et al., 2014). Each hybridization event was followed by chromosome doubling, thus producing a series of allopolyploid genomes. Each diploid donor has 7 pairs of chromosomes, which resulted in 21 pairs of chromosomes in common wheat.

The reference wheat cultivar “Chinese Spring”, was used for sequencing by the IWGSC. It was found out that the polyploidization events resulted in only mild genome changes and in wheat there is a high abundance of intrachromosomal gene duplications, which could be a mechanism for the global success of wheat (International Wheat Genome Sequencing Consortium et al., 2014). During the time, many translocations in wheat occurred, for example, in wheat chromosomes 4A, 5A, 7B (Devos et al., 1995; Ma et al., 2013) and pericentric inversions in chromosomes 2B, 4A, 4B, 5A (Conley et al., 2004; Linkiewicz et al., 2004; Miftahudin et al., 2004), 3B and 6B (Qi et al., 2006).

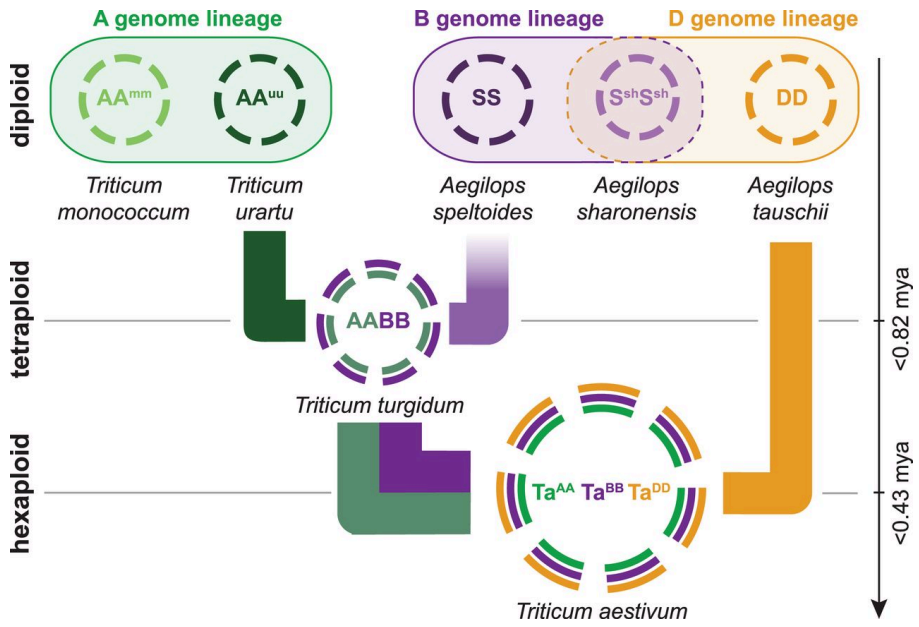


Figure 2: Schematic model of bread wheat origin (International Wheat Genome Sequencing Consortium et al., 2014).

Another important crop, rye (*Secale cereale*), is diploid (2n = 2x = 14), an allogamous, close relative of wheat with almost 8 Gbp genome size (Doležel et al., 2003). Rye contains non-coding repetitive DNA represented 84–92% of its genome (Bartoš et al., 2008; Flavell et al., 1974). It is well tolerant to frost, drought, and marginal soil fertility (Martis et al., 2013) and uses as resistance to pathogens and pests (Crespo-Herrera et al., 2017). The importance of rye is in the food and distilling industry (Balcerek et al., 2016; Johansson et al., 2018). Diploid barley genome (*Hordeum vulgare*, 2n = 2x = 14) has similar DNA repeat content as wheat. It has over 80% of non-coding repetitive DNA (Schulte et al., 2009; Wicker et al., 2009) with a genome size of more than 5 Gbp. Barley is important for its gene content such as genes for β-glucan biosynthesis (Burton et al., 2008) or powdery mildew resistance (Lyngkjaer et al., 2000). The value of barley lies in its use as fodder for animals, distilling, and brewing industry (Balcerek et al., 2016; Mayer et al., 2009).

3.1.1 Wild relatives of wheat

Domesticated cereals gradually lose functions of several genes, while genetic diversity in populations of wild relatives of crops still exists. Thus, the genes for important traits can be transferred from these populations. They are hidden in the tertiary gene pool where belongs, for instance, *Agropyron* (Dewey, 1984) or *Aegilops* genus (Molnár et al., 2016).

Agropyron species as a wild relative of bread wheat and a closest relative to barley is a source of biotic and abiotic stress resistance genes, for example, wheat streak mosaic viruses or barley yellow dwarf virus (Sharma et al., 1984), wheat rusts, powdery mildew (Said et al., 2018; Wu et al., 2006), drought (Dewey, 1984) and salinity tolerance (McGuire and Dvořák, 1981), or cold tolerance (Limin and Fowler, 1990). Addition and translocation lines for chromosome and chromosomal arms of *Agropyron* were developed (Chen et al., 1989; Han et al., 2014; Luan et al., 2010; Ochoa et al., 2015; Song et al., 2013; Wu et al., 2006).

Aegilops speltoides Tausch. (2n = 2x = 14), the closest relative to the wheat B genome (Dvořák et al., 1998), carries genes for tolerance against diseases such as powdery mildew, stem rust, leaf rust, genes responsible for grain hardness or heat and manganese toxicity tolerance (Kilian et al., 2011; Schneider et al., 2008). *Ae. cylindrica* Host. (2n = 4x = 28) and *Ae. triuncialis* L. (2n = 4x = 28) are a source of genes for tolerance to cold, salinity or drought, resistance to powdery mildew, rusts, barley yellow dwarf virus (Colmer et al., 2006; Friebe et al., 1996; Kilian et al., 2011; Schneider et al., 2008) and high grain content of micronutrient (Rawat et al., 2009). Many wheat-*Aegilops* additional lines were reported (Friebe et al., 1992, 1999; Wang et al., 2011). *Aegilops tauschii* genome showed much bigger polymorphic variation than its younger wheat D subgenome (Dvořák et al., 1998). That was the motivation to improve wheat by its ancestor genes (Chantret et al., 2005). To transfer *Aegilops tauschii* genetic material into bread wheat was constructed synthetic hexaploid wheat that was generated by artificial hybridization of *Triticum turgidum* with *Aegilops tauschii* (Mujeeb-Kazi et al., 1996). These synthetic lines carried genes for strong grain and stripe rust resistance (Yang et al., 2009).

3.1.1.1 Introgressions into wheat

Limited knowledge about the molecular organization of genomes of wild relatives of wheat, which could be an important source of genes, is a big obstacle in the introgression breeding. Since domestication and breeding crop have a narrow genetic base, alien introgression known as introgressive hybridization with wild relatives is used for spreading crop genetic base. Introgression is a transfer of genetic material from one species to another species, typically followed by hybridization and backcrossing to the parental species (Figure 3, Soltis, 2013). Cytogenetic approaches are used to

identify alien introgression lines, whereas marker systems have been used for alien introgression mapping (Olson et al., 2013). Many attempts to transfer alien chromatin into wheat were made with success (Friebe et al., 1996; Jiang et al., 1993; Kilian et al., 2011; Schneider et al., 2008), albeit the methods for development introgression lines and their selection and characterization are very time-consuming, laborious, low-throughput and had its own limitations such as low resolution (Lukaszewski et al., 2005). The major steps are detection, assignment, description, and positional assignment of genes in the genome. These are needed for the screening of agronomic traits and improving breeding programs (International Wheat Genome Sequencing Consortium et al., 2014).

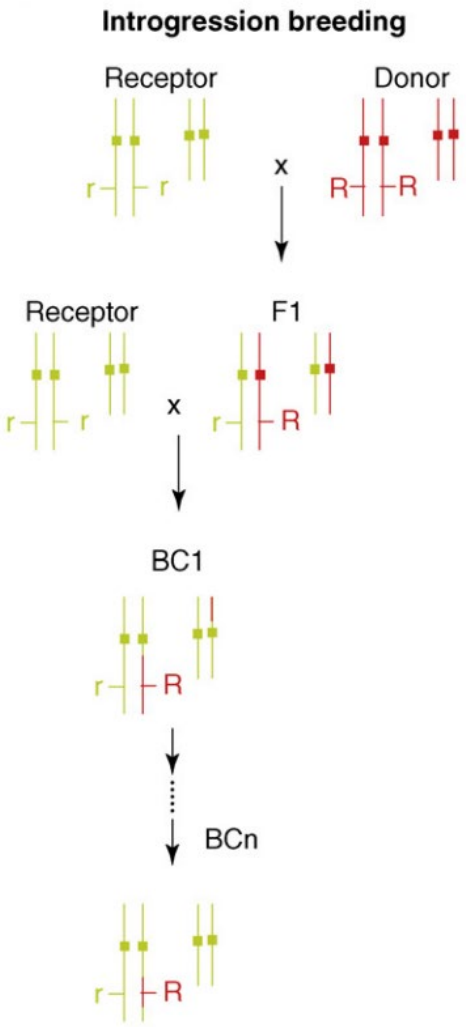


Figure 3: Introgression breeding (Jacobsen and Schouten, 2007).

Standard and very simple for breeders is control of introgression breeding with markers. In many crop species were identified large numbers of genetic markers.

Molnár et al. (2016) studied rearrangements between *Aegilops* and wheat using COS markers with known positions on wheat subgenomes. These markers could simplify work with alien gene introgressions. For accurate identification of the origin of introgressed segments, Aflitos et al. (2015) developed the "Introgression Browser" based on combining genotypic and phylogenetic data with the nucleotide or SNP accuracy. Abrouk et al. (2017) developed another *in silico* approach, rearrangement identification and characterization (RICh). Intending to better identification of the sequences to generate markers in the introgression segment, they combined shotgun sequences data of chromosome with introgression with the Genome Zipper (Mayer et al., 2011). It helps to validate the evolutionary rearrangements in modern wheat chromosomes, characterize introgression position, gene content, and gene order.

3.2 Grasses

Meadows, pastures, and lawns are one of the largest ecosystems worldwide. They are sustainable in various geographic and climatic variation (Reheul et al., 2010) and are important from the ecological, economical, and aesthetical points of view (Kopecký and Studer, 2014). Permanent lawns are suitable for biomass production (Searchinger et al., 2008). *Lolium perenne* L. has a high water-soluble carbohydrate content as a polymeric fructan that is easy to extract, split to the separate fructose units, and fermented to bio-ethanol (Farrar et al., 2012). Grasses serve as a readily available source of feed for ruminants that is an economic-friendly way for animals important in agriculture, such as cattle and other farm animals (Reheul et al., 2010). However, in the last decades, meadows are modified by farmers and declining alarmingly due to frequent conversion to arable land (Obermeier et al., 2020). Since the grasses are an important source of feed for farm animals, the need for understanding their genomes to support breeding of improved cultivars is essential. The aesthetical function of grasses is important in decorative areas in parks, around houses, as a recultivation, or in sports ground as a golf course or football field (Kopecký and Studer, 2014).

Loliinae subtribe is one of the largest of the *Poeae* tribe. Its genome size range from 2.6 to 11.8 Gbp (Zwyrtková et al., 2020). In their evolution, hybridization and polyploidization played an important role. Therefore, about 70% of the species are polyploid. The evolution of grasses was accompanied by frequent and repeated genome size gain and loss (Soreng et al., 2015). Grasses in this study range between diploid

and hexaploid level. The most important species in the mentioned ecosystems are *Festuca rubra* L., *Festuca pratensis* Huds., and *Festuca arundinacea* Schreb. On the other hand, the most occurring species in the production of temporary lawns all over the world is *Lolium perenne* L. and *Lolium multiflorum* Lam. (Kopecký and Studer, 2014). The basic chromosome number in fescues (*Festuca* L.) and ryegrasses (*Lolium* L.) is n = 7. *Festuca* is ranging between diploid and decaploid species, whereas *Lolium* comprises mainly diploid species. It was found out that the most used in agriculture is diploid, few autotetraploid, hexaploid, and octaploid *Festuca* species, and the best feed quality has tetraploid *Lolium* species (Loureiro et al., 2007). Production of Festulolium as a hybrid of *Festuca* and *Lolium* was powered by breeders and their efforts to have all properties in one species (Kopecký et al., 2008; Thomas et al., 2003). The kinship of the individual species is shown on the Figure 4.

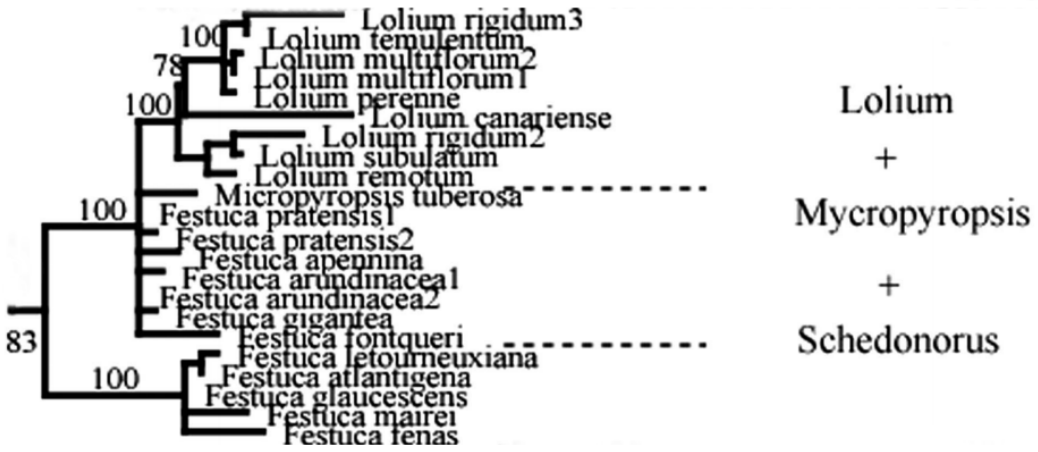


Figure 4: Part of phylogenetic tree of *Loliinae* derived by ITS sequences. Adapted from Catalán et al. (2004).

Another group that belongs to grasses is a complex of wheatgrass. In the Thesis, it had been worked with crested wheatgrass *Agropyron cristatum* belongs to the *Hordeinae* subtribe of the *Triticeae* tribe. Thus, from the important cereals, the close relative is *Hordeum vulgare* which is a phylogenetic member of the same subtribe. *Triticeae* tribe consists of the most important cereals and forage crops, i.e., wheat, barley, rye (Hsiao et al., 1995).

3.2.1 *Festuca* spp.

Each species grows in different specific habitat including wetlands, dry areas, cold and mild regions of the southern hemisphere as well. They are well adapted to extreme conditions in mountains, arctic and subantarctic regions (Inda et al., 2008), also for stages as alpine meadows, pastures, and forests (Šmarda et al., 2008). Individual species of this genus are different in morphology and ploidy. There is a high number of polyploids. Most polyploids belonging to the subgenus *Schedonorus* (Hand et al., 2010) are mostly allopolyploid (Loureiro et al., 2007). *Schedonorus* subgenus is a complex of species with various ploidy levels, which are a consequence of hybridization. To this subgenus belongs, for example, diploid *Festuca pratensis* Huds. (2n = 2x = 14, Figure 5), tetraploid *Festuca arundinacea* var. *glaucescens* Boiss. (2n = 4x = 28), tetraploid *Festuca mairei* (2n = 4x = 28), hexaploid *Festuca arundinacea* Schreb. (2n = 6x = 42), hexaploid *Festuca gigantea* (L.) Vill. (2n = 6x = 42) and others (Hand et al., 2010).

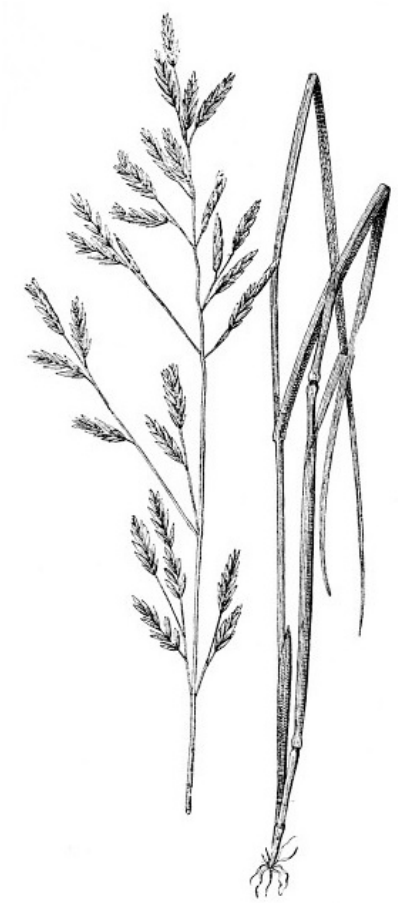


Figure 5: *Festuca pratensis* Huds. (Polívka, 1902).

3.2.2 *Lolium* spp.

Lolium L. is naturally diploid species (2n = 2x = 14) with a genome size 2 Gbp/1C (Hutchinson et al., 1979). They naturally cross each other, which leads to tetraploid representatives (Humphreys et al., 2010). Into main feed grasses is included *Lolium perenne* L. (2n = 4x = 28, Figure 6) and *Lolium multiflorum* L. (2n = 4x = 28). Their leaves and stalks are the most digestible for animals from all grass species (Frame, 1991). Besides, *Lolium perenne* is well stress-resistant (Humphreys et al., 2010).

Figure 6: *Lolium perenne* L. (Polívka, 1902).

3.2.3 *Agropyron cristatum*

Agropyron cristatum (crested wheatgrass, Figure 7) belongs to the family *Poaceae*. It is a perennial low maintenance grass, which can be grown permanently and needs a little additional treatment or special care, such as the application of fertilizers. It has been used for pasture, rangeland, and hayland. Therefore, it has economic importance as a forage, for example, for livestock (Han et al., 2019). It is grown worldwide, but mainly cultivated

in arid regions of the United States (Asay et al., 2003; Asay and Jensen, 1996). Wheatgrass has not been domesticated or bred, exhibits large genetic variation, and thus appears a potential source of new genes for wheat improvement (Han et al., 2014). Introgression of its genes to wheat can improve yields, increase resistance to diseases, and improve tolerance to drought (Said et al., 2018).

The most known are diploid, tetraploid, and hexaploid cultivars called Douglas, Ephraim, Parkway, Roadcrest, and Ruff. These cultivars differ in many aspects - any of them are more drought tolerant than others, another can accept more centimeters of annual precipitation. Due to this, each has its own special usage, for example, for turf or roadside applications. They can differ in morphology and productivity. Important for the Thesis is the diploid cultivar Parkway, which was released in 1969 by the Canada Department of Agriculture (Robins and Jensen, 2020). *Agropyron cristatum* has a basic P genome with 7 chromosomes. The genome size is relatively large. Diploid cultivar Parkway is represented by the estimated size of 1C ~ 6.3 Gbp (Said et al., 2018).

Figure 7: *Agropyron cristatum* L. (http://spuds.agron.ksu.edu/ksgrasskey/images/Agropyroncristatum.htm).

3.3 The organization and structure of plant genomes

Nuclear genome is represented by one copy of nuclear DNA. The amount of nuclear DNA of an organism is expressed in picograms or base pairs with the following relation: 1 pg = 0.978 Mbp (Doležel et al., 2003). For DNA amount in the haploid nucleus is known term C-value (Bennett and Smith, 1976; Swift, 1950). However, a higher C-value does not mean more complicated organisms. For that, there was a C-value paradox (Thomas, 1971), a phenomenon, which was then overcome with a C-value enigma. This newer phenomenon saying about uncertainties about genome size differences and DNA content variations (Gregory, 2001).

Genomes of eukaryotic organisms are organized into different numbers of chromosomes. The large genome of *Triticum aestivum* with its ~17 Gbp is about 34 times bigger than the genome of, for example, *Oryza sativa*, which has only ~0.5 Gbp. The main differences are in the contribution of repetitive sequences (Heslop-Harrison et al., 1997). Grass genomes have about 1C ~ 2–12 Gbp. The monoploid genome size of fescues and ryegrasses was determined using flow cytometry and accounted for 1Cx = 2.62–3.25 pg (Kopecký et al., 2010). The combination of polyploidization with various repetitive DNA sequences is an obstacle in genomic analyses. Genome structure and rearrangements in grass genomes were studied using next-generation sequencing. The genome of *Lolium perenne* (Studer et al., 2012) and chromosome 4F of *Festuca pratensis* Fure (Kopecký et al., 2013) has been sequenced to create virtual gene order when compared with smaller genome carriers such as *Oryza sativa*, *Sorghum bicolor*, *Brachypodium distachyon* or *Hordeum vulgare* (Kopecký et al., 2013; Studer et al., 2012). In chromosome 6P of *Agropyron cristatum,* genetic rearrangements were revealed by SSR and STS molecular markers and then compared with wheat subgenomes (Han et al., 2014).

The main morphological structures such as centromeres and telomeres have their predictable position. The centromeric part is located to the primary constriction, nucleolar organizer region is located to the secondary constriction, which is not always present. Using centromeric position was determined four types of chromosome morphology – metacentric, submetacentric, acrocentric, and telocentric chromosome. Telomeres are positioned at the ends of chromosomes. Plant genome comprises coding and non-coding parts. Centromeric, pericentromeric, telomeric, and subtelomeric parts of chromosomes are typical by repetitive DNA sequences. On the contrary, distal parts of chromosomes

are carriers of the majority of genes. Coding parts, genes, are obviously connected with its regulatory components. The average size of gene length with its regulation part covers 1–5 kbp (Kellogg and Bennetzen, 2004). Most genes are clustered in gene islands, regions in the genome with higher gene density, and, on the contrary, gene empty-spaces, where are located mainly non-coding sequences (Sandhu and Gill, 2002). There were described either few gene-rich islands (Erayman et al., 2004) or numerous gene-weaker islands (Brooks et al., 2002). In wheat, the definition of gene islands was provided by Choulet et al. (2010) as clusters of genes with a minimum of two genes separated from 50% by less than 43 kbp long intergenic distances. The highest gene density in gene-rich regions was observed in distal parts of chromosomes (Erayman et al., 2004). The rest genes are organized as isolated genes (Rustenholz et al., 2011).

3.3.1 Repetitive DNA sequences

The main parts of the non-coding genome regions are repetitive sequences. They are either dispersed, or tandemly organized with a high abundance in genomes. Formerly, Cot-based fractionation (Peterson et al., 2002) and methyl filtration (Palmer et al., 2003) were used to reduce genome complexity by removing repetitive parts and simplifying genome analyses.

The first group is represented mainly with autonomous or semiautonomous mobile elements, which are located between different repeat families or genes and have replication and moving ability. Transposable elements are the most abundant repetitive DNA sequences. In grass genomes, the elements take up more than half of their sizes (Heslop-Harrison and Schwarzacher, 2011). The most abundant in genomes are LTR retrotransposons and DNA transposons (Wicker and Keller, 2007). Schnable et al. (2009) described that hundreds of transposable elements, which were important in centromeric positioning on some chromosomes, make up ~85 % of the maize genome with the Copia elements in gene-rich regions, while Gypsy elements in gene-poor regions. Peterson-Burch et al. (2004) showed, that many retrotransposons of *Arabidopsis thaliana* localizing preferentially to pericentromeric regions, whereas retrotransposons of *Triticeae* are more dispersed along all chromosomes (Figure 8).

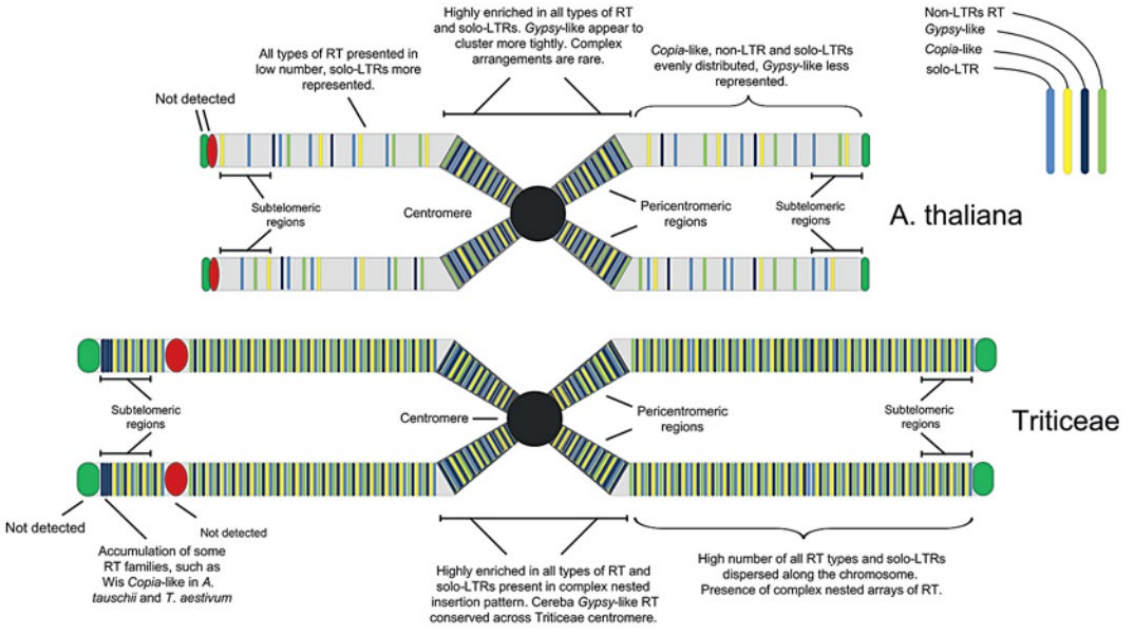


Figure 8: Distribution of retrotransposons in *Arabidopsis thaliana* and *Triticeae* tribe. Abbreviations: LTR (long terminal repeat), RT (retrotransposon) (Bento et al., 2013).

If retrotransposons are flanked with long terminal repeats (LTRs), they are classified in Copia or Gypsy group based on the organization of their coding composition make up of capsid protein, protease, integrase, reverse transcriptase, and RNase H genes. Terminal-repeat retrotransposons in miniature (TRIMs) and large retrotransposon derivatives (LARDs) are without coding sequences, therefore they are a nonautonomous group of retrotransposons. Nevertheless, retrotransposons cannot be always flanked with LTRs. In that case, they are classified as non-LTR retrotransposons diverged in short interspersed nuclear elements (SINEs) and long interspersed nuclear elements (LINEs). The first one contains Pol III promoter and a variable region, and the second one capsid protein, integrase, endonuclease, reverse transcriptase, and a variable region (Figure 9, Bento et al., 2013).

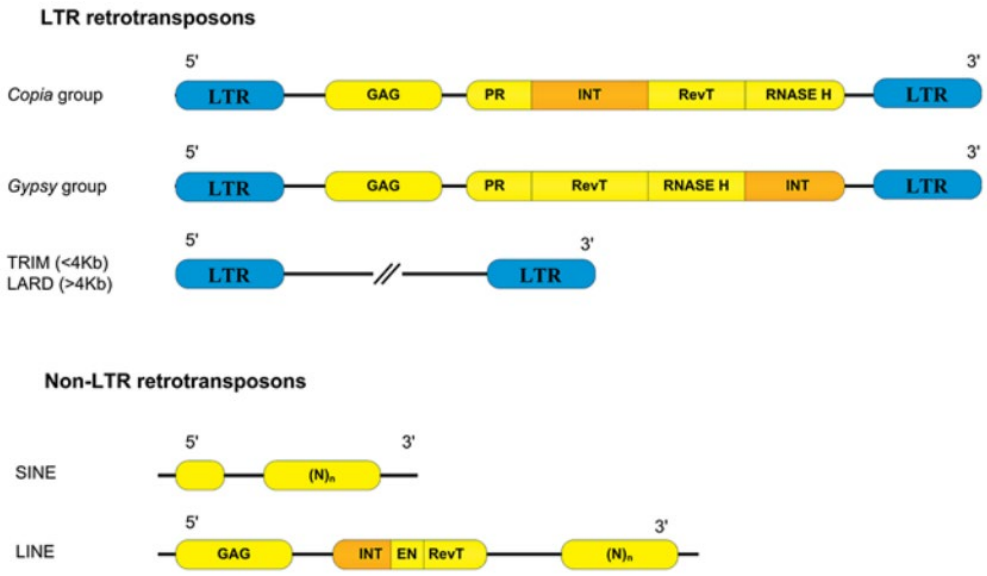


Figure 9: Sequence organisation of retrotransposons composed genes for long terminal repeat (LTR), capsid protein (GAG), protease (PR), integrase (INT), reverse transcriptase (RevT), RNase H (RNASE H), endonuclease (EN) and a variable regions $(N)_n$ (Bento et al., 2013).

The second group are tandemly organized repeats. This arrangement was discovered in telomeric sequences or at some eukaryotic genes such as rRNA genes (Lapitan, 1992) and are compounded from repetitive sequences located closely one by one (Flavell, 1980). Based on the lengths of these units are known microsatellites (1–20 bp), minisatellites (20–100 bp), and satellite DNA (Charlesworth et al., 1994). The last type was previously described as 100–400 bp long units, but even longer units are known (Hřibová et al., 2010; Kojima et al., 2002; Macas et al., 2009). They are obviously species-specific (Heslop-Harrison and Schwarzacher, 2011). Tandemly organized repeats are a fast-evolving part of eukaryotic genomes (Alexandrov et al., 2001; Vershinin et al., 1996) and are often chromosome position-specific. Typical tandemly organized repeat in centromeres was discovered in, for example, *Arabidopsis thaliana* (Heslop-Harrison et al., 1999; Kumekawa et al., 2001), *Oryza sativa* (Nonomura and Kurata, 2001) or *Zea mays* (Ananiev et al., 1998), whereas in other plant species studies not (Hřibová et al., 2010; Kim et al., 2002; Macas et al., 2007). In *Festuca* and *Lolium* species was not observed typical centromeric DNA sequence, nevertheless a specific type of Ty3/Gypsy family of LTR retrotransposons with chromodomain (Kiseleva et al., 2014; Neumann et al., 2011; Novák et al., 2014; Zwyrtková et al., 2020). The tandemly organized repeats

are of even megabase-long lengths, that are connected to sequencing problems (Alkhimova et al., 2004). Use of these sequences is as molecular and cytogenetic markers. The most abundant in *Agropyron cristatum* is Ty3/gypsy element Athila, which is of higher proportion than total of Ty1/copia retroelements (Table 1, Said et al., 2018).

Table 1: Repetitive DNA sequences observed in *Agropyron cristatum* (Said et al., 2018).

Repeat		Lineage/class	Alternative names	Proportion in the analyzed data [%]
LTR retroelements	*Ty1/copia*	Maximus-SIRE		4.42
		Angela		13.51
		TAR	Tont	0.56
		Tork	Tnt	0.01
		Ale	Hopscotch	0.05
		Bianca		0.02
		Total *Ty1/copia*		18.57
	Ty3/gypsy	Athila		22.96
		Chromovirideae		7.16
		Ogre-Tat		7.98
		Total *Ty3/gypsy*		38.10
Other	LINE			0.15
	DNA transposons			3.28
	Tandem repeats			1.62
	rRNA genes			0.21
Unclassified LTR elements				3.94
Unclassified repeats				6.71
Non-annotated sequences				8.36

3.3.2 Centromere regions

Centromeric regions are essential parts of chromosomes serving as an important domain in cell divisions, with kinetochore complex for sister chromatid cohesion, as a site for spindle microtubule attachment. It is located to primary constriction on monocentric chromosomes. Part of the whole genetic information lying in the centromeric regions may play a role in those processes where centromeres are involved (Presting et al., 1998). Ma et al. (2007) reviewed that size and sequence content varied between plant centromeres.

It is known that centromeric regions are often composed of repetitive sequences, and several of them have been physically mapped by fluorescence *in situ* hybridization. In *Arabidopsis*, ~50 kbp long arrays contain an average length of 180 bp or 360 bp

tandemly organized repeat unit (Martinez-Zapater et al., 1986). Maluszynska and Heslop-Harrison (1991) localized these sequences to paracentric regions of all chromosomes. In cereal centromeres, there are also repeat families occurred. In wheat, barley, rye, rice, or maize was localized the sequence family previously derived from *Brachypodium* containing similar region as CENP-B box, a human centromeric satellite DNA (Aragón-Alcaide et al., 1996; Willard, 1990).

The story of *the Cereba* element started, when Jiang et al. (1996) isolated the *pSau3A9* sequence from the genomic DNA of sorghum. It localized to the cereals centromeric regions (Chen et al., 1998; Presting et al., 1998). After a series of experiments, high homology was shown between *pSau3A9* and more than 800 bp long sequence in the barley genome. The sequence was searched against GenBank and found similarity with the integrase region of the retrotransposon *del* of *Lilium henryi,* related to retrotransposon Ty3/gypsy of yeast and *Drosophila* (Smyth et al., 1989). A study of Presting et al. (1998) established the centromeric barley sequence as the Ty3/gypsy retrotransposon-like sequence and named it *Cereba* as an abbreviation of *ce*ntromeric *re*troelement of *ba*rley.

On the contrary, under-represented or even absent in centromeres are retroelements from Ty1/copia family. These are more distributed along chromosomes (Pearce et al., 1996; Presting et al., 1998). Localization of Ty1/copia retroelements differs in the plant sphere. In *Vicia faba* they are mostly in euchromatin (Pearce et al., 1996), what differed from *Allium cepa* where have been observed in heterochromatin regions (Pich and Schubert, 1998), quite similarly as in the animal sphere – *Drosophila* (Carmena and González, 1995; Kumar, 1996). Heslop-Harrison et al. (1997) described Ty1/copia retroelements unspecifically in centromeric regions of *Arabidopsis thaliana*.

3.4 Genome sequencing strategies

There are more sequencing strategies, clone-by-clone sequencing, and whole-genome shotgun sequencing. The advances in plant genomics were fastly driven by the involvement of next-generation sequencing leading to lower prices and high-throughput organization. The former strategy was clone-by-clone sequencing (Figure 10A). The key term “minimum tilling path” (MTP) is the minimal set of clones to be sequenced based on overlapped large-insert DNA clones. Because of the small genome and importance in a plant kingdom, a full sequence of *Arabidopsis*

thaliana with a genome size ~150 Mbp was obtained by this technique (The Arabidopsis Genome Initiative, 2000). The second was crop sequencing, rice genome with a little bit higher genome size ~400 Mbp (Matsumoto et al., 2005) and then maize with ~2.5 Gbp genome (Schnable et al., 2009). A more recent strategy, whole-genome shotgun sequencing (Figure 10B), was used to acquire a sequence of poplar (Tuskan et al., 2006), grapevine (Jaillon et al., 2007), or sorghum (Paterson et al., 2009). The core lies in the comparison of each sequence read with the rest of the sequences, and extension of the reading sequence by overlapping resulted in a longer sequence of the genome of interest. *De novo* sequencing requires short reads for assembly, whereas simpler re-sequencing works with already existing reference genome sequence on which the reads are mapped. Thus, the quality of re-sequencing genomes is based on the quality of the reference genome. Redundancy of reads was an obstacle to both sequencing strategies mainly because of the polyploidy genomes (Doležel et al., 2014). Reduction of template complexity was desired.

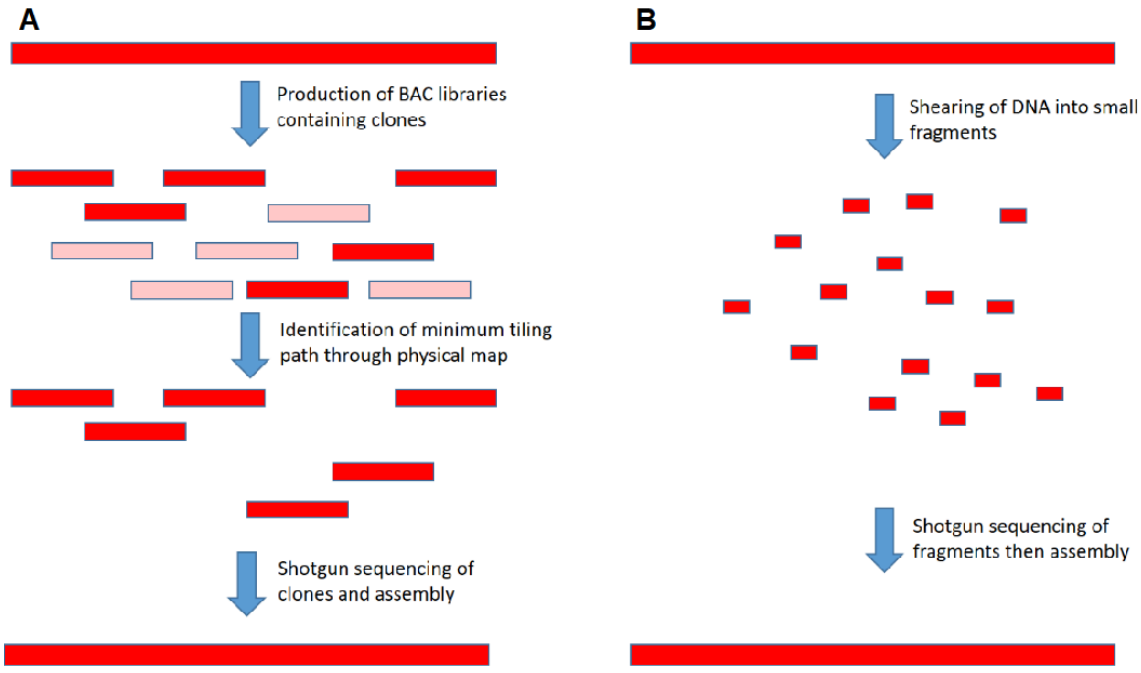


Figure 10: Schema of (A) clone-by-clone sequencing and (B) whole-genome shotgun sequencing. Adapted from Lowe (2018).

3.5 Genome complexity reduction

Important crops such as wheat, barley, oat, rye, and maize, have large genomes. The science community required to have tools for simplification of work with these big genomes. Until recently, sequencing these genomes was expensive and the analysis of whole genome sequences that are full of repetitive DNA was difficult. Another problem was the presence of homoeologous genomes in allopolyploids. There were many attempts to simplify genome complexity that led to more accurate analyses. It is needed to know the genome organization, to identify important genes and their function (Chen et al., 2019; Varshney et al., 2020). Sequencing of diploid progenitors of polyploid crops was a former way how to simplify gene cloning (Ling et al., 2018; Luo et al., 2017). High abundant repetitive DNA sequences of complex genomes were removed by Cot-based fractionation (Figure 11A, Peterson et al., 2002) and methyl filtration (Palmer et al., 2003). The exome capture method was successful in exons isolation from coding sequences and, therefore, reduced genome complexity by avoiding non-coding and surrounding DNA (Figure 11B, Hashmi et al., 2015). The problem of the mentioned methods was the impossibility of assembling longer data after sequencing, since lacking repeats. The beginning of flow cytometry with flow-sorting (Figure 11C, Doležel et al., 1992), which can split genomes into smaller parts - chromosomes, was a noteworthy event. Sorting genomes into individual chromosomes or chromosome arms and studying them happened a priceless part of everyday progress.

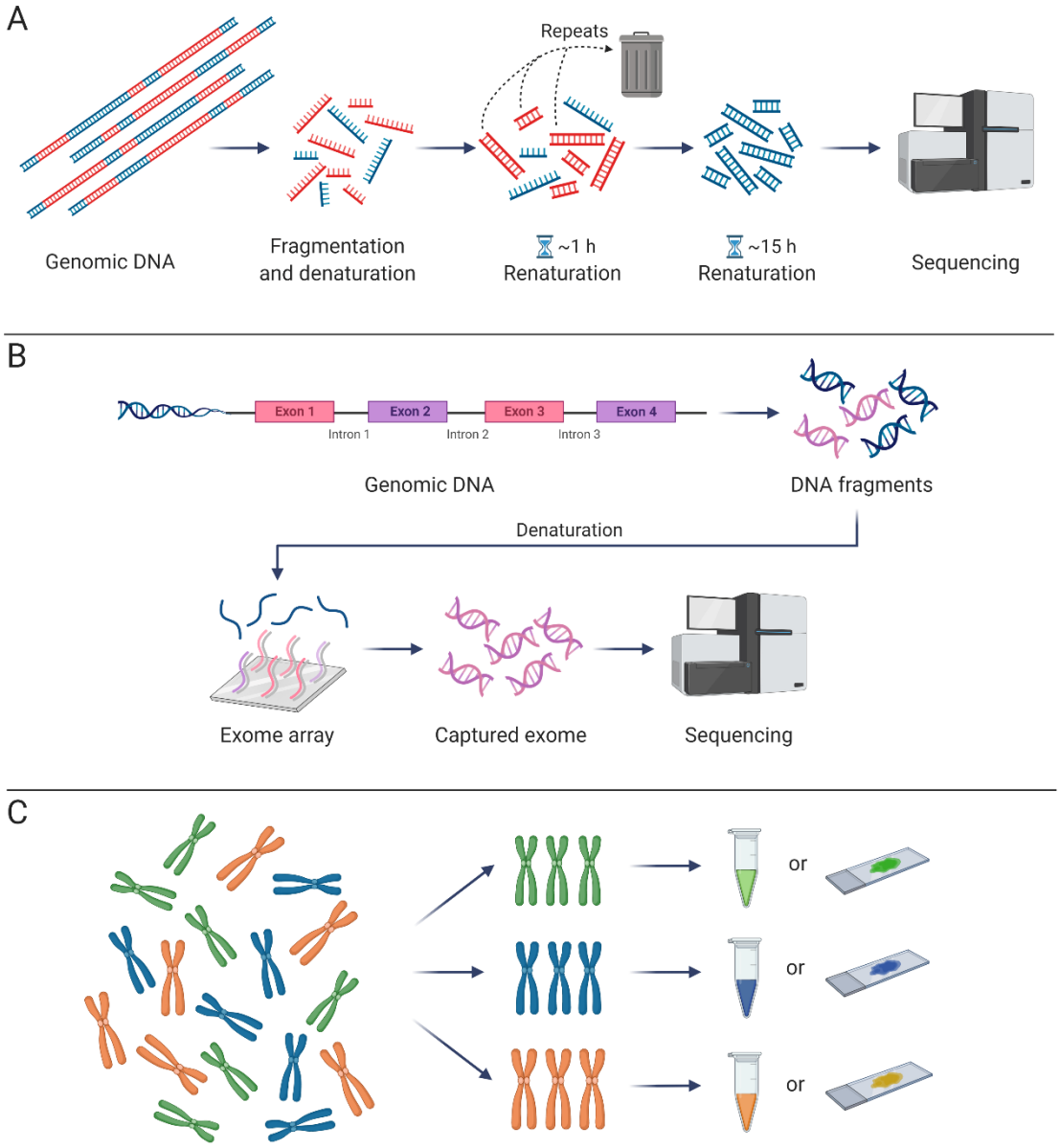


Figure 11: Genome complexity reduction methods: (A) Cot-based filtration, (B) Exome capture, (C) Flow-sorting (Zwyrtková et al., 2021).

3.5.1 Chromosome flow sorting

The requirement for isolation of chromosome-specific DNA is the mitotic metaphase stage of cells and the skill to collect selected chromosomes. A former method was micromanipulation where the chromosome of interest was picked-up out of others (Matsunaga et al., 1999; Schondelmaier et al., 1993). This method had its obstacles lying in low quality of isolated DNA and very low throughput (Hobza and Vyskot, 2007; Ma et al., 2010).

Doležel et al. (1992) developed an improved method for genome partitioning using flow-sorting of its individual chromosomes where each chromosome represents a smaller and complex genome part. Individual chromosomes are sorted based on the optical properties of one chromosome in comparison with others. Chromosomes in suspension are fluorochromatically stained and passed to a flow chamber with sheath fluid. Chamber structure leads the chromosomal suspension in a narrow stream. Chromosomes in a stream are encapsulated into single droplets and interact with laser beam. The pulses of scattered light and fluorescence are collected and converted to electrical pulses. Identification and sorting of a selected chromosome is based on different fluorescence intensity from others realizing between deflection plates and gathered in appropriate containers (Figure 12, Doležel et al., 2014). Sorted chromosomes can be purified and amplified. These chromosomes are intact and it is possible to use them for preparing high molecular weight (HMW) DNA (Šimková et al., 2003). The advantage of sorted chromosomes is avoiding assembly difficulties based on similarities of homoeologous chromosomes in polyploid species followed by, for example, targeting of selected chromosomes in introgression breeding approaches. This method has found its world-use in the greatest crop sequencing projects such as barley (Mayer et al., 2011), rye (Martis et al., 2013), chickpea (Varshney et al., 2013), wheat (International Wheat Genome Sequencing Consortium, 2018), or pea (Kreplak et al., 2019).

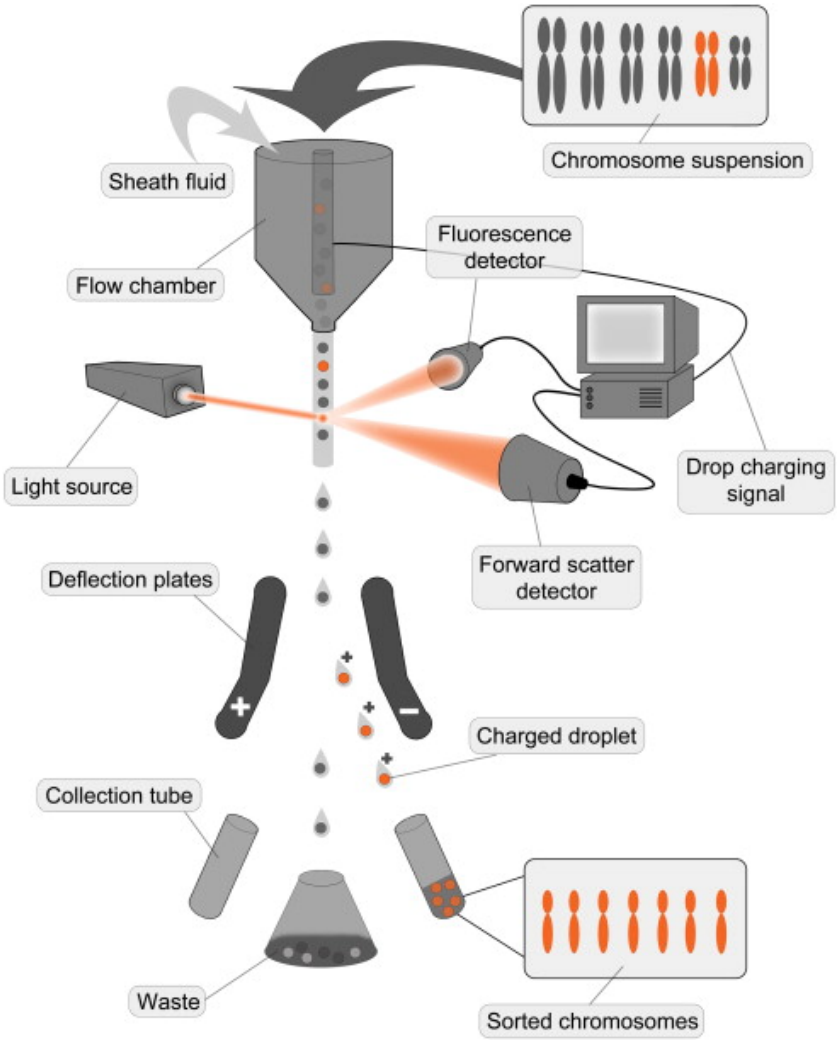


Figure 12: Flow-sorting mechanics (Doležel et al., 2014).

First sorting of plant chromosomes was reported with *Haplopappus gracilis* (de Laat and Blaas, 1984). Preparing of suspensions of intact chromosomes was an obstacle of the starting method. A high degree of mitotic metaphase synchrony and the possibility to obtain the intact chromosomes from rigid cell walls were the main issues (Doležel et al., 1994). Preparing samples from synchronized root tip meristems from young seedlings was the innovation. The former approach was based on removing cell walls by hydrolytic enzymes followed by lysing cells in a hypotonic buffer. The latter method was based on formaldehyde-fixed root tips followed by mechanical homogenization. This approach has been developed in *Vicia faba* (Doležel et al., 1992).

The problems with size-similarity of chromosomes in the sample or similar fluorescence of a dye bounded to the DNA were tried to overcome by measurements of fluorescence of dye binding to AT-rich regions and GC-rich regions. However, it was

unsuccessful (Lee et al., 2000, 1997; Lucretti and Doležel, 1997). Different approaches included the use of additional lines, chromosome translocation and deletion (Doležel and Lucretti, 1995), fluorescence *in situ* hybridization in suspension (FISHIS) with fluorescent-labeled microsatellites (Giorgi et al., 2013) or sorting of single chromosomes (Cápal et al., 2015). These days, chromosome flow-sorting has wide use in many important applications (Figure 13).

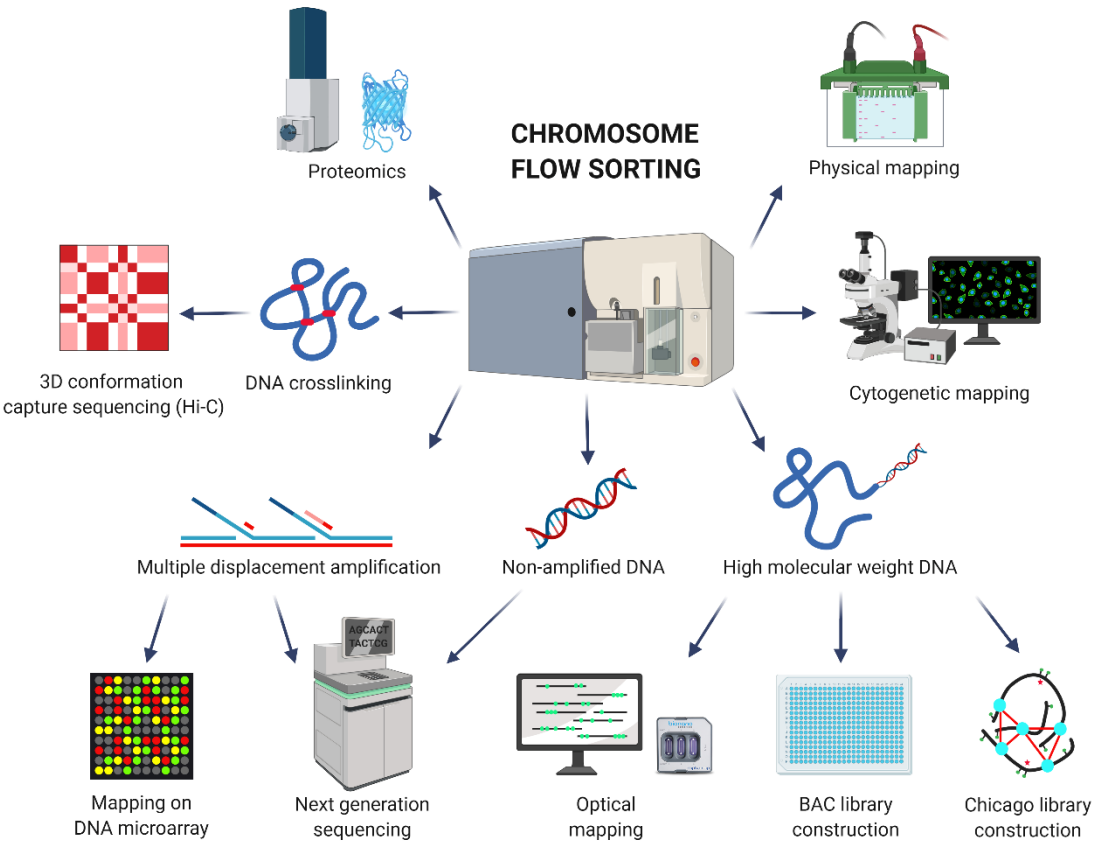


Figure 13: Applications working with flow-sorted chromosomes (Zwyrtková et al., 2021).

3.6 Use of flow-sorted chromosomes

Many uses of flow-sorted chromosomes were developed for genomics, molecular biology, or proteomics purposes (Figure 14; Zwyrtková et al., 2021). It has been already used in large sequencing projects such as hexaploid bread wheat sequencing (International Wheat Genome Sequencing Consortium et al., 2018, 2014), pea (Kreplak et al., 2019), rye (Martis et al., 2013) or barley (Mayer et al., 2011) sequencing.

Flow-sorted chromosomes are greatly used in the preparation of chromosome-specific DNA libraries. This approach is beneficial in multiplying chromosomal DNA through bacterial cultures and serve as a permanent source of chromosomal DNA. DNA libraries are divided into short-insert (<100 kbp) and long-insert (>100 kbp) types. The first short-insert library was generated by Wang et al. (1992) with flow-sorted chromosome 4A of *Triticum aestivum*. Telenius et al. (1992) added an amplification step and thus simplified the process. By this approach, the library of chromosome 2 of *Lycopersicon pennellii* (Arumuganathan et al., 1994) or the first chromosome-specific DNA library of the whole chromosomal set of *Vicia faba* (Macas et al., 1996), was performed. Higher amounts of high molecular weight DNA were required for long-insert libraries. Šimková et al. (2003) optimized the protocol for DNA isolation to use it with chromosomes isolated from synchronized root tips as described by Doležel et al. (1992). Šafář et al. (2004) developed a protocol for DNA libraries cloned in bacterial artificial chromosome (BAC) vectors originally for chromosome 3B of *Triticum aestivum* following by Janda et al. (2006, 2004) who performed a library of pooled 1D, 4D and 6D chromosomes as well as for short arm of chromosome 1B of *Triticum aestivum*. Until now, four wheat chromosomes and thirty-four chromosomal arms (International Wheat Genome Sequencing Consortium, 2018; Šafář et al., 2010) and one short arm of chromosome 1R of *Secale cereale* (Šimková et al., 2008a) were used for the construction of specific BAC libraries.

Flow-sorted chromosomes are also used in the methods for the development of different types of DNA markers, its physical mapping, chromosome sequencing, or gene cloning.

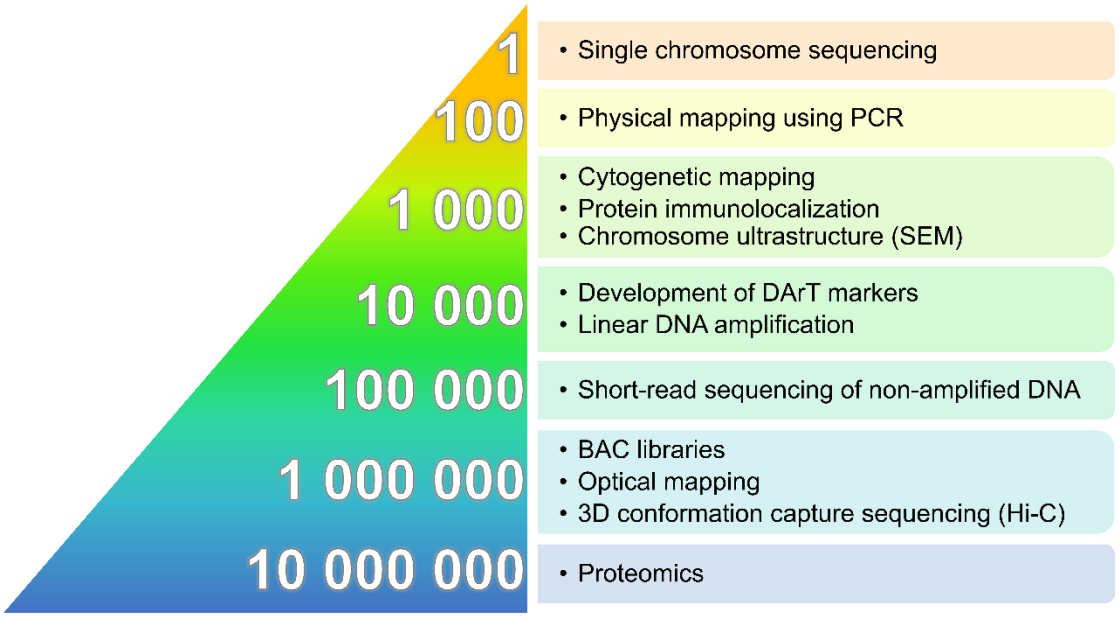


Figure 14: Chromosomal numbers needed for various approaches. Abbreviations: BAC (bacterial artificial chromosome), DArT (diversity arrays technology), PCR (polymerase chain reaction), SEM (scanning electron microscopy). (Zwyrtková et al., 2021).

3.6.1 Chromosome genomics

Dividing the genome into smaller parts, chromosomes, is an elegant solution to simplify the sequencing and establishment of assembly (Figure 15). This strategy is commonly used for barley, rye, wheat. Another reducing method, chromosome microdissection, had lower throughput than flow-sorting (Hobza and Vyskot, 2007). Nevertheless, chromosome genomics has many limiting factors and it has the potential to be improved all the time. To describe that, fluorescent *in situ* hybridization in suspension (FISHIS, Giorgi et al., 2013), a method served for labeling of repetitive DNA on chromosomes before flow cytometry analysis worked only for some species (Doležel et al., 2014). There were many attempts to isolate chromosomes. Stubblefield and Oro (1982) tried to separate chromosomes by gradient centrifugation, Dudin et al. (1988) by capturing on magnetic beads. However, flow-sorting was the most optimal one (Doležel et al., 1994). If normal sorting is unable to perform, another tactic is to sort single copies of chromosomes followed by whole-genome amplification of sorted DNA. The DNA from single copies, which was obtained after flow-sorting was used mainly in human genomics. Among the plants, sorting and amplification of wheat chromosome 3B

produced micrograms of DNA. The protocol allowed to be used for plant species suitable for flow-sorting. Single-chromosome sequencing was useful for the identification of missing genes in reference, to develop chromosome-specific markers, to verify contigs in pseudomolecules, or to validate assemblies (Cápal et al., 2015).

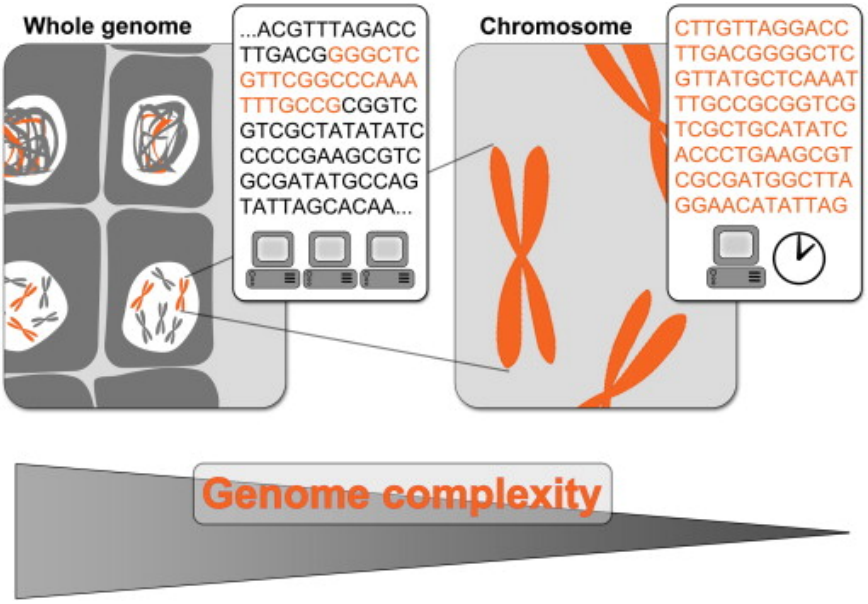


Figure 15: Schematic representation of advantages of chromosome genomics (Doležel et al., 2014).

3.6.2 Molecular markers and their development

Molecular markers as a tool for genetic variability analyses in the agricultural sphere are used for many years. Marker techniques and usable markers must be (i) reliable – to be very close to the targeted loci or even be inside, (ii) polymorphic – to be variable among genotypes, (iii) cheap, fast and simple – ideally starting with a low genetic material input. Molecular marker techniques can be categorized into three categories. Firstly, non-PCR-based markers that are based on hybridization, secondly, the large group of PCR-based markers, and lastly, sequence-based markers (Garrido-Cardenas et al., 2018). To these groups belong many types of markers – SSR (Taheri et al., 2018), SNP (Close et al., 2009), COS (Said et al., 2019), RFLP (Arumuganathan et al., 1994), RAPD (Williams et al., 1990), DArT (Kopecký et al., 2009), IT (Wang et al., 2017), RJM (Mazaheri et al., 2014) and others.

There are other ways for marker development. Chromosomal DNA can be cloned into a vector and selected clones are sequenced. The clones could be selected randomly

or purposefully that increasing the yield. Additionally, chromosomal DNA could be immediately sequenced and then processed *in silico* (Figure 16).

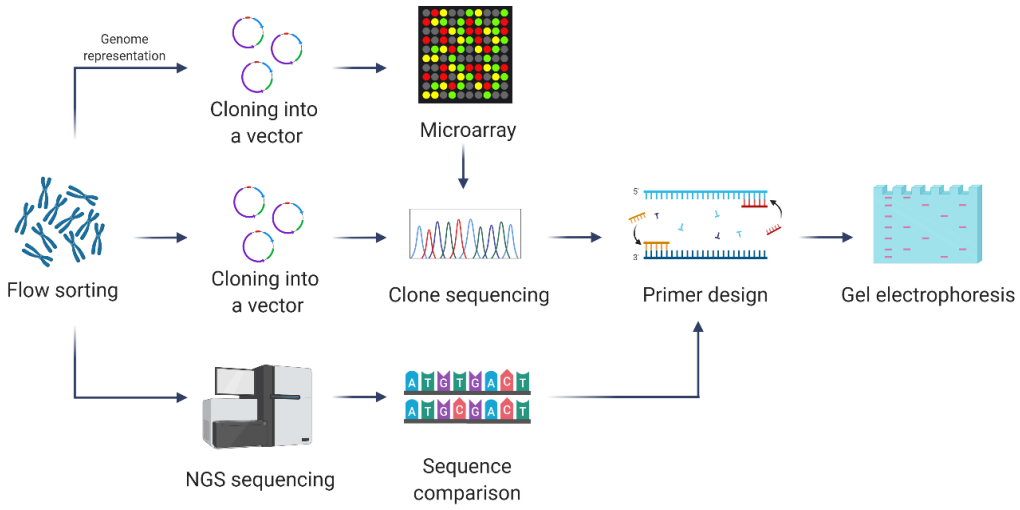


Figure 16: Different ways of marker development (Zwyrtková et al., 2021).

The majority of complex genomes were represented by non-coding repetitive DNA sequences. Meanwhile, this type of sequences is problematic in sequencing, they are a good source of markers. Concerning the fact that most plant DNA markers are developed to use in breeding programs, the development of markers distributed randomly along chromosomes is not the best. The most important regions are obviously these which including genes of agronomic interest (Požárková et al., 2002). Using various markers, molecular maps for many plant species were constructed, such as annual crops (Loarce et al., 1996), fruit trees (Foolad et al., 1995), forest species (Byrne et al., 1995), and legume crops (soybean – Shoemaker and Specht, 1995; pea – Timmerman-Vaughan et al., 1996; common bean – Freyre et al., 1998; chickpea – Ratnaparkhe et al., 1998; lentil – Ford et al., 1999; or broad bean – Román et al., 2004).

Flow-sorted chromosomes or chromosomal arms were suitable for the development of specific microsatellite markers. It was demonstrated that intra-genomic length variation of microsatellites correlates with polymorphism what could be useful in marker development. In cereals, it was recommended AG, AAG, AC, AAC, AT, and AAT microsatellite motifs (Kofler et al., 2008). Using of high-throughput molecular marker technologies led to more densely genetic maps. It was possible to integrate individual linkage maps into consensus genetic maps. This facilitated mapping of a high number of loci and the use of markers across germplasm (Close et al., 2009). PCR-based

and cytogenetic markers could be developed with the help of draft sequences. (Xiao et al., 2017). Next-generation sequencing (NGS) is a fast, relatively cheap, and high-throughput approach for marker development (Imelfort et al., 2009). The first study aimed to develop single-chromosome-arm-specific microsatellite markers by NGS technology was working with wheat. Markers gained from the study were used in the facilitation of QTL mapping, map-based gene cloning, and integration of genetic and physical maps (Nie et al., 2012).

3.6.2.1 Molecular non-PCR-based markers

Group of molecular non-PCR-based markers is based on hybridization. Significant representatives are RFLP markers (restriction fragment length polymorphism). DNA fragments that are digested using restriction endonuclease are electrophoretically separated and blotted onto a nylon membrane as a Southern blotting technique followed by detection by hybridization of a labeled probe (Williams, 1989). From different sequences, various and characteristic pattern after digestion is obtained. It can be studied from SNPs to larger DNA deletions or insertions, many samples in one experiment. However, the technique is expensive and laborious with the requirement of a high amount of input DNA of known sequence (Garrido-Cardenas et al., 2018). For the first time, Arumuganathan et al. (1994) developed the RFLP markers from the isolated chromosome of tomato. Thus, new markers have been developed.

The progress was the development of DArT microarrays (diversity array technology). It is based on hybridization and the advantages are in its high-throughput analysis and possibility to use input material without known sequence (Jaccoud et al., 2001). In this technique, probes are bounded on a solid microarray, and samples are fluorescently labeled. Detected signals finished in a hybridization map (Bartoš et al., 2011; Garrido-Cardenas et al., 2018). For genetic diversity studies, Kopecký et al. (2009) used the DArT array for *Festuca* and *Lolium* complex, Akbari et al. (2006) for wheat, or Wenzl et al. (2004) for barley.

3.6.2.2 Molecular PCR-based markers

Molecular PCR-based markers belong to a highly popular technique, because of its simplicity. RAPD markers (random amplified polymorphic DNA) are based on different sizes of products after PCR with arbitrary primers and genomic DNA with an unknown

sequence. On the contrary, the PCR itself has to be optimized with a quality template. The technique finished with an electrophoretic separation aimed to show genetic diversity among samples by polymorphisms. The problem of the technique is low reproducibility (Bardakci, 2001).

AFLP markers (amplified length polymorphism) is referred to be a combination of RAPD and RFLP markers. DNA is digested with restriction enzymes and subsequent amplification by PCR is selective. Beforehand, the sequence of DNA is not known. Fragments are separated and visualized with electrophoresis approaches and used for genetic diversity studies (Mueller and Wolfenbarger, 1999).

With the aim to find a new way of analyzing complex genomes, unique junctions generated by transposable element insertions were used in marker development (Bennetzen, 2000). Barley genome is full of transposable elements, therefore using repeat junctions were developed a molecular marker platform. To detect repeat junctions in the genome, survey sequencing data was used (Mazaheri et al., 2014). A quite similar approach is intron targeting marker development based on intron polymorphism (Poczai et al., 2013). The insertions, deletions, and substitutions are more common in introns (Kimura, 1983). Wang et al. (2017) isolated the short arm of chromosome 4V of *Haynaldia villosa* (syn. *Dasypyrum villosum*), which is a wild relative of wheat with many valuable genes, e.g., high grain protein content, high tiller ability, tolerance to abiotic stresses, resistance against powdery mildew (Grądzielewska, 2006), or wheat yellow mosaic virus (Zhang et al., 2005).

Simple sequence repeat markers (SSRs)

Microsatellites are distributed in all eukaryotic genomes (Cregan et al., 1994; Ellegren, 1993; Hamada et al., 1982). They are composed of tandemly organized repetitions of 1–5 bp length and their abundances are different among and within individual species (Lagercrantz et al., 1993). Based on the number of repeats, the microsatellites show length polymorphism (Kostia et al., 1995) which can be caused by imprecisely DNA replication (Grover and Sharma, 2016). They are often used as genetic markers called simple sequence repeats (SSRs, Powell et al., 1996) or short tandem repeats (STRs, Garrido-Cardenas et al., 2018).

Historically, there were two main approaches to isolate them. Firstly, by searching for known DNA sequences saved in databases (Akagi et al., 1996; Becker and Heun,

1995; Chin et al., 1996) and secondly, by a screening of genomic libraries (Guilford et al., 1997; Ma et al., 1996; Panaud et al., 1996). The second approach was limited by the low abundance of most microsatellites in eukaryotic genomes (Koblížková et al., 1998). The solution is an enrichment of the DNA libraries for that microsatellite motif (Edwards et al., 1996; Karagyozov et al., 1993; Kijas et al., 1994; Lyall et al., 1993). This strategy is based on the subtraction of genomic DNA fragments by hybridization followed by reuptake with solid supports with oligonucleotides complementary to the chosen sequence (Koblížková et al., 1998). In a study of Garbus et al. (2015), transposable elements identified on chromosome 4D of *Triticum aestivum* were analyzed, sorted into classes and families. Di/tri-nucleotides were the most abundant microsatellites and it was created a list of SSR motifs.

Development of SSR markers using draft assemblies based on next-generation sequencing results in cheaper discovering. SSRs are easy to use as PCR-based markers (Córdoba et al., 2010) with known DNA sequence. The primers for PCR, which are designed for the flanking loci, can be fluorescently or radioactively labeled, or without any labeling. Results of the analysis are read with the help of electrophoresis or, e.g., laser detection, based on the prior labeling method. They are co-dominant and world-widely used (Garrido-Cardenas et al., 2018) for diversity characterization, for the integration of genetic and physical maps (Córdoba et al., 2010), to order scaffolds to longer assemblies (Ren et al., 2012), for genetic variability studies (Christelová et al., 2017; Komínková et al., 2016; Nyine et al., 2017), or identification of alien chromosomal segments which were introgressed into common wheat or in control of transgenics, e.g., in transfer genes for resistances or tolerances. The SSRs are one of the most informative markers with high reproducibility and transferability. Nevertheless, only a limited number of SSR markers are known (Taheri et al., 2018). Thus, better knowledge and discovering of novel SSRs is desired.

3.6.2.3 Molecular sequence-based markers

SNP markers (single nucleotide polymorphism), as a representative of molecular sequence-based markers, rely on the comparison of two sequences where a single-nucleotide mismatch is observed, obviously in various members of one species (Garrido-Cardenas et al., 2018). A high-throughput SNP genotyping platform was a big step for the development of high-density genetic maps of barley (Close et al., 2009). It is useful for marker-assisted breeding, to support association genetic analysis, map-based cloning, anchoring DNA sequence scaffolds, and the physical map to the genetic map (Muñoz-Amatriaín et al., 2011). Large sets of SNP markers were developed (Allen et al., 2013; Krasileva et al., 2013; Saintenac et al., 2013). Using NGS technology, Tiwari et al. (2014) developed the first alien chromosome-specific SNP markers from a flow-sorted chromosome 5M of *Aegilops geniculata* further used for mapping of alien introgression in wheat breeding programs, mapping and cloning new genes from the wild relatives. Among cultivars of wheat is a low polymorphism level that was a challenge in developing a specific genetic map. Based on sequencing, Shatalina et al. (2013) developed 70 SNP markers for coding regions of sorted chromosome 3B of winter wheat cultivars Arina and Forno.

4 Results

The focus of the Thesis was (i) the comparative analyses of repetitive sequences and (ii) simple sequence repeat markers development in selected *Poaceae* representatives. More precisely, representatives of *Festuca*, *Lolium*, and *Agropyron* genus were included in the study.

4.1 Comparative analyses of repetitive elements and identification of centromeric element in representatives of fescues and ryegrasses

To perform comparative analyses of repetitive sequences content of ten closely related species of *Festuca* and *Lolium* genus from the *Loliinae* tribe, the study was started with *in silico* analysis of Illumina whole-genome sequencing data with small coverage, where tandemly organized repetitive DNA sequences were *de novo* identified using RepeatExplorer pipeline (Novák et al., 2013). Major repeat families were reconstructed and quantified between several species. The advantage of the pipeline was 3D graphs visualization by program SeqGrapheR, which enables us to visualize the distribution of reads. The program was used to align reads from species of interest, to discover specific contigs, and design specific primers for probes preparation, which was later used in cytogenetic and molecular methods, such as Southern hybridization that is based on hybridization of the probe with the cleavage genomic DNA fixed on a nylon membrane. Assessment of sequence diversity of repeats between and within selected species was performed. Retrotransposons were the most abundant type of repetitive sequences in studied genomes. One type of Ty3/gypsy class, Athila elements, represented about 30% of *Lolium* genomes what was about five times higher proportion compared to *Festuca* genomes.

It was shown that mentioned grasses did not contain a typical satellite repeat in centromeres, which was described in maize or rice (Bao et al., 2006; Cheng et al., 2002; Jiang et al., 2003), but another repetitive element of Ty3/gypsy family, novel LTR

element Fesreba. Its relationships with the Cereba element were confirmed by colocalization with histone variant CENH3 in cytogenetic experiments. The novel element can play a role in the function of centromere in studied species.

4.2 Genome organization and molecular chromosome structure of *Agropyron cristatum*

Firstly, the genomic DNA of *Agropyron cristatum* was flow-sorted into individual chromosomes and amplified using single-chromosome and whole-genome amplification methods (Cápal et al., 2015; Šimková et al., 2008b). The genetic material was used for the library preparation with TruSeq® DNA PCR-free High Throughput Library Prep Kit (Illumina, San Diego, USA) and sequenced on Illumina instruments MiSeq and HiSeq to the final coverage of 23.5–42.8× along all seven chromosomes. Quality-filtered and trimmed data were assembled into scaffolds by Ray *de novo* assembler (Boisvert et al., 2010). Scaffolds equal to or bigger than 500 nucleotides were selected, and subsequent chromosome-specific assemblies served for simple sequence repeat (SSR) markers development.

For SSR marker development it was used MIcroSAtellite identification tool (MISA, Thiel et al., 2003) and subsequent primer designing program Primer3 (Untergasser et al., 2012). It was identified more than 134 thousand SSR markers in several type classes, where minimal repeat was 12 bp long. The final output files from MISA were input files for Primer3. Chromosome or chromosome-arm specific primers were designed for 250 SSR markers. Each one was verified on chromosome-specific and chromosome-arm-specific addition lines for *Agropyron* using molecular methods. Out of them, 72 were specific for chromosome or chromosomal arms, another 16 were mapped on more chromosomes, and 71 were mapped on all *Agropyron* chromosomes. The rest ones showed no polymorphic pattern between *A. cristatum* and *Triticum aestivum*. With the help of newly developed markers, comparative analyses of individual *Agropyron* chromosomes with wheat subgenomes were performed to study their orthologous relationships. Chromosome or chromosome-arm specific SSR markers will be useful for the characterization of introgression lines or alien gene transfer, especially for breeders and their wheat improvement programs.

A study of repetitive and genic sequence contribution in the genome of *A. cristatum* revealed repetitive elements of the Athila lineage of Ty3/gypsy were

the most represented in the genome. The highest amount was on chromosome 7P. The contribution of all repetitive elements in the genome was about 80% in almost all chromosomes. Overall, the most genic sequences were identified on chromosome 4P. Out of agronomic important genes were in the *Agropyron* genome identified sequences of *HvCSlF* genes involved in β-glucan biosynthesis and genes for the main hardness locus, bread making quality genes puroindoline A and B, yield vernalization locus *VRN2*, leaf rust resistance gene or salt tolerance genes.

4.3 Original papers

4.2.1 Comparative analyses of DNA repeats and identification of a novel Fesreba centromeric element in fescues and ryegrasses

(Appendix I)

4.2.2 Chromosome genomics uncovers plant genome organization and function

(Appendix II)

4.2.3 Draft sequencing crested wheat grass chromosomes identifies evolutionary structural changes, genes and develop SSR markers

(This manuscript will be submitted soon to a journal with an impact factor, therefore, the manuscript is not a part of Appendices.)

4.3.1 Comparative analyses of DNA repeats and identification of a novel Fesreba centromeric element in fescues and ryegrasses

Zwyrtková, J., Němečková, A., Čížková, J., Holušová, K., Kapustová, V., Svačina, R., Kopecký, D., Till, B. J., Doležel, J., Hřibová, E.

BMC Plant Biology, 20: 280, 2020
doi: 10.1186/s12870-020-02495-0
IF_{2019}: 3.497

Abstract:

Cultivated grasses are an important source of food for domestic animals worldwide. Increased knowledge of their genomes can speed up the development of new cultivars with better quality and greater resistance to biotic and abiotic stresses. The most widely grown grasses are tetraploid ryegrass species (*Lolium*) and diploid and hexaploid fescue species (*Festuca*). In this work, we characterized repetitive DNA sequences and their contribution to genome size in five fescue and two ryegrass species as well as one fescue and two ryegrass cultivars.

Partial genome sequences produced by Illumina sequencing technology were used for genome-wide comparative analyses with the RepeatExplorer pipeline. Retrotransposons were the most abundant repeat type in all seven grass species. The Athila element of the Ty3/gypsy family showed the most striking differences in copy number between fescues and ryegrasses. The sequence data enabled the assembly of the long terminal repeat (LTR) element Fesreba, which is highly enriched in centromeric and (peri)centromeric regions in all species. A combination of fluorescence in situ hybridization (FISH) with a probe specific to the Fesreba element and immunostaining with centromeric histone H3 (CENH3) antibody showed their co-localization and indicated a possible role of Fesreba in centromere function.

Comparative repeatome analyses in a set of fescues and ryegrasses provided new insights into their genome organization and divergence, including the assembly of the LTR element Fesreba. A new LTR element Fesreba was identified and found in abundance in centromeric regions of the fescues and ryegrasses. It may play a role in the function of their centromeres.

4.3.2 Chromosome genomics uncovers plant genome organization and function

Zwyrtková, J., Šimková, H., Doležel, J.

Biotechnology Advances, 46: 107659, 2021
doi: 10.1016/j.biotechadv.2020.107659
IF_{2019}: 10.744

Abstract:
The identification of causal genomic loci and their interactions underlying various traits in plants has been greatly aided by progress in understanding the organization of the nuclear genome. This provides clues to the responses of plants to environmental stimuli at the molecular level. Apart from other uses, these insights are needed to fully explore the potential of new breeding techniques that rely on genome editing. However, genome analysis and sequencing is not straightforward in the many agricultural crops and their wild relatives that possess large and complex genomes. Chromosome genomics streamlines this task by dissecting the genome to single chromosomes whose DNA is then used instead of nuclear DNA. This results in a massive and lossless reduction in DNA sample complexity, reduces the time and cost of the experiment, and simplifies data interpretation. Flow cytometric sorting of condensed mitotic chromosomes makes it possible to purify single chromosomes in large quantities, and as the DNA remains intact this process can be coupled successfully with many techniques in molecular biology and genomics. Since the first experiments with flow cytometric sorting in the late 1980s, numerous applications have been developed, and chromosome genomics has been having a significant impact in many areas of research, including the sequencing of complex genomes of important crops and gene cloning. This review discusses these applications, describes their contribution to advancements in plant genome analysis and gene cloning, and outlines future directions.

4.4 Conference presentations

3.3.1 Comparative analysis of repetitive parts of genomes in eight *Loliinae* representatives

(poster presentation; Appendix III)

3.3.2 Comparative analysis of repetitive DNA in eight representatives of fescues and ryegrasses

(poster presentation; Appendix IV)

4.4.1 Comparative analysis of repetitive parts of genomes in eight *Loliinae* representatives

Zwyrtková, J., Kapustová, V., Doležel, J., Hřibová, E.

In: Abstracts of the International Conference "Plant Genome Evolution"
P1.24, Sitges, Spain, 2017

Abstract:

Loliinae subtribe is one the largest among *Poeae* (family *Poaceae*). *Loliinae* is represented by species with complex genomes with monoploid genome size range between 1.58 – 4.03 pg. Thanks to frequent and repeated genome size gain and loss events in the evolution of grasses about 70% of the species are polyploid. In our study, we have focused on a group of 'broad-leaved' *Festuca*, including different types of fescues and their sister genus ryegrass (*Lolium* ssp.). Six fescue species (*Festuca arundinacea, F. gigantea, F. glaucescens, F. mairei* and two cultivars of *F. pratensis*) and two ryegrass species (*Lolium multiflorum, L. perenne*) with different ploidy levels. Nuclear genomes of studied species contained from 3 to 10 Gbp and included large amount of repetitive DNA. Repetitive fraction of genomes evolved much faster than coding parts, which makes it suitable tool for genetic diversity analysis or genome evolution studies. In the present study, we used partial Illumina genome sequencing data for de-novo repeat identification and characterization using application of graph-based clustering done by RepeatExplorer pipeline to determine contribution of repetitive DNA to nuclear genome size, to assess inter-genomic differences and to identify new cytogenetic landmarks. The most abundant repetitive elements across all species under study were retrotransposons. Out of them, Ty3/Gypsy elements were the most frequent, followed by Ty1/Copia and DNA transposons. LINE elements were less common in their nuclear genomes. High percentage of similarity was found between two groups of grasses – *Festuca* and *Lolium*. The lowest level of similarity in DNA repetitive elements was shown in *F. mairei* and *F. glaucescens*, as compared to other tested fescues. Dotter analysis identified large number of tandem organized repeats which can be used as robust cytogenetic landmarks. The work provides detailed information about genome composition and variation in species of *Festuca* and *Lolium*.

4.4.2 Comparative analysis of repetitive DNA in eight representatives of fescues and ryegrasses

Hřibová, E., Zwyrtková, J., Holušová, K., Čížková, J., Kopecký, D., Doležel, J.

In: Book of Abstracts of the International Conference on Polyploidy, Hybridization and Biodiversity
P. 70. Rovinj, Croatia, 2016

Abstract:

Fescues (*Festuca* spp.) and ryegrasses (*Lolium* spp.) are economically important grass species widely distributed over all continents except Antarctica. Fescues are used for turf and high-quality forage for livestock, while ryegrasses are the most important forage and turf grass species. Both genera include diploid and polyploid species and their artificial hybrids called Festuloliums, which may differ in the contribution of parental genomes, combine superior agronomical traits from both parents. Fescues and ryegrasses have large nuclear genomes (1C ranging from about 3,000 Mbp to 10,000 Mbp), which consist mainly from repetitive DNA. This class of DNA sequences evolves more rapidly than coding sequences and can be used to analyze genetic diversity and study processes accompanying speciation and genome evolution. The present work aims to characterize repetitive part of nuclear genomes of seven species of ryegrass and fescue (*L. multiflorum, L. perenne, F. arundinaceae, F. gigantea, F. glaucescens, F. mairei* and two cultivars of *F. pratensis*) and identify the contribution of DNA repeats to their genome diversity. Partial genome sequence data were obtained by illumina sequencing and used for all-to-all comparisons. The RepeatExplorer pipeline was used for genome-wide comparative analyses, to estimate genomic abundance, identify orthologous repeat families and to estimate inter-genomic differences. Retrotransposons were the most abundant repeat type identified in all species. Out of them, Ty3/gypsy elements were the most frequent (more than 3.5 and 6 times more abundant than Ty1/copia in fescues and ryegrasses, respectively). DNA transposons and LINE elements were less frequent, but relatively high number of different tandem repeats was identified in nuclear genomes of all species. In general, high number of similarity hits was found between the species of *Festuca* and *Lolium*. Within the fescues, *F. mairei* and *F. glaucescens* showed the lowest similarity of DNA repeats as compared to other species of *Festuca*. This work provides detailed insights into the genome composition and its variation in fescues

and ryegrasses; it will be a useful resource for repeat masking during genome sequence analyses and facilitate development of new cytogenetic markers suitable for studying of genome organization by FISH.

5 Conclusions

I Comparative analyses of repetitive elements and identification of centromeric element in representatives of fescues and ryegrasses

Repetitive sequences content of ten *Festuca* and *Lolium* species were compared. The most abundant repeat in *Lolium*, Athila element of the Ty3/gypsy family, was about 5 times more frequent than in *Festuca* species. A novel centromeric element called Fesreba has been observed and confirmed with the help of molecular methods, mainly by a combination of immunostaining and fluorescence *in situ* hybridization. These comparative analyses confirmed the close relationships between fescues and ryegrasses. The work has been published in BMC Plant Biology (Zwyrtková et al., 2020).

II Genome organization and molecular chromosome structure of *Agropyron cristatum*

The genome of *Agropyron cristatum* has been dissected by flow-sorting into individual chromosomes, which were sequenced and assembled. The resulting draft sequences were used to develop chromosome and chromosome-arm specific SSR markers, which were successfully verified by molecular methods. Comparative analyses of individual *Agropyron* chromosomes with wheat subgenomes were performed to study their orthologous relationships. Repetitive and genic sequence content was characterized and some agronomically important genes were identified. The results of the study will support the use of *A. cristatum* in wheat introgression breeding and the manuscript will be submitted to a journal with an impact factor.

III Comparison of chromosome genomic approaches

The review (Zwyrtková et al., 2021) focuses on experimental approaches to genome complexity reduction with the aim to simplify genome sequencing and gene cloning in plants with large genomes. The potential of flow cytometric chromosome sorting is discussed and overview of the range of applications in molecular biology and genomics is provided. The review has been published in Biotechnology Advances.

6 List of abbreviations

1C	Holoploid genome
1Cx	Monoploid genome
A	Adenine
AFLP	Amplified fragment length polymorphism
BAC	Bacterial artificial chromosome
bp	Base pair
C	Cytosine
CENH3	Centromeric histone H3
CENP-B	Centromere protein B
COS	Conserved ortholog set
DArT	Diversity arrays technology
DNA	Deoxyribonucleic acid
FISH	Fluorescence *in situ* hybridization
FISHIS	Fluorescence *in situ* hybridization in suspension
G	Guanine
Gbp	Gigabase pair
HMW	High molecular weight
HvCSlF	*Hordeum vulgare* cellulose synthase-like gene
IT	Intron targeting
IWGSC	International Wheat Genome Sequencing Consortium
kbp	Kilobase pair
LARD	Large retrotransposon derivative
LINE	Long interspersed nuclear element
LTR	Long terminal repeat
Mbp	Megabase pair
MISA	Microsatellite identification tool
MTP	Minimum tilling path
NGS	Next-generation sequencing

P	Basic *Agropyron* genome
PCR	Polymerase chain reaction
pg	Picogram
Pol	Polymerase
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RICh	Rearrangement identification and characterization
RJM	Repeat-junction marker
rRNA	Ribosomal ribonucleic acid
RT	Reverse transcriptase
SINE	Short interspersed nuclear element
SNP	Single nucleotide polymorphism
SSR	Simple sequence repeat
STR	Short tandem repeat
STS	Sequence-tagged site
T	Thymine
TRIM	Terminal-repeat retrotransposons in miniature
VRN2	Vernalization gene

7 References

Abrouk, M., Balcárková, B., Šimková, H., Komínkova, E., Martis, M.M., Jakobson, I., Timofejeva, L., Rey, E., Vrána, J., Kilian, A., Järve, K., Doležel, J., Valárik, M., 2017. The *in silico* identification and characterization of a bread wheat/*Triticum militinae* introgression line. Plant Biotechnol. J. 15, 249–256. https://doi.org/10.1111/pbi.12610

Aflitos, S.A., Sanchez-Perez, G., De Ridder, D., Fransz, P., Schranz, M.E., De Jong, H., Peters, S.A., 2015. Introgression browser: High-throughput whole-genome SNP visualization. Plant J. 82, 174–182. https://doi.org/10.1111/tpj.12800

Akagi, H., Yokozeki, Y., Inagaki, A., Fujimura, T., 1996. Microsatellite DNA markers for rice chromosomes. Theor. Appl. Genet. 93, 1071–1077. https://doi.org/10.1007/BF00230127

Akbari, M., Wenzl, P., Caig, V., Carling, J., Xia, L., Yang, S., Uszynski, G., Mohler, V., Lehmensiek, A., Kuchel, H., Hayden, M.J., Howes, N., Sharp, P., Vaughan, P., Rathmell, B., Huttner, E., Kilian, A., 2006. Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theor. Appl. Genet. 113, 1409–1420. https://doi.org/10.1007/s00122-006-0365-4

Akhunov, E.D., Sehgal, S., Liang, H., Wang, S., Akhunova, A.R., Kaur, G., Li, W., Forrest, K.L., See, D., Šimková, H., Ma, Y., Hayden, M.J., Luo, M., Faris, J.D., Doležel, J., Gill, B.S., 2013. Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. Plant Physiol. 161, 252–265. https://doi.org/10.1104/pp.112.205161

Alexandrov, I., Kazakov, A., Tumeneva, I., Shepelev, V., Yurov, Y., 2001. Alpha-satellite DNA of primates: Old and new families. Chromosoma. https://doi.org/10.1007/s004120100146

Alkhimova, O.G., Mazurok, N.A., Potapova, T.A., Zakian, S.M., Heslop-Harrison, J.S., Vershinin, A. V., 2004. Diverse patterns of the tandem repeats organization in rye chromosomes. Chromosoma 113, 42–52. https://doi.org/10.1007/s00412-004-0294

Allen, A.M., Barker, G.L.A., Wilkinson, P., Burridge, A., Winfield, M., Coghill, J., Uauy, C., Griffiths, S., Jack, P., Berry, S., Werner, P., Melichar, J.P.E., Mcdougall, J., Gwilliam, R., Robinson, P., Edwards, K.J., 2013. Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat (*Triticum aestivum* L.). Plant Biotechnol. J. 11, 279–295. https://doi.org/10.1111/pbi.12009

Ananiev, E. V., Phillips, R.L., Rines, H.W., 1998. Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. Proc. Natl. Acad. Sci. U. S. A. 95, 13073–13078. https://doi.org/10.1073/pnas.95.22.13073

Aragón-Alcaide, L., Miller, T., Schwarzacher, T., Reader, S., Moore, G., 1996. A cereal centromeric sequence. Chromosoma 105, 261–268. https://doi.org/10.1007/BF02524643

Arumuganathan, K., Martin, G.B., Telenius, H., Tanksley, S.D., Earle, E.D., 1994. Chromosome 2-specific DNA clones from flow-sorted chromosomes of tomato. MGG Mol. Gen. Genet. 242, 551–558. https://doi.org/10.1007/BF00285278

Asay, K.H., Chatterton, N.J., Jensen, K.B., Jones, T.A., Waldron, B.L., Horton, W.H., 2003. Breeding Improved Grasses for Semiarid Rangelands. Arid L. Res. Manag. 17, 469–478. https://doi.org/10.1080/713936115

Asay, K.H., Jensen, K.B., 1996. Wheatgrasses. Cool. Forage Grasses, Agronomy Monographs. https://doi.org/doi:10.2134/agronmonogr34.c22

Balcerek, M., Pielech-Przybylska, K., Dziekońska-Kubczak, U., Patelski, P., Strak, E., 2016. Fermentation results and chemical composition of agricultural distillates obtained from rye and barley grains and the corresponding malts as a source of amylolytic enzymes and starch. Molecules 21. https://doi.org/10.3390/molecules21101320

Bao, W., Zhang, W., Yang, Q., Zhang, Y., Han, B., Gu, M., Xue, Y., Cheng, Z., 2006. Diversity of centromeric repeats in two closely related wild rice species, *Oryza officinalis* and *Oryza rhizomatis*. Mol. Genet. Genomics 275, 421–430. https://doi.org/10.1007/s00438-006-0103-2

Bardakci, F., 2001. Random Amplified Polymorphic DNA (RAPD) Markers. Turkish J. Biol. 25, 185–196.

Bartoš, J., Paux, E., Kofler, R., Havránková, M., Kopecký, D., Suchánková, P., Šafář, J., Šimková, H., Town, C.D., Lelley, T., Feuillet, C., Doležel, J., 2008. A first survey of the rye (*Secale cereale*) genome composition through BAC end sequencing of the

short arm of chromosome 1R. BMC Plant Biol. 8. https://doi.org/10.1186/1471-2229-8-95

Bartoš, J., Sandve, S.R., Kölliker, R., Kopecký, D., Christelová, P., Stočes, Š., Østrem, L., Larsen, A., Kilian, A., Rognli, O.A., Doležel, J., 2011. Genetic mapping of DArT markers in the *Festuca-Lolium* complex and their use in freezing tolerance association analysis. Theor. Appl. Genet. 122, 1133–1147. https://doi.org/10.1007/s00122-010-1518-z

Becker, J., Heun, M., 1995. Barley microsatellites: allele variation and mapping. Plant Mol. Biol. 27, 835–845. https://doi.org/10.1007/BF00020238

Bennett, M.D., Smith, J.B., 1976. Nuclear dna amounts in angiosperms. Philos. Trans. R. Soc. Lond. B. Biol. Sci. https://doi.org/10.1098/rstb.1976.0044

Bennetzen, J.L., 2000. Transposable element contributions to plant gene and genome evolution. Plant Mol. Biol. 42, 251–269. https://doi.org/10.1023/A:1006344508454

Bento, M., Tomás, D., Viegas, W., Silva, M., 2013. Retrotransposons represent the most labile fraction for genomic rearrangements in polyploid plant species. Cytogenet. Genome Res. 140, 286–294. https://doi.org/10.1159/000353308

Boisvert, S., Laviolette, F., Corbeil, J., 2010. Ray: Simultaneous assembly of reads from a mix of high-throughput sequencing technologies. J. Comput. Biol. 17, 1401–1415. https://doi.org/10.1089/cmb.2009.0238

Brooks, S.A., Huang, L., Gill, B.S., Fellers, J.P., 2002. Analysis of 106 kb of contiguous DNA sequence from the D genome of wheat reveals high gene density and a complex arrangement of genes related to disease resistance. Genome 45, 963–972. https://doi.org/10.1139/g02-049

Burton, R.A., Jobling, S.A., Harvey, A.J., Shirley, N.J., Mather, D.E., Bacic, A., Fincher, G.B., 2008. The genetics and transcriptional profiles of the cellulose synthase-like *HvCslF* gene family in barley. Plant Physiol. 146, 1821–1833. https://doi.org/10.1104/pp.107.114694

Byrne, M., Murrell, J.C., Allen, B., Moran, G.F., 1995. An integrated genetic linkage map for eucalypts using RFLP, RAPD and isozyme markers. Theor. Appl. Genet. 91, 869–875. https://doi.org/10.1007/BF00223894

Cápal, P., Blavet, N., Vrána, J., Kubaláková, M., Doležel, J., 2015. Multiple displacement amplification of the DNA from single flow-sorted plant chromosome. Plant J. 84, 838–844. https://doi.org/10.1111/tpj.13035

Carmena, M., González, C., 1995. Transposable elements map in a conserved pattern of

distribution extending from beta-heterochromatin to centromeres in *Drosophila melanogaster*. Chromosoma 103, 676–684. https://doi.org/10.1007/BF00344228

Catalán, P., Torrecilla, P., López Rodríguez, J.Á., Olmstead, R.G., 2004. Phylogeny of the festucoid grasses of subtribe *Loliinae* and allies (*Poeae, Pooideae*) inferred from ITS and trnL-F sequences. Mol. Phylogenet. Evol. 31, 517–541. https://doi.org/10.1016/j.ympev.2003.08.025

Chantret, N., Salse, J., Sabot, F., Rahman, S., Bellec, A., Laubin, B., Dubois, I., Dossat, C., Sourdille, P., Joudrier, P., Gautier, M.F., Cattolico, L., Beckert, M., Aubourg, S., Weissenbach, J., Caboche, M., Bernard, M., Leroy, P., Chalhoub, B., 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). Plant Cell 17, 1033–1045. https://doi.org/10.1105/tpc.104.029181

Charlesworth, B., Sniegowski, P., Stephan, W., 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature. https://doi.org/10.1038/371215a0

Chen, C.C., Chen, C.M., Yang, J.T., Kao, Y.Y., 1998. Localization of a repetitive DNA sequence to the primary constrictions of maize pachytene chromosomes. Chromosom. Res. 6, 236–238.

Chen, K., Wang, Y., Zhang, R., Zhang, H., Gao, C., 2019. CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. Annu. Rev. Plant Biol. 70, 667–697. https://doi.org/10.1146/annurev-arplant-050718-100049

Chen, Q., Jahier, J., Cauderon, Y., 1989. Production and cytogenetical studies of hybrids between *Triticum aestivum* L. Thell and *Agropyron cristatum* (L.) Gaertn. Comptes rendus l'Académie des Sci. Série 3, Sci. la vie 308, 425–430.

Cheng, Z., Dong, F., Langdon, T., Ouyang, S., Buell, C.R., Gu, M., Blattner, F.R., Jiang, J., 2002. Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell 14, 1691–1704. https://doi.org/10.1105/tpc.003079

Chin, E.C.L., Senior, M.L., Shu, H., Smith, J.S.C., 1996. Maize simple repetitive DNA sequences: Abundance and allele variation. Genome 39, 866–873. https://doi.org/10.1139/g96-109

Choulet, F., Wicker, T., Rustenholz, C., Paux, E., Salse, J., Leroy, P., Schlub, S., le Paslier, M.C., Magdelenat, G., Gonthier, C., Couloux, A., Budak, H., Breen, J., Pumphrey, M., Liu, S., Kong, X., Jia, J., Gut, M., Brunel, D., Anderson, J.A., Gill, B.S., Appels, R., Keller, B., Feuillet, C., 2010. Megabase level sequencing reveals

contrasted organization and evolution patterns of the wheat gene and transposable element spaces. Plant Cell 22, 1686–1701. https://doi.org/10.1105/tpc.110.074187

Christelová, P., De Langhe, E., Hřibová, E., Čížková, J., Sardos, J., Hušáková, M., Van den houwe, I., Sutanto, A., Kepler, A.K., Swennen, R., Roux, N., Doležel, J., 2017. Molecular and cytological characterization of the global *Musa* germplasm collection provides insights into the treasure of banana diversity. Biodivers. Conserv. 26, 801–824. https://doi.org/10.1007/s10531-016-1273-9

Close, T.J., Bhat, P.R., Lonardi, S., Wu, Y., Rostoks, N., Ramsay, L., Druka, A., Stein, N., Svensson, J.T., Wanamaker, S., Bozdag, S., Roose, M.L., Moscou, M.J., Chao, S., Varshney, R.K., Szucs, P., Sato, K., Hayes, P.M., Matthews, D.E., Kleinhofs, A., Muehlbauer, G.J., DeYoung, J., Marshall, D.F., Madishetty, K., Fenton, R.D., Condamine, P., Graner, A., Waugh, R., 2009. Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics 10. https://doi.org/10.1186/1471-2164-10-582

Colmer, T.D., Flowers, T.J., Munns, R., 2006. Use of wild relatives to improve salt tolerance in wheat, in: Journal of Experimental Botany. Oxford University Press, pp. 1059–1078. https://doi.org/10.1093/jxb/erj124

Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Mesfin, A., Trudeau-Spanjers, M., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Qi, L.L., Gill, B.S., Echalier, B., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvořák, J., Peng, J.H., Lapitan, N.L.V., Pathan, M.S., Nguyen, H.T., Ma, X.F., Miftahudin, Gustafson, J.P., Greene, R.A., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sidhu, D., Dilbirligi, M., Gill, K.S., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O., Anderson, J.A., 2004. A 2600-locus chromosome bin map of wheat homoeologous group 2 reveals interstitial gene-rich islands and colinearity with rice. Genetics 168, 625–637. https://doi.org/10.1534/genetics.104.034801

Córdoba, J.M., Chavarro, C., Rojas, F., Muñoz, C., Blair, M.W., 2010. Identification and Mapping of Simple Sequence Repeat Markers from Common Bean (*Phaseolus vulgaris* L.) Bacterial Artificial Chromosome End Sequences for Genome Characterization and Genetic-Physical Map Integration. Plant Genome 3, 154–165. https://doi.org/10.3835/plantgenome2010.06.0013

Cregan, P.B., Bhagwat, A.A., Akkaya, M.S., Rongwen, J., 1994. Microsatellite fingerprinting and mapping of soybean. Methods Mol. Cell. Biol. 5, 49–61.

Crespo-Herrera, L.A., Garkava-Gustavsson, L., Åhman, I., 2017. A systematic review of

rye (*Secale cereale* L.) as a source of resistance to pathogens and pests in wheat (*Triticum aestivum* L.). Hereditas. https://doi.org/10.1186/s41065-017-0033-5

D'Appolonia, B.L., Rayas-Duarte, P., 1994. Wheat carbohydrates: structure and functionality, in: Wheat. Springer US, pp. 107–127. https://doi.org/10.1007/978-1-4615-2672-8_8

de Laat, A.M.M., Blaas, J., 1984. Flow-cytometric characterization and sorting of plant chromosomes. Theor. Appl. Genet. 67, 463–467. https://doi.org/10.1007/BF00263414

Devos, K.M., Dubcovsky, J., Dvořák, J., Chinoy, C.N., Gale, M.D., 1995. Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor. Appl. Genet. 91, 282–288. https://doi.org/10.1007/BF00220890

Dewey, D.R., 1984. The Genomic System of Classification as a Guide to Intergeneric Hybridization with the Perennial *Triticeae*. Springer, Boston, MA, pp. 209–279. https://doi.org/10.1007/978-1-4613-2429-4_9

Doležel, J., Bartoš, J., Voglmayr, H., Greilhuber, J., 2003. Nuclear DNA content and genome size of trout and human. Cytometry. A. https://doi.org/10.1002/cyto.a.10013

Doležel, J., Číhalíková, J., Lucretti, S., 1992. A high-yield procedure for isolation of metaphase chromosomes from root tips of *Vicia faba* L. Planta 188, 93–98. https://doi.org/10.1007/BF00198944

Doležel, J., Lucretti, S., 1995. High-resolution flow karyotyping and chromosome sorting in *Vicia faba* lines with standard and reconstructed karyotypes. Theor. Appl. Genet. 90, 797–802. https://doi.org/10.1007/BF00222014

Doležel, J., Lucretti, S., Schubert, I., 1994. Plant chromosome analysis and sorting by flow cytometry. CRC. Crit. Rev. Plant Sci. 13, 275–309. https://doi.org/10.1080/07352689409701917

Doležel, J., Vrána, J., Cápal, P., Kubaláková, M., Burešová, V., Šimková, H., 2014. Advances in plant chromosome genomics. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2013.12.011

Dudin, G., Steegmayer, E.W., Vogt, P., Schnitzer, H., Diaz, E., Howell, K.E., Cremer, T., Cremer, C., 1988. Sorting of chromosomes by magnetic separation. Hum. Genet. 80, 111–116. https://doi.org/10.1007/BF00702851

Dvořák, J., Luo, M.C., Yang, Z.L., 1998. Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing

and cross- fertilizing *Aegilops* species. Genetics 148, 423–434.

Edwards, K.J., Barker, J.H.A., Daly, A., Jones, C., Karp, A., 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. Biotechniques 20, 758–760. https://doi.org/10.2144/96205bm04

Ellegren, H., 1993. Abundant (A)n·(T)n mononucleotide repeats in the pig genome: linkage mapping of the porcine *APOB*, *FSA*, *ALOX12*, *PEPN* and *RLN* loci. Anim. Genet. 24, 367–372. https://doi.org/10.1111/j.1365-2052.1993.tb00342.x

Erayman, M., Sandhu, D., Sidhu, D., Dilbirligi, M., Baenziger, P.S., Gill, K.S., 2004. Demarcating the gene-rich regions of the wheat genome. Nucleic Acids Res. 32, 3546–3565. https://doi.org/10.1093/nar/gkh639

Farrar, K., Bryant, D.N., Turner, L., Gallagher, J.A., Thomas, A., Farrell, M., Humphreys, M.O., Donnison, I.S., 2012. Breeding for Bio-ethanol Production in *Lolium perenne* L.: Association of Allelic Variation with High Water-Soluble Carbohydrate Content. Bioenergy Res. 5, 149–157. https://doi.org/10.1007/s12155-011-9156-0

Feldman, M., Liu, B., Segal, G., Abbo, S., Levy, A.A., Vega, J.M., 1997. Rapid elimination of low-copy DNA sequences in polyploid wheat: A possible mechanism for differentiation of homoeologous chromosomes. Genetics 147, 1381–1387.

Flavell, R., 1980. The Molecular Characterization and Organization of Plant Chromosomal DNA Sequences. Annu. Rev. Plant Physiol. 31, 569–596. https://doi.org/10.1146/annurev.pp.31.060180.003033

Flavell, R.B., Bennett, M.D., Smith, J.B., Smith, D.B., 1974. Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem. Genet. 12, 257–269. https://doi.org/10.1007/BF00485947

Foolad, M.R., Arulsekar, S., Becerra, V., Bliss, F.A., 1995. A genetic map of *Prunus* based on an interspecific cross between peach and almond. Theor. Appl. Genet. 91, 262–269. https://doi.org/10.1007/BF00220887

Ford, R., Pang, E.C.K., Taylor, P.W.J., 1999. Genetics of resistance to ascochyta blight (*Ascochyta lentis*) of lentil and the identification of closely linked RAPD markers. Theor. Appl. Genet. 98, 93–98. https://doi.org/10.1007/s001220051044

Frame, J., 1991. Herbage production and quality of a range of secondary grass species at five rates of fertilizer nitrogen application. Grass Forage Sci. 46, 139–151. https://doi.org/10.1111/j.1365-2494.1991.tb02216.x

Freyre, R., Skroch, P.W., Geffroy, V., Adam-Blondon, A.F., Shirmohamadali, A., Johnson, W.C., Llaca, V., Nodari, R.O., Pereira, P.A., Tsai, S.M., Tohme, J., Dron,

M., Nienhuis, J., Vallejos, C.E., Gepts, P., 1998. Towards an integrated linkage map of common bean. 4. Development of a core linkage map and alignment of RFLP maps. Theor. Appl. Genet. 97, 847–856. https://doi.org/10.1007/s001220050964

Friebe, B., Jiang, J., Raupp, W.J., McIntosh, R.A., Gill, B.S., 1996. Characterization of wheat-alien translocations conferring resistance to diseases and pests: Current status. Euphytica 91, 59–87. https://doi.org/10.1007/BF00035277

Friebe, B., Schubert, V., Blüthner, W.D., Hammer, K., 1992. C-banding pattern and polymorphism of *Aegilops caudata* and chromosomal constitutions of the amphiploid *T. aestivum* - *Ae. caudata* and six derived chromosome addition lines. Theor. Appl. Genet. 83, 589–596. https://doi.org/10.1007/BF00226902

Friebe, B.R., Tuleen, N.A., Gill, B.S., 1999. Development and identification of a complete set of *Triticum aestivum* - *Aegilops geniculata* chromosome addition lines. Genome 42, 374–380. https://doi.org/10.1139/g99-011

Garbus, I., Romero, J.R., Valarik, M., Vanžurová, H., Karafiátová, M., Cáccamo, M., Doležel, J., Tranquilli, G., Helguera, M., Echenique, V., 2015. Characterization of repetitive DNA landscape in wheat homeologous group 4 chromosomes. BMC Genomics 16. https://doi.org/10.1186/s12864-015-1579-0

Garrido-Cardenas, J.A., Mesa-Valle, C., Manzano-Agugliaro, F., 2018. Trends in plant research using molecular markers. Planta. https://doi.org/10.1007/s00425-017-2829-y

Giorgi, D., Farina, A., Grosso, V., Gennaro, A., Ceoloni, C., Lucretti, S., 2013. FISHIS: Fluorescence *In Situ* Hybridization in Suspension and Chromosome Flow Sorting Made Easy. PLoS One 8, e57994. https://doi.org/10.1371/journal.pone.0057994

Grądzielewska, A., 2006. The genus *Dasypyrum* - Part 2. *Dasypyrum villosum* - A wild species used in wheat improvement. Euphytica 152, 441–454. https://doi.org/10.1007/s10681-006-9245-x

Gregory, T.R., 2001. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. Biol. Rev. 76, 65–101. https://doi.org/10.1111/j.1469-185X.2000.tb00059.x

Grover, A., Sharma, P.C., 2016. Development and use of molecular markers: Past and present. Crit. Rev. Biotechnol. https://doi.org/10.3109/07388551.2014.959891

Guilford, P., Prakash, S., Zhu, J.M., Rikkerink, E., Gardiner, S., Bassett, H., Forster, R., 1997. Microsatellites in *Malus* X domestica (apple): abundance, polymorphism and cultivar identification. Theor. Appl. Genet. 94, 249–254.

https://doi.org/10.1007/s001220050407

Hamada, H., Petrino, M.G., Kakunaga, T., 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. U. S. A. 79, 6465–6469. https://doi.org/10.1073/pnas.79.21.6465

Han, H., Bai, L., Su, J., Zhang, J., Song, L., Gao, A., Yang, X., Li, X., Liu, W., Li, L., 2014. Genetic rearrangements of six wheat-*Agropyron cristatum* 6P addition lines revealed by molecular markers. PLoS One 9. https://doi.org/10.1371/journal.pone.0091066

Han, H., Liu, W., Zhang, J., Zhou, S., Yang, X., Li, X., Li, L., 2019. Identification of P genome chromosomes in *Agropyron cristatum* and wheat-*A. cristatum* derivative lines by FISH. Sci. Rep. 9. https://doi.org/10.1038/s41598-019-46197-6

Hand, M.L., Cogan, N.O., Stewart, A. V., Forster, J.W., 2010. Evolutionary history of tall fescue morphotypes inferred from molecular phylogenetics of the *Lolium-Festuca* species complex. BMC Evol. Biol. 10, 303. https://doi.org/10.1186/1471-2148-10-303

Hashmi, U., Shafqat, S., Khan, F., Majid, M., Hussain, H., Kazi, A.G., John, R., Ahmad, P., 2015. Plant exomics: Concepts, applications and methodologies in crop improvement. Plant Signal. Behav. https://doi.org/10.4161/15592324.2014.976152

Heslop-Harrison, J.S., Brandes, A., Taketa, S., Schmidt, T., Vershinin, A. V., Alkhimova, E.G., Kamm, A., Doudrick, R.L., Schwarzacher, T., Katsiotis, A., Kubis, S., Kumar, A., Pearce, S.R., Flavell, A.J., Harrison, G.E., 1997. The chromosomal distributions of Ty1-copia group retrotransposable elements in higher plants and their implications for genome evolution. Genetica 100, 197–204. https://doi.org/10.1023/a:1018337831039

Heslop-Harrison, J.S., Murata, M., Ogura, Y., Schwarzacher, T., Motoyoshi, F., 1999. Polymorphisms and genomic organization of repetitive DNA from centromeric regions of *Arabidopsis* chromosomes. Plant Cell 11, 31–42. https://doi.org/10.1105/tpc.11.1.31

Heslop-Harrison, J.S.P., Schwarzacher, T., 2011. Organisation of the plant genome in chromosomes. Plant J. 66, 18–33. https://doi.org/10.1111/j.1365-313X.2011.04544.x

Hobza, R., Vyskot, B., 2007. Laser Microdissection-Based Analysis of Plant Sex Chromosomes. Methods Cell Biol. https://doi.org/10.1016/S0091-679X(06)82015

Hřibová, E., Neumann, P., Matsumoto, T., Roux, N., Macas, J., Doležel, J., 2010. Repetitive part of the banana (*Musa acuminata*) genome investigated by low-depth 454 sequencing. BMC Plant Biol. 10, 204. https://doi.org/10.1186/1471-2229-10-204

Hsiao, C., Chatterton, N.J., Asay, K.H., Jensen, K.B., 1995. Phylogenetic relationships of the monogenomic species of the wheat tribe, *Triticeae* (*Poaceae*), inferred from nuclear rDNA (internal transcribed spacer) sequences. Genome 38, 211–223. https://doi.org/10.1139/g95-026

Humphreys, M., Feuerstein, U., Vandewalle, M., Baert, J., 2010. Ryegrasses, in: Fodder Crops and Amenity Grasses. Springer New York, New York, NY, pp. 211–260. https://doi.org/10.1007/978-1-4419-0760-8_10

Hussain, A., Larsson, H., Kuktaite, R., Johansson, E., 2010. Mineral composition of organically grown wheat genotypes: Contribution to daily minerals intake. Int. J. Environ. Res. Public Health 7, 3442–3456. https://doi.org/10.3390/ijerph7093442

Hutchinson, J., Rees, H., Seal, A.G., 1979. An assay of the activity of supplementary DNA in *Lolium*. Heredity (Edinb). 43, 411–421. https://doi.org/10.1038/hdy.1979.91

Imelfort, M., Duran, C., Batley, J., Edwards, D., 2009. Discovering genetic polymorphisms in next-generation sequencing data: Review article. Plant Biotechnol. J. https://doi.org/10.1111/j.1467-7652.2009.00406.x

Inda, L.A., Segarra-Moragues, J.G., Müller, J., Peterson, P.M., Catalán, P., 2008. Dated historical biogeography of the temperate *Loliinae* (*Poaceae*, *Pooideae*) grasses in the northern and southern hemispheres. Mol. Phylogenet. Evol. 46, 932–957. https://doi.org/10.1016/j.ympev.2007.11.022

International Wheat Genome Sequencing Consortium, Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C.J., Choulet, F., Distelfeld, A., Poland, J., Ronen, G., Sharpe, A.G., Barad, O., Baruch, K., Keeble-Gagnère, G., Mascher, M., Ben-Zvi, G., Josselin, A.-A., Himmelbach, A., Balfourier, F., Gutierrez-Gonzalez, J., Hayden, M., Koh, C., Muehlbauer, G., Pasam, R.K., Paux, E., Rigault, P., Tibbits, J., Tiwari, V., Spannagl, M., Lang, D., Gundlach, H., Haberer, G., Mayer, K.F.X., Ormanbekova, D., Prade, V., Šimková, H., Wicker, T., Swarbreck, D., Rimbert, H., Felder, M., Guilhot, N., Kaithakottil, G., Keilwagen, J., Leroy, P., Lux, T., Twardziok, S., Venturini, L., Juhász, A., Abrouk, M., Fischer, I., Uauy, C., Borrill, P., Ramirez-Gonzalez, R.H., Arnaud, D., Chalabi, S., Chalhoub,

B., Cory, A., Datla, R., Davey, M.W., Jacobs, J., Robinson, S.J., Steuernagel, B., van Ex, F., Wulff, B.B.H., Benhamed, M., Bendahmane, A., Concia, L., Latrasse, D., Bartoš, J., Bellec, A., Berges, H., Doležel, J., Frenkel, Z., Gill, B., Korol, A., Letellier, T., Olsen, O.-A., Singh, K., Valárik, M., van der Vossen, E., Vautrin, S., Weining, S., Fahima, T., Glikson, V., Raats, D., Číhalíková, J., Toegelová, H., Vrána, J., Sourdille, P., Darrier, B., Barabaschi, D., Cattivelli, L., Hernandez, P., Galvez, S., Budak, H., Jones, J.D.G., Witek, K., Yu, G., Small, I., Melonek, J., Zhou, R., Belova, T., Kanyuka, K., King, R., Nilsen, K., Walkowiak, S., Cuthbert, R., Knox, R., Wiebe, K., Xiang, D., Rohde, A., Golds, T., Čížková, J., Akpinar, B.A., Biyiklioglu, S., Gao, L., N'Daiye, A., Kubaláková, M., Šafář, J., Alfama, F., Adam-Blondon, A.-F., Flores, R., Guerche, C., Loaec, M., Quesneville, H., Condie, J., Ens, J., Maclachlan, R., Tan, Y., Alberti, A., Aury, J.-M., Barbe, V., Couloux, A., Cruaud, C., Labadie, K., Mangenot, S., Wincker, P., Kaur, G., Luo, M., Sehgal, S., Chhuneja, P., Gupta, O.P., Jindal, S., Kaur, P., Malik, P., Sharma, P., Yadav, B., Singh, N.K., Khurana, J.P., Chaudhary, C., Khurana, P., Kumar, V., Mahato, A., Mathur, S., Sevanthi, A., Sharma, N., Tomar, R.S., Holušová, K., Plíhal, O., Clark, M.D., Heavens, D., Kettleborough, G., Wright, J., Balcárková, B., Hu, Y., Salina, E., Ravin, N., Skryabin, K., Beletsky, A., Kadnikov, V., Mardanov, A., Nesterov, M., Rakitin, A., Sergeeva, E., Handa, H., Kanamori, H., Katagiri, S., Kobayashi, F., Nasuda, S., Tanaka, T., Wu, J., Cattonaro, F., Jiumeng, M., Kugler, K., Pfeifer, M., Sandve, S., Xun, X., Zhan, B., Batley, J., Bayer, P.E., Edwards, D., Hayashi, S., Tulpová, Z., Visendi, P., Cui, L., Du, X., Feng, K., Nie, X., Tong, W., Wang, L., 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361, eaar7191. https://doi.org/10.1126/science.aar7191

International Wheat Genome Sequencing Consortium, Lukaszewski, A.J., Alberti, A., Sharpe, A., Kilian, A., Stanca, A.M., Keller, B., Clavijo, B.J., Friebe, B., Gill, B., Wulff, B., Chapman, B., Steuernagel, B., Feuillet, C., Viseux, C., Pozniak, C., Rokhsar, D.S., Klassen, D., Edwards, D., Akhunov, E., Paux, E., Alfama, F., Choulet, F., Kobayashi, F., Muehlbauer, G.J., Quesneville, H., Šimková, H., Rimbert, H., Gundlach, H., Budak, H., Sakai, H., Handa, H., Kanamori, H., Batley, J., Vrána, J., Rogers, J., Číhalíková, J., Doležel, J., Chapman, J., Poland, J.A., Wu, J., Khurana, J., Wright, J., Bader, K.C., Eversole, K., Barry, K., McLay, K., Mayer, K.F.X., Singh, K., Clissold, L., Pingault, L., Couderc, L., Cattivelli, L., Spannagl, M., Kubaláková, M., Caccamo, M., Mascher, M., Bellgard, M., Pfeifer, M.,

Zytnicki, M., Febrer, M., Alaux, M., Martis, M.M., Loaec, M., Colaiacovo, M., Singh, N.K., Glover, N., Guilhot, N., Stein, N., Olsen, O.A., Maclachlan, P.R., Chhuneja, P., Wincker, P., Sourdille, P., Faccioli, P., Ramirez-Gonzalez, R.H., Waugh, R., Šperková, R., Knox, R., Appels, R., Sharma, S., Ayling, S., Praud, S., Wang, S., Lien, S., Sandve, S.R., Matsumoto, T., Endo, T.R., Itoh, T., Nussbaumer, T., Wicker, T., Tanaka, T., Scholz, U., Barbe, V., Jamilloux, V., Ogihara, Y., Dubská, Z., 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345. https://doi.org/10.1126/science.1251788

Jaccoud, D., Peng, K., Feinstein, D., Kilian, A., 2001. Diversity Arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res. 29, e25–e25. https://doi.org/10.1093/nar/29.4.e25

Jacobsen, E., Schouten, H.J., 2007. Cisgenesis strongly improves introgression breeding and induced translocation breeding of plants. Trends Biotechnol. https://doi.org/10.1016/j.tibtech.2007.03.008

Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyère, C., Billault, A., Segurens, B., Gouyvenoux, M., Ugarte, E., Cattonaro, F., Anthouard, V., Vico, V., Del Fabbro, C., Alaux, M., Di Gaspero, G., Dumas, V., Felice, N., Paillard, S., Juman, I., Moroldo, M., Scalabrin, S., Canaguier, A., Le Clainche, I., Malacrida, G., Durand, E., Pesole, G., Laucou, V., Chatelet, P., Merdinoglu, D., Delledonne, M., Pezzotti, M., Lecharny, A., Scarpelli, C., Artiguenave, F., Pè, M.E., Valle, G., Morgante, M., Caboche, M., Adam-Blondon, A.F., Weissenbach, J., Quétier, F., Wincker, P., 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449, 463–467. https://doi.org/10.1038/nature06148

Janda, J., Bartoš, J., Šafář, J., Kubaláková, M., Valárik, M., Číhalíková, J., Šimková, H., Caboche, M., Sourdille, P., Bernard, M., Chalhoub, B., Doležel, J., 2004. Construction of a subgenomic BAC library specific for chromosomes 1D, 4D and 6D of hexaploid wheat. Theor. Appl. Genet. 109, 1337–1345. https://doi.org/10.1007/s00122-004-1768-8

Janda, J., Šafář, J., Kubaláková, M., Bartoš, J., Kovářová, P., Suchánková, P., Pateyron, S., Číhalíková, J., Sourdille, P., Šimková, H., Faivre-Rampant, P., Hřibová, E., Bernard, M., Lukaszewski, A., Doležel, J., Chalhoub, B., 2006. Advanced resources for plant genomics: A BAC library specific for the short arm of wheat chromosome

1B. Plant J. 47, 977–986. https://doi.org/10.1111/j.1365-313X.2006.02840.x

Jiang, J., Birchler, J.A., Parrott, W.A., Dawe, R.K., 2003. A molecular view of plant centromeres. Trends Plant Sci. 8, 570–575. https://doi.org/10.1016/j.tplants.2003.10.011

Jiang, J., Friebe, B., Gill, B.S., 1993. Recent advances in alien gene transfer in wheat. Euphytica 73, 199–212. https://doi.org/10.1007/BF00036700

Jiang, J., Nasuda, S., Dong, F., Scherrer, C.W., Woo, S.S., Wing, R.A., Gill, B.S., Ward, D.C., 1996. A conserved repetitive DNA element located in the centromeres of cereal chromosomes. Proc. Natl. Acad. Sci. U. S. A. 93, 14210–14213. https://doi.org/10.1073/pnas.93.24.14210

Johansson, D.P., Gutiérrez, J.L.V., Landberg, R., Alminger, M., Langton, M., 2018. Impact of food processing on rye product properties and their *in vitro* digestion. Eur. J. Nutr. 57, 1651–1666. https://doi.org/10.1007/s00394-017-1450-y

Karagyozov, L., Kalcheva, I.D., Chapman, V.M., 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Res. 21, 3911–3912. https://doi.org/10.1093/nar/21.16.3911

Kellogg, E.A., Bennetzen, J.L., 2004. The evolution of nuclear genome structure in seed plants. Am. J. Bot. https://doi.org/10.3732/ajb.91.10.1709

Kijas, J.M., Fowler, J.C., Garbett, C.A., Thomas, M.R., 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. Biotechniques 16, 656–60, 662.

Kilian, B., Mammen, K., Millet, E., Sharma, R., Graner, A., Salamini, F., Hammer, K., Özkan, H., 2011. *Aegilops*, in: Wild Crop Relatives: Genomic and Breeding Resources. Springer Berlin Heidelberg, pp. 1–76. https://doi.org/10.1007/978-3-642-14228-4_1

Kim, N.S., Armstrong, K.C., Fedak, G., Ho, K., Park, N.I., 2002. A microsatellite sequence from the rice blast fungus (*Magnaporthe grisea*) distinguishes between the centromeres of *Hordeum vulgare* and *H. bulbosum* in hybrid plants. Genome 45, 165–174. https://doi.org/10.1139/g01-129

Kimura, M., 1983. Rare variant alleles in the light of the neutral theory. Mol. Biol. Evol. 1, 84–93. https://doi.org/10.1093/oxfordjournals.molbev.a040305

Kiseleva, A. V., Kirov, I. V., Khrustaleva, L.I., 2014. Chromosomal organization of centromeric Ty3/gypsy retrotransposons in *Allium cepa* L. and *Allium fistulosum* L. Russ. J. Genet. 50, 586–592. https://doi.org/10.1134/S102279541404005X

Koblížková, A., Doležel, J., Macas, J., 1998. Subtraction with 3' modified oligonucleotides eliminates amplification artifacts in DNA libraries enriched for microsatellites. Biotechniques 25, 32–38. https://doi.org/10.2144/98251bm06

Kofler, R., Bartoš, J., Gong, L., Stift, G., Suchánková, P., Šimková, H., Berenyi, M., Burg, K., Doležel, J., Lelley, T., 2008. Development of microsatellite markers specific for the short arm of rye (*Secale cereale* L.) chromosome 1. Theor. Appl. Genet. 117, 915–926. https://doi.org/10.1007/s00122-008-0831-2

Kojima, K.K., Kubo, Y., Fujiwara, H., 2002. Complex and Tandem Repeat Structure of Subtelomeric Regions in the Taiwan Cricket, *Teleogryllus taiwanemma*. J. Mol. Evol. 54, 474–485. https://doi.org/10.1007/s0023901-0038-5

Komínková, E., Dreiseitl, A., Malečková, E., Doležel, J., Valárik, M., 2016. Genetic diversity of *Blumeria graminis* f. sp. hordei in Central Europe and its comparison with Australian population. PLoS One 11. https://doi.org/10.1371/journal.pone.0167099

Kopecký, D., Bartoš, J., Lukaszewski, A.J., Baird, J.H., Černoch, V., Kölliker, R., Rognli, O.A., Blois, H., Caig, V., Lübberstedt, T., Studer, B., Shaw, P., Doležel, J., Kilian, A., 2009. Development and mapping of DArT markers within the *Festuca - Lolium* complex. BMC Genomics 10, 473. https://doi.org/10.1186/1471-2164-10-473

Kopecký, D., Havránková, M., Loureiro, J., Castro, S., Lukaszewski, A.J., Bartoš, J., Kopecká, J., Doležel, J., 2010. Physical distribution of homoeologous recombination in individual chromosomes of *Festuca pratensis* in *Lolium multiflorum*. Cytogenet. Genome Res. 129, 162–172. https://doi.org/10.1159/000313379

Kopecký, D., Lukaszewski, A.J., Doležel, J., 2008. Cytogenetics of Festulolium (*Festuca* x *Lolium* hybrids). Cytogenet. Genome Res. https://doi.org/10.1159/000121086

Kopecký, D., Martis, M., Číhalíková, J., Hřibová, E., Vrána, J., Bartoš, J., Kopecká, J., Cattonaro, F., Stočes, Š., Novák, P., Neumann, P., Macas, J., Šimková, H., Studer, B., Asp, T., Baird, J.H., Navrátil, P., Karafiátová, M., Kubaláková, M., Šafář, J., Mayer, K., Doležel, J., 2013. Flow sorting and sequencing meadow fescue chromosome 4F. Plant Physiol. 163, 1323–1337. https://doi.org/10.1104/pp.113.224105

Kopecký, D., Studer, B., 2014. Emerging technologies advancing forage and turf grass genomics. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2013.11.010

Kostia, S., Varvio, S.L., Vakkari, P., Pulkkinen, P., 1995. Microsatellite sequences in a conifer, *Pinus sylvestris*. Genome 38, 1244–1248. https://doi.org/10.1139/g95-163

Krasileva, K. V., Buffalo, V., Bailey, P., Pearce, S., Ayling, S., Tabbita, F., Soria, M., Wang, S., Akhunov, E., Uauy, C., Dubcovsky, J., 2013. Separating homeologs by phasing in the tetraploid wheat transcriptome. Genome Biol. 14, R66. https://doi.org/10.1186/gb-2013-14-6-r66

Kreplak, J., Madoui, M.A., Cápal, P., Novák, P., Labadie, K., Aubert, G., Bayer, P.E., Gali, K.K., Syme, R.A., Main, D., Klein, A., Bérard, A., Vrbová, I., Fournier, C., d'Agata, L., Belser, C., Berrabah, W., Toegelová, H., Milec, Z., Vrána, J., Lee, H.T., Kougbeadjo, A., Térézol, M., Huneau, C., Turo, C.J., Mohellibi, N., Neumann, P., Falque, M., Gallardo, K., McGee, R., Tar'an, B., Bendahmane, A., Aury, J.M., Batley, J., Le Paslier, M.C., Ellis, N., Warkentin, T.D., Coyne, C.J., Salse, J., Edwards, D., Lichtenzveig, J., Macas, J., Doležel, J., Wincker, P., Burstin, J., 2019. A reference genome for pea provides insight into legume genome evolution. Nat. Genet. 51, 1411–1422. https://doi.org/10.1038/s41588-019-0480-1

Kumar, A., 1996. The adventures of the Ty1-copia group of retrotransposons in plants. Trends Genet. https://doi.org/10.1016/0168-9525(96)81393-X

Kumekawa, N., Hosouchi, T., Tsuruoka, H., Kotani, H., 2001. The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 4. DNA Res. 8, 285–290. https://doi.org/10.1093/dnares/8.6.285

Lagercrantz, U., Ellegren, H., Andersson, L., 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res. 21, 1111–1115. https://doi.org/10.1093/nar/21.5.1111

Lapitan, N.L.V., 1992. Organization and evolution of higher plant nuclear genomes. Genome. https://doi.org/10.1139/g92-028

Lee, J.-H., Arumuganathan, K., Chung, Y.-S., Kim, K.-Y., Chung, W.-B., Bae, K.-S., Kim, D.-H., Chung, D.-S., Kwon, O.-C., 2000. Flow Cytometric Analysis and Chromosome Sorting of Barley (*Hordeum vulgare* L.). Mol. Cells 10, 619–625. https://doi.org/10.1007/s10059-000-0619-y

Lee, J.H., Arumuganathan, K., Yen, Y., Kaeppler, S., Kaeppler, H., Baenziger, P.S., 1997. Root tip cell cycle synchronization and metaphase-chromosome isolation suitable for flow sorting in common wheat (*Triticum aestivum* L.). Genome 40, 633–638. https://doi.org/10.1139/g97-083

Limin, A.E., Fowler, D.B., 1990. An interspecific hybrid and amphiploid produced from *Triticum aestivum* crosses with *Agropyron cristatum* and *Agropyron desertorum*. Genome 33, 581–584. https://doi.org/10.1139/g90-085

Linder, H.P., Lehmann, C.E.R., Archibald, S., Osborne, C.P., Richardson, D.M., 2018. Global grass (*Poaceae*) success underpinned by traits facilitating colonization, persistence and habitat transformation. Biol. Rev. 93, 1125–1144. https://doi.org/10.1111/brv.12388

Ling, H.Q., Ma, B., Shi, X., Liu, H., Dong, L., Sun, H., Cao, Y., Gao, Q., Zheng, S., Li, Ye, Yu, Y., Du, H., Qi, M., Li, Yan, Lu, H., Yu, H., Cui, Y., Wang, N., Chen, C., Wu, H., Zhao, Y., Zhang, J., Li, Yiwen, Zhou, W., Zhang, B., Hu, W., Van Eijk, M.J.T., Tang, J., Witsenboer, H.M.A., Zhao, S., Li, Z., Zhang, A., Wang, D., Liang, C., 2018. Genome sequence of the progenitor of wheat A subgenome *Triticum urartu*. Nature 557, 424–428. https://doi.org/10.1038/s41586-018-0108-0

Linkiewicz, A.M., Qi, L.L., Gill, B.S., Ratnasiri, A., Echalier, B., Chao, S., Lazo, G.R., Hummel, D.D., Anderson, O.D., Akhunov, E.D., Dvořák, J., Pathan, M.S., Nguyen, H.T., Peng, J.H., Lapitan, N.L.V., Miftahudin, Gustafson, J.P., La Rota, C.M., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Sandhu, D., Bondareva, S.N., Gill, K.S., Conley, E.J., Anderson, J.A., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O., Dubcovsky, J., 2004. A 2500-locus bin map of wheat homoeologous group 5 provides insights on gene distribution and colinearity with rice. Genetics 168, 665–676. https://doi.org/10.1534/genetics.104.034835

Loarce, Y., Hueros, G., Ferrer, E., 1996. A molecular linkage map of rye. Theor. Appl. Genet. 93, 1112–1118. https://doi.org/10.1007/BF00230133

Loureiro, J., Kopecký, D., Castro, S., Santos, C., Silveira, P., 2007. Flow cytometric and cytogenetic analyses of Iberian Peninsula *Festuca* spp. Plant Syst. Evol. 269, 89–105. https://doi.org/10.1007/s00606-007-0564-8

Lowe, J.W.E., 2018. Sequencing through thick and thin: Historiographical and philosophical implications. Stud. Hist. Philos. Sci. Part C Stud. Hist. Philos. Biol. Biomed. Sci. 72, 10–27. https://doi.org/10.1016/j.shpsc.2018.10.007

Luan, Y., Wang, X., Liu, W., Li, C., Zhang, J., Gao, A., Wang, Y., Yang, X., Li, L., 2010. Production and identification of wheat-*Agropyron cristatum* 6P translocation lines. Planta 232, 501–510. https://doi.org/10.1007/s00425-010-1187-9

Lucretti, S., Doležel, J., 1997. Bivariate flow karyotyping in broad bean (*Vicia faba*). Cytometry 28, 236–242.

Lukaszewski, A.J., Lapinski, B., Rybka, K., 2005. Limitations of in situ hybridization with total genomic DNA in routine screening for alien introgressions in wheat. Cytogenet. Genome Res. 109, 373–377. https://doi.org/10.1159/000082422

Luo, M.C., Gu, Y.Q., Puiu, D., Wang, H., Twardziok, S.O., Deal, K.R., Huo, N., Zhu, T., Wang, L., Wang, Y., McGuire, P.E., Liu, S., Long, H., Ramasamy, R.K., Rodriguez, J.C., Van Sonny, L., Yuan, L., Wang, Z., Xia, Z., Xiao, L., Anderson, O.D., Ouyang, S., Liang, Y., Zimin, A. V., Pertea, G., Qi, P., Bennetzen, J.L., Dai, X., Dawson, M.W., Müller, H.G., Kugler, K., Rivarola-Duarte, L., Spannagl, M., Mayer, K.F.X., Lu, F.H., Bevan, M.W., Leroy, P., Li, P., You, F.M., Sun, Q., Liu, Z., Lyons, E., Wicker, T., Salzberg, S.L., Devos, K.M., Dvoák, J., 2017. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. Nature 551, 498–502. https://doi.org/10.1038/nature24486

Lyall, J.E.W., Brown, G.M., Furlong, R.A., Ferguson-Smith, M.A., Affara, N.A., 1993. A method for creating chromosome-specific plasmid libraries enriched in clones containing [CA]n microsatellite repeat sequences directly from flow-sorted chromosomes. Nucleic Acids Res. 21, 4641–4642. https://doi.org/10.1093/nar/21.19.4641

Lyngkjaer, M.F., Newton, A.C., Atzema, J.L., Baker, S.J., 2000. The Barley mlo-gene: An important powdery mildew resistance source. Agronomie 20, 745–756. https://doi.org/10.1051/agro:2000173

Ma, J., Stiller, J., Berkman, P.J., Wei, Y., Rogers, J., Feuillet, C., Dolezel, J., Mayer, K.F., Eversole, K., Zheng, Y.-L., Liu, C., 2013. Sequence-Based Analysis of Translocations and Inversions in Bread Wheat (*Triticum aestivum* L.). PLoS One 8, e79329. https://doi.org/10.1371/journal.pone.0079329

Ma, J., Wing, R.A., Bennetzen, J.L., Jackson, S.A., 2007. Plant centromere organization: a dynamic structure with conserved functions. Trends Genet. https://doi.org/10.1016/j.tig.2007.01.004

Ma, L., Xiao, Y., Huang, H., Wang, Q., Rao, W., Feng, Y., Zhang, K., Song, Q., 2010. Direct determination of molecular haplotypes by chromosome microdissection. Nat. Methods 7, 299–301. https://doi.org/10.1038/nmeth.1443

Ma, Z.Q., Röder, M., Sorrells, M.E., 1996. Frequencies and sequence characteristics of di-, tri-, and tetra-nucleotide microsatellites in wheat. Genome 39, 123–130. https://doi.org/10.1139/g96-017

Macas, J., Gualberti, G., Nouzová, M., Samec, P., Lucretti, S., Doležel, J., 1996. Construction of chromosome-specific DNA libraries covering the whole genome of field bean (*Vicia faba* L.). Chromosom. Res. 4, 531–539. https://doi.org/10.1007/BF02261781

Macas, J., Koblížková, A., Navrátilová, A., Neumann, P., 2009. Hypervariable 3′ UTR region of plant LTR-retrotransposons as a source of novel satellite repeats. Gene 448, 198–206. https://doi.org/10.1016/j.gene.2009.06.014

Macas, J., Neumann, P., Navrátilová, A., 2007. Repetitive DNA in the pea (*Pisum sativum* L.) genome: Comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. BMC Genomics 8, 427. https://doi.org/10.1186/1471-2164-8-427

Maluszynska, J., Heslop-Harrison, J.S., 1991. Localization of tandemly repeated DMA sequences in *Arabidopsis thaliana*. Plant J. 1, 159–166. https://doi.org/10.1111/j.1365-313X.1991.00159.x

Martinez-Zapater, J.M., Estelle, M.A., Somerville, C.R., 1986. A highly repeated DNA sequence in *Arabidopsis thaliana*. MGG Mol. Gen. Genet. 204, 417–423. https://doi.org/10.1007/BF00331018

Martis, M.M., Zhou, R., Haseneyer, G., Schmutzer, T., Vrána, J., Kubaláková, M., König, S., Kugler, K.G., Scholz, U., Hackauf, B., Korzun, V., Schön, C.C., Doležel, J., Bauer, E., Mayer, K.F.X., Stein, N., 2013. Reticulate evolution of the rye genome. Plant Cell 25, 3685–3698. https://doi.org/10.1105/tpc.113.114553

Matsumoto, T., Wu, J., Kanamori, H., Katayose, Y., Fujisawa, M., Namiki, N., Mizuno, H., Yamamoto, K., Antonio, B.A., Baba, T., Sakata, K., Nagamura, Y., Aoki, H., Arikawa, K., Arita, K., Bito, T., Chiden, Y., Fujitsuka, N., Fukunaka, R., Hamada, M., Harada, C., Hayashi, A., Hijishita, S., Honda, M., Hosokawa, S., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Ito, K., Ito, S., Ito, T., Ito, Yuichi, Ito, Yukiyo, Iwabuchi, A., Kamiya, K., Karasawa, W., Kurita, K., Katagiri, S., Kikuta, A., Kobayashi, H., Kobayashi, N., MacHita, K., Maehara, T., Masukawa, M., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nagata, Y., Naito, S., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Meguro, A., Negishi, M., Ohta, I., Ohta, T., Okamoto, M., Ono, N., Saji, S., Sakaguchi, M., Sakai, K., Shibata, M., Shimokawa, T., Song, J., Takazaki, Y., Terasawa, K., Tsugane, M., Tsuji, K., Ueda, S., Waki, K., Yamagata, H., Yamamoto, M., Yamamoto, S., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H., Yano, M., Sasaki, T., Yuan, Q., Ouyang, S., Liu, J., Jones, K.M., Gansberger, K., Moffat, K., Hill, J., Bera, J., Fadrosh, D., Jin, S., Johri, S., Kim, M., Overton, L., Reardon, M., Tsitrin, T., Vuong, H., Weaver, B., Ciecko, A., Tallon, L., Jackson, J., Pai, G., Van Aken, S., Utterback, T., Reidmuller, S., Feldblyum, T., Hsiao, J., Zismann, V., Iobst, S., De Vazeille, A.R.,

Buell, C.R., Ying, K., Li, Y., Lu, T., Huang, Y., Zhao, Q., Feng, Q., Zhang, L., Zhu, J., Weng, Q., Mu, J., Lu, Y., Fan, D., Liu, Y., Guan, J., Zhang, Yujun, Yu, S., Liu, X., Zhang, Yu, Hong, G., Han, B., Choisne, N., Demange, N., Orjeda, G., Samain, S., Cattolico, L., Pelletier, E., Couloux, A., Segurens, B., Wincker, P., D'Hont, A., Scarpelli, C., Weissenbach, J., Salanoubat, M., Quetier, F., Yu, Y., Kim, H.R., Rambo, T., Currie, J., Collura, K., Luo, M., Yang, T.J., Ammiraju, J.S.S., Engler, F., Soderlund, C., Wing, R.A., Palmer, L.E., De La Bastide, M., Spiegel, L., Nascimento, L., Zutavern, T., O'Shaughnessy, A., Dike, S., Dedhia, N., Preston, R., Balija, V., McCombie, W.R., Chow, T.Y., Chen, H.H., Chung, M.C., Chen, C.S., Shaw, J.F., Wu, H.P., Hsiao, K.J., Chao, Y.T., Chu, M.K., Cheng, C.H., Hour, A.L., Lee, P.F., Lin, S.J., Lin, Y.C., Liou, J.Y., Liu, S.M., Hsing, Y.I., Raghuvanshi, S., Mohanty, A., Bharti, A.K., Gaur, A., Gupta, V., Kumar, D., Ravi, V., Vij, S., Kapur, A., Khurana, Parul, Khurana, Paramjit, Khurana, J.P., Tyagi, A.K., Gaikwad, K., Singh, A., Dalal, V., Srivastava, S., Dixit, A., Pal, A.K., Ghazi, I.A., Yadav, M., Pandit, A., Bhargava, A., Sureshbabu, K., Batra, K., Sharma, T.R., Mohapatra, T., Singh, N.K., Messing, J., Nelson, A.B., Fuks, G., Kavchok, S., Keizer, G., Llaca, E.L.V., Song, R., Tanyolac, B., Young, S., Ho, K., Hahn, J.H., Sangsakoo, G., Vanavichit, A., De Mattos, L.A.T., Zimmer, P.D., Malone, G., Dellagostin, O., De Oliveira, A.C., Bevan, M., Bancroft, I., Minx, P., Cordum, H., Wilson, R., Cheng, Z., Jin, W., Jiang, J., Leong, S.A., Iwama, H., Gojobori, T., Itoh, T., Niimura, Y., Fujii, Y., Habara, T., Sakai, H., Sato, Y., Wilson, G., Kumar, K., McCouch, S., Juretic, N., Hoen, D., Wright, S., Bruskiewich, R., Bureau, T., Miyao, A., Hirochika, H., Nishikawa, T., Kadowaki, K.I., Sugiura, M., Burr, B., 2005. The map-based sequence of the rice genome. Nature 436, 793–800. https://doi.org/10.1038/nature03895

Matsunaga, S., Kawano, S., Michimoto, T., Higashiyama, T., Nakao, S., Sakai, A., Kuroiwa, T., 1999. Semi-Automatic Laser Beam Microdissection of the Y Chromosome and Analysis of Y Chromosome DNA in a Dioecious Plant, *Silene latifolia*. Plant Cell Physiol. 40, 60–68. https://doi.org/10.1093/oxfordjournals.pcp.a029475

Mayer, K.F.X., Martis, M., Hedley, P.E., Šimková, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., Roessner, S., Gundlach, H., Kubaláková, M., Suchánková, P., Murat, F., Felder, M., Nussbaumer, T., Graner, A., Salse, J., Endo, T., Sakai, H., Tanaka, T., Itoh, T., Sato, K., Platzer, M., Matsumoto, T., Scholz, U., Doležel, J.,

Waugh, R., Stein, N., 2011. Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell 23, 1249–1263. https://doi.org/10.1105/tpc.110.082537

Mayer, K.F.X., Taudien, S., Martis, M., Šimková, H., Suchánková, P., Gundlach, H., Wicker, T., Petzold, A., Felder, M., Steuernagel, B., Scholz, U., Graner, A., Platzer, M., Doležel, J., Stein, N., 2009. Gene content and virtual gene order of barley chromosome 1H. Plant Physiol. 151, 496–505. https://doi.org/10.1104/pp.109.142612

Mazaheri, M., Kianian, P.M.A., Mergoum, M., Valentini, G.L., Seetan, R., Pirseyedi, S.M., Kumar, A., Gu, Y.Q., Stein, N., Kubaláková, M., Doležel, J., Denton, A.M., Kianian, S.F., 2014. Transposable Element Junctions in Marker Development and Genomic Characterization of Barley. Plant Genome 7, plantgenome2013.10.0036. https://doi.org/10.3835/plantgenome2013.10.0036

McGuire, P.E., Dvořák, J., 1981. High Salt-Tolerance Potential in Wheatgrasses. Crop Sci. 21, 702–705. https://doi.org/10.2135/cropsci1981.0011183x002100050018x

Miftahudin, Ross, K., Ma, X.F., Mahmoud, A.A., Layton, J., Rodriguez Milla, M.A., Chikmawati, T., Ramalingam, J., Feril, O., Pathan, M.S., Surlan Momirovic, G.S., Kim, S., Chema, K., Fang, P., Haule, L., Struxness, H., Birkes, J., Yaghoubian, C., Skinner, R., McAllister, J., Nguyen, V., Qi, L.L., Echalier, B., Gill, B.S., Linkiewicz, A.M., Dubcovsky, J., Akhunov, E.D., Dvořák, J., Dilbirligi, M., Gill, K.S., Peng, J.H., Lapitan, N.L.V., Bermudez-Kandianis, C.E., Sorrells, M.E., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Lazo, G.R., Chao, S., Anderson, O.D., Gonzalez-Hernandez, J., Conley, E.J., Anderson, J.A., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O., Nguyen, H.T., Gustafson, J.P., 2004. Analysis of expressed sequence tag loci on wheat chromosome group 4. Genetics 168, 651–663. https://doi.org/10.1534/genetics.104.034827

Molnár, I., Vrána, J., Burešová, V., Cápal, P., Farkas, A., Darkó, É., Cseh, A., Kubaláková, M., Molnár-Láng, M., Doležel, J., 2016. Dissecting the U, M, S and C genomes of wild relatives of bread wheat (*Aegilops* spp.) into chromosomes and exploring their synteny with wheat. Plant J. 88, 452–467. https://doi.org/10.1111/tpj.13266

Mueller, U.G., Wolfenbarger, L.L.R., 1999. AFLP genotyping and fingerprinting. Trends Ecol. Evol. https://doi.org/10.1016/S0169-5347(99)01659-6

Mujeeb-Kazi, A., Rosas, V., Roldan, S., 1996. Conservation of the genetic variation of

Triticum tauschii (Coss.) Schmalh. (*Aegilops squarrosa* auct. non L.) in synthetic hexaploid wheats (*T. turgidum* L. s.lat. × *T. tauschii*; 2n = 6x = 42, AABBDD) and its potential utilization for wheat improvement. Genet. Resour. Crop Evol. 43, 129–134. https://doi.org/10.1007/BF00126756

Muñoz-Amatriaín, M., Moscou, M.J., Bhat, P.R., Svensson, J.T., Bartoš, J., Suchánková, P., Šimková, H., Endo, T.R., Fenton, R.D., Lonardi, S., Castillo, A.M., Chao, S., Cistué, L., Cuesta-Marcos, A., Forrest, K.L., Hayden, M.J., Hayes, P.M., Horsley, R.D., Makoto, K., Moody, D., Sato, K., Vallés, M.P., Wulff, B.B.H., Muehlbauer, G.J., Doležel, J., Close, T.J., 2011. An Improved Consensus Linkage Map of Barley Based on Flow-Sorted Chromosomes and Single Nucleotide Polymorphism Markers. Plant Genome 4, 238–249. https://doi.org/10.3835/plantgenome2011.08.0023

Neumann, P., Navrátilová, A., Koblížková, A., Kejnovsk, E., Hřibová, E., Hobza, R., Widmer, A., Doležel, J., MacAs, J., 2011. Plant centromeric retrotransposons: A structural and cytogenetic perspective. Mob. DNA. https://doi.org/10.1186/1759-8753-2-4

Nie, X., Li, B., Wang, L., Liu, P., Biradar, S.S., Li, T., Doležel, J., Edwards, D., Luo, M., Weining, S., 2012. Development of chromosome-arm-specific microsatellite markers in *Triticum aestivum* (*Poaceae*) using NGS technology. Am. J. Bot. 99. https://doi.org/10.3732/ajb.1200077

Nonomura, K.I., Kurata, N., 2001. The centromere composition of multiple repetitive sequences on rice chromosome 5. Chromosoma 110, 284–291. https://doi.org/10.1007/s004120100148

Novák, P., Hřibová, E., Neumann, P., Koblížková, A., Doležel, J., Macas, J., 2014. Genome-wide analysis of repeat diversity across the family *Musaceae*. PLoS One 9, e98918. https://doi.org/10.1371/journal.pone.0098918

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. https://doi.org/10.1093/bioinformatics/btt054

Nyine, M., Uwimana, B., Swennen, R., Batte, M., Brown, A., Christelová, P., Hřibová, E., Lorenzen, J., Doleziel, J., 2017. Trait variation and genetic diversity in a banana genomic selection training population. PLoS One 12. https://doi.org/10.1371/journal.pone.0178734

Obermeier, M.M., Gnädinger, F., Durai Raj, A.C., Obermeier, W.A., Schmid, C.A.O., Balàzs, H., Schröder, P., 2020. Under temperate climate, the conversion of grassland to arable land affects soil nutrient stocks and bacteria in a short term. Sci. Total Environ. 703, 135494. https://doi.org/10.1016/j.scitotenv.2019.135494

Ochoa, V., Madrid, E., Said, M., Rubiales, D., Cabrera, A., 2015. Molecular and cytogenetic characterization of a common wheat-*Agropyron cristatum* chromosome translocation conferring resistance to leaf rust. Euphytica 201, 89–95. https://doi.org/10.1007/s10681-014-1190-5

Olson, E.L., Rouse, M.N., Pumphrey, M.O., Bowden, R.L., Gill, B.S., Poland, J.A., 2013. Introgression of stem rust resistance genes *SrTA10187* and *SrTA10171* from *Aegilops tauschii* to wheat. Theor. Appl. Genet. 126, 2477–2484. https://doi.org/10.1007/s00122-013-2148-z

Palmer, L.E., Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V.S., Nascimento, L.U., Dike, S., De La Bastide, M., Martienssen, R.A., McCombie, W.R., 2003. Maize Genome Sequencing by Methylation Filtration. Science 302, 2115–2117. https://doi.org/10.1126/science.1091265

Panaud, O., Chen, X., McCouch, S.R., 1996. Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). Mol. Gen. Genet. MGG 252, 597–607. https://doi.org/10.1007/bf02172406

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schmutz, J., Spannagl, M., Tang, H., Wang, X., Wicker, T., Bharti, A.K., Chapman, J., Feltus, F.A., Gowik, U., Grigoriev, I. V., Lyons, E., Maher, C.A., Martis, M., Narechania, A., Otillar, R.P., Penning, B.W., Salamov, A.A., Wang, Y., Zhang, L., Carpita, N.C., Freeling, M., Gingle, A.R., Hash, C.T., Keller, B., Klein, P., Kresovich, S., McCann, M.C., Ming, R., Peterson, D.G., Mehboob-Ur-Rahman, Ware, D., Westhoff, P., Mayer, K.F.X., Messing, J., Rokhsar, D.S., 2009. The *Sorghum bicolor* genome and the diversification of grasses. Nature 457, 551–556. https://doi.org/10.1038/nature07723

Paterson, A.H., Bowers, J.E., Peterson, D.G., Estill, J.C., Chapman, B.A., 2003. Structure and evolution of cereal genomes. Curr. Opin. Genet. Dev. https://doi.org/10.1016/j.gde.2003.10.002

Pearce, S.R., Harrison, G., Li, D., Heslop-Harrison, J.S., Kumar, A., Flavell, A.J., 1996.

The Ty1-copia group retrotransposons in *Vicia* species: Copy number, sequence heterogeneity and chromosomal localisation. Mol. Gen. Genet. 250, 305–315. https://doi.org/10.1007/BF02174388

Peterson-Burch, B.D., Nettleton, D., Voytas, D.F., 2004. Genomic neighborhoods for *Arabidopsis* retrotransposons: a role for targeted integration in the distribution of the *Metaviridae*. Genome Biol. 5, R78. https://doi.org/10.1186/gb-2004-5-10-r78

Peterson, D.G., Wessler, S.R., Paterson, A.H., 2002. Efficient capture of unique sequences from eukaryotic genomes. Trends Genet. https://doi.org/10.1016/S0168-9525(02)02764-6

Pich, U., Schubert, I., 1998. Terminal heterochromatin and alternative telomeric sequences in *Allium cepa*. Chromosom. Res. 6, 315–321. https://doi.org/10.1023/A:1009227009121

Poczai, P., Varga, I., Laos, M., Cseh, A., Bell, N., Valkonen, J.P.T., Hyvönen, J., 2013. Advances in plant gene-targeted and functional markers: A review. Plant Methods. https://doi.org/10.1186/1746-4811-9-6

Polívka, F., 1902. Názorná květena zemé koruny české. Svazek IV.

Powell, W., Machray, G.C., Proven, J., 1996. Polymorphism revealed by simple sequence repeats. Trends Plant Sci. https://doi.org/10.1016/1360-1385(96)86898-1

Požárková, D., Koblížková, A., Román, B., Torres, A.M., Lucretti, S., Lysák, M., Doležel, J., Macas, J., 2002. Development and characterization of microsatellite markers from chromosome 1-specific DNA libraries of *Vicia faba*. Biol. Plant. 45, 337–345. https://doi.org/10.1023/A:1016253214182

Presting, G.G., Malysheva, L., Fuchs, J., Schubert, I., 1998. A Ty3/gypsy retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. Plant J. 16, 721–728. https://doi.org/10.1046/j.1365-313X.1998.00341.x

Qi, L., Friebe, B., Gill, B.S., 2006. Complex genome rearrangements reveal evolutionary dynamics of pericentromeric regions in the *Triticeae*. Genome 49, 1628–1639. https://doi.org/10.1139/G06-123

Ratnaparkhe, M.B., Santra, D.K., Tullu, A., Muehlbauer, F.J., 1998. Inheritance of inter-simple-sequence-repeat polymorphisms and linkage with a fusarium wilt resistance gene in chickpea. Theor. Appl. Genet. 96, 348–353. https://doi.org/10.1007/s001220050747

Rawat, N., Tiwari, V.K., Singh, N., Randhawa, G.S., Singh, K., Chhuneja, P., Dhaliwal,

H.S., 2009. Evaluation and utilization of *Aegilops* and wild *Triticum* species for enhancing iron and zinc content in wheat. Genet. Resour. Crop Evol. 56, 53–64. https://doi.org/10.1007/s10722-008-9344-8

Reheul, D., De Cauwer, B., Cougnon, M., 2010. The Role of Forage Crops in Multifunctional Agriculture, in: Fodder Crops and Amenity Grasses. Springer New York, pp. 1–12. https://doi.org/10.1007/978-1-4419-0760-8_1

Ren, Y., Zhao, H., Kou, Q., Jiang, J., Guo, S., Zhang, H., Hou, W., Zou, X., Sun, H., Gong, G., Levi, A., Xu, Y., 2012. A high resolution genetic map anchoring scaffolds of the sequenced watermelon genome. PLoS One 7, e29453. https://doi.org/10.1371/journal.pone.0029453

Robins, J.G., Jensen, K.B., 2020. Breeding of the crested wheatgrass complex (*Agropyron* spp.) for north american temperate rangeland agriculture and conservation. Agronomy. https://doi.org/10.3390/agronomy10081134

Román, B., Satovic, Z., Pozarkova, D., Macas, J., Dolezel, J., Cubero, J.I., Torres, A.M., 2004. Development of a composite map in *Vicia faba*, breeding applications and future prospects. Theor. Appl. Genet. 108, 1079–1088. https://doi.org/10.1007/s00122-003-1515-6

Rustenholz, C., Choulet, F., Laugier, C., Šafář, J., Šimková, H., Doležel, J., Magni, F., Scalabrin, S., Cattonaro, F., Vautrin, S., Bellec, A., Bergès, H., Feuillet, C., Paux, E., 2011. A 3,000-Loci transcription map of chromosome 3B unravels the structural and functional features of gene islands in hexaploid wheat. Plant Physiol. 157, 1596–1608. https://doi.org/10.1104/pp.111.183921

Šafář, J., Bartoš, J., Janda, J., Bellec, A., Kubaláková, M., Valárik, M., Pateyron, S., Weiserová, J., Tušková, R., Číhalíková, J., Vrána, J., Šimková, H., Faivre-Rampant, P., Sourdille, P., Caboche, M., Bernard, M., Doležel, J., Chalhoub, B., 2004. Dissecting large and complex genomes: Flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J. 39, 960–968. https://doi.org/10.1111/j.1365-313X.2004.02179.x

Šafář, J., Šimková, H., Kubaláková, M., Číhalíková, J., Suchánková, P., Bartoš, J., Doležel, J., 2010. Development of chromosome-specific BAC resources for genomics of bread wheat. Cytogenet. Genome Res. 129, 211–223. https://doi.org/10.1159/000313072

Said, M., Hřibová, E., Danilova, T. V., Karafiátová, M., Čížková, J., Friebe, B., Doležel, J., Gill, B.S., 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. Theor. Appl. Genet. 131, 2213–2227. https://doi.org/10.1007/s00122-018-3148-9

Said, M., Parada, A.C., Gaál, E., Molnár, I., Cabrera, A., Doležel, J., Vrána, J., 2019. Uncovering homeologous relationships between tetraploid *Agropyron cristatum* and bread wheat genomes using COS markers. Theor. Appl. Genet. 132, 2881–2898. https://doi.org/10.1007/s00122-019-03394-1

Saintenac, C., Jiang, D., Wang, S., Akhunov, E., 2013. Sequence-based mapping of the polyploid wheat genome. G3 Genes, Genomes, Genet. 3, 1105–1114. https://doi.org/10.1534/g3.113.005819

Sandhu, D., Gill, K.S., 2002. Gene-containing regions of wheat and the other grass genomes. Plant Physiol. https://doi.org/10.1104/pp.010745

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A., Minx, P., Reily, A.D., Courtney, L., Kruchowski, S.S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S.M., Belter, E., Du, F., Kim, K., Abbott, R.M., Cotton, M., Levy, A., Marchetto, P., Ochoa, K., Jackson, S.M., Gillam, B., Chen, W., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., Cordes, M., Kohlberg, S., Sgro, J., Delgado, B., Mead, K., Chinwalla, A., Leonard, S., Crouse, K., Collura, K., Kudrna, D., Currie, J., He, R., Angelova, A., Rajasekar, S., Mueller, T., Lomeli, R., Scara, G., Ko, A., Delaney, K., Wissotski, M., Lopez, G., Campos, D., Braidotti, M., Ashley, E., Golser, W., Kim, H., Lee, S., Lin, J., Dujmic, Z., Kim, W., Talag, J., Zuccolo, A., Fan, C., Sebastian, A., Kramer, M., Spiegel, L., Nascimento, L., Zutavern, T., Miller, B., Ambroise, C., Muller, S., Spooner, W., Narechania, A., Ren, L., Wei, S., Kumari, S., Faga, B., Levy, M.J., McMahan, L., Van Buren, P., Vaughn, M.W., Ying, K., Yeh, C.T., Emrich, S.J., Jia, Y., Kalyanaraman, A., Hsia, A.P., Barbazuk, W.B., Baucom, R.S., Brutnell, T.P., Carpita, N.C., Chaparro, C., Chia, J.M., Deragon, J.M., Estill, J.C., Fu, Y., Jeddeloh, J.A., Han, Y., Lee, H., Li, P., Lisch, D.R., Liu, S., Liu, Z., Nagel, D.H., McCann, M.C., Sanmiguel, P., Myers, A.M., Nettleton, D., Nguyen, J., Penning, B.W., Ponnala, L., Schneider, K.L., Schwartz, D.C., Sharma, A., Soderlund, C., Springer, N.M., Sun, Q., Wang, H., Waterman, M., Westerman, R., Wolfgruber, T.K., Yang, L., Yu, Y., Zhang, L., Zhou, S., Zhu, Q.,

Bennetzen, J.L., Dawe, R.K., Jiang, J., Jiang, N., Presting, G.G., Wessler, S.R., Aluru, S., Martienssen, R.A., Clifton, S.W., McCombie, W.R., Wing, R.A., Wilson, R.K., 2009. The B73 maize genome: Complexity, diversity, and dynamics. Science 326, 1112–1115. https://doi.org/10.1126/science.1178534

Schneider, A., Molnár, I., Molnár-Láng, M., 2008. Utilisation of *Aegilops* (goatgrass) species to widen the genetic diversity of cultivated wheat. Euphytica. https://doi.org/10.1007/s10681-007-9624-y

Schondelmaier, J., Martin, R., Jahoor, A., Houben, A., Graner, A., Koop, H.-U., Herrmann, R.G., Jung, C., 1993. Microdissection and microcloning of the barley (*Hordeum vulgare* L.) chromosome 1HS. Theor. Appl. Genet. 86, 629–636. https://doi.org/10.1007/BF00838719

Schulte, D., Close, T.J., Graner, A., Langridge, P., Matsumoto, T., Muehlbauer, G., Sato, K., Schulman, A.H., Waugh, R., Wise, R.P., Stein, N., 2009. The International Barley Sequencing Consortium - At the threshold of efficient access to the barley genome. Plant Physiol. https://doi.org/10.1104/pp.108.128967

Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D., Yu, T.H., 2008. Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. Science 319, 1238–1240. https://doi.org/10.1126/science.1151861

Sharma, H.C., Gill, B.S., Uyemoto, J.K., 1984. High Levels of Resistance in *Agropyron* Species to Barley Yellow Dwarf and Wheat Streak Mosaic Viruses. J. Phytopathol. 110, 143–147. https://doi.org/10.1111/j.1439-0434.1984.tb03402.x

Shatalina, M., Wicker, T., Buchmann, J.P., Oberhaensli, S., Šimková, H., Doležel, J., Keller, B., 2013. Genotype-specific SNP map based on whole chromosome 3B sequence information from wheat cultivars Arina and Forno. Plant Biotechnol. J. 11, 23–32. https://doi.org/10.1111/pbi.12003

Shoemaker, R.C., Specht, J.E., 1995. Integration of the soybean molecular and classical genetic linkage groups. Crop Sci. 35, 436–446. https://doi.org/10.2135/cropsci1995.0011183X003500020027x

Šimková, H., Číhalíková, J., Vrána, J., Lysák, M.A., Doležel, J., 2003. Preparation of HMW DNA from plant nuclei and chromosomes isolated from root tips. Biol. Plant. 46, 369–373. https://doi.org/10.1023/A:1024322001786

Šimková, H., Šafář, J., Suchánková, P., Kovářová, P., Bartoš, J., Kubaláková, M., Janda, J., Číhalíková, J., Mago, R., Lelley, T., Doležel, J., 2008a. A novel resource for

genomics of *Triticeae*: BAC library specific for the short arm of rye (*Secale cereale* L.) chromosome 1R (1RS). BMC Genomics 9, 237. https://doi.org/10.1186/1471-2164-9-237

Šimková, H., Svensson, J.T., Condamine, P., Hřibová, E., Suchánková, P., Bhat, P.R., Bartoš, J., Šafář, J., Close, T.J., Doležel, J., 2008b. Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genomics 9, 294. https://doi.org/10.1186/1471-2164-9-294

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. https://doi.org/10.1093/aob/mcm307

Smyth, D.R., Kalitsis, P., Joseph, J.L., Sentry, J.W., 1989. Plant retrotransposon from *Lilium henryi* is related to Ty3 of yeast and the gypsy group of *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 86, 5015–5019. https://doi.org/10.1073/pnas.86.13.5015

Soltis, D.E., Bell, C.D., Kim, S., Soltis, P.S., 2008. Origin and early evolution of angiosperms. Ann. N. Y. Acad. Sci. https://doi.org/10.1196/annals.1438.005

Soltis, P.S., 2013. Hybridization in Plants, in: Encyclopedia of Biodiversity: Second Edition. Elsevier Inc., pp. 166–176. https://doi.org/10.1016/B978-0-12-384719-5.00202-1

Song, L., Jiang, L., Han, H., Gao, A., Yang, X., Li, L., Liu, W., 2013. Efficient Induction of Wheat-*Agropyron cristatum* 6P Translocation Lines and GISH Detection. PLoS One 8. https://doi.org/10.1371/journal.pone.0069501

Soreng, R.J., Peterson, P.M., Romaschenko, K., Davidse, G., Zuloaga, F.O., Judziewicz, E.J., Filgueiras, T.S., Davis, J.I., Morrone, O., 2015. A worldwide phylogenetic classification of the *Poaceae* (*Gramineae*). J. Syst. Evol. https://doi.org/10.1111/jse.12150

Stubblefield, E., Oro, J., 1982. The isolation of specific chicken macrochromosomes by zonal centrifugation and flow sorting. Cytometry 2, 273–281. https://doi.org/10.1002/cyto.990020502

Studer, B., Byrne, S., Nielsen, R.O., Panitz, F., Bendixen, C., Islam, M.S., Pfeifer, M., Lübberstedt, T., Asp, T., 2012. A transcriptome map of perennial ryegrass (*Lolium perenne* L.). BMC Genomics 13, 1–13. https://doi.org/10.1186/1471-2164-13-140

Swift, H., 1950. The constancy of desoxyribose nucleic acid in plant nuclei. Proc. Natl. Acad. Sci. U. S. A. 36, 643–654. https://doi.org/10.1073/pnas.36.11.643

Taheri, S., Abdullah, T.L., Yusop, M.R., Hanafi, M.M., Sahebi, M., Azizi, P., Shamshiri,

R.R., 2018. Mining and development of novel SSR markers using Next Generation Sequencing (NGS) data in plants. Molecules. https://doi.org/10.3390/molecules23020399

Telenius, H., Carter, N.P., Bebb, C.E., Nordenskjöld, M., Ponder, B.A.J., Tunnacliffe, A., 1992. Degenerate oligonucleotide-primed PCR: General amplification of target DNA by a single degenerate primer. Genomics 13, 718–725. https://doi.org/10.1016/0888-7543(92)90147-K

The Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796–815. https://doi.org/10.1038/35048692

Thiel, T., Michalek, W., Varshney, R.K., Graner, A., 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor. Appl. Genet. 106, 411–422. https://doi.org/10.1007/s00122-002-1031-0

Thomas, C.A., 1971. The genetic organization of chromosomes. Annu. Rev. Genet. https://doi.org/10.1146/annurev.ge.05.120171.001321

Thomas, H.M., Morgan, W.G., Humphreys, M.W., 2003. Designing grasses with a future - Combining the attributes of *Lolium* and *Festuca*. Euphytica. https://doi.org/10.1023/A:1025694819031

Timmerman-Vaughan, G.M., McCallum, J.A., Frew, T.J., Weeden, N.F., Russell, A.C., 1996. Linkage mapping of quantitative trait loci controlling seed weight in pea (*Pisum sativum* L.). Theor. Appl. Genet. 93, 431–439. https://doi.org/10.1007/BF00223187

Tiwari, V.K., Wang, S., Sehgal, S., Vrána, J., Friebe, B., Kubaláková, M., Chhuneja, P., Doležel, J., Akhunov, E., Kalia, B., Sabir, J., Gill, B.S., 2014. SNP Discovery for mapping alien introgressions in wheat. BMC Genomics 15, 273. https://doi.org/10.1186/1471-2164-15-273

Tuskan, G.A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, M., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R.R., Bhalerao, R.P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., Busov, V., Campbell, M., Carlson, J., Chalot, M., Chapman, J., Chen, G.L., Cooper, D., Coutinho, P.M., Couturier, J., Covert, S., Cronk, Q., Cunningham, R., Davis, J., Degroeve, S., Dŕjardin, A., DePamphilis, C., Detter, J., Dirks, B., Dubchak, I., Duplessis, S., Ehlting, J., Ellis, B., Gendler, K., Goodstein, D., Gribskov, M.,

Grimwood, J., Groover, A., Gunter, L., Hamberger, B., Heinze, B., Helariutta, Y., Henrissat, B., Holligan, D., Holt, R., Huang, W., Islam-Faridi, N., Jones, S., Jones-Rhoades, M., Jorgensen, R., Joshi, C., Kangasjärvi, J., Karlsson, J., Kelleher, C., Kirkpatrick, R., Kirst, M., Kohler, A., Kalluri, U., Larimer, F., Leebens-Mack, J., Leplé, J.C., Locascio, P., Lou, Y., Lucas, S., Martin, F., Montanini, B., Napoli, C., Nelson, D.R., Nelson, C., Nieminen, K., Nilsson, O., Pereda, V., Peter, G., Philippe, R., Pilate, G., Poliakov, A., Razumovskaya, J., Richardson, P., Rinaldi, C., Ritland, K., Rouzé, P., Ryaboy, D., Schmutz, J., Schrader, J., Segerman, B., Shin, H., Siddiqui, A., Sterky, F., Terry, A., Tsai, C.J., Uberbacher, E., Unneberg, P., Vahala, J., Wall, K., Wessler, S., Yang, G., Yin, T., Douglas, C., Marra, M., Sandberg, G., Van De Peer, Y., Rokhsar, D., 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313, 1596–1604. https://doi.org/10.1126/science.1128691

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3—new capabilities and interfaces. Nucleic Acids Res. 40, e115–e115. https://doi.org/10.1093/nar/gks596

Van De Peer, Y., Maere, S., Meyer, A., 2009. The evolutionary significance of ancient genome duplications. Nat. Rev. Genet. https://doi.org/10.1038/nrg2600

Varshney, R.K., Sinha, P., Singh, V.K., Kumar, A., Zhang, Q., Bennetzen, J.L., 2020. 5Gs for crop genetic improvement. Curr. Opin. Plant Biol. https://doi.org/10.1016/j.pbi.2019.12.004

Varshney, R.K., Song, C., Saxena, R.K., Azam, S., Yu, S., Sharpe, A.G., Cannon, S., Baek, J., Rosen, B.D., Tar'an, B., Millan, T., Zhang, X., Ramsay, L.D., Iwata, A., Wang, Y., Nelson, W., Farmer, A.D., Gaur, P.M., Soderlund, C., Penmetsa, R.V., Xu, C., Bharti, A.K., He, W., Winter, P., Zhao, S., Hane, J.K., Carrasquilla-Garcia, N., Condie, J.A., Padhyaya, H.D., Luo, M.C., Thudi, M., Gowda, C.L.L., Singh, N.P., Lichtenzveig, J., Gali, K.K., Rubio, J., Nadarajan, N., Dolezel, J., Bansal, K.C., Xu, X., Edwards, D., Zhang, G., Kahl, G., Gil, J., Singh, K.B., Datta, S.K., Jackson, S.A., Wang, J., Cook, D.R., 2013. Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nat. Biotechnol. 31, 240–246. https://doi.org/10.1038/nbt.2491

Vershinin, A. V., Alkhimova, E.G., Heslop-Harrison, J.S., 1996. Molecular diversification of tandemly organized DNA sequences and heterochromatic chromosome regions in some *Triticeae* species. Chromosom. Res. 4, 517–525.

https://doi.org/10.1007/BF02261779

Wang, H., Dai, K., Xiao, J., Yuan, C., Zhao, R., Doležel, J., Wu, Y., Cao, A., Chen, P., Zhang, S., Wang, X., 2017. Development of intron targeting (IT) markers specific for chromosome arm 4VS of *Haynaldia villosa* by chromosome sorting and next-generation sequencing. BMC Genomics 18. https://doi.org/10.1186/s12864-017-3567-z

Wang, M.L., Leitch, A.R., Schwarzacher, T., Heslop-Harrison, J.S., Moore, G., 1992. Construction of a chromosome-enriched *Hpa*II library from flow-sorted wheat chromosomes. Nucleic Acids Res. 20, 1897–1901. https://doi.org/10.1093/nar/20.8.1897

Wang, S., Yin, L., Tanaka, H., Tanaka, K., Tsujimoto, H., 2011. Wheat-*Aegilops* chromosome addition lines showing high iron and zinc contents in grains. Breed. Sci. 61, 189–195. https://doi.org/10.1270/jsbbs.61.189

Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinhofs, A., Kilian, A., 2004. Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proc. Natl. Acad. Sci. U. S. A. 101, 9915–9920. https://doi.org/10.1073/pnas.0401076101

Wicker, T., Keller, B., 2007. Genome-wide comparative analysis of copia retrotransposons in *Triticeae*, rice, and *Arabidopsis* reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. Genome Res. 17, 1072–1081. https://doi.org/10.1101/gr.6214107

Wicker, T., Taudien, S., Houben, A., Keller, B., Graner, A., Platzer, M., Stein, N., 2009. A whole-genome snapshot of 454 sequences exposes the composition of the barley genome and provides evidence for parallel evolution of genome size in wheat and barley. Plant J. 59, 712–722. https://doi.org/10.1111/j.1365-313X.2009.03911.x

Willard, H.F., 1990. Centromeres of mammalian chromosomes. Trends Genet. 6, 410–416. https://doi.org/10.1016/0168-9525(90)90302-M

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., Tingey, S. V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18, 6531–6535. https://doi.org/10.1093/nar/18.22.6531

Williams, R.C., 1989. Restriction fragment length polymorphism (RFLP). Am. J. Phys. Anthropol. 32, 159–184. https://doi.org/10.1002/ajpa.1330320508

Wu, J., Yang, X., Wang, H., Li, H., Li, L., Li, X., Liu, W., 2006. The introgression of chromosome 6P specifying for increased numbers of florets and kernels from

Agropyron cristatum into wheat. Theor. Appl. Genet. 114, 13–20. https://doi.org/10.1007/s00122-006-0405-0

Xiao, J., Dai, K., Fu, L., Vrána, J., Kubaláková, M., Wan, W., Sun, H., Zhao, J., Yu, C., Wu, Y., Abrouk, M., Wang, H., Doležel, J., Wang, X., 2017. Sequencing flow-sorted short arm of *Haynaldia villosa* chromosome 4V provides insights into its molecular structure and virtual gene order. BMC Genomics 18. https://doi.org/10.1186/s12864-017-4211-7

Yang, W., Liu, D., Li, J., Zhang, L., Wei, H., Hu, X., Zheng, Y., He, Z., Zou, Y., 2009. Synthetic hexaploid wheat and its utilization for wheat genetic improvement in China. J. Genet. Genomics 36, 539–546. https://doi.org/10.1016/S1673-8527(08)60145-9

Zhang, Q., Li, Q., Wang, X., Wang, H., Lang, S., Wang, Y., Wang, S., Chen, P., Liu, D., 2005. Development and characterization of a *Triticum aestivum-Haynaldia villosa* translocation line T4VS·4DL conferring resistance to wheat spindle streak mosaic virus. Euphytica 145, 317–320. https://doi.org/10.1007/s10681-005-1743-8

Zwyrtková, J., Němečková, A., Čížková, J., Holušová, K., Kapustová, V., Svačina, R., Kopecký, D., Till, B.J., Doležel, J., Hřibová, E., 2020. Comparative analyses of DNA repeats and identification of a novel Fesreba centromeric element in fescues and ryegrasses. BMC Plant Biol. 20, 280. https://doi.org/10.1186/s12870-020-02495-0

Zwyrtková, J., Šimková, H., Doležel, J., 2021. Chromosome genomics uncovers plant genome organization and function. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2020.107659

http://www.fao.org/faostat/en/#rankings/countries_by_commodity
[accessed 15. 2. 2021]

http://spuds.agron.ksu.edu/ksgrasskey/images/Agropyroncristatum.htm
[accessed 15. 2. 2021]

http://www.wheatgenome.org
[accessed 15. 2. 2021]

8 Papers and posters not included in the Thesis

These papers and posters are not included in the Appendices of the Thesis since they are not focused on its main topic.

Original papers

Kopecký, D., Talukder, S.K., Zwyrtková, J., Trammell, M., Doležel, J., Saha, M.C., 2019. Inter-morphotype hybridization in tall fescue (*Festuca arundinacea* Schreb.): exploration of meiotic irregularities and potential for breeding. Euphytica 215. https://doi.org/10.1007/s10681-019-2419-0
IF_{2019}: 1.614

Nowicka, A., Tokarz, B., Zwyrtková, J., Dvořák Tomaštíková, E., Procházková, K., Ercan, U., Finke, A., Rozhon, W., Poppenberger, B., Otmar, M., Niezgodzki, I., Krečmerová, M., Schubert, I., Pecinka, A., 2020. Comparative analysis of epigenetic inhibitors reveals different degrees of interference with transcriptional gene silencing and induction of DNA damage. Plant J. 102, 68–84. https://doi.org/10.1111/tpj.14612
IF_{2019}: 6.141

Makowski, W., Królicka, A., Nowicka, A., Zwyrtková, J., Tokarz, B., Pecinka, A., Banasiuk, R., Tokarz, K.M., 2021. Transformed tissue of *Dionaea muscipula* J. Ellis as a source of biologically active phenolic compounds with bactericidal properties. Appl. Microbiol. Biotechnol.
https://doi.org/10.1007/s00253-021-11101-8
IF_{2019}: 3.530

Published abstract – poster presentation

Abrouk, M., Zwyrtková, J., Christelová, P., Rey, E., Komínková, E., Jakobson, I., Vrána, J., Järve, K., Valárik, M., Doležel, J.: Characterization of a recombination hotspot in wheat using flow-sorted pollen nuclei and digital PCR. – In: Abstracts of the International Conference "Plant and Animal Genome XXV". P0843. Scherago International, Inc., San Diego, California, USA, 2017.

9 List of appendices

Original papers

Appendix I Comparative analyses of DNA repeats and identification of a novel Fesreba centromeric element in fescues and ryegrasses

Appendix II Chromosome genomics uncovers plant genome organization and function

Conference presentations

Appendix III Comparative analysis of repetitive parts of genomes in eight *Loliinae* representatives (poster presentation)

Appendix IV Comparative analysis of repetitive DNA in eight representatives of fescues and ryegrasses (poster presentation)

APPENDIX I

Comparative analyses of DNA repeats and identification of a novel Fesreba centromeric element in fescues and ryegrasses

Zwyrtková, J., Němečková, A., Čížková, J., Holušová, K., Kapustová, V., Svačina, R., Kopecký, D., Till, B. J., Doležel, J., Hřibová, E.

BMC Plant Biology, 20: 280, 2020

doi: 10.1186/s12870-020-02495-0

IF_{2019}: 3.497

Zwyrtková et al. BMC Plant Biology (2020) 20:280
https://doi.org/10.1186/s12870-020-02495-0

BMC Plant Biology

RESEARCH ARTICLE Open Access

Comparative analyses of DNA repeats and identification of a novel Fesreba centromeric element in fescues and ryegrasses

Jana Zwyrtková[1], Alžběta Němečková[1], Jana Čížková[1], Kateřina Holušová[1], Veronika Kapustová[1], Radim Svačina[1], David Kopecký[1], Bradley John Till[2], Jaroslav Doležel[1] and Eva Hřibová[1*]

Abstract

Background: Cultivated grasses are an important source of food for domestic animals worldwide. Increased knowledge of their genomes can speed up the development of new cultivars with better quality and greater resistance to biotic and abiotic stresses. The most widely grown grasses are tetraploid ryegrass species (*Lolium*) and diploid and hexaploid fescue species (*Festuca*). In this work, we characterized repetitive DNA sequences and their contribution to genome size in five fescue and two ryegrass species as well as one fescue and two ryegrass cultivars.

Results: Partial genome sequences produced by Illumina sequencing technology were used for genome-wide comparative analyses with the RepeatExplorer pipeline. Retrotransposons were the most abundant repeat type in all seven grass species. The Athila element of the Ty3/gypsy family showed the most striking differences in copy number between fescues and ryegrasses. The sequence data enabled the assembly of the long terminal repeat (LTR) element Fesreba, which is highly enriched in centromeric and (peri)centromeric regions in all species. A combination of fluorescence in situ hybridization (FISH) with a probe specific to the Fesreba element and immunostaining with centromeric histone H3 (CENH3) antibody showed their co-localization and indicated a possible role of Fesreba in centromere function.

Conclusions: Comparative repeatome analyses in a set of fescues and ryegrasses provided new insights into their genome organization and divergence, including the assembly of the LTR element Fesreba. A new LTR element Fesreba was identified and found in abundance in centromeric regions of the fescues and ryegrasses. It may play a role in the function of their centromeres.

Keywords: *Festuca*, *Lolium*, Illumina sequencing, Repetitive DNA, Centromere organization

* Correspondence: hribova@ueb.cas.cz
[1]Institute of Experimental Botany, Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, CZ-77900 Olomouc, Czech Republic
Full list of author information is available at the end of the article

Background

Grasses (Poaceae) are an important source of food for domestic animals worldwide and perform important ecological and environmental functions. The tribe Poeae is the largest tribe in family Poaceae, and species from its largest subtribe, Loliinae, grow in a range of habitats, including wetlands, dry areas, and regions with cold and temperate climates; some are even well adapted to the extreme conditions of mountain, arctic, and sub-Antarctic regions [1]. The subtribe Loliinae comprises a cosmopolitan genus *Festuca* and its satellite genera [2, 3]. *Festuca* is the largest genus of the family Poaceae, containing more than 600 species, and Torrecilla and Catalán [4] discriminated its two main evolutionary lines: broad leaved and fine leaved (Fig. 1). Broad-leaved *Festuca* species (hereafter "fescues") include the subgenus Schedonorus, which gave rise to *Lolium* species (hereafter "ryegrasses"), a sister group of fescues (Fig. 1) [1]. The evolution of grasses, including *Loliinae*, has been accompanied by frequent polyploidization and hybridization events, and about 70% of grass species are polyploid [6]. The species of Loliinae have large genomes ranging from 2.6 Gbp/1C to 11.8 Gbp/1C [7, 8].

This study focuses on species from the subgenus Schedonorus, a complex of species with various ploidy levels [7, 9] that includes important species widely used for forage and turf. Although some Schedonorus species are diploid, such as *Festuca pratensis* Huds. (2n = 2x = 14) and *Lolium multiflorum* Lam. (2n = 2x = 14) and *L. perenne* L. (2n = 2x = 14), the majority of species are allopolyploid [10, 11], including tetraploids *F. glaucescens* Boiss. (2n = 4x = 28) and *F. mairei* St. Yves (2n = 4x = 28) and hexaploids *F. arundinacea* Schreb. (2n = 6x = 42) and *F. gigantea* (L.) Vill. (2n = 6x = 42) [3, 11]. Fescues are more tolerant than ryegrasses of abiotic stresses, provide high-quality forage for livestock, and are grown especially for turf purposes. In contrast, ryegrasses are characterized by high yield and excellent nutritional value and are mostly cultivated as pasture. Artificial intergeneric hybrids of fescue and ryegrass species have been developed to combine the most favorable characteristics of both genera [12–14].

Although fescues and ryegrasses have been intensively studied, their evolution and the origin of most polyploid representatives remain obscure [11, 15, 16]. Like in other species with large genomes, the nuclear genomes of fescues and ryegrasses include a large number and variety of repetitive DNA sequences [17, 18]. Their amplification in the genome, accompanied by interspecific hybridization and polyploidization, has expanded the genome size [19–24]. However, these processes have likely been counterbalanced by recombination-based mechanisms that have removed substantial parts of nuclear genomes [25–27].

Repetitive DNA elements may play different roles in a nuclear genome. Tandem organized ribosomal RNA genes and telomeric sequences are the key components of nucleolar organizing regions and chromosome

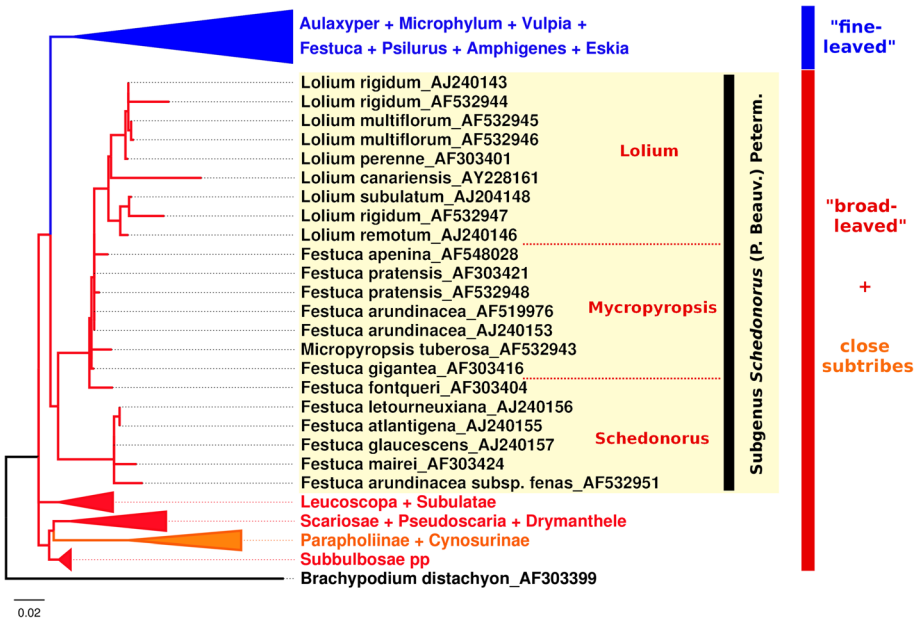


Fig. 1 Phylogenetic tree of *Loliinae* subtribe. Phylogeny of subtribe Loliinae with *Brachypodium distachyon* was used as an outgroup. The tree was constructed from ITS sequence regions of *Loliinae* species and *B. distachyon* with PhyML implemented in SeaView [5]. Detailed phylogeny of subgenus Schedonorus is depicted and shows the relationships of fescue and ryegrass species in this lineage (highlighted in light yellow)

termini, respectively. Centromeric regions in *Arabidopsis*, *Brachypodium*, rice, and maize are partly formed from specific satellite DNAs with ~ 130 bp long units [28–31], whereas in other plant species, including cereals, these regions are formed by large blocks of Ty3/gypsy retrotransposons containing chromodomain [29, 32–34]. In *F. pratensis*, a putative long terminal repeat (LTR) element localizing preferentially to centromeric regions has been identified [35]. In addition to elucidating the molecular organization of chromosome domains, characterization of repetitive parts of nuclear genomes helps in the development of cytogenetic markers [21, 35, 36]. Repetitive DNA sequences are also used extensively to study genetic diversity and processes of genome evolution and speciation [37–40].

The main goal of the present work was to elucidate the repetitive landscape and its impact on genome size and genome divergence in closely related land grasses, including natural polyploid species. We characterized repetitive DNA sequences in the nuclear genomes of 10 representatives of fescues and ryegrasses. We performed global analyses of repetitive DNA sequences and characterized their abundance and variability after partial Illumina sequencing. Moreover, we characterized and assembled the DNA sequence of an LTR element that was highly enriched in centromeric and (peri)centromeric chromosome regions in all 10 genotypes. Co-localization of the centromere-specific histone H3 variant CENH3 with the LTR element indicated its role in centromere function.

Results

Genome size estimation

Flow cytometric analysis of propidium iodide–stained nuclei was conducted to estimate nuclear DNA content (Fig. 2). Because of the large differences in genome size between the species analyzed, two internal reference standards were used: *Pisum sativum* cv. Ctirad (2C = 9.09 pg DNA) [41] and *Secale cereale* cv. Dankovske (2C = 16.19 pg DNA) [41]. All histograms of relative DNA content represented two dominant peaks corresponding to G1 nuclei of the sample and the standard. The 2C nuclear DNA content thus determined ranged from 5.32 pg in *L. multiflorum* to 20.17 pg in *F. gigantea*. The monoploid genome (1Cx) ranged in size from 2.43 in *F. mairei* to 3.36 pg in *F. gigantea* (Table 1). The remaining representatives of fescues and ryegrasses had similar 1Cx sizes (~ 2.7 Gb).

Repeat composition and comparative analyses of repetitive DNA sequences

Interspecific comparisons, reconstruction, and quantification of major repeat families were performed with the RepeatExplorer pipeline [42]. The process, which involved grouping orthologous repeat families from all analyzed species in the same cluster, facilitated the assembly, identification, and quantification of individual repeat elements.

In all accessions, LTR retroelements were the most abundant component of the nuclear genome (Table 2, Fig. 3). Ty3/gypsy elements were more than 4 times more abundant than Ty1/copia retrotransposons (Table 2). The biggest difference in copy number between fescues and ryegrasses was for an LTR element from the Athila clade. Whereas the nuclear genomes of both *Lolium* species were enriched for the element, which accounted for ~ 25–30% of their genomes, the orthologous Athila element accounted for only ~ 5–7% of the nuclear genomes of fescues (Table 2). A relatively large part of the genomes was represented by unclassified LTR sequences, which indicates a high frequency of unique LTR sequences. DNA transposons and long interspersed nuclear element (LINE) elements were found in low copy numbers, and tandem repeats accounted for 1.5% to more than 8% of the genome sequences (Table 2, Fig. 3).

Comparative analyses with RepeatExplorer showed that most clusters of orthologous repeat families contained reads from all accessions and that a large number of similar sequences were present in fescues and ryegrasses. Among the fescues, *F. mairei* and *F. glaucescens* showed the lowest similarity in DNA repeats. The composition as well as the abundance of DNA repeats in ryegrasses were highly conserved. Tandem organized repeats were the most diverged elements among the fescues and ryegrasses studied, and some of the repeats were species specific (Fig. 4, Additional file 1: Table S1). In addition to tandem repeats, some small sequence clusters contained reads from only a few species. Species-specific variants of the majority of repetitive elements within and between fescues and ryegrasses were identified only after detailed analyses of individual repeat clusters with SeqGrapheR (Fig. 5a–c). Detailed analyses revealed the presence of species-specific DNA contigs, which may be used to develop molecular and cytogenetic markers.

To confirm the differences determined in silico, we analyzed selected repetitive DNA elements using Southern hybridization. We designed specific probes for those DNA repeats that seemed to have species-specific variants. A probe for the Ty3/gypsy Athila element that was reconstructed in cluster CL1 and showed the largest copy number variation between fescues and ryegrasses (Table 2) gave strong hybridization signals on genomic DNA from ryegrasses but no or weak signals on DNA from fescues (Fig. 5d). Similarly, a probe for the Ty3/gypsy Athila element that was reconstructed in cluster

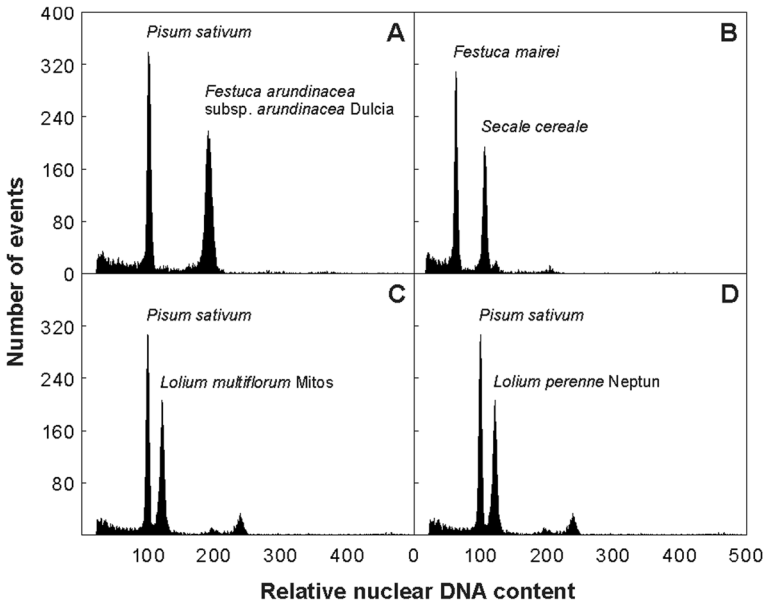


Fig. 2 Estimation of nuclear genome size. Histograms of propidium iodide–stained nuclei corresponding to relative nuclear DNA content obtained after flow cytometric analyses of *F. arundinacea* subsp. *arundinacea* (**a**), *F. mairei* (**b**), *Lolium multiflorum* cv. Mitos (**c**), and *L. perenne* cv. Neptun (**d**). *Pisum sativum* cv. Ctirad (2C = 9.09 pg) and *Secale cereale* cv. Dankovske (2C = 16.19 pg), respectively, were used as internal reference standards. The ratio of relative G1 peak positions was used to calculate DNA amounts of the fescue and ryegrass accessions

CL38 and contained mostly *Festuca* sequence reads (Fig. 5b) provided strong visible signals only with fescue genomic DNA (Fig. 5e). Finally, Southern hybridization was performed with a probe for the Ty3/gypsy Ogre-Tat retrotransposon, identified in cluster CL20. The probe, which was designed from contigs representing fescues (Fig. 5c), provided strong hybridization signals on all fescues analyzed and low intensity signals on ryegrasses (Fig. 5f). In general, the signal intensities obtained after Southern hybridization corresponded to the copy numbers identified in silico.

Centromere composition

Partial genome sequence data obtained using Illumina sequencing technology made it possible to reconstruct nearly complete centromeric LTR elements in all 10 accessions of fescues and ryegrasses. Detailed characterization of the element called Fesreba confirmed that it belongs to the Ty3/gypsy Chromoviridae lineage. Phylogenetic analyses of its reverse transcriptase (RT) domain showed a close relationship with the Cereba element (Fig. 6), which was identified earlier in barley (*Hordeum vulgare*) [43].

Southern hybridization with a probe for the RT domain of Fesreba and a probe for its LTR region [35] showed their presence in all fescues and ryegrasses included in this work (Additional file 2: Fig. S1). Similar hybridization patterns indicated sequence conservation between Fesreba repetitive DNA elements in these species. The results were supported by in silico data, which showed high similarity at the DNA sequence level (most abundant copies of Fesreba shared at least 92% similarity at the DNA level within and between fescues and ryegrasses) but lower abundance in ryegrasses. To confirm the differences in Fesreba copy number, we performed quantification for the RT domain and LTR sequence using droplet digital polymerase chain reaction (ddPCR). The results confirmed a two-fold higher copy number of Fesreba in fescues compared to ryegrasses (Additional file 3: Table S2). The assay also showed that the majority of genotypes analyzed contained 5 to 50 times more copies of the LTR region of Fesreba than its coding region (Additional file 3: Table S2).

To confirm preferential localization of Fesreba to centromeric chromosome regions, we conducted fluorescence in situ hybridization (FISH) on mitotic metaphase plates with probes derived from its RT domain and LTR region. In all fescues and ryegrasses, both probes localized preferentially to centromeric regions of all chromosomes (Fig. 7). Whereas the hybridization signals of the RT domain were observed almost exclusively in centromeric regions, a probe derived from the non-coding LTR region resulted in stronger signals in centromeric and/or pericentromeric regions and weak signals along the chromosomal arms, as shown previously in *F. pratensis* [35]. Weak signals of the LTR part of Fesreba

Table 1 Flow cytometric estimation of nuclear genome size

Species	Accession name	Code	Ploidy level	2C nuclear DNA content		Monoploid genome size (1Cx)	
				Mean [pg]	± SD	[pg]	[Mbp]
Festuca pratensis	Fure	FPF	2n = 2x = 14	6.4	0.04	3.2	3130
Festuca pratensis	Westa	FPW	2n = 4x = 28	12.79	0.09	3.2	3127
Festuca arundinacea ssp. arundinacea	Dulcia	FAR	2n = 6x = 42	16.85	0.24	2.81	2747
Festuca arundinacea ssp. glaucescens	–	FGL	2n = 4x = 28	10.79	0.07	2.7	2638
Festuca gigantea	GR 11759	FGI	2n = 6x = 42	20.17	0.14	3.36	3288
Festuca mairei	GR 610941	FMA	2n = 4x = 28	9.73	0.05	2.43	2379
Lolium multiflorum	Lm2	LM2	2n = 2x = 14	5.32	0.03	2.66	2601
Lolium multiflorum	Mitos	LMM	2n = 4x = 28	11.13	0.05	2.78	2721
Lolium perenne	GR 3320	LP2	2n = 2x = 14	5.54	0.03	2.77	2709
Lolium perenne	Neptun	LPN	2n = 4x = 28	10.94	0.15	2.74	2675

in distal parts of chromosomes indicate the presence of unique LTRs spread over the genome and correspond to a higher copy number of the LTR non-coding part of Fesreba compared to its coding sequence.

In addition to the fescues and ryegrasses included in this study, FISH was performed with the same probes on mitotic metaphase plates from related grass species, oat, barley, rye, bread wheat, and *Aegilops tauschii*. High homology of the RT coding domain resulted in successful in situ localization in all species. However, the probe specific to the LTR region of Fesreba provided visible signals only in *A. sativa* (Additional file 4: Fig. S2). Finally, immunostaining with the centromere-specific histone H3 variant CENH3 [44] in combination with FISH with probes for the RT domain and LTR region of Fesreba resulted in overlapping signals in all fescues and ryegrasses studied (Fig. 8, Additional file 5: Fig. S3).

Discussion

Because of genome shock, the 1Cx size of polyploid species is often, but not always, lower than that of their progenitors [25, 45]. In this study, we performed comparative analyses of repeatomes and analyzed the impact of DNA repeats on genome size in a set of *Festuca* and

Table 2 Proportion of repetitive DNA sequences identified de novo

Repeat		Lineage/class	Proportion of repeat in monoploid genomes [%]									
			FPF	FPW	FAR	FGI	FGL	FMA	LM2	LMM	LP2	LPN
LTR retroelements	***Ty1/Copia***	Maximus-SIRE	1.72	1.65	1.69	1.78	1.84	1.93	0.89	0.87	1.16	1.25
		Angela	4.43	4.53	3.33	4.86	2.83	2.54	3.63	3.32	4.52	4.13
		TAR (Tont)	0.3	0.27	0.28	0.30	0.31	0.34	0.28	0.25	0.24	0.25
		Tork (Tnt)	0.05	0.04	0.05	0.05	0.05	0.06	0.07	0.07	0.08	0.07
		Ale (Hopscotch)	0.1	0.07	0.07	0.07	0.04	0.03	0.22	0.22	0.14	0.14
		Ivana-Oryoco	0.05	0.05	0.03	0.07	0.02	0.02	0.03	0.02	0.01	0.02
		Total *Ty1/Copia*	**6.65**	**6.61**	**5.45**	**7.13**	**5.09**	**4.92**	**5.12**	**4.75**	**6.15**	**5.86**
	Ty3/Gypsy	Athila	6.32	6.88	6.73	6.02	4.96	5.56	25.69	23.54	30.33	24.4
		Chromovirideae	9.6	9.57	7.97	7.40	7.35	6.17	7.11	6.63	7.49	6.97
		Ogre-Tat	12.61	12.03	8.65	8.40	6.76	4.22	5.10	5.20	5.83	6.68
		Total *Ty3/Gypsy*	**28.53**	**28.48**	**23.35**	**21.82**	**19.07**	**15.95**	**37.90**	**35.37**	**43.65**	**38.05**
Unclassified LTR elements			**5.51**	**5.15**	**6.35**	**4.43**	**7.14**	**5.35**	**4.55**	**4.14**	**5.54**	**5.15**
Other	**LINE**		**0.26**	**0.27**	**0.29**	**0.37**	**0.27**	**0.23**	**0.34**	**0.31**	**0.20**	**0.23**
	DNA transposons		**2.35**	**2.16**	**1.95**	**1.81**	**1.44**	**1.45**	**2.38**	**2.25**	**2.08**	**2.15**
	Tandem repeats		**5.52**	**5.53**	**3.41**	**14.63**	**2.55**	**3.63**	**8.67**	**9.86**	**4.20**	**4.99**
	rRNA genes		**1.13**	**1.07**	**0.57**	**0.50**	**0.43**	**0.56**	**1.48**	**2.03**	**1.23**	**2.10**
Unclassified repeats			**13.79**	**13.94**	**10.82**	**12.76**	**9.39**	**8.29**	**10.04**	**9.86**	**8.51**	**9.02**

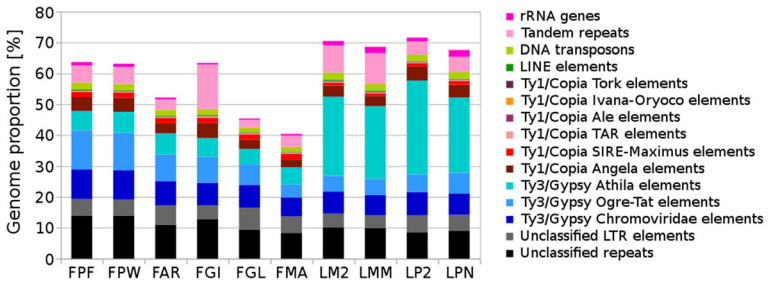
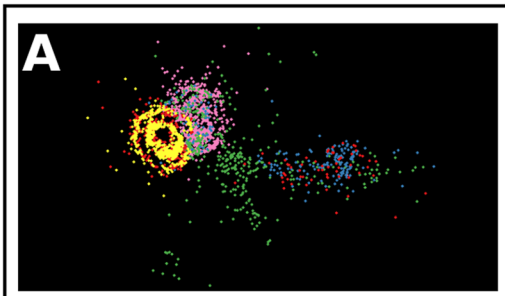


Fig. 3 Genome proportion of the most abundant DNA repeats. The genome proportion of individual repeat types was obtained as a ratio of reads specific to individual repeat types to all reads used for clustering analyses by the RepeatExplorer pipeline. Diploid *Festuca pratensis* cv. Fure (FPF); tetraploid *F. pratensis* cv. Westa (FPW); hexaploid *F. arundinacea* subsp. *arundinacea* (FAR); hexaploid *F. gigantea* (FGI); tetraploid *F. glaucescens* (FGL); tetraploid *F. mairei* (FMA); diploid cv. Kuri1 of *L. multiflorum* (LM2); tetraploid cv. Mitos of *L. multiflorum* (LMM); diploid *L. perenne* (LP2); tetraploid cv. Neptun of *L. perenne* (LPN)

Lolium species differing in ploidy. The set comprised hexaploids *F. arundinacea* subsp. *arundinacea* and *F. gigantea*; tetraploids *F. glaucescens* and *F. mairei*; and artificial autotetraploids *F. pratensis* cv. Westa, *L. multiflorum* cv. Mitos, and *L. perenne* cv. Neptun developed in breeding programs. We estimated nuclear DNA amounts using flow cytometry, and a test of normality confirmed that the data set had a normal distribution. Our results suggest possible genome changes in hexaploid *F. arundinacea* and tetraploid ryegrasses compared to their probable progenitors. Although the differences in the 1Cx size of natural polyploid *F. arundinacea* and its probable parents (*F. pratensis* and *F. glaucescens*) are small, they are statistically significant ($P < 0.01$). The same is true for tetraploid ryegrass cultivars obtained after polyploidization. Genome downsizing was detected in the case of *F. arundinacea* (~ 2% difference between expected and estimated values) and tetraploid *L. perenne* (~ 1% decrease). In the tetraploid cultivar of *L. multiflorum*, a slight increase in genome size (~ 4%) was



Accession name	Prefix	Colour	Genome (1Cx) proportion
F. pratensis (2x)	FPF	Pink	0.11 %
F. pratensis (4x)	FPW	Pink	0.07 %
F. gigantea (6x)	FGI	Pink	0.09 %
F. arundinacea (6x)	FAR	Blue	0.09 %
F. mairei (4x)	FMA	Green	0.20 %
F. glaucescens (4x)	FGL	Red	Not present as tandem repeat
L. perenne (2x)	LP2	Yellow	0.06 %
L. perenne (4x)	LPN	Yellow	0.13 %
L. multiflorum (2x)	LM2	Yellow	0.01 %
L. multiflorum (4x)	LMM	Yellow	0.09 %

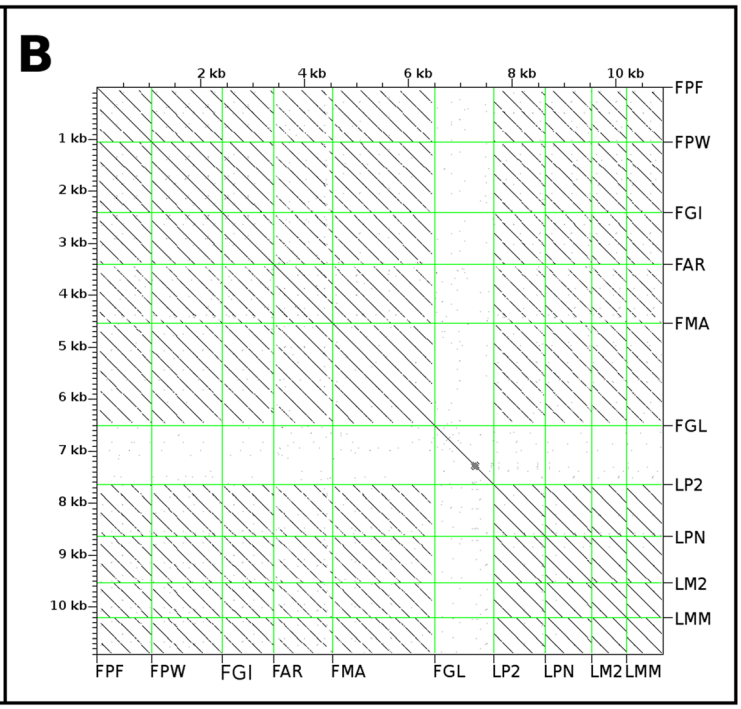


Fig. 4 Tandem organized repeat sequences identified in cluster CL102. **a** Graphical layout of cluster CL102. **b** Dot-plot analyses show the presence of homologous tandem organized units (parallel lines) of DNA repeats identified in cluster CL102 in all species except *F. glaucescens*, in which the assembled sequence contigs did not represent tandem organized sequences

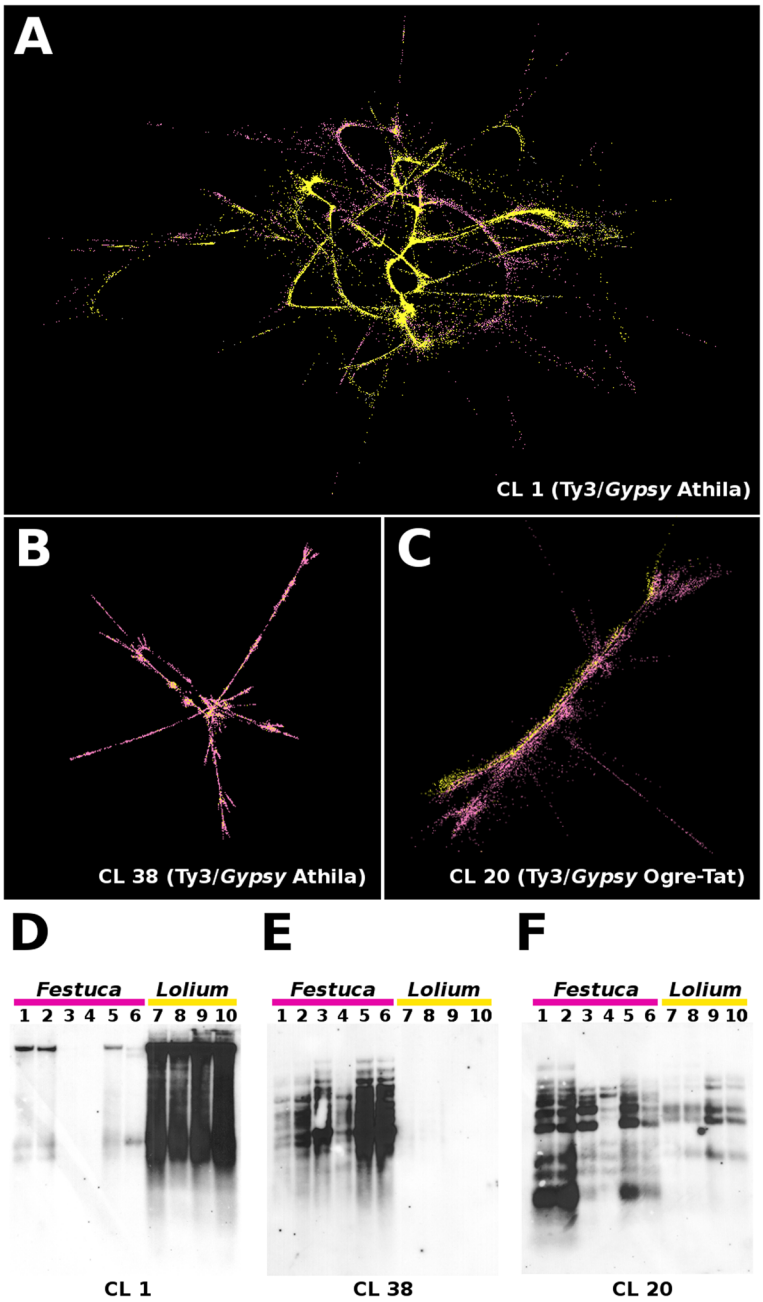


Fig. 5 Graphical layouts of selected DNA repeats and their validation by Southern blotting. Graphical layouts were obtained after clustering analyses done by RepeatExplorer. **a** Cluster CL1. **b** Cluster CL38 containing the Ty3/gypsy Athila element. **c** Cluster CL20 containing Ty3/gypsy Ogre/Tat elements. Sequencing reads from *Festuca* species are in pink, whereas sequencing reads from *Lolium* species are in yellow. **d**–**f** Validation of clustering results by Southern hybridization with sequences derived from clusters CL1, CL38, and CL20. Lanes contained genomic DNA digested by *Hae*III restriction endonuclease. Lane 1: *F. pratensis* cv. Fure (2n = 2x = 14); lane 2: *F. pratensis* cv. Westa (2n = 4x = 28); lane 3: *F. arundinacea* subsp. *arundinacea* (2n = 6x = 42); lane 4: *F. gigantea* (2n = 6x = 42); lane 5: *F. glaucescens* (2n = 4x = 28); lane 6: *F. mairei* (2n = 4x = 28); lane 7: *L. multiflorum* cv. Mitos (2n = 4x = 28); lane 8: *L. multiflorum* cv. Kuri1 (2n = 2x = 14); lane 9: *L. perenne* cv. Neptun (2n = 4x = 28); lane 10: *L. perenne* GR 3320 (2n = 2x = 14)

detected, corresponding with Kopecký et al. [8]. In the case of tetraploid fescue cultivars obtained after polyploidization, no statistically significant difference in 1Cx value was found ($P > 0.01$).

DNA retrotransposons are major contributors to the variation in nuclear genomes in plants [24, 46, 47]. Various approaches and tools have been developed to study these important parts of nuclear genomes, one of them

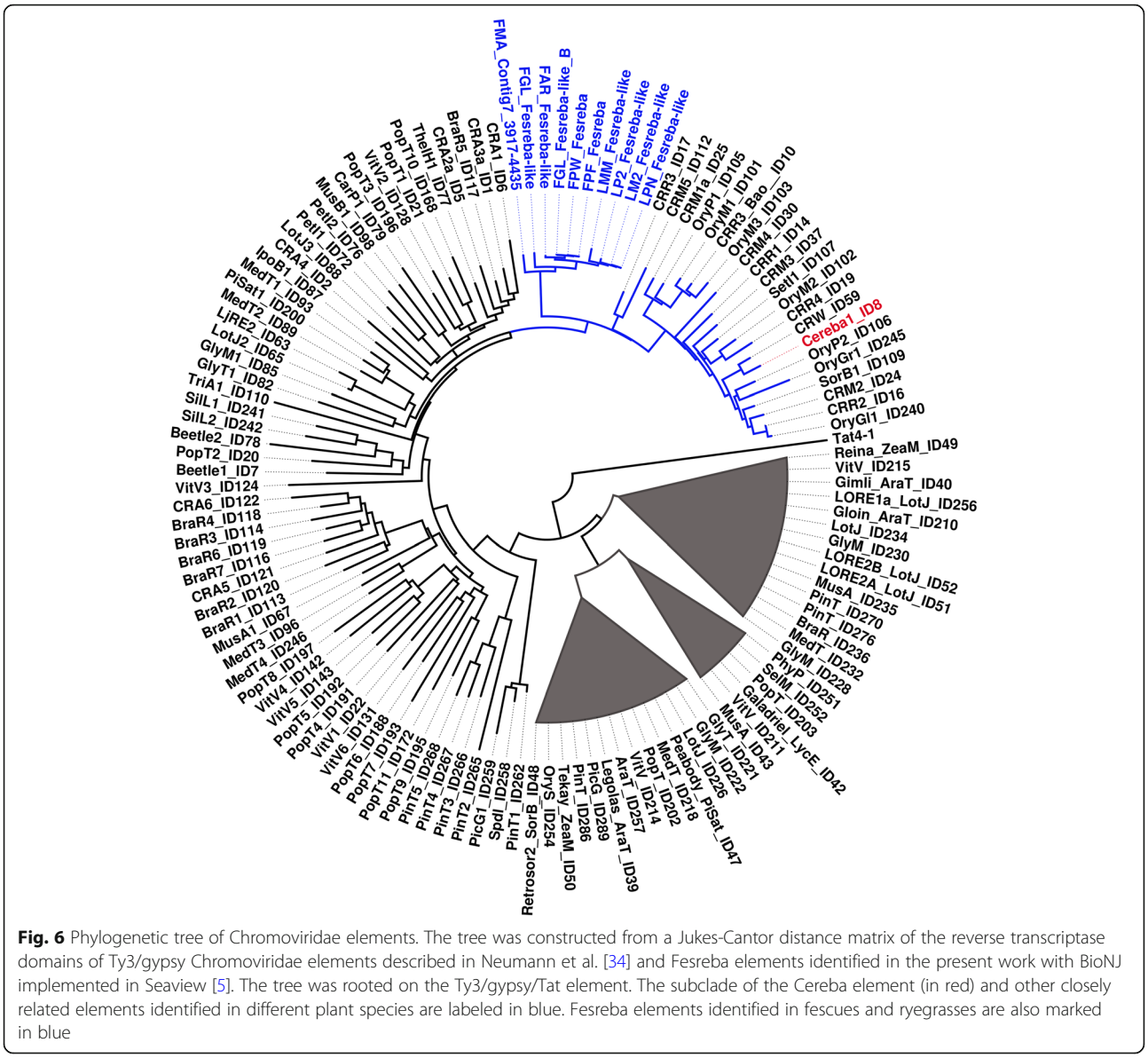


Fig. 6 Phylogenetic tree of Chromoviridae elements. The tree was constructed from a Jukes-Cantor distance matrix of the reverse transcriptase domains of Ty3/gypsy Chromoviridae elements described in Neumann et al. [34] and Fesreba elements identified in the present work with BioNJ implemented in Seaview [5]. The tree was rooted on the Ty3/gypsy/Tat element. The subclade of the Cereba element (in red) and other closely related elements identified in different plant species are labeled in blue. Fesreba elements identified in fescues and ryegrasses are also marked in blue

being RepeatExplorer, which facilitates de novo repeat identification and characterization [42, 48]. The pipeline uses graph-based clustering and analyzes next-generation sequencing data to reconstruct and characterize DNA repeats in a particular species or to compare DNA repeat composition in different genotypes [23, 24, 49–51]. The pipeline has been frequently used to reconstruct DNA repeats in diversity studies, to create repeat databases for repeat masking [19, 46, 48], and to identify tandem organized repeats suitable as probes for molecular cytogenetics [35, 51–53].

Our work revealed that Ty3/gypsy elements had the highest impact on genome size in fescues and ryegrasses. Ty3/gypsy elements are also abundant in other *Poaceae* species, including wheat, rice, maize, and barley [8, 54–56]. In barley, about 50% of the genome is made up of 15 high-copy transposable element (TE) families, with elements of the Angela lineage (Ty1/copia family) being the most abundant and representing almost 14% of the genome [56]. The Ty3/gypsy superfamily is 1.5-fold more abundant than the Ty1/copia superfamily [56].

Festuca and *Lolium* genera comprise closely related complexes of species, and thus a high homology of DNA repeats was observed in this study. The main difference was the copy number. In *Lolium* species the Ty3/gypsy Athila LTR retroelement accounted for ~ 25% of the nuclear genomes, whereas in fescues it accounted for ~ 0.7% in tetraploids *F. glaucescens* and *F. mairei* and for ~ 6% in other fescues analyzed. This indicates a burst of Athila LTR element linked with *Lolium* speciation.

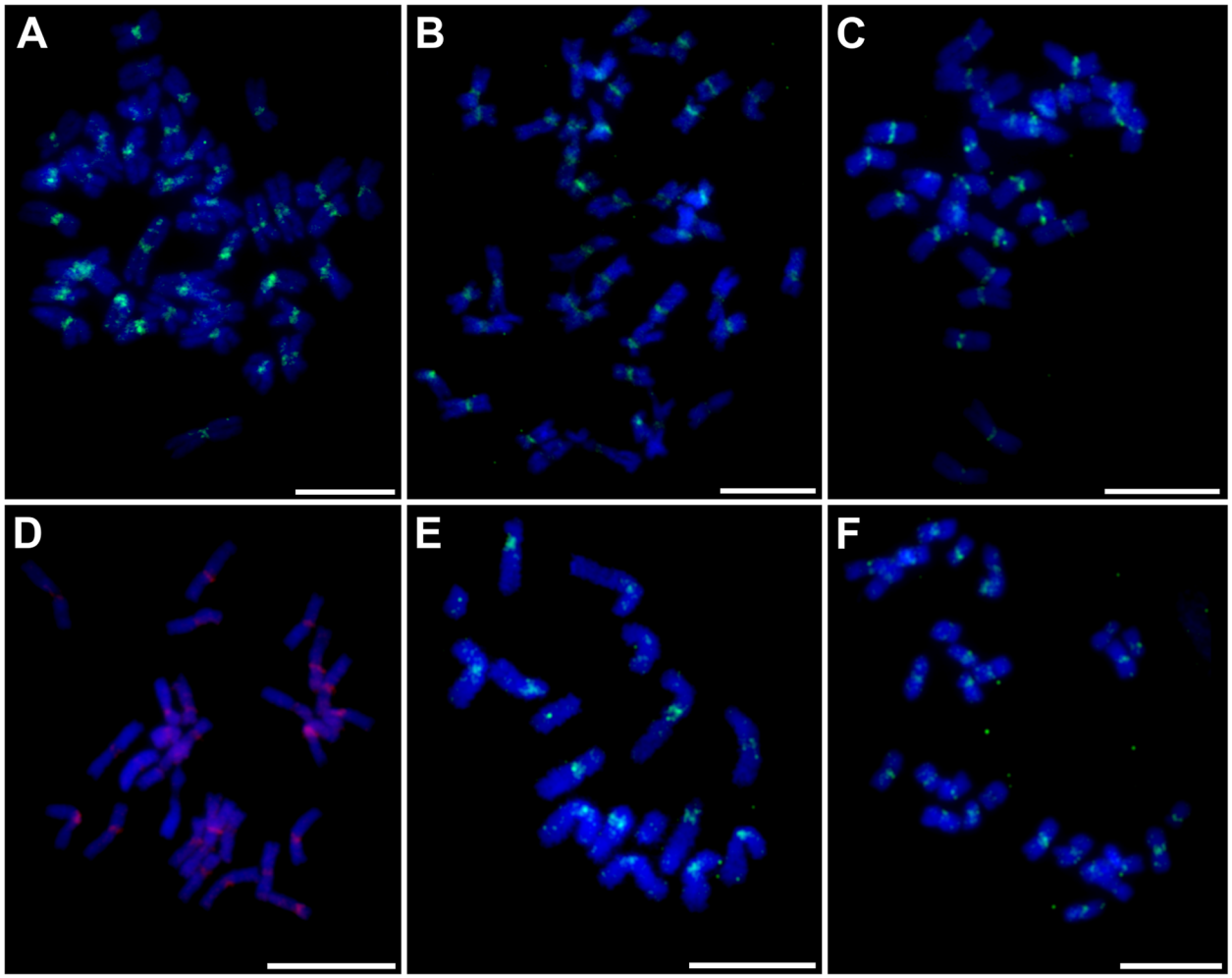


Fig. 7 Localization of the centromeric LTR retrotransposon Fesreba on mitotic metaphase chromosomes. Localization was performed in *Festuca* and *Lolium* species with fluorescence in situ hybridization (yellow-green or violet signals) with a probe for the reverse transcriptase domain of the Fesreba element. **a** *F. arundinacea* subsp. *arundinacea* (2n = 6x = 42). **b** *F. gigantea* (2n = 6x = 42). **c** *F. mairei* (2n = 4x = 28). **d** *F. pratensis* cv. Westa (2n = 4x = 28). (E) *L. perenne* GR 3320 (2n = 2x = 14). **f** *L. multiflorum* cv. Mitos (2n = 4x = 28). Chromosomes were counterstained with DAPI (blue). The bar corresponds to 10 μm

Activation and integration of TEs (e.g., as a result of environment change) may lead to a rapid burst of the Athila element in a species-specific manner [46, 47, 57] and impact evolution and speciation [46, 58]. In some species, a rapid increase in the number of lineage-specific retroelements can also result in significant genome upsizing [24, 58–60], which was not observed in the fescues and ryegrasses included in our study.

Species-specific DNA elements identified in this work were represented by tandem organized repeats (Additional file 1: Table S1). Unique tandem repeats are also found in other plant species, and thanks to their genus or species specificity they have been widely used in molecular cytogenetics (e.g., to identify chromosomes using FISH) [61–64]. Tandem repeats originally identified in *F. pratensis* chromosome 4F are useful as probes for FISH to identify individual chromosomes of the species [18, 35] and in comparative karyotype analyses of its cultivars. The present work resulted in the identification of other putative tandem organized repeats, either genus or species specific (Additional file 1: Table S1). These observations expand the number of potential cytogenetic markers for comparative karyotyping and identification of chromosomes in other fescue and ryegrass species.

Although the sequencing of *F. pratensis* chromosome 4F revealed a relatively high number of tandem repeats, none of them localized to chromosome centromeric regions [18, 35]. However, the mapping of other types of

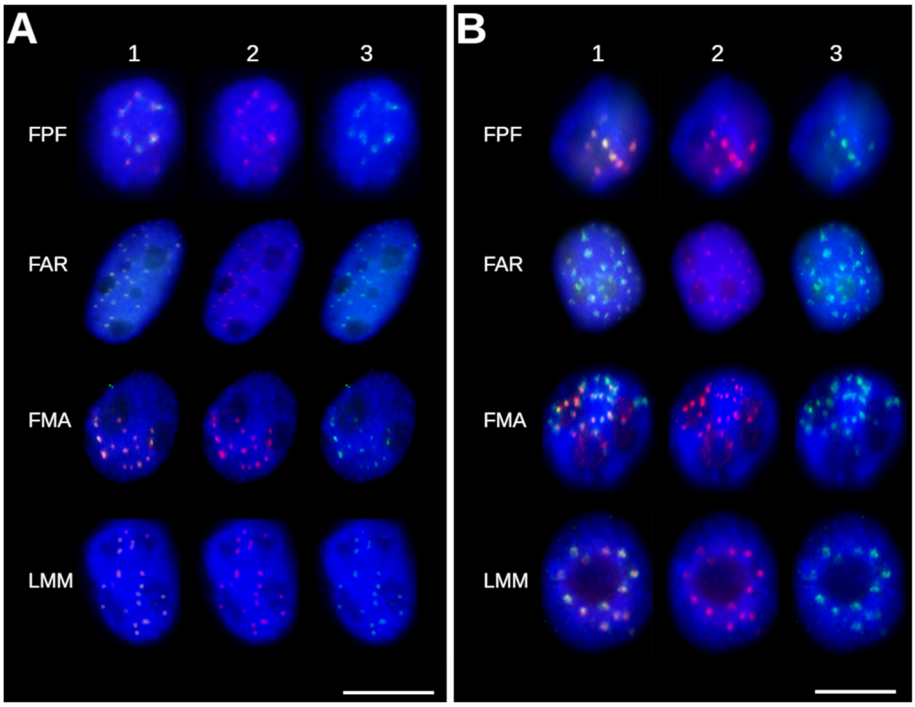


Fig. 8 Co-localization of CENH3 with the Fesreba element in *Festuca* and *Lolium*. Combination of immunolocalization of the histone H3 variant CENH3 (red) and FISH on interphase nuclei with probes for: **a** the reverse transcriptase (RT) domain (green); and **b** the non-coding LTR part of the Fesreba element (green). *F. pratensis* cv. Fure (FPF); *F. arundinacea* subsp. *arundinacea* (FAR); *F. mairei* (FMA); *L. multiflorum* cv. Mitos (LMM). Column 1 shows merged images, column 2 shows CENH3 signals (red), and column 3 shows FISH signals corresponding to the Fesreba element. In all accessions, the signals of CENH3 and FISH probes are overlapping. Nuclei were counterstained with DAPI (blue). The bar corresponds to 10 μm

DNA repeats on mitotic metaphase chromosomes showed preferential localization of one uncharacterized DNA element CL38 to centromeric regions of *F. pratensis* chromosomes [35]. In this work, the entire DNA element homologous to the CL38 repeat was reconstructed and its nature was clarified. Phylogenetic analyses of its coding domains (Fig. 6) confirmed close relationships with other plant centromeric elements of Ty3/gypsy Chromoviridae lineage, such as Cereba-like elements [43]. Preferential localization of the Cereba element to centromeric regions of barley chromosomes was shown by Hudakova et al. [33], and more complex study of centromere-specific elements belonginging to the lineage of Centromeric retrotransposons in maize (CRM) of the Ty3/gypsy family in larger set of plant species followed [20, 34]. These studies imply a role for TEs at the structural level and their impact on centromere structure. Li et al. [65] showed that the Cereba element was strongly associated with the histone H3 variant CENH3, which plays a role in centromere function. Co-localization of the centromere-specific element Fesreba, reconstructed in this work with histone CENH3 (Fig. 8, Additional file 5: Fig. S3), indicates a role for this element in the function of fescue and ryegrass centromeres as well.

Conclusions

Partial sequencing of genomes of 10 fescues and ryegrasses revealed various types of retrotransposons as the most abundant repeat. These comparative repeatome analyses increase knowledge of genome organization in fescues and ryegrasses and confirm close relationships between *Festuca* and *Lolium*. The most striking difference was observed for the Athila element, which was ~ 5 times more abundant in *Lolium* than *Festuca*. Highly diverged DNA repeats were represented by tandem organized repeats, which are candidates for species-specific cytogenetic markers. In addition to tandem repeats, other species-specific variants of the majority of repetitive DNA sequences within and between fescues and ryegrasses were identified. A nearly complete LTR element Fesreba was assembled and found to be highly enriched in centromeric and (peri)centromeric chromosome regions in all species. A combination of FISH with a probe for Fesreba and immunostaining with CENH3 antibody showed their co-

localization and indicated a possible role of Fesreba in centromere function.

Methods

Plant material

Lolium perenne GR3320 (2n = 2x = 14), *Festuca arundinacea subsp. arundinacea* (2n = 6x = 42), *Festuca gigantea* GR11759 (2n = 6x = 42), and *Festuca mairei* GR610941 (2n = 4x = 28) were obtained as seeds from the Leibniz Institute of Plant Genetics and Crop Plant Research (Gatersleben, Germany) gene bank. Seeds of *Festuca pratensis* cv. Fure (2n = 2x = 14) were obtained from Dr. Arild Larson (Graminor, Norway). *Lolium perenne* cv. Neptun (2n = 4x = 28), *Lolium multiflorum* cv. Kuri1 (2n = 2x = 14), and two commercially available cultivars, *Lolium. multiflorum* cv. Mitos (2n = 4x = 28) and *Festuca pratensis* cv. Westa (2n = 4x = 28), were obtained from Dr. Vladimír Černoch (DLF Seeds, Czech Republic). *Festuca glaucescens* genotype C-45 (2n = 4x = 28) was obtained from Seed Bank, W. Reg. P. I. Station, Pullman, WA.

Seeds of barley (*Hordeum vulgare*) cv. Morex, rye (*Secale cereale*) cv. Dánkowskie Diament, and oat (*Avena sativa*) cv. Atego were obtained from the Leibniz Institute of Plant Genetics and Crop Plant Research gene bank. Seeds of *Triticum aestivum* cv. Chinese Spring were obtained from Professor Takashi R. Endo (Kyoto University, Japan), and seeds of *Aegilops tauschii* (lineage AL 8/78; collected by V. Jaaska, University of Estonia, Tartu, Estonia) were provided by Dr. Valárik (Institute of Experimental Botany, Czech Republic). Seeds of pea (*Pisum sativum* cv. Ctirad) and rye (*Secale cereale* cv. Dankovske), which served as internal reference standards in flow cytometric analyses, were provided by one of us (JD) and are available at the Institute of Experimental Botany, Czech Republic (https://olomouc.ueb.cas.cz/en/technology/flow-cytometry-1/reference-dna-standards).

Estimation of nuclear genome size

Nuclear DNA amounts were determined according to Doležel et al. [66] following the two-step procedure of Otto [67] with modifications. Samples of isolated nuclei stained with propidium iodide were analyzed with a Sysmex CyFlow Space flow cytometer (Sysmex Partec, Münster, Germany) equipped with a 532 nm laser. Two reference standards were used to estimate DNA amounts in absolute units. Pea (*Pisum sativum* cv. Ctirad; 2C = 9.09 pg DNA) [41] served as an internal standard for estimating DNA content in all accessions except *F. mairei*, for which rye (*Secale cereale* cv. Dankovske; 2C = 16.19 pg DNA) [41] was used. Three plants were measured per accession, and each plant was analyzed three times on three different days. At least 5000 nuclei per sample were analyzed. Nuclear amounts were calculated from measurements of individual samples as follows: 2C nuclear DNA content (pg) = 2C nuclear DNA content of reference standard × sample G_1 peak mean / standard G_1 peak mean.

Mean nuclear DNA content (2C) was estimated for each plant, with 1 pg DNA equal to 0.978×10^9 bp [68]. The statistical significance of the differences between 1Cx sizes was determined with one-way ANOVA. Analyses were conducted with NCSS 97 (Statistical Solutions, Cork, Ireland). The significance level $\alpha = 0.01$ was used.

Phylogenetic analyses

Phylogenetic analyses of Loliinae subtribe were based on data published by Catalán et al. [3]. Sequences of ITS regions were downloaded from the NCBI GenBank (GB codes: AF303401–407, AF303410–416, AF303418–419, AF303421–425, AF303428, AF478475–476, AF478478–491, AF478493, AF478498–499, AF519975–981, AF519983, AF532937, AF532939–948, AF532951–952, AF532954, AF532956–960, AF532962–963, AF543514, AF548028, AJ240143, AJ240146, AJ240148, AJ240153, AJ240155–157, AJ240160, AJ240162, AY099007, AY118087–088, AY118090–092, AY118094–096, AY228161). *Brachypodium distachyon* (GB code AF303339) was used as an outgroup species. Multiple sequence alignment was done with MAFFT v7.029 (--localpair --maxiterate 1000) [69], and phylograms were constructed with PhyML 3.0 [70] implemented in SeaView v5.0.2 [5]. Approximate likelihood ratio tests [71] were performed to assess branch support. Final phylogenetic trees were depicted with FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

Illumina sequencing and data analyses

Genomic DNA was isolated with the NucleoSpin PlantII kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's recommendations and used to prepare Illumina libraries with a Nextera® DNA Sample Preparation Kit (Illumina, San Diego, CA, USA). Briefly, 50 ng DNA was fragmented, purified, and amplified according to the protocol. The DNA concentration in individual libraries was measured with a Qubit fluorometer, adjusted to an equal molar concentration, and pooled prior to sequencing. DNA sequencing was done with an Illumina MiSeq with either single or paired-end sequencing to produce up to 500 bp reads. Sequence reads were deposited in the Sequence Read Archive (BioProject ID: PRJNA601325, accessions SAMN13866227, SAMN13866228, SAMN13866229, SAMN13866230, SAMN13866231, SAMN13866232, SAMN13866233, SAMN13866234, SAMN13866235, SAMN13866236).

Illumina reads were trimmed for adapters and for quality with the FASTX-Toolkit (–q 20 -p 90; http://hannonlab.cshl.edu/fastx_toolkit/index.html). Detailed characterization of repeat families was performed with a stand-alone version

of the RepeatExplorer pipeline [37] running on an IBM server with 16 processors, 100 Gb RAM, and 17 Tb disk space. In the first step, comparative analyses of repetitive parts of the genomes were performed with the RepeatExplorer pipeline according to Novák et al. [49]. Random data sets represented the same amount of reads (0.5× coverage of individual accessions) were used to reconstruct repetitive elements using a graph-based method according Novák et al. [48]. The RepeatExplorer pipeline led to the characterization of assembled sequences using different tools (e.g., BLASTN and BLASTX, phylogenetic analysis) [37, 48]. Tandem organized repeats were identified with Dotter [72].

In the second step, the RepeatExplorer pipeline was applied to a merged data set containing all species marked by specific prefixes to perform comparative analyses [49]. The results of the clustering were then used to create repetitive databases. Databases of Illumina reads were deposited in the Sequence Read Archive (accessions: SRX7566047–SRX7566056). Assembled contigs from different types of repetitive DNA elements are publicly available online (https://olomouc.ueb.cas.cz/en/content/dna-repeats) and in the Dryad digital repository (doi:https://doi.org/10.5061/dryad.xksn02vch).

Southern hybridization

Genomic DNA corresponding to 3×10^6 copies of a 1Cx nuclear genome was digested by *Hae*III enzyme (New England Biolabs, Ipswich, MA, USA). DNA fragments were size-fractionated by electrophoresis in 1.2% agarose gel and then transferred onto Hybond™ N+ nylon membranes (GE Healthcare, Chicago, IL, USA). Probes were prepared with *F. pratensis* genomic DNA as a template and polymerase chain reaction (PCR) with biotin-labeled dUTP (Roche, Mannheim, Germany) and specific primers (Additional file 6: Table S3, Additional file 7: Fig. S4). Southern hybridization was performed at 68 °C overnight, and hybridization signals were detected with a Chemiluminescent Nucleic Acid Module (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's recommendations with 90% stringency. Hybridization signals were visualized with chemiluminiscent substrate on Medical X-Ray Film Blue (Agfa Healthcare, Mortsel, Belgium).

ddPCR

Based on the assembled DNA contigs from the Fesreba retrotransposon, two restriction endonucleases with unique restriction sites in the retrotransposon (*Hpa*I and *Hpa*II) were identified and used for further analyses. Briefly, 3 μg genomic DNA was digested according to the manufacturer's recommendations (Bio-Rad Laboratories, Hercules, CA, USA) and then diluted 1000-fold to reach a starter concentration of 0.06 ng/μl. ddPCR was performed on a QX200 Droplet Digital PCR machine (Bio-Rad Laboratories) following the manufacturer's recommendations with EvaGreen Supermix (Bio-Rad Laboratories), template DNA, and specific primers for Fesreba (Additional file 6: Table S3). Three independent replicates were performed for every accession analyzed.

Cytogenetic mapping and immunostaining

Cytogenetic mapping of selected repeats was done by FISH on mitotic metaphase plates. Chromosome spreads were prepared according to Křivánková et al. [35], and immunostaining was performed according to Neumann et al. [73]. Root tips were collected in ice water for 28 h; washed in LB01 buffer [74]; fixed in 3.7% formaldehyde for 25 min; and digested using 2% cellulose, 2% pectinase, and 2% cytohelicase in 1× phosphate-buffered saline (PBS) for 90 min at 37 °C. After the coverslip was removed, the preparations were washed in 1× PBS and then in PBS-Triton buffer (1× PBS, 0.5% Triton X-100, pH 7.4) for 25 min and then again in 1× PBS. For incubation with anti-grass CENH3 primary antibody [75], the slides were washed in PBS-Tween buffer (1× PBS, 0.1% Tween 20, pH 7.4) for 25 min and then incubated with anti-grass CENH3 primary antibody (diluted 1:200 in PBS-Tween) overnight at 4 °C. The next day the slides were washed in 1× PBS, CENH3 antibody was detected using the anti-Rabbit Alexa Fluor 546 secondary antibody (Thermo Fisher Scientific/Invitrogen) diluted 1:250 in PBS-Tween buffer for 1 h at room temperature, and washed 1× PBS. Before FISH, immunofluorescent signals were stabilized with ethanol:acetic acid (3:1) fixative and 3.7% formaldehyde for 10 min at room temperature. FISH was performed after three washes in 1× PBS.

Probes for FISH, derived from the RT and LTR regions of the Fesreba element, were labeled with digoxigenin-11-dUTP or biotin-16-dUTP (Roche Applied Science) using PCR with specific primers (Additional file 6: Table S3). FISH and detection of hybridization sites were performed according to Křivánková et al. [35]. The chromosomes were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) and mounted in Vectashield (Vector Laboratories). The slides were examined with an Axio Imager.Z2 microscope (Carl Zeiss, Oberkochen, Germany) equipped with a Cool Cube 1 (Metasystems, Altlussheim, Germany) camera, and images were prepared with ISIS 5.4.7 (Metasystems). Final adjustments were made to figures in Adobe Photoshop 12.0.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12870-020-02495-0.

Additional file 1: Table S1. List of clusters containing putative tandem repeats identified in *Festuca* and *Lolium*.

Additional file 2: Fig. S1. Southern blots for the RT domain and non-coding LTR part of the Fesreba element. Southern blots were made with probes for the reverse transcriptase domain (A) and non-coding LTR region (B) of the Fesreba element. Lanes contained genomic DNA digested by *Hae*III restriction endonuclease. Lane 1: diploid *F. pratensis* cv. Fure; lane 2: tetraploid *F. pratensis* cv. Westa; lane 3: hexaploid *F. arundinacea* subsp. *arundinacea*; lane 4: hexaploid *F. gigantea*; lane 5: tetraploid *F. glaucescens*; lane 6: tetraploid *F. mairei*; lane 7: tetraploid *L. multiflorum* cv. Mitos; lane 8: diploid *L. multiflorum* cv. Kuri1; lane 9: tetraploid *L. perenne* cv. Neptun; lane 10: diploid *L. perenne*.

Additional file 3: Table S2. Representation of the RT domain and non-coding part of the LTR region of the Fesreba element estimated by ddPCR. Copy numbers of the reverse transcriptase (RT) domain and non-coding part of the LTR region of the Fesreba element were estimated with droplet digital PCR. Values are averages of three independent experiments with standard deviations.

Additional file 4: Fig. S2. Localization of the centromeric LTR retrotransposon Fesreba on mitotic chromosomes with fluorescence in situ hybridization. Mitotic metaphase plates were hybridized with a probe for the reverse transcriptase domain of the Fesreba element (A, C, E, G, I) and with a combination of probes for the non-coding LTR part of the Fesreba element and 45S rDNA, which served as control (B, D, F, H, J). (A, B) *Avena sativa* cv. Atego (2n = 2x = 14). (C, D) *Secale cereale* cv. Dánkowskie Diament (2n = 2x = 14). (E, F) *Hordeum vulgare* cv. Morex (2n = 2x = 14). (G, H) *Triticum aestivum* cv. Chinese Spring (2n = 6x = 42). (I, J) *Aegilops tauschii* (2n = 2x = 14). Signals corresponding to 45S rDNA loci are marked by arrows. Hybridization signals of a probe for the LTR region of the Fesreba element were absent in all related species (D, F, H, J) except *A. sativa* (B). Chromosomes were counterstained with DAPI (blue). The bar corresponds to 10 μm.

Additional file 5: Fig. S3. Co-localization of CENH3 with the Fesreba element in three *Festuca* and three *Lolium* species. Immunolocalization of the histone H3 variant CENH3 (red) and FISH with probes for the reverse transcriptase (RT) domain and non-coding LTR part of the Fesreba element (green). *F. gigantea* (FGI); *F. glaucescens* (FGL); *F. pratensis* Westa (FPW); *L. multiflorum* Lm2 (LM2); *L. perenne* Neptun (LP2); *L. perenne* (LPN). Column 1 shows merged images, column 2 shows CENH3 signals (red), and column 3 shows FISH signals corresponding to the Fesreba element. In all accessions, the signals of CENH3 and FISH probes are overlapping. Nuclei were counterstained with DAPI (blue). The bar corresponds to 10 μm.

Additional file 6: Table S3. Primers used for PCR amplification of DNA repeats.

Additional file 7: Fig. S4. Original images of Southern hybridization depicted in Fig. 5 and Additional file 2: Fig. S1, respectively. Original images of Southern hybridization with sequences derived from cluster CL1 (A), cluster CL38 (B), and cluster CL20 (C) and with sequences for the reverse transcriptase domain (D) and non-coding LTR region (E) of the Fesreba element. Lanes contained genomic DNA digested by *Hae*III restriction endonuclease. Lane 1: diploid *F. pratensis* cv. Fure; lane 2: tetraploid *F. pratensis* cv. Westa; lane 3: hexaploid *F. arundinacea* subsp. *arundinacea*; lane 4: hexaploid *F. gigantea*; lane 5: tetraploid *F. glaucescens*; lane 6: tetraploid *F. mairei*; lane 7: tetraploid *L. multiflorum* cv. Mitos; lane 8: diploid *L. multiflorum* cv. Kuri1; lane 9: tetraploid *L. perenne* cv. Neptun; lane 10: diploid *L. perenne*.

Abbreviations

1C: Holoploid genome; 1Cx: Monoploid genome; 2C: Nuclear DNA amount in G1 nucleus prior to DNA replication; 4F: Chromosome 4 of *Festuca pratensis* cv. Fure; bp: Base pairs; CENH3: Centromeric histone H3; CL: Cluster of orthologous sequences obtained by RepeatExplorer analysis; CRM: Centromeric retrotransposon in maize; Cy3: Cy3 fluorescent dye; DAPI: 4′,6-diamidino-2-phenylindole; ddPCR: Droplet digital polymerase chain reaction; DNA: Deoxyribonucleic acid; dUTP: 2′-deoxyuridine 5′-triphosphate; FAR: *Festuca arundinacea* Schreb. subsp. *arundinacea*; FGI: *Festuca gigantea* L. GR11759; FGL: *Festuca arundinacea* Schreb. subsp. *glaucescens*; FISH: Fluorescence in situ hybridization; FITC: Fluorescein isothiocyanate; FMA: *Festuca mairei* GR610941; FPF: *Festuca pratensis* Huds. cv. Fure; FPW: *Festuca pratensis* Huds. cv. Westa; G1: G1 phase of cell cycle; Gbp: Gigabase pairs; ID: Identity number; LINE: Long interspersed nuclear element; LM2: *Lolium multiflorum* cv. Kuri1; LMM: *Lolium multiflorum* Lam. cv. Mitos; LP2: *Lolium perenne* L. GR3320; LPN: *Lolium perenne* L. cv. Neptun; LTR: Long terminal repeat; μl: Microliter; NCBI: National Center for Biotechnology Information; ng: Nanogram; PBS: Phosphate-buffered saline; PCR: Polymerase chain reaction; pg: Picogram; pH: Potential of hydrogen; rDNA: Ribosomal DNA, DNA with ribosomal RNA genes; rRNA: Ribosomal RNA, RNA involved in structure of ribosomes and proteosynthesis; RT: Reverse transcriptase; SSC: Saline sodium citrate; TE: Transposable element

Acknowledgments

We thank H. Tvardíková and E. Jahnová for technical support, P. Navrátil for technological support, and V. Černoch (DLF Seeds, Czech Republic) and M. Valárik (Institute of Experimental Botany, Czech Republic) for providing plant material.

Authors' contributions

JZ prepared DNA for sequencing, analyzed Illumina sequence data, and performed DNA repeat reconstruction and further analyses of repeats. JC and DK performed flow cytometric estimation of genome size; KH and BJT performed Illumina sequencing; and JZ, DK, VK, and AN performed cytogenetic analyses, including immuno-FISH. RS and JZ performed ddPCR and interpreted the results. EH and JD made intellectual contributions to the study and revised the manuscript critically for important intellectual content. All authors read and approved the final manuscript.

Funding

This work was supported by the ERDF project "Plants as a Tool for Sustainable Global Development" (No. CZ.02.1.01/0.0/0.0/16_019/0000827). Computational support was provided by the National Grid Infrastructure MetaCentrum (Grant No. LM2010005 under the program Projects of Large Infrastructure for Research, Development, and Innovations). The funders had no role in the study design, data analysis and interpretation, or writing of the manuscript.

Availability of data and materials

All relevant supporting data sets are included in the article and its additional files. The data sets supporting the conclusions in this article are available in the Sequence Read Archive (accessions: SRX7566047–SRX7566056) and Dryad repository (doi:https://doi.org/10.5061/dryad.xksn02vch; https://datadryad.org/stash/share/8pm4qJ41tIJaXNd7EYkan125DKR-Vi8BINbF4HSobVg).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Institute of Experimental Botany, Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, CZ-77900 Olomouc, Czech Republic. [2]Centro de Genómica Nutricional Agroacuícola, Las Heras 350, Temuco, Chile.

Received: 16 January 2020 Accepted: 15 June 2020
Published online: 17 June 2020

References

1. Inda LA, Segarra-Moragues JG, Müller J, Peterson PM, Catalán P. Dated historical biogeography of the temperate *Loliinae* (*Poaceae*, Pooideae) grasses in the northern and southern hemispheres. Mol Phylogenet Evol. 2008;46:932–57. https://doi.org/10.1016/j.ympev.2007.11.022.
2. Watson L, Dawitz MJ. The grass genera of the world. Wallingford: C. A. B. International; 1992.
3. Catalán P, Torrecilla P, López Rodriguez JA, Olmstead RG. Phylogeny of the festucoid grasses of subtribe *Loliinae* and allies (Poaea, Pooideae) inferred

from ITS and trnL-F sequences. Mol Phylogenet Evol. 2004;31(2):517–41. https://doi.org/10.1016/j.ympev.2003.08.025.
4. Torrecilla P, Catalán P. Phylogeny of broad-leaved and fine-leaved *Festuca* lineages (*Poaceae*) based on nuclear ITS sequences. Syst Bot. 2002;27(2):241–52. https://doi.org/10.1043/0363-6445-27.2.241.
5. Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol. 2010;27(2):221–4. https://doi.org/10.1093/molbev/msp259.
6. Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Zuloaga FO, Judziewicz EJ, et al. A worldwide phylogenetic classification of the *Poaceae* (*Gramineae*). J Syst Evol. 2015;53:117–37. https://doi.org/10.1111/jse.12150.
7. Šmarda P, Bureš P, Horová L, Foggi B, Rossi G. Genome size and GC content evolution of *Festuca*: ancestral expansion and subsequent reduction. Ann Bot. 2008;101(3):421–33. https://doi.org/10.1093/aob/mcm307.
8. Kopecký D, Havránková M, Loureiro J, Castro S, Lukaszewski AJ, Bartoš J, et al. Physical distribution of homoeologous recombination in individual chromosomes of *Festuca pratensis* in *Lolium multiflorum*. Cytogenet Genome Res. 2010;129(1–3):162–72. https://doi.org/10.1159/000313379.
9. Kopecký D, Lukaszewski AJ, Doležel J. Cytogenetics of *Festulolium* (*Festuca* x *Lolium* hybrids). Cytogenet Genome Res. 2008;120(3–4):370–83. https://doi.org/10.1159/000121086.
10. Loureiro J, Kopecký D, Castro S, Santos C, Silveira P. Flow cytometric and cytogenetic analyses of Iberian Peninsula *Festuca* spp. Plant Syst Evol. 2007;269:89–105. https://doi.org/10.1007/s00606-007-0564-8.
11. Hand ML, Cogan NO, Stewart AV, Forster JW. Evolutionary history of tall fescue morphotypes inferred from molecular phylogenetics of the *Lolium-Festuca* species complex. BMC Evol Biol. 2010;10:303. https://doi.org/10.1186/1471-2148-10-303.
12. Humphreys J, Harper JA, Armstead IP, Humhreys MW. Introgression-mapping of genes for drought resistance transferred from *Festuca arundinaceae* var. *glaucescens* into *Lolium multiflorum*. Theor Appl Genet. 2005;110:579–87. https://doi.org/10.1007/s00122-004-1879-2.
13. Kopecký D, Bartoš J, Christelová P, Černoch V, Kilian A, Doležel J. Genomic constitution of *Festuca* x *Lolium* hybrids revealed by the DArTFest array. Theor Appl Genet. 2011;122(2):355–63. https://doi.org/10.1007/s00122-010-1451-1.
14. Kosmala A, Zwierzykowski Z, Gasior D, Rapacz M, Zwierzykowska E, Humphreys MW. GISH/FISH mapping of genes for freezing tolerance transferred from *Festuca pratensis* to *Lolium multiflorum*. Heredity. 2006;96:243–51. https://doi.org/10.1038/sj.hdy.6800787.
15. Ezquerro-López D, Kopecký D, Inda Luis Á. Cytogenetic relationships within the Maghrebian clade of *Festuca* subgen. *Schedonorus* (*Poaceae*), using flow cytometry and FISH. Anales del Jardín Botánico de Madrid. 2017;74(1):e052. doi:https://doi.org/10.3989/ajbm.2455.
16. Czaban A, Sharma S, Byrne SL, Spannagl M, Mayer KF, Asp T. Comparative transcriptome analysis within the *Lolium/Festuca* species complex reveals high sequence conservation. BMC Genomics. 2015;16(1):249. https://doi.org/10.1186/s12864-015-1447-y.
17. Byrne SL, Nagy I, Pfeifer M, Armstead I, Swain S, Studer B, et al. A synteny-based draft genome sequence of the forage grass *Lolium perenne*. Plant J. 2015;84:816–26. https://doi.org/10.1111/tpj.13037.
18. Kopecký D, Martis M, Číhalíková J, Hřibová E, Vrána J, Bartoš J, et al. Flow sorting and sequencing meadow fescue chromosome 4F. Plant Physiol. 2013;163(3):1323–37. https://doi.org/10.1104/pp.113.224105.
19. SanMiguel P, Bennetzen JL. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann Bot. 1998;82(1):37–44. https://doi.org/10.1006/anbo.1998.0746.
20. Wicker T, Gundlach H, Spannagl M, Uauy C, Borrill P, Ramírez-González RH, et al. Impact of transposable elements on genome structure and evolution in bread wheat. Genome Biol. 2018;19(1):103. https://doi.org/10.1186/s13059-018-1479-0.
21. Hřibová E, Neumann P, Matsumoto T, Roux N, Macas J, Doležel J. Repetitive part of the banana (*Musa acuminata*) genome investigated by low-depth 454 sequencing. BMC Plant Biol. 2010;10:204. https://doi.org/10.1186/1471-2229-10-204.
22. Piednoel M, Aberer AJ, Schneeweiss GM, Macas J, Novák P, Gundlach H, et al. Next-generation sequencing reveals the impact of repetitive DNA across phylogenetically closely related genomes of Orobanchaceae. Mol Biol Evol. 2012;29(11):3601–11. https://doi.org/10.1093/molbev/mss168.
23. Dodsworth S, Chase MW, Kelly LJ, Leitch IJ, Macas J, Novák P, et al. Genomic repeat abundances contain phylogenetic signal. Syst Biol. 2015;64(1):112–26. https://doi.org/10.1093/sysbio/syu080.
24. Macas J, Novák P, Pellicer J, Čížková J, Koblížková A, Neumann P, et al. In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the vegume vribe Fabeae. PLoS One. 2015;10(11):e0143424. https://doi.org/10.1371/journal.pone.0143424.
25. Leitch IJ, Bennett MD. Genome downsizing in polyploid plants. Biol J Linn Soc. 2004;82:651–63. https://doi.org/10.1111/j.1095-8312.2004.00349.x.
26. Mandáková T, Joly S, Krzywinski M, Mummenhoff K, Lysák MA. Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. Plant Cell. 2010;22(7):2277–90. https://doi.org/10.1105/tpc.110.074526.
27. Renny-Byfield S, Kovařík A, Kelly LJ, Macas J, Novák P, Chase MW, et al. Diploidization and genome size change in allopolyploids is associated with differential dynamics of low- and high-copy sequences. Plant J. 2013;74(5):829–39. https://doi.org/10.1111/tpj.12168.
28. Ananiev EV, Phillips RL, Rines HW. Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. Proc. Natl Acad. Sci. USA. 1998;95:13073–8. https://doi.org/10.1073/pnas.95.22.13073.
29. Kumekawa N, Hosouchi T, Tsuruoka H, Kotani H. The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 4. DNA Res. 2001;8:285–90. https://doi.org/10.1093/dnares/8.6.285.
30. Li Y, Zuo S, Zhang Z, Li Z, Han J, Chu Z, et al. Centromeric DNA characterization in the model grass *Brachypodium distachyon* provides insights on the evolution of the genus. Plant J. 2018;93(6):1088–101. https://doi.org/10.1111/tpj.13832.
31. Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, et al. Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell. 2002;14(8):1691–704. https://doi.org/10.1105/tpc.003079.
32. Gorinšek B, Gubenšek F, Kordiš D. Phylogenomic analysis of chromoviruses. Cytogenet Genome Res. 2005;110(1–4):543–52. https://doi.org/10.1159/00008487.
33. Hudakova S, Michalek W, Presting GG, ten Hoopen R, dos Santos K, Jasencakova Z, et al. Sequence organization of barley centromeres. Nucleic Acids Res. 2001;29:5029–35. https://doi.org/10.1093/nar/29.24.5029.
34. Neumann P, Navrátilová A, Koblížková A, Kejnovský E, Hřibová E, Hobza R, et al. Plant centromeric retrotransposons: a structural and cytogenetic perspective. Mob DNA. 2011;2(1):4. https://doi.org/10.1186/1759-8753-2-4.
35. Křivánková A, Kopecký D, Stočes Š, Doležel J, Hřibová E. Repetitive DNA: A versatile tool for karyotyping in *Festuca pratensis* Huds. Cytogenet Genome Res 2017;151(2):96–105. doi:https://doi.org/10.1159/000462915.
36. Paux E, Roger D, Badaeva E, Gay G, Bernard M, Sourdille P, et al. Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. Plant J. 2006;48(3):463–74. https://doi.org/10.1111/j.1365-313X.2006.02891.x.
37. Fu K, Guo Z, Zhang X, Fan Y, Wu W, Li D, et al. Insight into the genetic variability analysis and cultivar identification of tall fescue by using SSR markers. Hereditas. 2016;153:9. https://doi.org/10.1186/s41065-016-0013-1.
38. Koo DH, Nam YW, Choi D, Bang JW, de Jong H, Hur Y. Molecular cytogenetic mapping of *Cucumis sativus* and *C. melo* using highly repetitive DNA sequences. Chromosom Res. 2010;18(3):325–36. https://doi.org/10.1007/s10577-010-9116-0.
39. Mehrotra S, Goyal V. Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. Genom Proteom Bioinf. 2014;12(4):164–71. https://doi.org/10.1016/j.gpb.2014.07.003.
40. Nybom H, Weising K1, Rotter B. DNA fingerprinting in botany: past, present, future. Investig Genet. 2014;5(1):1. doi:https://doi.org/10.1186/2041-2223-5-1.
41. Doležel J, Greilhuber J, Lucretti S, Meister A, Lysák MA, Nardi L, et al. Plant genome size estimation by flow Cytometry: inter-laboratory comparison. Ann Bot. 1998;82:17–26. https://doi.org/10.1093/aob/mci005.
42. Novák P, Neumann P, Pech J, Steinhaisl J, Macas J. RepeatExplorer: a galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics. 2013;29(6):792–3. https://doi.org/10.1093/bioinformatics/btt054.
43. Presting GG, Malysheva L, Fuchs J, Schubert I. A Ty3/gypsy retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. Plant J. 1998;16(6):721–8. https://doi.org/10.1046/j.1362-313x.1998.00341.x.
44. Sanei M, Pickering R, Kumke K, Nasuda S, Houben A. Loss of centromeric histone H3 (CENH3) from cetromeres precedes uniparental chromosome

elimination in interspecific barley hybrids. Proc Natl Acad Sci U S A. 2011; 108:E498–505. https://doi.org/10.1073/pnas.1103190108.
45. Murray BG, De Lange PJ, Ferguson AR. Nuclear DNA variation, chromosome numbers and polyploidy in the endemic and indigenous grass flora of New Zealand. Ann Bot. 2005;96(7):1293–305. https://doi.org/10.1093/aob/mci281.
46. Zhang Q-J, Gao L-l. Rapid and recent evolution of LTR retrotransposons drives rice genome evolution during the speciation of AA-genome *Oryza* species. G3 (Bethesda). 2017;7(6). doi:https://doi.org/10.1534/g3.116.037572.
47. McCann J, Macas J, Novák P, Stuessy TF, Villasenor JL, Weiss-Schneweiss H. Differential genome size and repetitive DNA evolution in diploid species of *Melampodium* sect. *Melampodium* (Asteraceae). Front. Plant Sci. 2020;11:362. https://doi.org/10.3389/fpls.2020.00362.
48. Novák P, Neumann P, Macas J. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. BMC Bioinformatics. 2010;11:378. https://doi.org/10.1186/1471-2105-11-378.
49. Novák P, Hřibová E, Neumann P, Koblížková A, Doležel J, Macas J. Genome-wide analysis of repeat diversity across the family *Musaceae*. PLoS One. 2014;9(6):e98918. https://doi.org/10.1371/journal.pone.0098918.
50. McCann J, Jang TS, Macas J, Schneeweiss GM, Matzke NJ, Novák P, et al. Dating the species network: Allopolyploidy and repetitive DNA evolution in American daisies (*Melampodium* sect. *Melampodium*, *Asteraceae*). Syst Biol. 2018;67(6):1010–24. https://doi.org/10.1093/sysbio/syy024.
51. Renny-Byfield S, Kovařík A, Chester M, Nichols RA, Macas J, Novák P, et al. Independent, rapid and targeted loss of highly repetitive DNA in natural and synthetic allopolyploids of *Nicotiana tabacum*. PLoS One. 2012;7(5): e36963. https://doi.org/10.1371/journal.pone.0036963.
52. Macas J, Kejnovský E, Neumann P, Novák P, Koblížková A, Vyskot B. Next generation sequencing-based analysis of repetitive DNA in the model dioecious plant *Silene latifolia*. PLoS One. 2011;6(11):e27335. https://doi.org/10.1371/journal.pone.0027335.
53. Said M, Hřibová E, Danilova TV, Karafiátová M, Čížková J, Friebe B, et al. 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. Theor Appl Genet. 2018;131(10): 2213–27. https://doi.org/10.1007/s00122-018-3148-9.
54. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, et al. The B73 maize genome: complexity, diversity, and dynamics. Science. 2009; 326(5956):1112–5. https://doi.org/10.1126/science.1178534.
55. International Rice Genome Sequencing Project, Sasaki T. The map-based sequence of the rice genome. Nature. 2005;436(7052):793–800. https://doi.org/10.1038/nature03895.
56. International Barley Genome Sequencing Consortium, Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, et al. A physical, genetic and functional sequence assembly of the barley genome. Nature. 2012; 491(7426):711–6. https://doi.org/10.1038/nature11543.
57. Grandbastien MA, Audeon C, Bonnivard E, Casacuberta JM, Chalhoub B, Costa A-PP, et al. Stress activation and genomic impact of Tnt1 retrotransposon in *Solanaceae*. Cytogenet Genome Res. 2005;110(1–4):229–41. https://doi.org/10.1159/000084957.
58. Lee J, Waminal NE, Choi HI, Perumal S, Lee SC, Nguyen VB, et al. Rapid amplification of four retrotransposon families promoted speciation and genome size expansion in the genus *Panax*. Sci Rep. 2017;7(1):17986. https://doi.org/10.1038/s41598-017-08194-5.
59. Bennetzen JL, Wang H. The contribution of transposable elements to the structure, function, and evolution of plant genomes. Annu Rev Plant Biol. 2014;65:505–30. https://doi.org/10.1146/annurev-arplant-050213-035811.
60. Kelly LJ, Renny-Byfield S, Pellicer J, Macas J, Novák P, Neumann P, et al. Analysis of the giant genomes of *Fritillaria* (*Liliaceae*) indicates that a lack of DNA removal characterizes extreme expansion in genome size. New Phytol. 2015;208(2):596–607. https://doi.org/10.1111/nph.13471.
61. Hřibová E, Doleželová M, Town CD, Macas J, Doležel J. Isolation and characterization of the highly repeated fraction of the banana genome. Cytogenet Genome Res. 2007;119(3–4):268–74. https://doi.org/10.1159/000112073.
62. Macas J, Neumann P, Navrátilová A. Repetitive DNA In the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. BMC Genomics 2007; 8:427. doi:https://doi.org/10.1186/1471-2164-8-427.
63. Badaeva ED, Amosova AV, Goncharov NP, Macas J, Ruban AS, Grechishnikova IV, et al. A set of cytogenetic markers allows the precise identification of all A-genome chromosomes in diploid and polyploid wheat. Cytogenet Genome Res. 2015;146(1):71–9. https://doi.org/10.1159/000433458.
64. Koo DH, Tiwari VK, Hřibová E, Doležel J, Friebe B, Gill BS. Molecular cytogenetic mapping of satellite DNA sequences in *Aegilops geniculata* and wheat. Cytogenet Genome Res. 2016;148(4):314–21. https://doi.org/10.1159/000447471.
65. Li B, Choulet F, Heng Y, Hao W, Paux E, Liu Z, et al. Wheat centromeric retrotransposons: the new ones take a major role in centromeric structure. Plant J. 2013;73(6):952–65. https://doi.org/10.1111/tpj.12086.
66. Doležel J, Greilhuber J, Suda J. Estimation of nuclear DNA content in plants using flow cytometry. Nature Prot. 2007;2(9):2233–44. https://doi.org/10.1038/nprot.2007.310.
67. Otto F. DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. In: Crissman HA, Darzynkiewicz Z, editors. Methods in cell biology, Vol 33, pp 105–110. New York: Acad Press; 1990.
68. Doležel J, Bartoš J, Voglmayr H, Greilhuber J. Nuclear DNA content and genome size of trout and human. Cytometry A. 2003;51:127–8. https://doi.org/10.1002/cyto.a.10013.
69. Katoh K, Toh H. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinf. 2008;9:286–98. https://doi.org/10.1093/bib/bbn013.
70. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59(3):307–21. https://doi.org/10.1093/sysbio/syq010.
71. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst Biol. 2006;55:539–52. https://doi.org/10.1080/10635150600755453.
72. Sonnhammer EL, Durbin R. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene. 1995; 167(1–2):GC1–10. https://doi.org/10.1016/0378-1119(95)00714-8.
73. Neumann P, Schubert V, Fuková I, Manning JE, Houben A, Macas J. Epigenetic histone marks of extended meta-polycentric centromeres of *Lathyrus* and *Pisum* chromosomes. Front Plant Sci. 2016;7:234. https://doi.org/10.3389/fpls.2016.00234.
74. Doležel J, Binarová P, Lucretti S. Analysis of nuclear DNA content in plant cells by flow cytometry. Biol Plantarum. 1989;31:113–20. https://doi.org/10.1007/BF02907241.
75. Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM, et al. Sequencing of a rice centromere uncovers active genes. Nat Genet. 2004; 36(2):138–45. https://doi.org/10.1038/ng1289.

Publisher's Note

APPENDIX II

Chromosome genomics uncovers plant genome organization and function

Zwyrtková, J., Šimková, H., Doležel, J.

Biotechnology Advances, 46: 107659, 2021

doi: 10.1016/j.biotechadv.2020.107659

IF_{2019}: 10.744

Biotechnology Advances 46 (2021) 107659

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Chromosome genomics uncovers plant genome organization and function

Jana Zwyrtková, Hana Šimková, Jaroslav Doležel *

Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, CZ-77900 Olomouc, Czech Republic

ARTICLE INFO

Keywords:
Chromosome flow sorting
Chromosome genomics
DNA library
Flow cytometry
Gene cloning
Molecular marker
Next-generation sequencing
Plant chromosome

ABSTRACT

The identification of causal genomic loci and their interactions underlying various traits in plants has been greatly aided by progress in understanding the organization of the nuclear genome. This provides clues to the responses of plants to environmental stimuli at the molecular level. Apart from other uses, these insights are needed to fully explore the potential of new breeding techniques that rely on genome editing. However, genome analysis and sequencing is not straightforward in the many agricultural crops and their wild relatives that possess large and complex genomes. Chromosome genomics streamlines this task by dissecting the genome to single chromosomes whose DNA is then used instead of nuclear DNA. This results in a massive and lossless reduction in DNA sample complexity, reduces the time and cost of the experiment, and simplifies data interpretation. Flow cytometric sorting of condensed mitotic chromosomes makes it possible to purify single chromosomes in large quantities, and as the DNA remains intact this process can be coupled successfully with many techniques in molecular biology and genomics. Since the first experiments with flow cytometric sorting in the late 1980s, numerous applications have been developed, and chromosome genomics has been having a significant impact in many areas of research, including the sequencing of complex genomes of important crops and gene cloning. This review discusses these applications, describes their contribution to advancements in plant genome analysis and gene cloning, and outlines future directions.

1. Introduction

The breeding of crops resistant to diseases and pests and adapted to climate change is critical to securing enough food for humankind in an environmentally sustainable manner. New breeding techniques may play an important role in this endeavor, as they promise to develop improved cultivars in a faster and more precise manner than conventional breeding (Aglawe et al., 2018; Lusser et al., 2011; Seyran and Craig, 2018). To use these techniques efficiently, researchers must have prior knowledge of genome organization, identify relevant genes, and understand their function (Chen et al., 2019; Collard and Mackill, 2008; Varshney et al., 2020). However, this is not easy to achieve for the many agricultural crops and their relatives that have large and complex genomes (Pellicer and Leitch, 2020).

Early attempts to simplify genome analysis and gene discovery used methyl filtration (Palmer et al., 2003) and Cot-based fractionation (Peterson et al., 2002) to remove the repetitive parts of genomes to reduce their complexity (Fig. 1A). Later, exome capture methods were designed (Albert et al., 2007; Gnirke et al., 2009) to isolate protein-coding regions of genes in a genome and avoid non-coding and repetitive DNA (Fig. 1B). Both approaches eliminated large genome fractions and did not allow the assembly of longer stretches of DNA. To overcome this disadvantage and support the analysis of polyploid genomes, such as bread wheat, some authors sequenced genomes of diploid progenitors of crops as surrogates to facilitate gene mapping and cloning (Ling et al., 2018; Luo et al., 2017). Yet another, radically different, way of simplifying genome analysis and gene cloning is to isolate natural genome subunits – chromosomes (Doležel et al., 2012). Chromosomes are smaller and more defined parts of the genome, and dissecting a genome to its chromosomes reduces its complexity in a lossless manner (Fig. 1C). DNA prepared from individual chromosomes offers targeted approaches to gene mapping and cloning and avoids problems due to DNA sequence similarity between homologous chromosomes in polyploids.

Isolating chromosomal DNA requires cells in the metaphase of mitosis and the ability to collect chromosomes of interest. This can be done by picking up single chromosomes by micromanipulation (Matsunaga et al., 1999; Schondelmaier et al., 1993). Although this approach allows precise selection of chromosomes of interest, it results in low-

* Corresponding author.
E-mail addresses: zwyrtkova@ueb.cas.cz (J. Zwyrtková), simkovah@ueb.cas.cz (H. Šimková), dolezel@ueb.cas.cz (J. Doležel).

https://doi.org/10.1016/j.biotechadv.2020.107659
Received 4 July 2020; Received in revised form 10 October 2020; Accepted 13 November 2020
Available online 28 November 2020

throughput extraction of inferior-quality DNA (Hobza and Vyskot, 2007; Ma et al., 2010). A high-throughput alternative is chromosome sorting by flow cytometry. This method analyzes light scatter and the fluorescence properties of chromosomes at rates of 10^3 to 10^4/s during their movement in a single file within a narrow stream of liquid. If the optical properties of a chromosome differ from those of other chromosomes, that chromosome can be discriminated and purified in large quantities. Flow-sorted chromosomes are mechanically intact, and high-molecular-weight DNA can be prepared from them (Šimková et al., 2003).

De Laat and Blaas (1984) were the first to report flow cytometric sorting of plant chromosomes, purifying both chromosome types from *Haplopappus gracilis*. However, a broader use of the method was hampered by difficulties preparing suspensions of intact chromosomes suitable for flow cytometry. The two main problems were achieving a high degree of mitotic metaphase synchrony and releasing intact chromosomes from cells with rigid walls (Doležel et al., 1994). Several different biological systems were tested for chromosome isolation (Fig. 2). The quality of suspensions of intact chromosomes prepared from *in vitro* cultured cells of *H. gracilis* (de Laat and Blaas, 1984) and some other species (Arumuganathan et al., 1994; Wang et al., 1992) and from cultures of leaf mesophyll protoplasts (Conia et al., 1989, 1987) did not meet expectations. Faced with these difficulties, Doležel et al. (1992) pioneered a different approach and prepared chromosome samples from synchronized root tip meristems of young seedlings. Whereas previous authors had released chromosomes by lysing cells in a hypotonic buffer after removing cell walls with hydrolytic enzymes, Doležel et al. (1992) released chromosomes by mechanical homogenization of root tips fixed mildly by formaldehyde. The method was originally developed in field bean (*Vicia faba*) and to date has been used for chromosome sorting in 29 species (Supplementary Table 1). In some cases, genetically transformed 'hairy' root cultures were used as an alternative to young seedlings to obtain actively growing root tips (Veuskens et al., 1995; Neumann et al.,

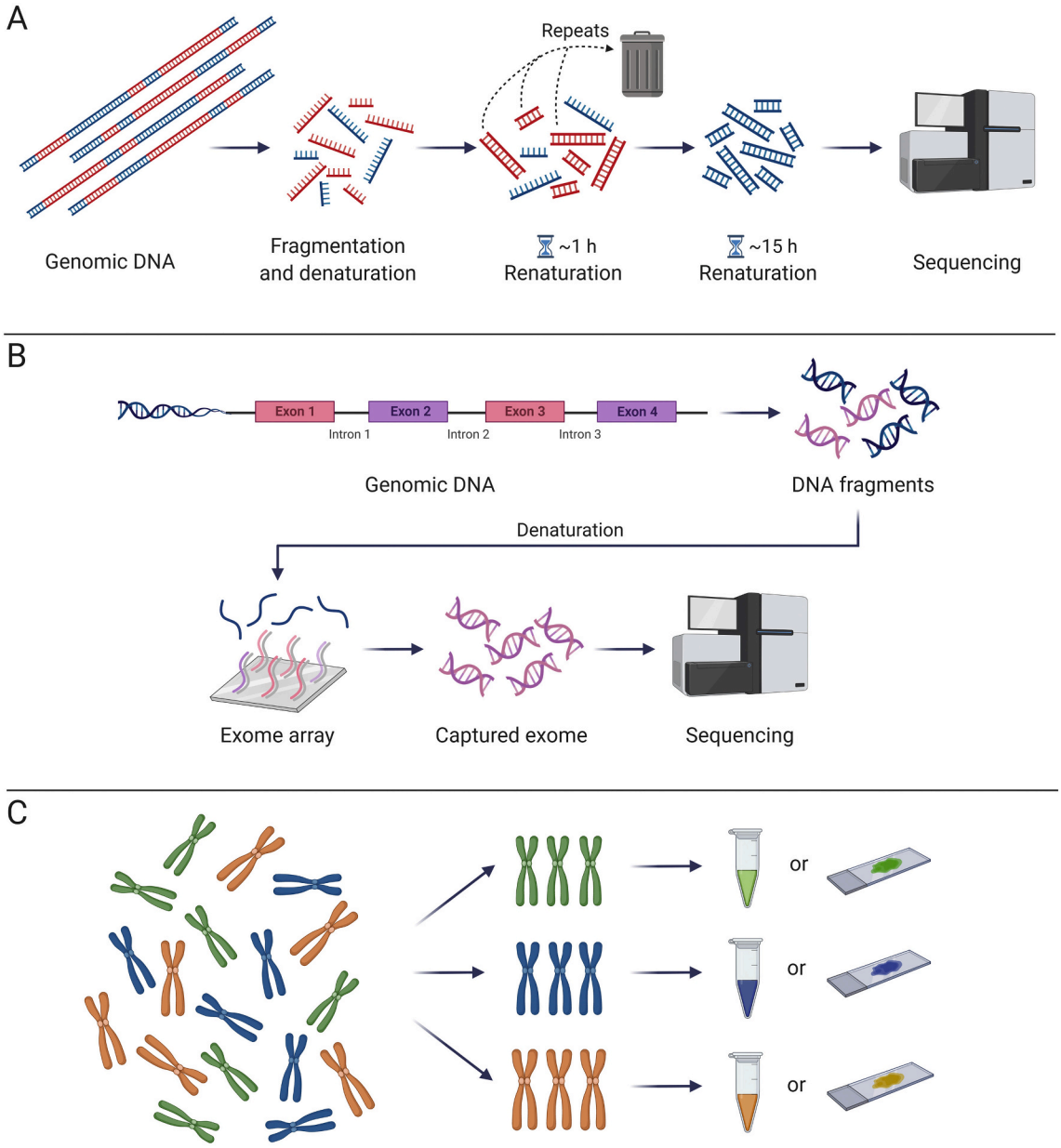


Fig. 1. Methods for genomic DNA complexity reduction. (A) So-called Cot-based fractionation takes advantage of different renaturation kinetics of repetitive DNA sequences and low-copy "genic" sequences to isolate the latter. (B) The exome-capture technique isolates protein-coding regions of genes by hybridizing sheared genomic DNA to oligonucleotides on a microarray or on beads in solution. Both approaches reduce sample complexity at the expense of eliminating a significant part of a genome. In contrast, chromosome sorting (C) reduces DNA sample complexity in a lossless manner by dissecting a genome into single chromosomes.

1998).

Another obstacle in applying chromosome flow sorting to plant genetics and genomics has been discriminating individual chromosome types. Many plant species have two or more chromosomes with similar DNA content that cannot be resolved based on size or the fluorescence of a dye bound to the DNA. Simultaneous measurement of the fluorescence of two dyes, one of them binding to AT-rich DNA regions and the other to GC-rich regions, which has been the gold standard in humans and animals (Dixon et al., 2009; Gray et al., 1979; Silverman et al., 1995), failed to improve chromosome discrimination in plants (Lee et al., 2000, 1997; Lucretti and Doležel, 1997). Alternative and successful approaches suitable for dissecting plant genomes include the use of chromosome deletion, translocation or addition lines (Doležel and Lucretti, 1995), fluorescent labeling of microsatellites using fluorescence *in situ* hybridization in suspension (FISHIS) (Giorgi et al., 2013), and single chromosome sorting (Cápal et al., 2015).

Since the first experiments demonstrated the possibility of isolating particular chromosomes by flow cytometric sorting, numerous uses of flow-sorted chromosomes in molecular biology, genomics and proteomics have been developed (Fig. 3). These applications were facilitated by the ability to purify chromosomes in quantities required for particular applications (Fig. 4) and made the sorted chromosomes an invaluable resource for important endeavors, including large international projects on sequencing complex plant genomes (International Wheat Genome Sequencing Consortium, 2018, 2014; Kreplak et al., 2019; Martis et al., 2013; Mayer et al., 2011). This review discusses these applications (Supplementary Table 2) and outlines their contribution to advancements in plant genome analysis and gene cloning (Fig. 5). Readers interested in flow cytometric chromosome sorting technology itself are referred to specialized publications (Doležel et al., 1999; Vrána et al., 2016).

2. Chromosome-specific DNA libraries

The isolation of chromosomes in large quantities by flow cytometry may take several days, as the process involves not only flow sorting itself but also the preparation of biological material, induction of cell cycle synchrony, accumulation of dividing cells in the metaphase of mitosis, and preparation of a suspension of intact chromosomes (Doležel et al., 1999; Vrána et al., 2016). If compatible with downstream applications, repeated chromosome sorting may be avoided by storing and multiplying chromosomal DNA in a form of DNA library that can be used as a permanent source of chromosomal DNA.

2.1. Short-insert DNA libraries

The first plant chromosome-enriched DNA library was generated by Wang et al. (1992) from 100,000 copies of flow-sorted chromosome 4A of bread wheat (*Triticum aestivum*). The library was constructed in a pUBs1 vector; half of the DNA clones were unique or low copy, and about half detected sequences on chromosome 4A. The task of constructing short-insert DNA libraries was simplified by a protocol for amplifying chromosomal DNA (Telenius et al., 1992), which reduced the number of chromosomes needed to only several hundred to 1000. Using this approach, Arumuganathan et al. (1994) prepared a library from 1000 copies of tomato (*Lycopersicon pennellii*) chromosome 2 using a Bluescript plasmid vector; the size of cloned inserts ranged from 250 to 1600 bp. Macas et al. (1996) were the first to construct a complete set of chromosome-specific DNA libraries covering a whole plant genome. The libraries, cloned in a plasmid vector, were constructed from seven chromosomes flow-sorted from two translocation lines of field bean. They comprised about 10^6 clones each, and their average insert sizes ranged from 310 to 487 bp. In a similar work, Požárková et al. (2002) constructed a library from chromosome 1 isolated from field bean translocation line JF. DNA amplified from the samples of 250 flow-sorted chromosomes was subjected to one or two rounds of hybridization-based enrichment for microsatellite (GA)n or (GATA)n and cloned into a plasmid vector. The frequency of positive clones increased 14.5-fold and 245-fold after one and two rounds of enrichment for (GA)n, respectively.

2.2. Long-insert DNA libraries

Some projects require large-insert DNA libraries, which are, however, more difficult to construct from chromosomes because of the need for microgram amounts of high-molecular-weight DNA. This translates to millions of copies of chromosomes that need to be flow-sorted. The preparation of high-molecular-weight DNA from flow-sorted chromosomes became possible after Šimková et al. (2003) modified the original protocol for chromosome isolation from synchronized root tips (Doležel et al., 1992). This progress enabled Šafář et al. (2004) to develop a low-input protocol for the construction of DNA libraries cloned in bacterial artificial chromosome (BAC) vector. The authors demonstrated its utility by constructing a BAC library from chromosome 3B of bread wheat. Following this, Janda et al. (2004) generated a BAC library from a pool of bread wheat chromosomes 1D, 4D, and 6D. The next advance was made by Janda et al. (2006), who constructed a BAC library from the short arm of bread wheat chromosome 1B (1BS) and thus demonstrated

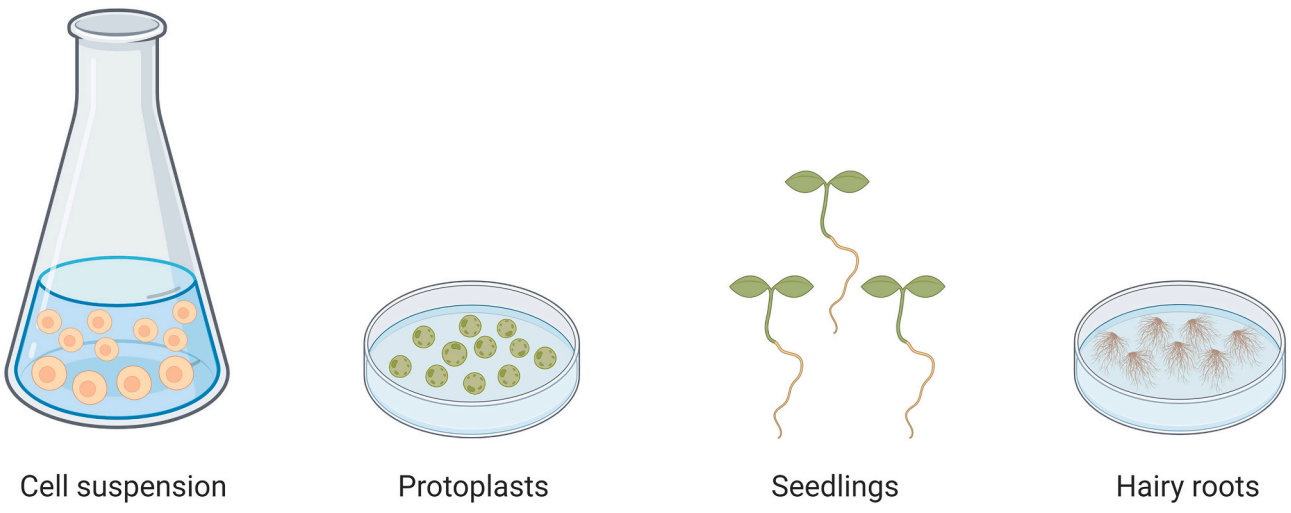


Fig. 2. Biological systems used for chromosome isolation. From left: Suspension-cultured cells, protoplasts isolated from leaf mesophyll, root tips of young seedlings, root tips of hairy root cultures.

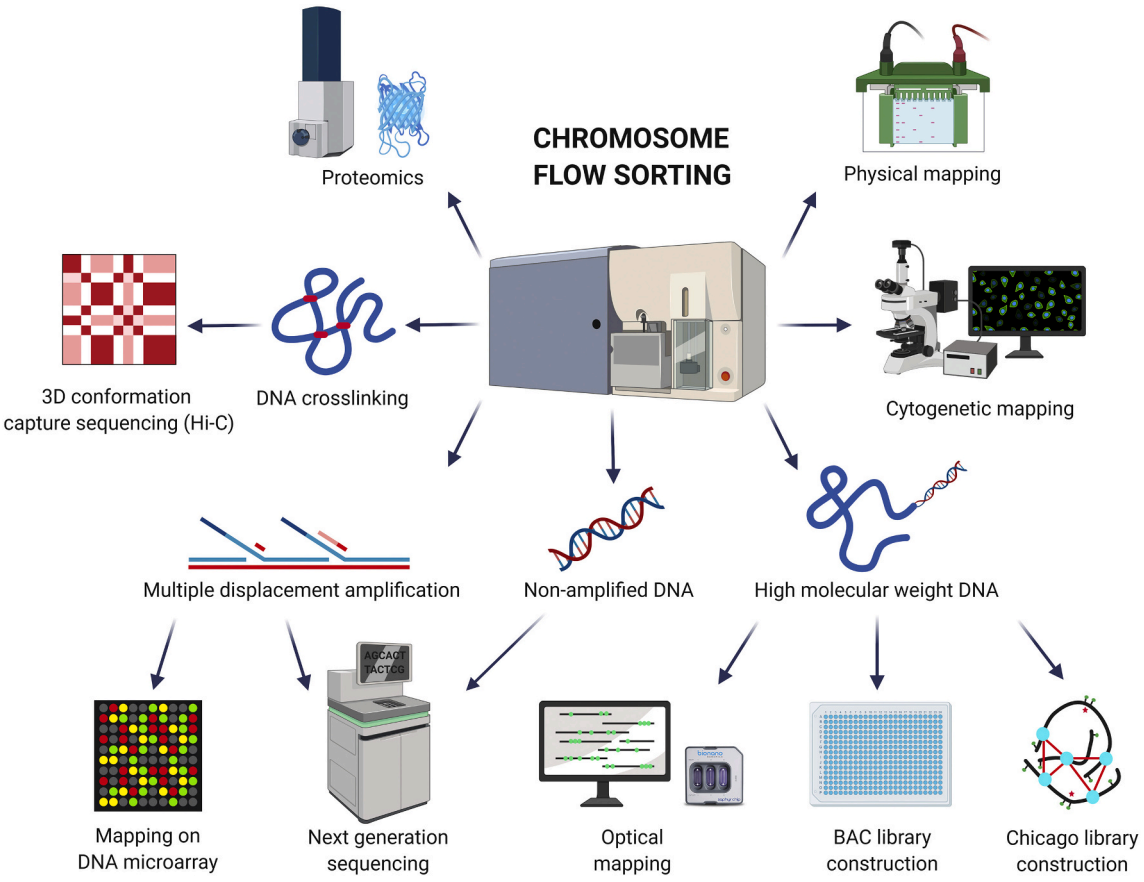


Fig. 3. Overview of methods for which flow-sorted chromosomes, their DNA, or proteins have been used.

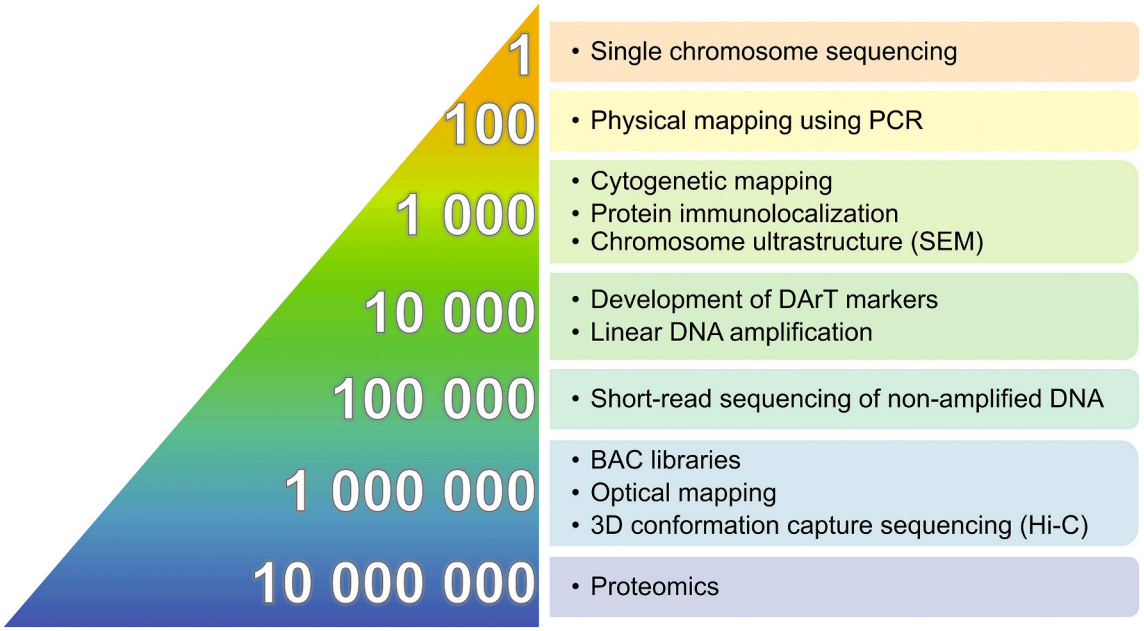


Fig. 4. The number of chromosomes needed for various downstream applications. Abbreviations: BAC (bacterial artificial chromosome), DArT (diversity arrays technology), PCR (polymerase chain reaction), SEM (scanning electron microscopy).

that even chromosome arm-specific BAC libraries can be constructed. To date, BAC libraries have been generated from 34 bread wheat chromosome arms and four chromosomes, including a group of 1D, 4D, and 6D chromosomes (International Wheat Genome Sequencing Consortium, 2018; Šafář et al., 2010) and two BAC libraries from the short arm of rye (*Secale cereale*) chromosome 1R (Šimková et al., 2008a). While early BAC libraries had an average insert size of <100 kb, those constructed with improved protocol comprise 85–95% of clones with inserts

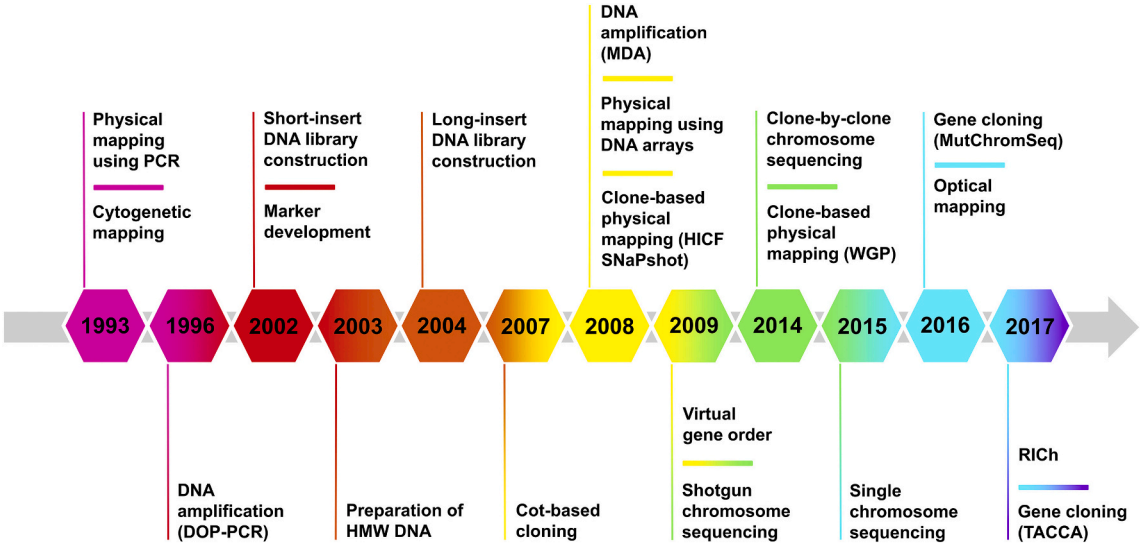


Fig. 5. Timeline of applications for which flow-sorted chromosomes have been used. Abbreviations: DNA (deoxyribonucleic acid), DOP (degenerate oligonucleotide-primed PCR), HICF (high information content fingerprinting), HMW (high molecular weight), MDA (multiple displacement amplification), PCR (polymerase chain reaction), RICh (rearrangement identification and characterization), TACCA (targeted chromosome-based cloning *via* long-range assembly), WGP (whole genome profiling).

>100 kb (https://olomouc.ueb.cas.cz/en/resources/dna-libraries, 2020).

3. Targeted development of DNA markers

DNA markers are an important resource for the development of genetic maps, the anchoring of physical maps, gene cloning, as well as marker-assisted selection (Powell et al., 1996). If a goal is to saturate the genetic map at a particular genome region, marker development may be simplified by targeting the relevant chromosome (Fig. 6). Restriction fragment length polymorphism (RFLP) markers were the first markers to be developed from isolated chromosomes (Arumuganathan et al., 1994). In this work, RFLP markers were developed from chromosome 2 of L. *pennellii* and used to saturate an RFLP linkage map of the species. However, as RFLP analysis requires large amounts of DNA and is slow and cumbersome, RFLP markers were soon replaced by other marker types.

3.1. Microsatellite (SSR) markers

Macas et al. (1996) showed that short-insert DNA libraries prepared from flow-sorted chromosomes and cloned into a plasmid vector were suitable for isolating microsatellite clones. To increase the microsatellite yield, Koblížková et al. (1998) optimized a protocol for enriching libraries for microsatellite motifs, and Požárková et al. (2002) used it to develop SSR markers from field bean chromosome 1. The markers were used to assign two genetic linkage groups to the chromosome and were also included in a set of markers used by Román et al. (2004) to construct a composite map of field bean. The short arm of rye chromosome 1 (1RS) carries important genes (McIntosh et al., 1988; Wricke and

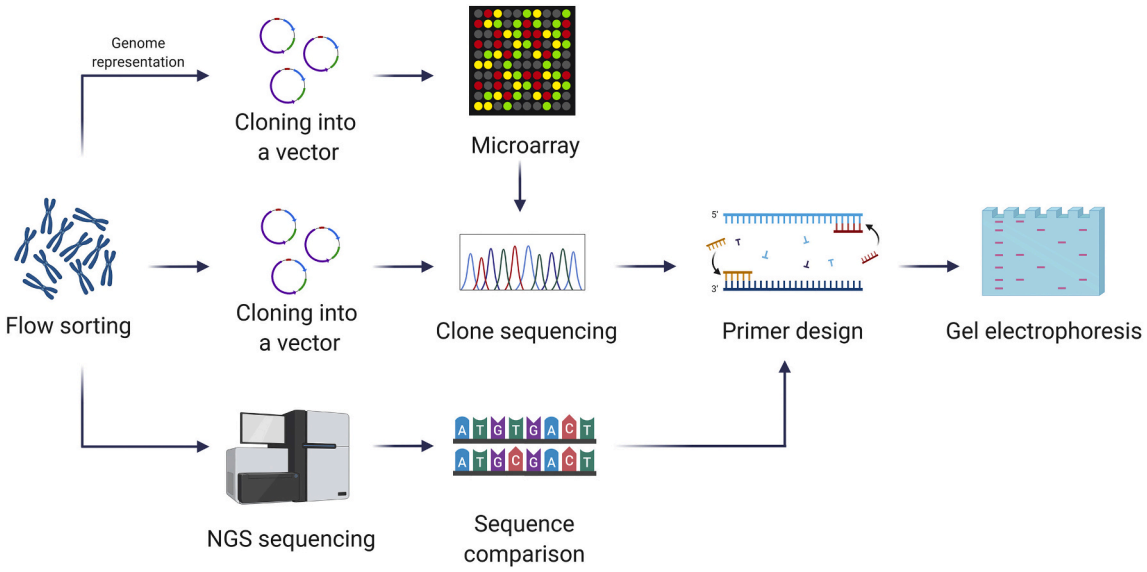


Fig. 6. Targeted development of DNA markers from a chromosome of interest. DNA from flow-sorted chromosomes may be cloned in a suitable vector, and randomly selected clones are sequenced. The yield of putative marker loci can be increased by prior selection of clones harboring particular DNA sequences. Alternatively, DNA of sorted chromosomes is shotgun-sequenced, and putative marker loci are identified *in silico*.

Wehling, 1985) and is present in many cultivars of bread wheat as 1BL·1RS and 1AL·1RS wheat-rye translocations (Baum and Appels, 1991). To saturate the 1RS genetic map, Kofler et al. (2008) flow-sorted 1RS telosomes and constructed four libraries enriched for microsatellite motifs AG, AAG, AC, and AAC. After sequencing 1290 clones, they developed 57 1RS-specific SSR markers. The potential of next-generation sequencing for high-throughput *in silico* identification of microsatellite loci on a particular chromosome was demonstrated by Nie et al. (2012). After sequencing DNA from the long arm of bread wheat chromosome 7D (7DL) on an Illumina HiSeq2000 platform, they identified 16,315 candidate markers from 161,061 assembled sequence contigs. Out of 33 randomly selected primer pairs, 26 produced PCR amplicons of the expected size, with 18 of them showing polymorphism in a panel of 20 cultivars of bread wheat.

3.2. Markers based on transposable element insertions

Bartoš et al. (2008) showed that chromosome arm-specific BAC libraries facilitate targeted development of DNA markers. After sequencing BAC ends of 1536 clones from a 1RS-specific BAC library (Šimková et al., 2008a), they identified transposable element junctions and developed insertion site-based polymorphism (ISBP) markers. Out of 64 primer pairs tested, 17 (26.6%) were found specific for 1RS. Next-generation sequencing of flow-sorted chromosomes was also useful in the development of chromosome-specific repeat-junction markers. Mazaheri et al. (2014) used shotgun sequences obtained from barley (*Hordeum vulgare*) chromosome 1H and from the arms of the remaining six barley chromosomes (Mayer et al., 2011) to identify a total of 9,881,561 transposable element junctions, from which they designed 400,538 repeat-junction markers across the genome, yielding 39 markers/Mb on average. More than 94% of 96 tested markers amplified one, or in a few cases two amplicons, and about 90% of them were genome specific. This marker platform supported the assembly of the barley genome.

3.3. Single nucleotide polymorphism (SNP) markers

The chromosome-centric approach has also been used to develop SNP markers. To construct genotype-specific genetic maps of chromosome 3B from bread wheat cultivars Arina and Forno, Shatalina et al. (2013) sequenced by Illumina technology to 10× coverage chromosome 3B isolated from both cultivars. Relying on a synteny with the *Brachypodium* genome and on gene annotation, they identified sequences close to coding regions and used them to develop 70 SNP-based markers. This work increased the number of markers 4-fold compared to the previous map of chromosome 3B based on SSR and RFLP markers. Staňková et al. (2015) used sequences from the flow-sorted chromosome arm to develop 11 SNP markers associated with the gene region for Russian wheat aphid resistance on wheat chromosome arm 7DS. Introgression of important genes from wild relatives to agricultural crops has been hampered by linkage drag and difficulties eliminating unwanted chromatin transferred simultaneously with the genes (Rey et al., 2015). To develop markers to monitor introgressions from the short arm of chromosome $5M^g$ ($5M^gS$) of *Aegilops geniculata* to bread wheat, Tiwari et al. (2014) sequenced by Illumina technology DNA from the $5M^gS$ arm, which they flow-sorted from a bread wheat-*Ae. geniculata* disomic chromosome substitution line. This led to the discovery of 2178 unique $5M^gS$-specific SNPs and the identification of a single SNP tagging chromosome T5DL·5DS-$5M^gS$, with the smallest introgression carrying leaf rust *Lr57* and stripe rust *Yr40* resistance genes. Nsabiyera et al. (2020) sequenced wheat chromosome 4B, which they flow-sorted from two parental lines of a recombinant inbred line population used to map the leaf rust resistance gene *Lr49*. They used the sequences from the flow-sorted 4B chromosomes and high-density SNP genotyping to fine-map the *Lr49* locus as a starting point to develop a diagnostic marker for use in breeding and to clone the resistance gene.

3.4. Intron length polymorphism markers

Haynaldia villosa is another important source of genes for improving bread wheat, and its chromosome 4 V carries a wheat yellow mosaic and a take-all resistance gene. Wang et al. (2017) flow-sorted the short arm of 4 V from a ditelosomic addition wheat-*H. villosa* line, assembled 4VS reads obtained by Illumina sequencing, and compared the assembly to the available assemblies of bread wheat and *Ae. tauschii* homoeologous chromosome arms. This allowed for the identification of exon-exon junctions and the localization of introns on the 4VS arm. A total of 359 intron length polymorphism markers were designed, 232 (64.62%) of which were specific for 4VS chromatin and suitable for tracing it in the background of bread wheat.

3.5. Diversity arrays technology (DArT) markers

Unlike the majority of DNA markers, the development of DArT markers (Jaccoud et al., 2001) did not require prior knowledge of DNA sequences. Wenzl et al. (2010) demonstrated that a chromosome-enriched DArT array can be developed directly from only a few nanograms of chromosomal DNA. Of 711 polymorphic markers derived from bread wheat chromosome 3B, 553 (78%) mapped to the chromosome, and even higher efficiency (87%) was observed for the short arm of bread wheat chromosome 1B (1BS). Linkage analyses showed that 553 of the 711 polymorphic 3B-derived markers (78%) mapped to chromosome 3B, and 59 of the 68 polymorphic 1BS-derived markers (87%) mapped to chromosome 1BS, confirming the efficiency of the chromosome-centric approach.

4. Physical mapping

4.1. Anchoring of genes and transgenes to chromosomes with PCR

Early work by Macas et al. (1993) demonstrated the potential of chromosome sorting for physical gene mapping, including the mapping of genes that cannot be mapped genetically because of a lack of allelic variants. Using chromosomes sorted from field bean (Lucretti et al., 1993), Macas et al. (1993) localized seed storage and non-storage protein genes to particular chromosomes using PCR with gene-specific primers. Lysák et al. (1999) confirmed the suitability of this approach for physically mapping genetic markers in barley and showed that PCR with translocation chromosomes flow-sorted from cytogenetic stocks permitted physical mapping at the subchromosomal level (Fig. 7A). Until a study by Neumann et al. (2002), the assignment of genetic linkage groups IV and VII of pea (*Pisum sativum*) to particular chromosomes was ambiguous. These authors flow-sorted chromosomes from translocation line JI148 and used them as templates for PCR to localize genetic markers selected from the two linkage groups. By assigning the linkage groups IV and VII to chromosomes 4 and 7, respectively, this work completed the assignment of all genetic linkage groups to individual chromosomes in pea. Using PCR with DNA amplified from single chromosomes that were flow-sorted from three transgenic lines of bread wheat as a template, Cápal et al. (2016) confirmed chromosomes carrying a transgene.

4.2. Physical mapping of conserved orthologous set (COS) markers with PCR

To investigate the structure and evolution of the C genome of *Aegilops*, an important source of genes for improving bread wheat, Molnár et al. (2015) flow-sorted chromosomes from *Ae. markgrafii* ($2n = 2x = 14$; CC) and its allotetraploid hybrids *Ae. triuncialis* ($2n = 4x = 28$; $U^tU^tC^tC^t$) and *Ae. cylindrica* ($2n = 4x = 28$; $D^CD^CC^CC^C$). DNA of sorted chromosomes was used for PCR with COS markers covering all seven wheat homoeologous chromosome groups. Various degrees of homology were observed between bread wheat and C-

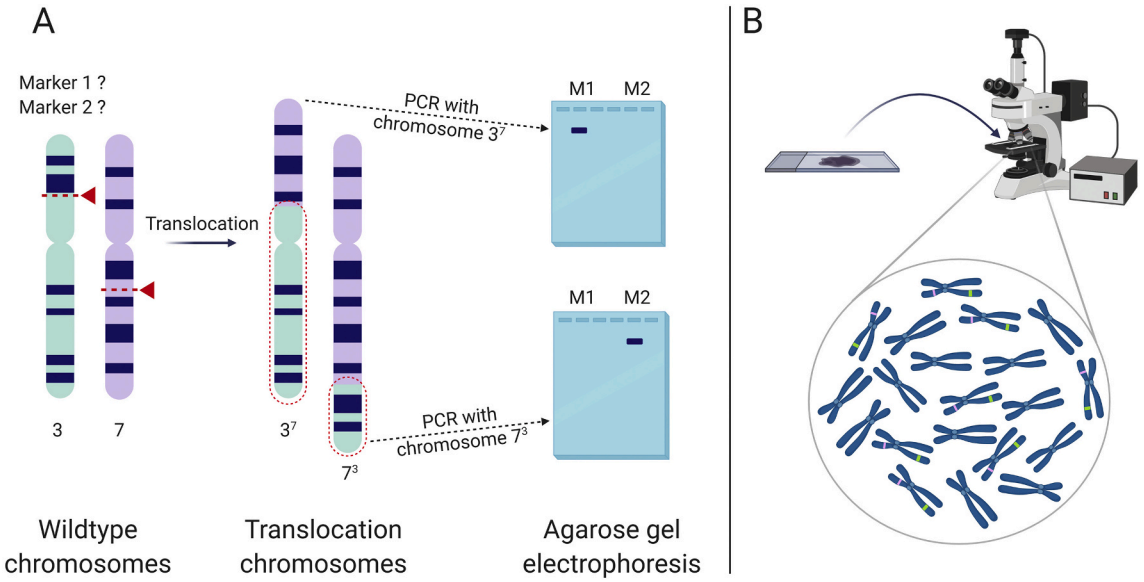


Fig. 7. The use of flow-sorted chromosomes for physical mapping of DNA sequences using PCR and microscopy. (A) Mapping two DNA markers (1 and 2) from chromosome 3 to subchromosomal regions using PCR is facilitated by reciprocal translocation between chromosomes 3 and 7. This separates two regions of chromosome 3 on translocation chromosomes 3^7 and 7^3. PCR with DNA of the sorted translocation chromosomes as templates assigns each of the two markers to a different region of chromosome 3. (B) Large numbers of chromosomes are flow-sorted onto a microscope slide and used as template for fluorescence *in situ* hybridization (FISH). Many chromosomes concentrated on a small spot facilitate microscopic evaluation. As the preparation is free of cell walls and cytoplasmic remnants, and the chromosome can be longitudinally extended, FISH mapping can be done without nonspecific background signals and at a higher spatial resolution.

genome chromosomes, with chromosome 7C being the most rearranged relative to bread wheat. In similar work, Molnár et al. (2016) flow-sorted chromosomes from *Ae. umbellulata* (2n = 2x = 14; UU), *Ae. markgrafii* (2n = 2x = 14; CC), *Ae. speltoides* (2n = 2x = 14; SS), and *Ae. comosa* (2n = 2x = 14; MM). Mapping of COS markers revealed significant rearrangements in the U and C genomes, whereas the M and S genomes exhibited a structure similar to bread wheat. The results of both studies support alien introgression breeding of bread wheat.

4.3. Mapping of marker sequences on DNA arrays

A protocol for representative amplification of chromosomal DNA (Šimková et al., 2008b) provided an opportunity for high-throughput physical mapping on DNA arrays. Using DNA amplified from flow-sorted chromosomes and chromosome arms of barley and oligonucleotide arrays BOPA1 and BOPA2, Muñoz-Amatriaín et al. (2011) localized 2930 genes and improved marker order and the resolution of the barley consensus genetic map by almost 20%. The results confirmed the position of 2545 SNP-mapped loci and added chromosome or chromosome arm allocations to an additional 370 SNP loci. The availability of DNA from the arms of chromosomes 2H–7H enabled the delineation of pericentromeric regions for these chromosomes. Using flow-sorted chromosomes was a more robust approach to allocating SNP loci than using DNA isolated from disomic bread wheat-barley chromosome addition lines.

4.4. Cytogenetic mapping

Lucretti et al. (1993) showed that chromosomes sorted onto a microscope slide could be used as a template for localizing DNA sequences using fluorescence *in situ* hybridization (FISH). This observation stimulated the use of FISH to identify flow-sorted chromosomes based on chromosome-specific fluorescent labeling patterns of various types of repeats (Kubaláková et al., 2003, 2005; Molnár et al., 2016). Performing FISH on hundreds of chromosomes in a small area of a slide has an advantage over performing traditional FISH on preparations made following the traditional methods (Fig. 7B). Thus, after localizing GAA microsatellites on flow-sorted chromosomes by FISH, Kubaláková et al. (2002) not only described polymorphisms in banding pattern among the cultivars of bread wheat but also revealed unexpected intravarietal variation in the GAA banding pattern. Performing FISH on a large number of B chromosomes flow-sorted from a line of rye, Kubaláková et al. (2003) identified and characterized rare A–B translocations.

As thousands of chromosomes can be sorted in a short time, are free of cell walls and cytoplasmic remnants, and are longitudinally extended, FISH mapping is done without nonspecific background signals and at a higher spatial resolution. Using this approach, Janda et al. (2006) localized low-copy subclones shorter than 2 kb developed from a wheat 1BS-specific BAC library on flow-sorted 1BS arms. Cápal et al. (2016) used FISH with flow-sorted chromosomes to determine the chromosomal location of a barley sucrose transporter construct in three transgenic lines of bread wheat.

With the aim of increasing the spatial resolution of FISH, Valárik et al. (2004) developed a protocol for the longitudinal extension of chromosomes sorted onto a microscope slide more than 100 times the size of the original mitotic metaphase. The maximum observed sensitivity was 1 kb, and the spatial resolution of neighboring loci was improved from 5 to 10 Mb to 70 kb with mitotic metaphase spreads. An alternative protocol for hyperexpansion of flow-sorted chromosomes produced chromosomes more than 7-fold longer (Endo et al., 2014).

4.5. Clone-based physical mapping

Clone-by-clone sequencing was the method of choice for genome sequencing before the advent of next-generation sequencing technologies (International Rice Genome Sequencing Project and Sasaki, 2005; Schnable et al., 2009; The Arabidopsis Genome Initiative, 2000). This approach required physical contig maps developed by ordering large-insert DNA clones. The original plan to sequence the huge genome of hexaploid bread wheat envisaged constructing BAC-based physical maps for each of its 21 chromosomes and sequencing clones from minimum tiling paths. Paux et al. (2008) demonstrated the feasibility of developing such maps by constructing a physical map of the largest bread wheat chromosome, 3B. They fingerprinted a 3B BAC library created by

Šafář et al. (2004) with a High Information Content Fingerprinting SNaPshot protocol (Luo et al., 2003) and assembled 82% of the chromosome into 1036 contigs, which were anchored with 1443 genetic markers. The map also opened up avenues for the targeted development of molecular markers and map-based gene cloning. A similar approach was used to develop physical maps for chromosome arms of other bread wheat chromosomes (Akpinar et al., 2015b; Barabaschi et al., 2015; Belova et al., 2014; Breen et al., 2013; Feng et al., 2020; Holušová et al., 2017; Keeble-Gagnère et al., 2018; Lucas et al., 2013; Philippe et al., 2013; Raats et al., 2013; Salina et al., 2018; Tulpová et al., 2019a).

BAC-based physical maps from bread wheat chromosomes were also developed with whole genome profiling (WGP) technology (Van Oeveren et al., 2011). One of the advantages of WGP was the production of sequence tags, which allowed for *in silico* anchoring of almost 80% of the map against genetic markers with known sequences (Poursarebani et al., 2014). Thus, Poursarebani et al. (2014) constructed physical maps of both arms of bread wheat chromosome 6A, and subsequently physical maps were constructed after WGP of both arms of chromosome 6B (Kobayashi et al., 2015); chromosomes 2B, 2D, and 4B; and chromosome arms 5BL and 5DL (International Wheat Genome Sequencing Consortium, 2018).

5. Clone-by-clone chromosome sequencing

The availability of a clone-based physical map of bread wheat chromosome 3B (Paux et al., 2008) allowed Choulet et al. (2014) to produce a reference sequence of the first and largest chromosome of the species. They used a hybrid sequencing and BAC pooling strategy to sequence by Roche 454 technology 8452 BAC clones from the minimum tiling path. They then built a chromosome 3B pseudomolecule representing 774.4 Mb (93% of the complete sequence) by ordering 1358 scaffolds along the chromosome. Keeble-Gagnère et al. (2018) produced the assembly of wheat chromosome 7A after sequencing pools of 20 or more BAC clones from minimum tiling paths of both chromosome arms using Illumina paired-end technology. Tulpová et al. (2019a) used Illumina technology to complete paired-end sequencing of 4608 BAC clones from the minimum tiling path of bread wheat chromosome arm 7DS using a pooling strategy, in which each pool consisted of four non-overlapping clones. The assembly comprised 9063 scaffolds with an average size of 63 kb. BAC clones from chromosome (arm) BAC libraries were used for many other chromosomes, chromosome arms, or particular genome loci (Feng et al., 2020; International Wheat Genome Sequencing Consortium, 2018; Lu et al., 2018; Nesterov et al., 2016; Sergeeva et al., 2017; Shorinola et al., 2017). Researchers obtained the bread wheat reference genome RefSeq v1.0 (International Wheat Genome Sequencing Consortium, 2018) by combining available BAC assemblies with a whole genome short-read (Illumina) assembly. The sequences were ordered with information from the BAC-based physical maps; chromatin-conformation-capture-based technique Hi-C; and genetic, optical, and radiation-hybrid maps.

6. Shotgun chromosome sequencing

The advent of *de novo* genome assemblers suitable for short sequence reads from next-generation sequencing technologies made chromosome-based sequencing more accessible and attractive by avoiding the laborious construction of BAC libraries. High molecular weight DNA was not needed for the short-read technologies, and thus microgram amounts of DNA could be obtained after multiple displacement amplification (Šimková et al., 2008b). The few tens of thousands of chromosomes needed as a template, equal to tens of nanograms of DNA, could be isolated within 1 h of flow sorting. Amplified chromosomal DNA was found suitable for a range of downstream applications, including short-read next-generation sequencing (Mayer et al., 2009) and physical mapping on DNA arrays (Muñoz-Amatriaín et al., 2011). Further progress in the preparation of Illumina sequencing libraries and release of protocols requiring nanogram amounts of DNA, or less, made it possible to skip the chromosome amplification step, resulting in better assemblies compared to those from amplified DNA (Dracatos et al., 2019).

Difficulties discriminating and sorting a particular chromosome may compromise the potential of chromosome genomics in some species or their accessions. To overcome this bottleneck, Cápal et al. (2015) developed a protocol for amplifying DNA from single copies of chromosomes. In a pilot experiment, 1.4 μg amplified DNA was obtained on average from three single chromosomes of bread wheat chromosome 3B, with fragment sizes ranging from 300 bp to more than 20 kb; the majority of fragments were around 10 kb. Sequencing DNA pools from the three amplifications resulted in coverage of 60% of the 3B reference sequence and entire coverage of 30% of the genes on the chromosome, with only 16% of genes not mapped. This advance expanded the potential of chromosome genomics as it may be applied to any plant species from which chromosome samples suitable for flow cytometric sorting can be prepared. It also opened up new avenues for studies on chromosome structural heterozygosity and haplotype phasing in plants similar to those done in humans (Chen et al., 2017; Yang et al., 2011).

6.1. Draft sequences from particular chromosomes in agricultural crops

Mayer et al. (2009) were the first to demonstrate that low-pass (~1.3× coverage) next-generation sequencing (Roche 454 technology) of amplified chromosomal DNA is sufficient to characterize DNA repeats and identify genic sequences on a chromosome. By comparing barley chromosome 1H sequences with the genomes of rice (*Oryza sativa*) and *Sorghum* and with expressed sequence tags of barley and bread wheat, the authors identified more than 5000 genic sequences on the chromosome. Based on synteny-based analyses they constructed a virtual gene order map of barley chromosome 1H.

The ability to sequence DNA from a particular chromosome free of DNA from other chromosomes made it possible to study the molecular organization and evolution of rye B chromosomes, which remained a mystery until the work of Martis et al. (2012). This team sequenced by Roche 454 technology DNA amplified from flow-sorted B chromosomes to 0.9× coverage to characterize the molecular composition of these chromosomes and identify almost 5000 gene-derived sequences. This work traced the origin of rye B chromosome to fragments of A chromosomes and dated its origin ~1.1–1.3 Mya, close to the origin of the genus *Secale* ~ 1.7 Mya.

Because of its relative ease, chromosome shotgun sequencing was used extensively in species with large genomes for which whole genome assemblies were either not yet available or not available from a particular genotype. This cost-effective approach provided chromosome-specific and, in the case of polyploids, homoeologue-specific information. Thus, Wicker et al. (2011) sequenced the arms of bread wheat chromosome 1 group by Roche 454 technology, identified genic sequences, and analyzed their conservation between barley and bread wheat and between the subgenomes of bread wheat. Vitulo et al. (2011) sequenced bread wheat chromosome 5A by Roche 454 technology to identify transposable elements, genes, and miRNAs and inferred a virtual gene order. Chromosome 4A of bread wheat has a unique structure because of evolutionary changes (Devos et al., 1995). After sequencing flow-sorted arms of 4A by Roche 454 technology, Hernandez et al. (2012) constructed an ordered chromosome gene map and used it to localize translocation and inversion breakpoints that differentiate the chromosome from a putative ancestor of the tribe Triticeae.

The sequencing of flow-sorted chromosomes was instrumental in numerous other studies for unraveling molecular organization and evolution, characterizing repeat content, identifying miRNA and tRNA genes, and localizing genes underlying agronomically important traits for breeding programs (Akhunov et al., 2013; Akpinar et al., 2015c; Belova et al., 2013; Berkman et al., 2012, 2011; Garbus et al., 2015; Helguera et al., 2015; Kantar et al., 2012; Liu et al., 2016; Lucas et al., 2014; Sergeeva et al., 2014; Tanaka et al., 2013; Thind et al., 2018).

6.2. Draft chromosome sequences of crop wild relatives

Flow sorting was used extensively to sequence particular chromosomes from crop wild relatives. Akpinar et al. (2015a) sequenced by Roche 454 technology DNA amplified from chromosome 5D of *Ae. tauschii*, the D genome donor of bread wheat. Among other things, the authors characterized the DNA repeat content of the chromosome, predicted 6188 gene loci, and assessed the conservation of gene order between wheat and *Ae. tauschii*. A more distant relative of bread wheat, *Ae. geniculata* belongs to its tertiary gene pool and is an attractive source of important genes. To facilitate its use in introgression breeding, Tiwari et al. (2015) sequenced by Illumina technology chromosome 5 Mg purified from a bread wheat-*Ae. geniculata* disomic chromosome substitution line, constructed the chromosome virtual gene order, and used the data to develop SNP markers specific for chromosome 5 Mg. *H. villosa* is another member of the tertiary gene pool of bread wheat and a source of disease resistance genes. Xiao et al. (2017) flow-sorted and sequenced by Illumina technology the short arm of its chromosome 4 V (4VS) to provide an insight into its organization and facilitate the discovery of candidate genes. Apart from characterizing DNA repeats and precursors of miRNAs, and assessing synteny between genomes of homoeologous chromosome group 4 of Triticeae and other grasses, this work resulted in the identification of 1977 gene loci and a virtual gene order map comprising 735 gene loci on the 4VS arm.

6.3. Chromosome-based draft genome assemblies

Mayer et al. (2011) extended the chromosome-based approach to obtain a draft sequence of the whole barley genome and improved it by incorporating full-length cDNA and DNA hybridization microarray data. The assembly of more than 21,000 genes in a putative linear order ("genome zipper") enabled the analysis of structural conservation between model grass genomes of *Brachypodium*, rice and *Sorghum*, and barley and bread wheat. As only chromosome 1H could be discriminated from other barley chromosomes, chromosome arms 2HS to 7HL were flow-sorted from bread wheat-barley telosome addition lines. This turned out to be advantageous, as it made it possible to identify centromere positions and genes located in these regions. A similar approach based on sequencing amplified chromosomal DNA by Roche 454 technology was used by Martis et al. (2013) to produce a draft genome sequence of the ~8 Gb genome of rye. Whereas chromosome 1R could be flow-sorted directly from a rye accession, chromosomes 2R–7R were sorted from wheat-rye disomic chromosome addition lines. The work delivered a virtual linear gene order model comprising 22,426 genes representing 72% of the identified 31,008 genes. Among other things, the study recognized six major translocations that shaped the genome of modern rye compared to the genome of a putative ancestor of the tribe Triticeae.

Probably the best demonstration of the power of the chromosome-centric approach to sequencing complex genomes has been the development of an ordered and structured draft sequence of the ~17 Gb genome of hexaploid bread wheat by International Wheat Genome Sequencing Consortium (2014). To overcome difficulties discriminating individual chromosomes, they flow-sorted chromosome arms from telosomic lines of bread wheat. Unlike in the sequencing of chromosomes in barley and rye, amplified chromosomal DNA was sequenced by Illumina technology to a depth of 30× to 241×. A total of 124,201 gene loci were annotated and more than 75,000 genes were positioned along chromosomes after high-density wheat SNP mapping and based on synteny to sequenced grass genomes. Apart from delivering a draft genome for the important crop, the authors anchored more than 3.6 million marker loci to chromosome sequences, uncovered the molecular organization of the three homologous genomes, and described patterns in gene expression across the three genomes. The study also provided new insights into the phylogeny of hexaploid bread wheat.

7. Validation and improvement of whole genome assemblies

Despite the progress being made in whole genome sequencing technologies and sequence assembly algorithms, genome assemblies may be compromised by misplaced or chimeric contigs and scaffolds. To assess and validate the assembled pseudomolecules from the genome assemblies of kabuli and desi varieties of chickpea (*Cicer arietinum*), Ruperao et al. (2014) flow-sorted individual chromosomes from both varieties and sequenced amplified chromosomal DNA by Illumina technology. Mapping of the chromosome sequence reads to the available genome assemblies (Jain et al., 2013; Varshney et al., 2013) revealed short misassembled regions in the kabuli genome assembly and large-scale misassemblies in the desi genome assembly. The identification of the misplaced regions provided an opportunity to relocate these genomic regions to their correct pseudomolecules and produce improved genome assemblies.

The production of the reference genome of pea by Kreplak et al. (2019) was supported by the availability of chromosome-derived sequences. Because of difficulties discriminating pea chromosomes, the authors sequenced DNA amplified from single copies of particular chromosomes. Mapping chromosome-specific reads to the genome assembly identified scaffolds that contained contigs from different chromosomes and allowed them to be split into smaller scaffolds. In the same work, Kreplak et al. (2019) identified chromosome translocations discriminating cultivated *P. sativum* from its wild relatives *P. fulvum*, *P. sativum elatius*, and *P. sativum southern humile*. They did this by sequencing DNA amplified from single flow-sorted chromosomes from the three wild species and comparing the sequences to the genome assembly of *P. sativum* (Fig. 8).

7.1. Optical mapping

The assembly of complex genome regions is facilitated by optical mapping technology (Schwartz et al., 1993; Lam et al., 2012), which produces maps of short sequence motifs (restriction sites) from single long DNA molecules. Using the short arm of bread wheat chromosome 7D (7DS), Staňková et al. (2016) demonstrated the feasibility of generating optical maps from DNA of flow-sorted chromosomes (Fig. 9). The 7DS map comprised 371 contigs with an N50 of 1.3 Mb. The power of optical mapping was confirmed in this study with the identification of an array of tandem repeats about 800 kb long that could not be assembled by any other available sequencing technology. Tulpová et al. (2019a) used the 7DS optical map to validate and anchor a BAC-based physical map of the 7DS arm. Keeble-Gagnère et al. (2018) developed optical maps from flow-sorted arms of bread wheat chromosome 7A and combined them with sequence scaffolds of BAC clones to obtain 18 super-scaffolds across the chromosome. Optical maps constructed from the bread wheat chromosomes 7A, 7B, and 7D were also used to support the production of the reference genome of bread wheat (International Wheat Genome Sequencing Consortium, 2018).

8. Gene cloning

Chromosome genomics facilitates the cloning of genes from species and/or genotypes for which whole genome reference assemblies are not available (Bettgenhaeuser and Krattinger, 2019; Zhang et al., 2020). To discover molecular mechanism controlling the dominant, environmentally dependent chlorosis phenotype in wheat, Harrington et al. (2019) screened the TILLING population of the sequenced *T. turgidum* cv. Kronos and identified a single region on chromosome 3A: Yellow Early Senescence 1 (YES-1). Additional SNP markers for fine-mapping were identified after the sequencing of flow-sorted chromosome 3A. By mapping the reads against both the bread wheat reference genome and cv. Kronos assembly, the authors identified SNPs in non-coding regions within the YES-1 region, which enabled fine-mapping of the YES-1 locus to 4.3 Mb with 59 genes. To identify introgressed alien chromatin and

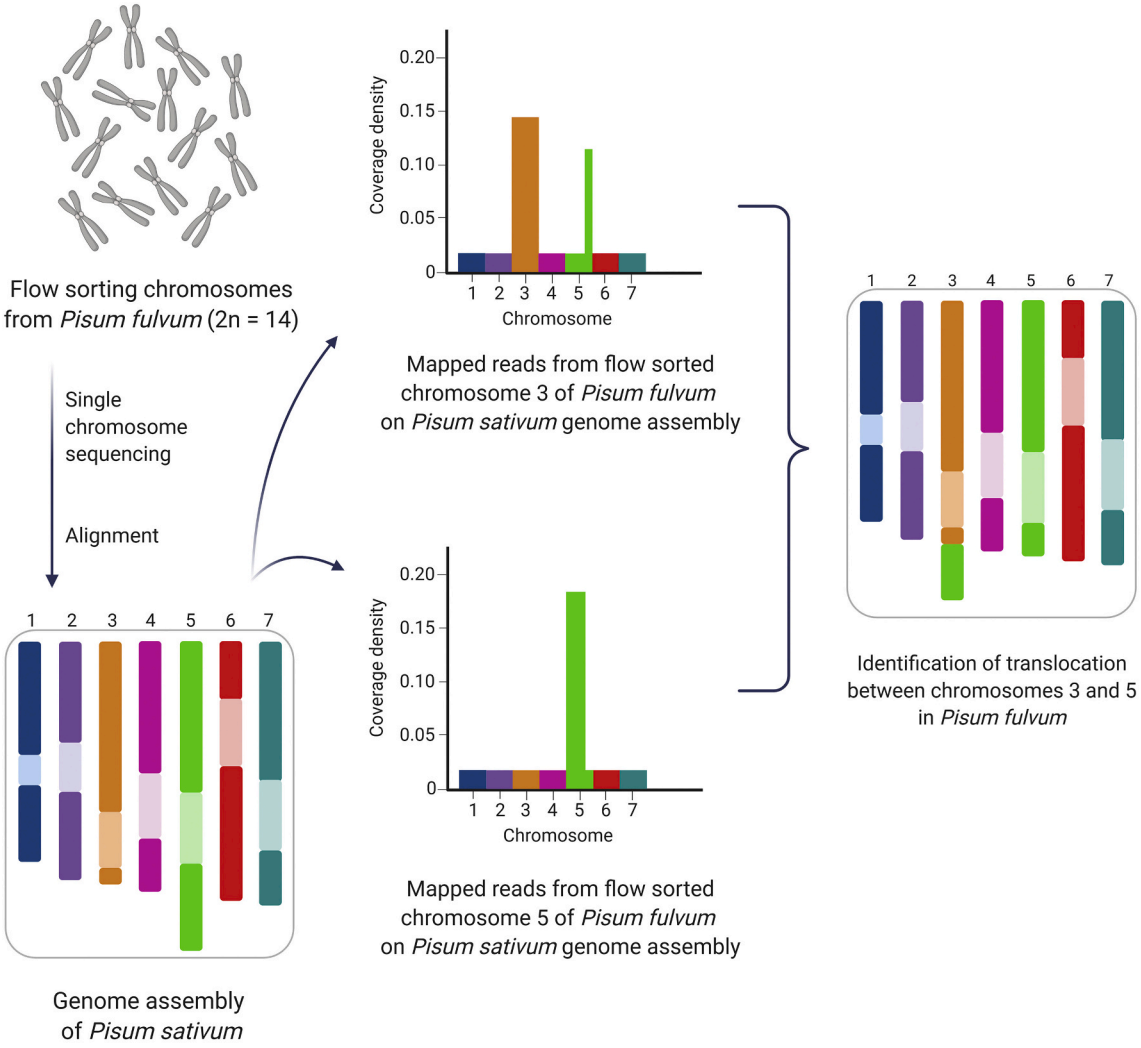


Fig. 8. Identification of chromosome translocation discriminating cultivated *Pisum sativum* from its wild relative, *P. fulvum*. DNA amplified from single flow-sorted *P. fulvum* chromosomes was sequenced and the reads were mapped to *P. sativum* pseudomolecules. Reads from a *P. fulvum* chromosome corresponding to *P. sativum* chromosome 3 were mapped on *P. sativum* pseudomolecules of chromosomes 3 and 5. This indicated a translocation between chromosomes 3 and 5 in *P. fulvum* as compared to *P. sativum*. The translocation was confirmed by the lack of mapping reads from *P. fulvum* chromosome corresponding to *P. sativum* chromosome 5 from ~465 Mb to its end (Kreplak et al., 2019, modified).

support cloning of introgressed genes, Abrouk et al. (2017) developed an *in silico* method called rearrangement identification and characterization (RICh). To characterize a chromosome segment transferred from *T. militinae* to the long arm of chromosome 4A (4AL) of bread wheat cv. Tähti, the team developed a virtual gene order of cv. Tähti chromosome 4A. A sequence comparison of the wild-type 4AL arm and 4AL arm with the introgression, which carried powdery mildew resistance gene *QPm-tut-4A*, identified alien chromatin with 169 putative genes originating from *T. militinae*. A similar approach was used by Bansal et al. (2020) to fine-map leaf rust and stripe rust resistance genes *Lr76* and *Yr70* introduced from *Ae. umbellulata* to bread wheat. These authors sequenced flow-sorted chromosomes 5 U of *Ae. umbellulata*, 5D from a bread wheat-*Ae. umbellulata* introgression line, and 5D from the recurrent parent. Using the sequences thus obtained, newly developed markers, and comparisons to the bread wheat reference genome (International Wheat Genome Sequencing Consortium, 2018), the authors delineated the introgression to a 9.47 Mb region, in which they identified candidates for *Lr76* and *Yr70* genes.

To clone Russian wheat aphid resistance gene *Dn2401*, mapped to the short arm of chromosome 7D of a resistant line CI2401 of bread wheat, Tulpová et al. (2019b) used several resources developed from flow-sorted chromosomes. These included shotgun sequences from short arms of homoeologous group 7 chromosomes of bread wheat cv. Chinese Spring, sequences of five BAC clones from a 7DS-specific library that spanned the *Dn2401* interval in Chinese Spring, and optical maps of 7DS arms from cv. Chinese Spring and from the CI2401 line. The complete sequence of the *Dn2401* region in cv. Chinese Spring together with optical maps of the susceptible and resistant lines facilitated the identification of six high-confidence genes, with *Epoxide hydrolase 2* being the most likely *Dn2401* candidate.

Thind et al. (2017) successfully cloned wheat leaf rust resistance gene *Lr22a* after producing a high-quality *de novo* assembly of chromosome 2D from resistant line CH Campala *Lr22a*. DNA of the flow-sorted chromosome was sequenced, and the resulting sequences were scaffolded with Chicago long-range linkage (Putnam et al., 2016). The assembly comprised 10,344 scaffolds with an N50 of 9.76 Mb, and the longest scaffold was 36.4 Mb. The development of additional markers reduced the *Lr22a* region to 438 kb and facilitated the identification of the *Lr22a* gene. The authors called this approach "targeted chromosome-based cloning via long-range assembly" (TACCA). The TACCA strategy was then used by Xing et al. (2018) to clone powdery mildew resistance gene *Pm21*, which was introduced to bread wheat from *H. villosa* chromosome 6 V. After identifying the candidate gene NLR1-V at the *Pm21* locus using multiple experimental methods, the authors flow-sorted,

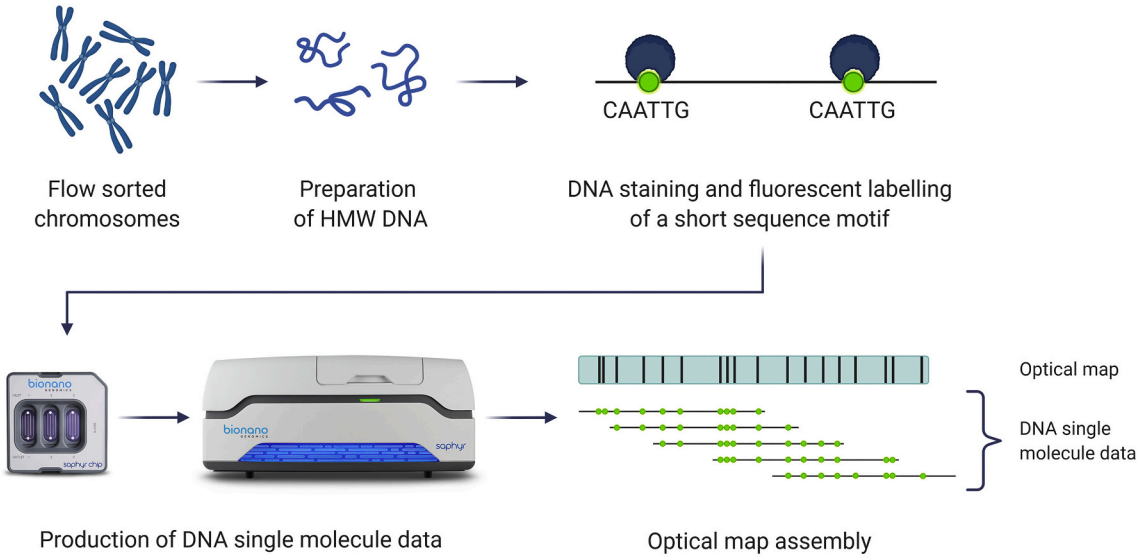


Fig. 9. Optical mapping is used to construct maps of short DNA sequence motifs based on microscopic images of long DNA molecules. Prior to analysis on a multichannel nanofluidic chip, the molecules are stained with a DNA fluorochrome and the sequence motif is labeled with a different fluorochrome. Optical map contigs, spanning hundreds of kilobases to more than one hundred megabases, bridge repeat-rich genome regions and serve as a clue to scaffold genome sequences and detect copy number and structural variation. Metaphase chromosomes purified by flow cytometry are a superior resource for the preparation of ultra-high-molecular-weight DNA from plant material. Moreover, construction of optical maps from flow-sorted chromosomes reduces DNA sample complexity, enabling targeted analyses in complex/polyploid genomes.

sequenced, assembled, and scaffolded translocation chromosome 6 *VS*·6AL *via* Chicago long-range linkage. This provided the physical structure of the *Pm21* locus and revealed mutations in NLR1-V orthologues in the subgenomes of bread wheat.

A gene cloning strategy that does not require high-resolution genetic mapping and does not exclude any DNA sequence from being targeted was developed by Sánchez-Martín et al. (2016). Mutant chromosome sequencing (MutChromSeq) involves comparing sequences from wild-type parental chromosome to chromosomes from several independently derived mutants to identify causative mutations in a candidate gene (Fig. 10). The authors demonstrated the benefits of this approach by re-cloning barley *Eceriferum-q* gene and *de novo* cloning wheat powdery mildew resistance gene *Pm2*. Ford et al. (2018) used MutChromSeq to identify a semi-dwarfism gene in bread wheat. Sequencing chromosome 6A from mutants of semi-dwarf *T. turgidum* cv. Icaro with a range of tall phenotypes revealed independent mutations in the coding region of *GA2oxA9* predicted to encode gibberellin 2-oxidase. The lower plant height was the result of increased expression of *GA2oxA9*, which led to a lower content of bioactive gibberellin. MutChromSeq was also used to clone the first leaf rust resistance gene *Rph1* from barley (Dracatos et al., 2019). Flow-sorted chromosome 2H from wild-type barley cv. Sudan and six mutants were sequenced and compared to identify a single gene candidate encoding a coiled-coil nucleotide binding site leucine-rich repeat (NLR) receptor protein. Recently, Hiebert et al. (2020) used MutChromSeq to isolate the *SuSr-D1* gene that suppresses resistance to stem rust in bread wheat by

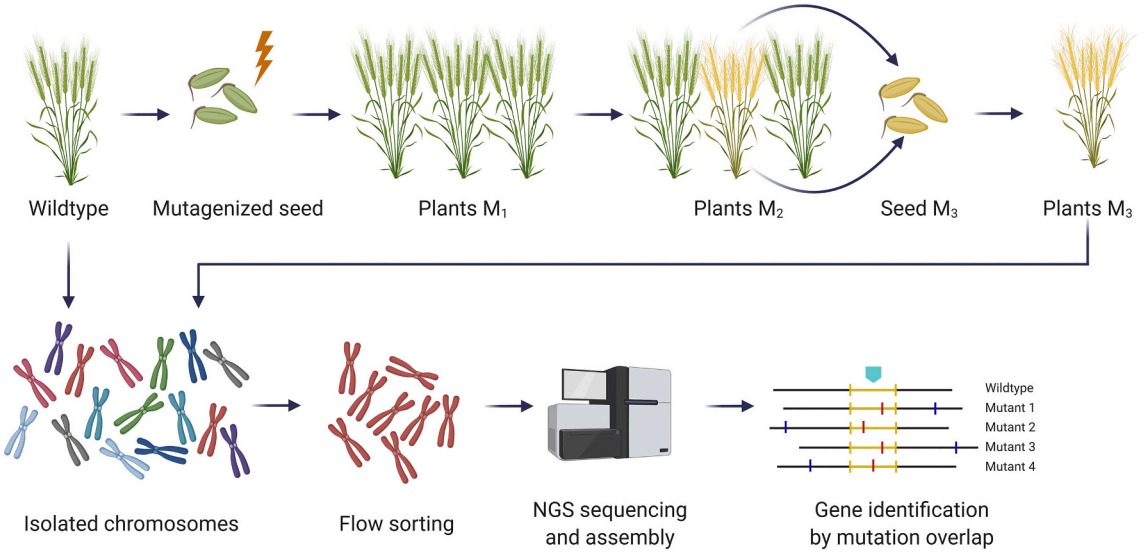


Fig. 10. The MutChromSeq approach towards gene cloning is suitable for any gene with a strong phenotype (Sánchez-Martín et al., 2016). The method compares sequences of a gene-bearing chromosome isolated from a wild type and several independently derived mutants, respectively, to identify causative mutations in a candidate gene. Its application does not require a prior assumption about the gene product and the only information needed is the chromosome location of the gene.

comparing sequences from chromosome 7D of a susceptible cv. Canthatch to those from resistant lines and mutants. The gene encodes the Med15b.D subunit of a mediator complex that functions as transcriptional coactivator, and its nonsense mutations result in the expression of stem rust resistance. To support attempts to improve the lysine content of barley for animal feed, Orman-Ligeza et al. (2020) used MutChromSeq to identify a gene underlying this trait. Comparison of sequences from chromosome 5H isolated from wild-type lines and from high-lysine mutants identified the *LYS3* gene, which encodes PBF transcription factor.

The results obtained with MutChromSeq show the superiority of the approach for gene cloning, provided the gene or a regulatory sequence has a strong phenotype and independent mutants either are available or can be developed. If these conditions are met, the whole project can be completed in less than 4 months (Sánchez-Martín et al., 2016) and at very reasonable cost. However, the TACCA approach, which relies on the production of a reference-quality *de novo* assembly of a chromosome, provides a straightforward alternative to map-based gene cloning. Its main advantage over a whole genome approach is in reducing complexity by targeting a particular chromosome or even chromosome arm. As it separates homoeologous sequences, it is particularly useful in polyploids such as hexaploid bread wheat.

9. Conclusions and perspectives

Since the first successful experiment on chromosome isolation by flow cytometric sorting, applications of chromosomal DNA samples have kept pace with progress in molecular biology and genomics. Thanks to this, chromosome genomics has been having a significant impact in many areas of research, including the sequencing of complex genomes of important crops and gene cloning. This has been achieved now that researchers have overcome two major bottlenecks of the technology: the ability to discriminate and sort individual chromosomes and the ability to deliver sufficient amounts of intact DNA. The former was overcome by sorting chromosomes from cytogenetic stocks with altered chromosome sizes (Doležel and Lucretti, 1995), by fluorescent-labeling some DNA repeats (Giorgi et al., 2013), and by sorting single chromosome copies (Cápal et al., 2015). The latter has been resolved by the development of protocols for representative amplification of chromosomal DNA (Šimková et al., 2008b) and preparing DNA sequencing libraries from small amounts of DNA.

A future challenge is to develop a procedure for fluorescent-labeling particular DNA sequences on chromosomes in suspensions prior to flow cytometry. Here CRISPR/Cas9-based methods (Ishii et al., 2019) hold great promise. Another challenge is to reduce the amount of DNA needed to produce reference-quality *de novo* assemblies. Here the use of long-read sequencing technologies such as those of Pacific Biosciences and Oxford Nanopore Technologies will be vital. For example, the release by Pacific Biosciences of a low-input protocol, which requires as little as 150 ng input DNA to produce a highly contiguous assembly for a genome up to 300 Mb (Kingan et al., 2019), offers a great opportunity. The possibility of coupling chromosome flow sorting with Oxford Nanopore sequencing was demonstrated recently by Yano et al. (2020). Finally, although single chromosome sequencing can be done after manual collection of single chromosome fractions, this approach is cumbersome and low-throughput. Coupling chromosome sorting with droplet-based microfluidic technology similar to that used by Nishikawa et al. (2015) would provide an opportunity to prepare multiplexed DNA sequencing libraries from thousands of single chromosomes for next-generation sequencing.

Currently the major applications of chromosome genomics involve gene cloning and related applications, such as the development of DNA markers from genome regions of interest. Chromosome-based approaches continue to be useful for sequencing particular chromosomes from species for which whole genome assemblies are not yet available, such as crop wild relatives. Because flow cytometric sorting purifies mitotic metaphase chromosomes, it has great potential for studies of their molecular organization and the three-dimensional arrangement of chromatin fibers in particular (Beseda et al., 2020). Various research projects to date have analyzed chromosomal DNA, while the identification of RNA in mitotic metaphase chromosomes remains unexplored. Finally, given that flow-sorted nuclei are suitable for proteomics (Petrovská et al., 2014), it may be expected that proteomic analysis of flow-sorted chromosomes will contribute to the characterization of proteins contributing to the structure and function of the condensed chromosomes in the metaphase.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biotechadv.2020.107659.

Author contributions

Jana Zwyrtková: Data curation; Writing - original draft; Visualization; Writing - review & editing. Hana Šimková: Writing - review & editing. Jaroslav Doležel: Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Writing - review & editing.

Declaration of Competing Interest

There are no conflicts of interest to declare.

Acknowledgments

This work was supported by the European Regional Development Fund project “Plants as a Tool for Sustainable Global Development” [project number CZ.02.1.01/0.0/0.0/16_019/0000827]. Figs. 1–3, 6–10 were created with BioRender.com.

References

Abrouk, M., Balcárková, B., Šimková, H., Komínkova, E., Martis, M.M., Jakobson, I., Timofejeva, L., Rey, E., Vrána, J., Kilian, A., Järve, K., Doležel, J., Valárik, M., 2017. The in silico identification and characterization of a bread wheat/*Triticum militinae* introgression line. Plant Biotechnol. J. 15, 249–256. https://doi.org/10.1111/pbi.12610.

Aglawe, S.B., Barbadikar, K.M., Mangrauthia, S.K., Madhav, M.S., 2018. New breeding technique “genome editing” for crop improvement: applications, potentials and challenges. 3 Biotech. 8, 336. https://doi.org/10.1007/s13205-018-1355-3.

Akhunov, E.D., Sehgal, S., Liang, H., Wang, S., Akhunova, A.R., Kaur, G., Li, W., Forrest, K.L., See, D., Šimková, H., Ma, Y., Hayden, M.J., Luo, M., Faris, J.D., Doležel, J., Gill, B.S., 2013. Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. Plant Physiol. 161, 252–265. https://doi.org/10.1104/pp.112.205161.

Akpinar, B.A., Lucas, S.J., Vrána, J., Doležel, J., Budak, H., 2015a. Sequencing chromosome 5D of *Aegilops tauschii* and comparison with its allopolyploid descendant bread wheat (*Triticum aestivum*). Plant Biotechnol. J. 13, 740–752. https://doi.org/10.1111/pbi.12302.

Akpinar, B.A., Magni, F., Yuce, M., Lucas, S.J., Šimková, H., Šafář, J., Vautrin, S., Bergès, H., Cattonaro, F., Doležel, J., Budak, H., 2015b. The physical map of wheat chromosome 5DS revealed gene duplications and small rearrangements. BMC Genomics 16, 453. https://doi.org/10.1186/s12864-015-1641-y.

Akpinar, B.A., Yuce, M., Lucas, S., Vrána, J., Burešová, V., Doležel, J., Budak, H., 2015c. Molecular organization and comparative analysis of chromosome 5B of the wild wheat ancestor *Triticum dicoccoides*. Sci. Rep. 5, 1–13. https://doi.org/10.1038/srep10763.

Albert, T.J., Molla, M.N., Muzny, D.M., Nazareth, L., Wheeler, D., Song, X., Richmond, T. A., Middle, C.M., Rodesch, M.J., Packard, C.J., Weinstock, G.M., Gibbs, R.A., 2007. Direct selection of human genomic loci by microarray hybridization. Nat. Methods 4, 903–905. https://doi.org/10.1038/nmeth1111.

Arumuganathan, K., Martin, G.B., Telenius, H., Tanksley, S.D., Earle, E.D., 1994. Chromosome 2-specific DNA clones from flow-sorted chromosomes of tomato. MGG Mol. Gen. Genet. 242, 551–558. https://doi.org/10.1007/BF00285278.

Bansal, M., Adamski, N.M., Toor, P.I., Kaur, S., Molnár, I., Holušová, K., Vrána, J., Doležel, J., Valárik, M., Uauy, C., Chhuneja, P., 2020. Aegilops umbellulata introgression carrying leaf rust and stripe rust resistance genes Lr76 and Yr70 located to 9.47-Mb region on 5DS telomeric end through a combination of chromosome sorting and sequencing. Theor. Appl. Genet. 133, 903–915. https://doi.org/10.1007/s00122-019-03514-x.

Barabaschi, D., Magni, F., Volante, A., Gadaleta, A., Šimková, H., Scalabrin, S., Prazzoli, M.L., Bagnaresi, P., Lacrima, K., Michelotti, V., Desiderio, F., Orrù, L., Mazzamurro, V., Fricano, A., Mastrangelo, A., Tononi, P., Vitulo, N., Jurman, I.,

Frenkel, Z., Cattonaro, F., Morgante, M., Blanco, A., Doležel, J., Delledonne, M., Stanca, A.M., Cattivelli, L., Valè, G., 2015. Physical mapping of bread wheat chromosome 5A: an integrated approach. Plant Genome. 8 https://doi.org/10.3835/plantgenome2015.03.0011.

Bartoš, J., Paux, E., Kofler, R., Havránková, M., Kopecký, D., Suchánková, P., Šafář, J., Šimková, H., Town, C.D., Lelley, T., Feuillet, C., Doležel, J., 2008. A first survey of the rye (*Secale cereale*) genome composition through BAC end sequencing of the short arm of chromosome 1R. BMC Plant Biol. 8 https://doi.org/10.1186/1471-2229-8-95.

Baum, M., Appels, R., 1991. Review: the cytogenetic and molecular architecture of chromosome 1R-one of the most widely utilized sources of alien chromatin in wheat varieties. Chromosoma 101, 1–10. https://doi.org/10.1007/BF00360680.

Belova, T., Zhan, B., Wright, J., Caccamo, M., Asp, T., Šimková, H., Kent, M., Bendixen, C., Panitz, F., Lien, S., Doležel, J., Olsen, O.A., Sandve, S.R., 2013. Integration of mate pair sequences to improve shotgun assemblies of flow-sorted chromosome arms of hexaploid wheat. BMC Genomics 14. https://doi.org/10.1186/1471-2164-14-222.

Belova, T., Grønvold, L., Kumar, A., Kianian, S., He, X., Lillemo, M., Springer, N.M., Lien, S., Olsen, O.A., Sandve, S.R., 2014. Utilization of deletion bins to anchor and order sequences along the wheat 7B chromosome. Theor. Appl. Genet. 127, 2029–2040. https://doi.org/10.1007/s00122-014-2358-z.

Berkman, P.J., Skarshewski, A., Lorenc, M.T., Lai, K., Duran, C., Ling, E.Y.S., Stiller, J., Smits, L., Imelfort, M., Manoli, S., Mckenzie, M., Kubaláková, M., Šimková, H., Batley, J., Fleury, D., Doležel, J., Edwards, D., 2011. Sequencing and assembly of low copy and genic regions of isolated *Triticum aestivum* chromosome arm 7DS. Plant Biotechnol. J. 9, 768–775. https://doi.org/10.1111/j.1467-7652.2010.00587.x.

Berkman, P.J., Skarshewski, A., Manoli, S., Lorenc, M.T., Stiller, J., Smits, L., Lai, K., Campbell, E., Kubaláková, M., Šimková, H., Batley, J., Doležel, J., Hernandez, P., Edwards, D., 2012. Sequencing wheat chromosome arm 7BS delimits the 7BS/4AL translocation and reveals homoeologous gene conservation. Theor. Appl. Genet. 124, 423–432. https://doi.org/10.1007/s00122-011-1717-2.

Beseda, T., Cápal, P., Kubalová, I., Schubert, V., Doležel, J., Šimková, H., 2020. Mitotic chromosome organization: general rules meet species-specific variability. Comput. Struct. Biotechnol. J. 18, 1311–1319. https://doi.org/10.1016/j.csbj.2020.01.006.

Bettgenhaeuser, J., Krattinger, S.G., 2019. Rapid gene cloning in cereals. Theor. Appl. Genet. 132, 699–711. https://doi.org/10.1007/s00122-018-3210-7.

Breen, J., Wicker, T., Shatalina, M., Frenkel, Z., Bertin, I., Philippe, R., Spielmeyer, W., Šimková, H., Šafář, J., Cattonaro, F., Scalabrin, S., Magni, F., Vautrin, S., Bergès, H., Paux, E., Fahima, T., Doležel, J., Korol, A., Feuillet, C., Keller, B., 2013. A physical map of the short arm of wheat chromosome 1A. PLoS One 8. https://doi.org/10.1371/journal.pone.0080272.

Cápal, P., Blavet, N., Vrána, J., Kubaláková, M., Doležel, J., 2015. Multiple displacement amplification of the DNA from single flow-sorted plant chromosome. Plant J. 84, 838–844. https://doi.org/10.1111/tpj.13035.

Cápal, P., Endo, T.R., Vrána, J., Kubaláková, M., Karafiátová, M., Komínková, E., Mora-Ramírez, I., Weschke, W., Doležel, J., 2016. The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat. Plant Methods 12. https://doi.org/10.1186/s13007-016-0124-8.

Chen, X., Yang, H., Wong, W.H., 2017. Phased genome sequencing through chromosome sorting. In: Methods in Molecular Biology. Humana Press Inc., pp. 171–188. https://doi.org/10.1007/978-1-4939-6750-6_10

Chen, K., Wang, Y., Zhang, R., Zhang, H., Gao, C., 2019. CRISPR/Cas genome editing and precision plant breeding in agriculture. Annu. Rev. Plant Biol. 70, 667–697. https://doi.org/10.1146/annurev-arplant-050718-100049.

Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., Pingault, L., Sourdille, P., Couloux, A., Paux, E., Leroy, P., Mangenot, S., Guilhot, N., Le Gouis, J., Balfourier, F., Alaux, M., Jamilloux, V., Poulain, J., Durand, C., Bellec, A., Gaspin, C., Safar, J., Doležel, J., Rogers, J., Vandepoele, K., Aury, J.M., Mayer, K., Berges, H., Quesneville, H., Wincker, P., Feuillet, C., 2014. Structural and functional partitioning of bread wheat chromosome 3B. Science (80-.) 345. https://doi.org/10.1126/science.1249721.

Collard, B.C.Y., Mackill, D.J., 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos. Trans. R. Soc. B Biol. Sci. 363, 557–572. https://doi.org/10.1098/rstb.2007.2170.

Conia, J., Bergounioux, C., Perennes, C., Muller, P., Brown, S., Gadal, P., 1987. Flow cytometric analysis and sorting of plant chromosomes from *Petunia hybrida* protoplasts. Cytometry 8, 500–508. https://doi.org/10.1002/cyto.990080511.

Conia, J., Muller, P., Brown, S., Bergounioux, C., Gadal, P., 1989. Monoparametric models of flow cytometric karyotypes with spreadsheet software. Theor. Appl. Genet. 77, 295–303. https://doi.org/10.1007/BF00266200.

de Laat, A.M.M., Blaas, J., 1984. Flow-cytometric characterization and sorting of plant chromosomes. Theor. Appl. Genet. 67, 463–467. https://doi.org/10.1007/BF00263414.

Devos, K.M., Dubcovsky, J., Dvořák, J., Chinoy, C.N., Gale, M.D., 1995. Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor. Appl. Genet. 91, 282–288. https://doi.org/10.1007/BF00220890.

Dixon, S.C., Miller, N.G.A., Carter, N.P., Tucker, E.M., 2009. Bivariate flow cytometry of farm animal chromosomes: a potential tool for gene mapping. Anim. Genet. 23, 203–210. https://doi.org/10.1111/j.1365-2052.1992.tb00132.x.

Doležel, J., Lucretti, S., 1995. High-resolution flow karyotyping and chromosome sorting in *Vicia faba* lines with standard and reconstructed karyotypes. Theor. Appl. Genet. 90, 797–802. https://doi.org/10.1007/BF00222014.

Doležel, J., Číhalíková, J., Lucretti, S., 1992. A high-yield procedure for isolation of metaphase chromosomes from root tips of *Vicia faba* L. Planta 188, 93–98. https://doi.org/10.1007/BF00198944.

Doležel, J., Lucretti, S., Schubert, I., 1994. Plant chromosome analysis and sorting by flow cytometry. CRC. Crit. Rev. Plant Sci. 13, 275–309. https://doi.org/10.1080/07352689409701917.

Doležel, J., Macas, J., Lucretti, S., 1999. Flow analysis and sorting of plant chromosomes. Curr. Protoc. Cytom. 9, 5.3.1–5.3.33. https://doi.org/10.1002/0471142956.cy0503s09.

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. https://doi.org/10.1007/s10142-012-0293-0.

Dracatos, P.M., Bartoš, J., Elmansour, H., Singh, D., Karafiátová, M., Zhang, P., Steuernagel, B., Svačina, R., Cobbin, J.C.A., Clark, B., Hoxha, S., Khatkar, M.S., Doležel, J., Wulff, B.B., Park, R.F., 2019. The coiled-coil NLR RPH1, confers leaf rust resistance in barley cultivar Sudan. Plant Physiol. 179, 1362–1372. https://doi.org/10.1104/pp.18.01052.

Endo, T.R., Kubaláková, M., Vrána, J., Doležel, J., 2014. Hyperexpansion of wheat chromosomes sorted by flow cytometry. Genes Genet. Syst. 89, 181–185. https://doi.org/10.1266/ggs.89.181.

Feng, K., Cui, L., Wang, L., Shan, D., Tong, W., Deng, P., Yan, Z., Wang, M., Zhan, H., Wu, X., He, W., Zhou, X., Ji, J., Zhang, G., Mao, L., Karafiátová, M., Šimková, H., Doležel, J., Du, X., Zhao, S., Luo, M.C., Han, D., Zhang, C., Kang, Z., Appels, R., Edwards, D., Nie, X., Weining, S., 2020. The improved assembly of 7DL chromosome provides insight into the structure and evolution of bread wheat. Plant Biotechnol. J. 18, 732–742. https://doi.org/10.1111/pbi.13240.

Ford, B.A., Foo, E., Sharwood, R., Karafiátová, M., Vrána, J., MacMillan, C., Nichols, D.S., Steuernagel, B., Uauy, C., Doležel, J., Chandler, P.M., Spielmeyer, W., 2018. Rht18 semidwarfism in wheat is due to increased GA 2-oxidaseA9 expression and reduced GA content. Plant Physiol. 177, 168–180. https://doi.org/10.1104/pp.18.00023.

Garbus, I., Romero, J.R., Valarik, M., Vanžurová, H., Karafiátová, M., Cáccamo, M., Doležel, J., Tranquilli, G., Helguera, M., Echenique, V., 2015. Characterization of repetitive DNA landscape in wheat homeologous group 4 chromosomes. BMC Genomics 16. https://doi.org/10.1186/s12864-015-1579-0.

Giorgi, D., Farina, A., Grosso, V., Gennaro, A., Ceoloni, C., Lucretti, S., 2013. FISHIS: fluorescence in situ hybridization in suspension and chromosome flow sorting made easy. PLoS One 8, e57994. https://doi.org/10.1371/journal.pone.0057994.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E.M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Russ, C., Gabriel, S., Jaffe, D.B., Lander, E.S., Nusbaum, C., 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat. Biotechnol. 27, 182–189. https://doi.org/10.1038/nbt.1523.

Gray, J.W., Langlois, R.G., Carrano, A.V., Burkhart-Schulte, K., Van Dilla, M.A., 1979. High resolution chromosome analysis: one and two parameter flow cytometry. Chromosoma 73, 9–27. https://doi.org/10.1007/BF00294840.

Harrington, S.A., Overend, L.E., Cobo, N., Borrill, P., Uauy, C., 2019. Conserved residues in the wheat (*Triticum aestivum*) NAM-A1 NAC domain are required for protein binding and when mutated lead to delayed peduncle and flag leaf senescence. BMC Plant Biol. 19 https://doi.org/10.1186/s12870-019-2022-5.

Helguera, M., Rivarola, M., Clavijo, B., Martis, M.M., Vanzetti, L.S., González, S., Garbus, I., Leroy, P., Šimková, H., Valárik, M., Caccamo, M., Doležel, J., Mayer, K.F.X., Feuillet, C., Tranquilli, G., Paniego, N., Echenique, V., 2015. New insights into the wheat chromosome 4D structure and virtual gene order, revealed by survey pyrosequencing. Plant Sci. 233, 200–212. https://doi.org/10.1016/j.plantsci.2014.12.004.

Hernandez, P., Martis, M., Dorado, G., Pfeifer, M., Gálvez, S., Schaaf, S., Jouve, N., Šimková, H., Valárik, M., Doležel, J., Mayer, K.F.X., 2012. Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. Plant J. 69, 377–386. https://doi.org/10.1111/j.1365-313X.2011.04808.x.

Hiebert, C.W., Moscou, M.J., Hewitt, T., Steuernagel, B., Hernández-Pinzón, I., Green, P., Pujol, V., Zhang, P., Rouse, M.N., Jin, Y., McIntosh, R.A., Upadhyaya, N., Zhang, J., Bhavani, S., Vrána, J., Karafiátová, M., Huang, L., Fetch, T., Doležel, J., Wulff, B.B. H., Lagudah, E., Spielmeyer, W., 2020. Stem rust resistance in wheat is suppressed by a subunit of the mediator complex. Nat. Commun. 11, 1123 https://doi.org/10.1038/s41467-020-14937-2.

Hobza, R., Vyskot, B., 2007. Laser microdissection-based analysis of plant sex chromosomes. Methods Cell Biol. 82, 433–453. https://doi.org/10.1016/S0091-679X(06)82015-7.

Holušová, K., Vrána, J., Šafář, J., Šimková, H., Balcárková, B., Frenkel, Z., Darrier, B., Paux, E., Cattonaro, F., Berges, H., Letellier, T., Alaux, M., Doležel, J., Bartoš, J., 2017. Physical map of the short arm of bread wheat chromosome 3D. Plant Genome. 10 https://doi.org/10.3835/plantgenome2017.03.0021.

https://olomouc.ueb.cas.cz/en/resources/dna-libraries, n.d. Chromosome-specific and genomic BAC resources of cereals [WWW Document]. URL https://olomouc.ueb.cas.cz/en/resources/dna-libraries (accessed 7.3.20).

International Rice Genome Sequencing Project T, 2005. The map-based sequence of the rice genome. Nature 436, 793–800. https://doi.org/10.1038/nature03895.

International Wheat Genome Sequencing Consortium (IWGSC), 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345, 1251788. https://doi.org/10.1126/science.1251788.

International Wheat Genome Sequencing Consortium, Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., Pozniak, C.J., Choulet, F., Distelfeld, A., Poland, J., Ronen, G., Sharpe, A.G., Barad, O., Baruch, K., Keeble-Gagnère, G., Mascher, M., Ben-Zvi, G., Josselin, A.-A., Himmelbach, A., Balfourier, F., Gutierrez-Gonzalez, J., Hayden, M., Koh, C., Muehlbauer, G., Pasam, R.K., Paux, E., Rigault, P., Tibbits, J., Tiwari, V., Spannagl, M., Lang, D., Gundlach, H., Haberer, G., Mayer, K.F.X., Ormanbekova, D., Prade, V., Šimková, H., Wicker, T., Swarbreck, D., Rimbert, H., Felder, M., Guilhot, N., Kaithakottil, G., Keilwagen, J., Leroy, P., Lux, T.,

Twardziok, S., Venturini, L., Juhász, A., Abrouk, M., Fischer, I., Uauy, C., Borrill, P., Ramirez-Gonzalez, R.H., Arnaud, D., Chalabi, S., Chalhoub, B., Cory, A., Datla, R., Davey, M.W., Jacobs, J., Robinson, S.J., Steuernagel, B., van Ex, F., Wulff, B.B.H., Benhamed, M., Bendahmane, A., Concia, L., Latrasse, D., Bartoš, J., Bellec, A., Berges, H., Doležel, J., Frenkel, Z., Gill, B., Korol, A., Letellier, T., Olsen, O.-A., Singh, K., Valárik, M., van der Vossen, E., Vautrin, S., Weining, S., Fahima, T., Glikson, V., Raats, D., Číhalíková, J., Toegelová, H., Vrána, J., Sourdille, P., Darrier, B., Barabaschi, D., Cattivelli, L., Hernandez, P., Galvez, S., Budak, H., Jones, J.D.G., Witek, K., Yu, G., Small, I., Melonek, J., Zhou, R., Belova, T., Kanyuka, K., King, R., Nilsen, K., Walkowiak, S., Cuthbert, R., Knox, R., Wiebe, K., Xiang, D., Rohde, A., Golds, T., Čížková, J., Akpinar, B.A., Biyiklioglu, S., Gao, L., N'Daiye, A., Kubaláková, M., Šafář, J., Alfama, F., Adam-Blondon, A.-F., Flores, R., Guerche, C., Loaec, M., Quesneville, H., Condie, J., Ens, J., Maclachlan, R., Tan, Y., Alberti, A., Aury, J.-M., Barbe, V., Couloux, A., Cruaud, C., Labadie, K., Mangenot, S., Wincker, P., Kaur, G., Luo, M., Sehgal, S., Chhuneja, P., Gupta, O.P., Jindal, S., Kaur, P., Malik, P., Sharma, P., Yadav, B., Singh, N.K., Khurana, J.P., Chaudhary, C., Khurana, P., Kumar, V., Mahato, A., Mathur, S., Sevanthi, A., Sharma, N., Tomar, R.S., Holušová, K., Plíhal, O., Clark, M.D., Heavens, D., Kettleborough, G., Wright, J., Balcárková, B., Hu, Y., Salina, E., Ravin, N., Skryabin, K., Beletsky, A., Kadnikov, V., Mardanov, A., Nesterov, M., Rakitin, A., Sergeeva, E., Handa, H., Kanamori, H., Katagiri, S., Kobayashi, F., Nasuda, S., Tanaka, T., Wu, J., Cattonaro, F., Jiumeng, M., Kugler, K., Pfeifer, M., Sandve, S., Xun, X., Zhan, B., Batley, J., Bayer, P.E., Edwards, D., Hayashi, S., Tulpová, Z., Visendi, P., Cui, L., Du, X., Feng, K., Nie, X., Tong, W., Wang, L., 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361. https://doi.org/10.1126/science.aar7191 eaar7191.

Ishii, T., Schubert, V., Khosravi, S., Dreissig, S., Metje-Sprink, J., Sprink, T., Fuchs, J., Meister, A., Houben, A., 2019. RNA-guided endonuclease – in situ labelling (RGEN-ISL): a fast CRISPR/Cas9-based method to label genomic sequences in various species. New Phytol. 222, 1652–1661. https://doi.org/10.1111/nph.15720.

Jaccoud, D., Peng, K., Feinstein, D., Kilian, A., 2001. Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Res. 29, e25. https://doi.org/10.1093/nar/29.4.e25.

Jain, M., Misra, G., Patel, R.K., Priya, P., Jhanwar, S., Khan, A.W., Shah, N., Singh, V.K., Garg, R., Jeena, G., Yadav, M., Kant, C., Sharma, P., Yadav, G., Bhatia, S., Tyagi, A. K., Chattopadhyay, D., 2013. A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). Plant J. 74, 715–729. https://doi.org/10.1111/tpj.12173.

Janda, J., Bartoš, J., Šafář, J., Kubaláková, M., Valárik, M., Číhalíková, J., Šimková, H., Caboche, M., Sourdille, P., Bernard, M., Chalhoub, B., Doležel, J., 2004. Construction of a subgenomic BAC library specific for chromosomes 1D, 4D and 6D of hexaploid wheat. Theor. Appl. Genet. 109, 1337–1345. https://doi.org/10.1007/s00122-004-1768-8.

Janda, J., Šafář, J., Kubaláková, M., Bartoš, J., Kovářová, P., Suchánková, P., Pateyron, S., Číhalíková, J., Sourdille, P., Šimková, H., Faivre-Rampant, P., Hřibová, E., Bernard, M., Lukaszewski, A., Doležel, J., Chalhoub, B., 2006. Advanced resources for plant genomics: a BAC library specific for the short arm of wheat chromosome 1B. Plant J. 47, 977–986. https://doi.org/10.1111/j.1365-313X.2006.02840.x.

Kantar, M., Akpinar, B.A., Valárik, M., Lucas, S.J., Doležel, J., Hernández, P., Budak, H., 2012. Subgenomic analysis of microRNAs in polyploid wheat. Funct. Integr. Genomics 12, 465–479. https://doi.org/10.1007/s10142-012-0285-0.

Keeble-Gagnère, G., Rigault, P., Tibbits, J., Pasam, R., Hayden, M., Forrest, K., Frenkel, Z., Korol, A., Huang, B.E., Cavanagh, C., Taylor, J., Abrouk, M., Sharpe, A., Konkin, D., Sourdille, P., Darrier, B., Choulet, F., Bernard, A., Rochfort, S., Dimech, A., Watson-Haigh, N., Baumann, U., Eckermann, P., Fleury, D., Juhasz, A., Boisvert, S., Nolin, M.A., Doležel, J., Šimková, H., Toegelová, H., Šafář, J., Luo, M.C., Câmara, F., Pfeifer, M., Isdale, D., Nyström-Persson, J., Koo, D.H., Tinning, M., Cui, D., Ru, Z., Appels, R., 2018. Optical and physical mapping with local finishing enables megabase-scale resolution of agronomically important regions in the wheat genome. Genome Biol. 19, 112. https://doi.org/10.1186/s13059-018-1475-4.

Kingan, S.B., Heaton, H., Cudini, J., Lambert, C.C., Baybayan, P., Galvin, B.D., Durbin, R., Korlach, J., Lawniczak, M.K.N., 2019. A high-quality de novo genome assembly from a single mosquito using pacbio sequencing. Genes (Basel) 10, 62. https://doi.org/10.3390/genes10010062.

Kobayashi, F., Wu, J., Kanamori, H., Tanaka, T., Katagiri, S., Karasawa, W., Kaneko, S., Watanabe, S., Sakaguchi, T., Hanawa, Y., Fujisawa, H., Kurita, K., Abe, C., Iehisa, J. C.M., Ohno, R., Šafář, J., Šimková, H., Mukai, Y., Hamada, M., Saito, M., Ishikawa, G., Katayose, Y., Endo, T.R., Takumi, S., Nakamura, T., Sato, K., Ogihara, Y., Hayakawa, K., Doležel, J., Nasuda, S., Matsumoto, T., Handa, H., 2015. A high-resolution physical map integrating an anchored chromosome with the BAC physical maps of wheat chromosome 6B. BMC Genomics 16, 495. https://doi.org/10.1186/s12864-015-1803-y.

Koblížková, A., Doležel, J., Macas, J., 1998. Subtraction with 3′ modified oligonucleotides eliminates amplification artifacts in DNA libraries enriched for microsatellites. Biotechniques 25, 32–38. https://doi.org/10.2144/98251bm06.

Kofler, R., Bartoš, J., Gong, L., Stift, G., Suchánková, P., Šimková, H., Berenyi, M., Burg, K., Doležel, J., Lelley, T., 2008. Development of microsatellite markers specific for the short arm of rye (*Secale cereale* L.) chromosome 1. Theor. Appl. Genet. 117, 915–926. https://doi.org/10.1007/s00122-008-0831-2.

Kreplak, J., Madoui, M.A., Cápal, P., Novák, P., Labadie, K., Aubert, G., Bayer, P.E., Gali, K.K., Syme, R.A., Main, D., Klein, A., Bérard, A., Vrbová, I., Fournier, C., d'Agata, L., Belser, C., Berrabah, W., Toegelová, H., Milec, Z., Vrána, J., Lee, H.T., Kougbeadjo, A., Térézol, M., Huneau, C., Turo, C.J., Mohellibi, N., Neumann, P., Falque, M., Gallardo, K., McGee, R., Tar'an, B., Bendahmane, A., Aury, J.M., Batley, J., Le Paslier, M.C., Ellis, N., Warkentin, T.D., Coyne, C.J., Salse, J., Edwards, D., Lichtenzveig, J., Macas, J., Doležel, J., Wincker, P., Burstin, J., 2019. A reference genome for pea provides insight into legume genome evolution. Nat. Genet. 51, 1411–1422. https://doi.org/10.1038/s41588-019-0480-1.

Kubaláková, M., Vrána, J., Číhalíková, J., Šimková, H., Doležel, J., 2002. Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). Theor. Appl. Genet. 104, 1362–1372. https://doi.org/10.1007/s00122-002-0888-2.

Kubaláková, M., Valárik, M., Bartoš, J., Vrána, J., Číhalíková, J., Molnár-Láng, M., Doležel, J., 2003. Analysis and sorting of rye (*Secale cereale* L.) chromosomes using flow cytometry. Genome 46, 893–905. https://doi.org/10.1139/g03-054.

Kubaláková, M., Kovářová, P., Suchánková, P., Číhalíková, J., Bartoš, J., Lucretti, S., Watanabe, N., Kianian, S.F., Doležel, J., 2005. Chromosome sorting in tetraploid wheat and its potential for genome analysis. Genetics 170, 823–829. https://doi.org/10.1534/genetics.104.039180.

Lam, E.T., Hastie, A., Lin, C., Ehrlich, D., Das, S.K., Austin, M.D., Deshpande, P., Cao, H., Nagarajan, N., Xiao, M., Kwok, P.Y., 2012. Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. Nat. Biotechnol. 30, 771–776. https://doi.org/10.1038/nbt.2303.

Lee, J.H., Arumuganathan, K., Yen, Y., Kaeppler, S., Kaeppler, H., Baenziger, P.S., 1997. Root tip cell cycle synchronization and metaphase-chromosome isolation suitable for flow sorting in common wheat (*Triticum aestivum* L.). Genome 40, 633–638. https://doi.org/10.1139/g97-083.

Lee, J.-H., Arumuganathan, K., Chung, Y.-S., Kim, K.-Y., Chung, W.-B., Bae, K.-S., Kim, D.-H., Chung, D.-S., Kwon, O.-C., 2000. Flow Cytometric analysis and chromosome sorting of barley (*Hordeum vulgare* L.). Mol. Cell 10, 619–625. https://doi.org/10.1007/s10059-000-0619-y.

Ling, H.Q., Ma, B., Shi, X., Liu, H., Dong, L., Sun, H., Cao, Y., Gao, Q., Zheng, S., Li, Ye, Yu, Y., Du, H., Qi, M., Li, Yan, Lu, H., Yu, H., Cui, Y., Wang, N., Chen, C., Wu, H., Zhao, Y., Zhang, J., Li, Yiwen, Zhou, W., Zhang, B., Hu, W., Van Eijk, M.J.T., Tang, J., Witsenboer, H.M.A., Zhao, S., Li, Z., Zhang, A., Wang, D., Liang, C., 2018. Genome sequence of the progenitor of wheat a subgenome *Triticum urartu*. Nature 557, 424–428. https://doi.org/10.1038/s41586-018-0108-0.

Liu, M., Stiller, J., Holušová, K., Vrána, J., Liu, D., Doležel, J., Liu, C., 2016. Chromosome-specific sequencing reveals an extensive dispensable genome component in wheat. Sci. Rep. 6, 36398 https://doi.org/10.1038/srep36398.

Lu, F.-H., McKenzie, N., Kettleborough, G., Heavens, D., Clark, M.D., Bevan, M.W., 2018. Independent assessment and improvement of wheat genome sequence assemblies using Fosill jumping libraries. Gigascience 7. https://doi.org/10.1093/gigascience/giy053.

Lucas, S.J., Akpinar, B.A., Kantar, M., Weinstein, Z., Aydinoğlu, F., Šafář, J., Šimková, H., Frenkel, Z., Korol, A., Magni, F., Cattonaro, F., Vautrin, S., Bellec, A., Bergès, H., Doležel, J., Budak, H., 2013. Physical mapping integrated with syntenic analysis to characterize the gene space of the long arm of wheat chromosome 1A. PLoS One 8, e59542. https://doi.org/10.1371/journal.pone.0059542.

Lucas, S.J., Akpinar, A.A., Šimková, H., Kubaláková, M., Doležel, J., Budak, H., 2014. Next-generation sequencing of flow-sorted wheat chromosome 5D reveals lineage-specific translocations and widespread gene duplications. BMC Genomics 15, 1080. https://doi.org/10.1186/1471-2164-15-1080.

Lucretti, S., Doležel, J., 1997. Bivariate flow karyotyping in broad bean (*Vicia faba*). Cytometry 28, 236–242.

Lucretti, S., Doležel, J., Schubert, I., Fuchs, J., 1993. Flow karyotyping and sorting of *Vicia faba* chromosomes. Theor. Appl. Genet. 85, 665–672. https://doi.org/10.1007/BF00225003.

Luo, M.C., Thomas, C., You, F.M., Hsiao, J., Ouyang, S., Buell, C.R., Malandro, M., McGuire, P.E., Anderson, O.D., Dvorak, J., 2003. High-throughput fingerprinting of bacterial artificial chromosomes using the SNaPshot labeling kit and sizing of restriction fragments by capillary electrophoresis. Genomics 82, 378–389. https://doi.org/10.1016/S0888-7543(03)00128-9.

Luo, M.C., Gu, Y.Q., Puiu, D., Wang, H., Twardziok, S.O., Deal, K.R., Huo, N., Zhu, T., Wang, L., Wang, Y., McGuire, P.E., Liu, S., Long, H., Ramasamy, R.K., Rodriguez, J. C., Van Sonny, L., Yuan, L., Wang, Z., Xia, Z., Xiao, L., Anderson, O.D., Ouyang, S., Liang, Y., Zimin, A.V., Pertea, G., Qi, P., Bennetzen, J.L., Dai, X., Dawson, M.W., Müller, H.G., Kugler, K., Rivarola-Duarte, L., Spannagl, M., Mayer, K.F.X., Lu, F.H., Bevan, M.W., Leroy, P., Li, P., You, F.M., Sun, Q., Liu, Z., Lyons, E., Wicker, T., Salzberg, S.L., Devos, K.M., Dvořák, J., 2017. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. Nature 551, 498–502. https://doi.org/10.1038/nature24486.

Lusser, M., Parisi, C., Plan, D., Rodríguez-Cerezo, E., 2011. New Plant Breeding Techniques State-of-the-Art and Prospects for Commercial Development, EUR 24760 EN-2011. Luxemb. Publ. Off. Eur. Union.

Lysák, M.A., Číhalíková, J., Kubaláková, M., Šimková, H., Künzel, G., Doležel, J., 1999. Flow karyotyping and sorting of mitotic chromosomes of barley (*Hordeum vulgare* L.). Chromosom. Res. 7, 431–444. https://doi.org/10.1023/A:1009293628638.

Ma, L., Xiao, Y., Huang, H., Wang, Q., Rao, W., Feng, Y., Zhang, K., Song, Q., 2010. Direct determination of molecular haplotypes by chromosome microdissection. Nat. Methods 7, 299–301. https://doi.org/10.1038/nmeth.1443.

Macas, J., Doležel, J., Lucretti, S., Pich, U., Meister, A., Fuchs, J., Schubert, I., 1993. Localization of seed protein genes on flow-sorted field bean chromosomes. Chromosom. Res. 1, 107–115. https://doi.org/10.1007/BF00710033.

Macas, J., Gualberti, G., Nouzová, M., Samec, P., Lucretti, S., Doležel, J., 1996. Construction of chromosome-specific DNA libraries covering the whole genome of field bean (*Vicia faba* L.). Chromosom. Res. 4, 531–539. https://doi.org/10.1007/BF02261781.

Martis, M.M., Klemme, S., Banaei-Moghaddam, A.M., Blattner, F.R., Macas, J., Schmutzer, T., Scholz, U., Gundlach, H., Wicker, T., Šímková, H., Novak, P., Neumann, P., Kubaláková, M., Bauer, E., Haseneyer, G., Fuchs, J., Doležel, J., Stein, N., Mayer, K.F.X., Houben, A., 2012. Selfish supernumerary chromosome

reveals its origin as a mosaic of host genome and organellar sequences. Proc. Natl. Acad. Sci. U. S. A. 109, 13343–13346. https://doi.org/10.1073/pnas.1204237109.
Martis, M.M., Zhou, R., Haseneyer, G., Schmutzer, T., Vrána, J., Kubaláková, M., König, S., Kugler, K.G., Scholz, U., Hackauf, B., Korzun, V., Schön, C.C., Doležel, J., Bauer, E., Mayer, K.F.X., Stein, N., 2013. Reticulate evolution of the rye genome. Plant Cell 25, 3685–3698. https://doi.org/10.1105/tpc.113.114553.
Matsunaga, S., Kawano, S., Michimoto, T., Higashiyama, T., Nakao, S., Sakai, A., Kuroiwa, T., 1999. Semi-automatic laser beam microdissection of the Y chromosome and analysis of Y chromosome DNA in a dioecious plant, *Silene latifolia*. Plant Cell Physiol. 40, 60–68. https://doi.org/10.1093/oxfordjournals.pcp.a029475.
Mayer, K.F.X., Taudien, S., Martis, M., Šimková, H., Suchánková, P., Gundlach, H., Wicker, T., Petzold, A., Felder, M., Steuernagel, B., Scholz, U., Graner, A., Platzer, M., Doležel, J., Stein, N., 2009. Gene content and virtual gene order of barley chromosome 1H. Plant Physiol. 151, 496–505. https://doi.org/10.1104/pp.109.142612.
Mayer, K.F.X., Martis, M., Hedley, P.E., Šimková, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., Roessner, S., Gundlach, H., Kubaláková, M., Suchánková, P., Murat, F., Felder, M., Nussbaumer, T., Graner, A., Salse, J., Endo, T., Sakai, H., Tanaka, T., Itoh, T., Sato, K., Platzer, M., Matsumoto, T., Scholz, U., Doležel, J., Waugh, R., Stein, N., 2011. Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell 23, 1249–1263. https://doi.org/10.1105/tpc.110.082537.
Mazaheri, M., Kianian, P.M.A., Mergoum, M., Valentini, G.L., Seetan, R., Pirseyedi, S.M., Kumar, A., Gu, Y.Q., Stein, N., Kubaláková, M., Doležel, J., Denton, A.M., Kianian, S. F., 2014. Transposable element junctions in marker development and genomic characterization of barley. Plant Genome. 7 https://doi.org/10.3835/plantgenome2013.10.0036.
McIntosh, R.A., Hart, G.E., Gale, M.D., 1988. Catalogue of gene symbols for wheat. Cereal Res. Comm. 16, 121–135.
Molnár, I., Vrána, J., Farkas, A., Kubaláková, M., Cseh, A., Molnár-Láng, M., Doležel, J., 2015. Flow sorting of C-genome chromosomes from wild relatives of wheat *Aegilops markgrafii*, *Ae. triuncialis* and *Ae. cylindrica*, and their molecular organization. Ann. Bot. 116, 189–200. https://doi.org/10.1093/aob/mcv073.
Molnár, I., Vrána, J., Burešová, V., Cápal, P., Farkas, A., Darkó, É., Cseh, A., Kubaláková, M., Molnár-Láng, M., Doležel, J., 2016. Dissecting the U, M, S and C genomes of wild relatives of bread wheat (*Aegilops* spp.) into chromosomes and exploring their synteny with wheat. Plant J. 88, 452–467. https://doi.org/10.1111/tpj.13266.
Muñoz-Amatriaín, M., Moscou, M.J., Bhat, P.R., Svensson, J.T., Bartoš, J., Suchánková, P., Šimková, H., Endo, T.R., Fenton, R.D., Lonardi, S., Castillo, A.M., Chao, S., Cistué, L., Cuesta-Marcos, A., Forrest, K.L., Hayden, M.J., Hayes, P.M., Horsley, R.D., Makoto, K., Moody, D., Sato, K., Vallés, M.P., Wulff, B.B.H., Muehlbauer, G.J., Doležel, J., Close, T.J., 2011. An improved consensus linkage map of barley based on flow-sorted chromosomes and single nucleotide polymorphism markers. Plant Genome. 4, 238–249. https://doi.org/10.3835/plantgenome2011.08.0023.
Nesterov, M.A., Afonnikov, D.A., Sergeeva, E.M., Miroshnichenko, L.A., Bragina, M.K., Bragin, A.O., Vasiliev, G.V., Salina, E.A., 2016. Identification of microsatellite loci based on BAC sequencing data and their physical mapping into the soft wheat 5B chromosome. Russ. J. Genet. Appl. Res. 6, 825–837. https://doi.org/10.1134/S2079059716070078.
Neumann, P., Lysák, M., Doležel, J., Macas, J., 1998. Isolation of chromosomes from *Pisum sativum* L. hairy root cultures and their analysis by flow cytometry. Plant Sci. 137, 205–215. https://doi.org/10.1016/S0168-9452(98)00141-1.
Neumann, P., Požárková, D., Vrána, J., Doležel, J., Macas, J., 2002. Chromosome sorting and PCR-based physical mapping in pea (*Pisum sativum* L.), 10, 63–67. https://doi.org/10.1023/A:1014274328269.
Nie, X., Li, B., Wang, L., Liu, P., Biradar, S.S., Li, T., Doležel, J., Edwards, D., Luo, M., Weining, S., 2012. Development of chromosome-arm-specific microsatellite markers in *Triticum aestivum* (Poaceae) using NGS technology. Am. J. Bot. 99, e369–71. https://doi.org/10.3732/ajb.1200077.
Nishikawa, Y., Hosokawa, M., Maruyama, T., Yamagishi, K., Mori, T., Takeyama, H., 2015. Monodisperse picoliter droplets for low-bias and contamination-free reactions in single-cell whole genome amplification. PLoS One 10, e0138733. https://doi.org/10.1371/journal.pone.0138733.
Nsabiyera, V., Baranwal, D., Qureshi, N., Kay, P., Forrest, K., Valárik, M., Doležel, J., Hayden, M.J., Bariana, H.S., Bansal, U.K., 2020. Fine mapping of *Lr49* using 90K SNP chip array and flow-sorted chromosome sequencing in wheat. Front. Plant Sci. 10, 1787. https://doi.org/10.3389/fpls.2019.01787.
Orman-Ligeza, B., Borrill, P., Chia, T., Chirico, M., Doležel, J., Drea, S., Karafiátová, M., Schatlowski, N., Solomon, C.U., Steuernagel, B., Wulff, B.B.H., Uauy, C., Trafford, K., 2020. *LYS3* encodes a prolamin-box-binding transcription factor that controls embryo growth in barley and wheat. J. Cereal Sci. 93, 102965. https://doi.org/10.1016/j.jcs.2020.102965, 2020 May.
Palmer, L.E., Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V.S., Nascimento, L.U., Dike, S., De La Bastide, M., Martienssen, R.A., McCombie, W.R., 2003. Maize genome sequencing by methylation filtration. Science 302, 2115–2117. https://doi.org/10.1126/science.1091265.
Paux, E., Sourdille, P., Salse, J., Saintenac, C., Choulet, F., Leroy, P., Korol, A., Michalak, M., Kianian, S., Spielmeyer, W., Lagudah, E., Somers, D., Kilian, A., Alaux, M., Vautrin, S., Bergès, H., Eversole, K., Appels, R., Šafář, J., Šimková, H., Doležel, J., Bernard, M., Feuillet, C., 2008. A physical map of the 1-gigabase bread wheat chromosome 3B. Science 322, 101–104. https://doi.org/10.1126/science.1161847.
Pellicer, J., Leitch, I.J., 2020. The plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. New Phytol. 226, 301–305. https://doi.org/10.1111/nph.16261.
Peterson, D.G., Wessler, S.R., Paterson, A.H., 2002. Efficient capture of unique sequences from eukaryotic genomes. Trends Genet. 18, 547–550. https://doi.org/10.1016/S0168-9525(02)02764-6.
Petrovská, B., Jeřábková, H., Chamrád, I., Vrána, J., Lenobel, R., Uřinovská, J., Šebela, M., Doležel, J., 2014. Proteomic analysis of barley cell nuclei purified by flow sorting. Cytogenet. Genome Res. 143, 78–86. https://doi.org/10.1159/000365311.
Philippe, R., Paux, E., Bertin, I., Sourdille, P., Choulet, F., Laugier, C., Šimková, H., Šafář, J., Bellec, A., Vautrin, S., Frenkel, Z., Cattonaro, F., Magni, F., Scalabrin, S., Martis, M.M., Mayer, K.F.X., Korol, A., Bergès, H., Doležel, J., Feuillet, C., 2013. A high density physical map of chromosome 1BL supports evolutionary studies, map-based cloning and sequencing in wheat. Genome Biol. 14 https://doi.org/10.1186/gb-2013-14-6-r64.
Poursarebani, N., Nussbaumer, T., Šimková, H., Šafář, J., Witsenboer, H., Van Oeveren, J., Doležel, J., Mayer, K.F.X., Stein, N., Schnurbusch, T., 2014. Whole-genome profiling and shotgun sequencing delivers an anchored, gene-decorated, physical map assembly of bread wheat chromosome 6A. Plant J. 79, 334–347. https://doi.org/10.1111/tpj.12550.
Powell, W., Machray, G.C., Proven, J., 1996. Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 1, 215–222. https://doi.org/10.1016/1360-1385(96)86898-1.
Požárková, D., Koblížková, A., Román, B., Torres, A.M., Lucretti, S., Lysák, M., Doležel, J., Macas, J., 2002. Development and characterization of microsatellite markers from chromosome 1-specific DNA libraries of *Vicia faba*. Biol. Plant. 45, 337–345. https://doi.org/10.1023/A:1016253214182.
Putnam, N.H., O'Connell, B.L., Stites, J.C., Rice, B.J., Blanchette, M., Calef, R., Troll, C.J., Fields, A., Hartley, P.D., Sugnet, C.W., Haussler, D., Rokhsar, D.S., Green, R.E., 2016. Chromosome-scale shotgun assembly using an *in vitro* method for long-range linkage. Genome Res. 26, 342–350. https://doi.org/10.1101/gr.193474.115.
Raats, D., Frenkel, Z., Krugman, T., Dodek, I., Sela, H., Šimková, H., Magni, F., Cattonaro, F., Vautrin, S., Bergès, H., Wicker, T., Keller, B., Leroy, P., Philippe, R., Paux, E., Doležel, J., Feuillet, C., Korol, A., Fahima, T., 2013. The physical map of wheat chromosome 1BS provides insights into its gene space organization and evolution. Genome Biol. 14, R138. https://doi.org/10.1186/gb-2013-14-12-r138.
Rey, E., Molnár, I., Doležel, J., 2015. Genomics of wild relatives and alien introgressions. In: Alien Introgression in Wheat. Springer International Publishing, pp. 347–381. https://doi.org/10.1007/978-3-319-23494-6_13.
Román, B., Satovic, Z., Požárková, D., Macas, J., Doležel, J., Cubero, J.I., Torres, A.M., 2004. Development of a composite map in *Vicia faba*, breeding applications and future prospects. Theor. Appl. Genet. 108, 1079–1088. https://doi.org/10.1007/s00122-003-1515-6.
Ruperao, P., Chan, C.K.K., Azam, S., Karafiátová, M., Hayashi, S., Čížková, J., Saxena, R. K., Šimková, H., Song, C., Vrána, J., Chitikineni, A., Visendi, P., Gaur, P.M., Millán, T., Singh, K.B., Taran, B., Wang, J., Batley, J., Doležel, J., Varshney, R.K., Edwards, D., 2014. A chromosomal genomics approach to assess and validate the desi and kabuli draft chickpea genome assemblies. Plant Biotechnol. J. 12, 778–786. https://doi.org/10.1111/pbi.12182.
Šafář, J., Bartoš, J., Janda, J., Bellec, A., Kubaláková, M., Valárik, M., Pateyron, S., Weiserová, J., Tušková, R., Číhalíková, J., Vrána, J., Šimková, H., Faivre-Rampant, P., Sourdille, P., Caboche, M., Bernard, M., Doležel, J., Chalhoub, B., 2004. Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J. 39, 960–968. https://doi.org/10.1111/j.1365-313X.2004.02179.x.
Šafář, J., Šimková, H., Kubaláková, M., Číhalíková, J., Suchánková, P., Bartoš, J., Doležel, J., 2010. Development of chromosome-specific BAC resources for genomics of bread wheat. Cytogenet. Genome Res. 129, 211–223. https://doi.org/10.1159/000313072.
Salina, E.A., Nesterov, M.A., Frenkel, Z., Kiseleva, A.A., Timonova, E.M., Magni, F., Vrána, J., Šafár, J., Šimková, H., Doležel, J., Korol, A., Sergeeva, E.M., 2018. Features of the organization of bread wheat chromosome 5BS based on physical mapping. BMC Genomics 19, 80. https://doi.org/10.1186/s12864-018-4470-y.
Sánchez-Martín, J., Steuernagel, B., Ghosh, S., Herren, G., Hurni, S., Adamski, N., Vrána, J., Kubaláková, M., Krattinger, S.G., Wicker, T., Doležel, J., Keller, B., Wulff, B.B.H., 2016. Rapid gene isolation in barley and wheat by mutant chromosome sequencing. Genome Biol. 17, 221. https://doi.org/10.1186/s13059-016-1082-1.
Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T.A., Minx, P., Reily, A.D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S.M., Belter, E., Du, F., Kim, K., Abbott, R.M., Cotton, M., Levy, A., Marchetto, P., Ochoa, K., Jackson, S.M., Gillam, B., Chen, W., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., Cordes, M., Kohlberg, S., Sgro, J., Delgado, B., Mead, K., Chinwalla, A., Leonard, S., Crouse, K., Collura, K., Kudrna, D., Currie, J., He, R., Angelova, A., Rajasekar, S., Mueller, T., Lomeli, R., Scara, G., Ko, A., Delaney, K., Wissotski, M., Lopez, G., Campos, D., Braidotti, M., Ashley, E., Golser, W., Kim, H., Lee, S., Lin, J., Dujmic, Z., Kim, W., Talag, J., Zuccolo, A., Fan, C., Sebastian, A., Kramer, M., Spiegel, L., Nascimento, L., Zutavern, T., Miller, B., Ambroise, C., Muller, S., Spooner, W., Narechania, A., Ren, L., Wei, S., Kumari, S., Faga, B., Levy, M.J., McMahan, L., Van Buren, P., Vaughn, M.W., Ying, K., Yeh, C.T., Emrich, S.J., Jia, Y., Kalyanaraman, A., Hsia, A.P., Barbazuk, W. B., Baucom, R.S., Brutnell, T.P., Carpita, N.C., Chaparro, C., Chia, J.M., Deragon, J. M., Estill, J.C., Fu, Y., Jeddeloh, J.A., Han, Y., Lee, H., Li, P., Lisch, D.R., Liu, S.,

Liu, Z., Nagel, D.H., McCann, M.C., Sanmiguel, P., Myers, A.M., Nettleton, D., Nguyen, J., Penning, B.W., Ponnala, L., Schneider, K.L., Schwartz, D.C., Sharma, A., Soderlund, C., Springer, N.M., Sun, Q., Wang, H., Waterman, M., Westerman, R., Wolfgruber, T.K., Yang, L., Yu, Y., Zhang, L., Zhou, S., Zhu, Q., Bennetzen, J.L., Dawe, R.K., Jiang, J., Jiang, N., Presting, G.G., Wessler, S.R., Aluru, S., Martienssen, R.A., Clifton, S.W., McCombie, W.R., Wing, R.A., Wilson, R.K., 2009. The B73 maize genome: complexity, diversity, and dynamics. Science 326, 1112–1115. https://doi.org/10.1126/science.1178534.

Schondelmaier, J., Martin, R., Jahoor, A., Houben, A., Graner, A., Koop, H.-U., Herrmann, R.G., Jung, C., 1993. Microdissection and microcloning of the barley (*Hordeum vulgare* L.) chromosome 1HS. Theor. Appl. Genet. 86, 629–636. https://doi.org/10.1007/BF00838719.

Schwartz, D., Li, X., Hernandez, L., Ramnarain, S., Huff, E., Wang, Y., 1993. Ordered restriction maps of *Saccharomyces cerevisiae* chromosomes constructed by optical mapping. Science 262, 110–114. https://doi.org/10.1126/science.8211116.

Sergeeva, E.M., Afonnikov, D.A., Koltunova, M.K., Gusev, V.D., Miroshnichenko, L.A., Vrána, J., Kubaláková, M., Poncet, C., Sourdille, P., Feuillet, C., Doležel, J., Salina, E. A., 2014. Common wheat chromosome 5B composition analysis using low-coverage 454 sequencing. Plant Genome. 7, 183 https://doi.org/10.3835/plantgenome2013.10.0031.

Sergeeva, E.M., Shcherban, A.B., Adonina, I.G., Nesterov, M.A., Beletsky, A.V., Rakitin, A.L., Mardanov, A.V., Ravin, N.V., Salina, E.A., 2017. Fine organization of genomic regions tagged to the 5S rDNA locus of the bread wheat 5B chromosome. BMC Plant Biol. 17 https://doi.org/10.1186/s12870-017-1120-5.

Seyran, E., Craig, W., 2018. New breeding techniques and their possible regulation. AgBioForum 21, 1–12.

Shatalina, M., Wicker, T., Buchmann, J.P., Oberhaensli, S., Šimková, H., Doležel, J., Keller, B., 2013. Genotype-specific SNP map based on whole chromosome 3B sequence information from wheat cultivars Arina and Forno. Plant Biotechnol. J. 11, 23–32. https://doi.org/10.1111/pbi.12003.

Shorinola, O., Balcárková, B., Hyles, J., Tibbits, J.F.G., Hayden, M.J., Holušová, K., Valárik, M., Distelfeld, A., Torada, A., Barrero, J.M., Uauy, C., 2017. Haplotype analysis of the pre-harvest sprouting resistance locus Phs-A1 reveals a causal role of TAMKK3-a in global germplasm. Front. Plant Sci. 8, 1555 https://doi.org/10.3389/fpls.2017.01555.

Silverman, G.A., Schneider, S.S., Massa, H.F., Flint, A., Lalande, M., Leonard, J.C., Overhauser, J., van den Engh, G., Trask, B.J., 1995. The 18q- syndrome: analysis of chromosomes by bivariate flow karyotyping and the PCR reveals a successive set of deletion breakpoints within 18q21.2-q22.2. Am. J. Hum. Genet. 56, 926–937.

Šimková, H., Číhalíková, J., Vrána, J., Lysák, M.A., Doležel, J., 2003. Preparation of HMW DNA from plant nuclei and chromosomes isolated from root tips. Biol. Plant. 46, 369–373. https://doi.org/10.1023/A:1024322001786.

Šimková, H., Šafář, J., Suchánková, P., Kovářová, P., Bartoš, J., Kubaláková, M., Janda, J., Číhalíková, J., Mago, R., Lelley, T., Doležel, J., 2008a. A novel resource for genomics of Triticeae: BAC library specific for the short arm of rye (*Secale cereale* L.) chromosome 1R (1RS). BMC Genomics 9, 237. https://doi.org/10.1186/1471-2164-9-237.

Šimková, H., Svensson, J.T., Condamine, P., Hřibová, E., Suchánková, P., Bhat, P.R., Bartoš, J., Šafář, J., Close, T.J., Doležel, J., 2008b. Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genomics 9, 294. https://doi.org/10.1186/1471-2164-9-294.

Staňková, H., Valárik, M., Lapitan, N.L.V., Berkman, P.J., Batley, J., Edwards, D., Luo, M. C., Tulpová, Z., Kubaláková, M., Stein, N., Doležel, J., Šimková, H., 2015. Chromosomal genomics facilitates fine mapping of a Russian wheat aphid resistance gene. Theor. Appl. Genet. 128, 1373–1383. https://doi.org/10.1007/s00122-015-2512-2.

Staňková, H., Hastie, A.R., Chan, S., Vrána, J., Tulpová, Z., Kubaláková, M., Visendi, P., Hayashi, S., Luo, M., Batley, J., Edwards, D., Doležel, J., Šimková, H., 2016. BioNano genome mapping of individual chromosomes supports physical mapping and sequence assembly in complex plant genomes. Plant Biotechnol. J. 14, 1523–1531. https://doi.org/10.1111/pbi.12513.

Tanaka, T., Kobayashi, F., Joshi, G.P., Onuki, R., Sakai, H., Kanamori, H., Wu, J., Šimková, H., Nasuda, S., Endo, T.R., Hayakawa, K., Doležel, J., Ogihara, Y., Itoh, T., Matsumoto, T., Handa, H., 2013. Next-generation survey sequencing and the molecular organization of wheat chromosome 6B. DNA Res. 21, 103–114. https://doi.org/10.1093/dnares/dst041.

Telenius, H., Carter, N.P., Bebb, C.E., Nordenskjöld, M., Ponder, B.A.J., Tunnacliffe, A., 1992. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. Genomics 13, 718–725. https://doi.org/10.1016/0888-7543(92)90147-K.

The Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815. https://doi.org/10.1038/35048692.

Thind, A.K., Wicker, T., Šimková, H., Fossati, D., Moullet, O., Brabant, C., Vrána, J., Doležel, J., Krattinger, S.G., 2017. Rapid cloning of genes in hexaploid wheat using cultivar-specific long-range chromosome assembly. Nat. Biotechnol. 35, 793–796. https://doi.org/10.1038/nbt.3877.

Thind, A.K., Wicker, T., Müller, T., Ackermann, P.M., Steuernagel, B., Wulff, B.B.H., Spannagl, M., Twardziok, S.O., Felder, M., Lux, T., Mayer, K.F.X., Keller, B., Krattinger, S.G., 2018. Chromosome-scale comparative sequence analysis unravels molecular mechanisms of genome dynamics between two wheat cultivars. Genome Biol. 19, 104. https://doi.org/10.1186/s13059-018-1477-2.

Tiwari, V.K., Wang, S., Sehgal, S., Vrána, J., Friebe, B., Kubaláková, M., Chhuneja, P., Doležel, J., Akhunov, E., Kalia, B., Sabir, J., Gill, B.S., 2014. SNP discovery for mapping alien introgressions in wheat. BMC Genomics 15, 273. https://doi.org/10.1186/1471-2164-15-273.

Tiwari, V.K., Wang, S., Danilova, T., Koo, D.H., Vrána, J., Kubaláková, M., Hribova, E., Rawat, N., Kalia, B., Singh, N., Friebe, B., Doležel, J., Akhunov, E., Poland, J., Sabir, J.S.M., Gill, B.S., 2015. Exploring the tertiary gene pool of bread wheat: sequence assembly and analysis of chromosome 5Mg of *Aegilops geniculata*. Plant J. 84, 733–746. https://doi.org/10.1111/tpj.13036.

Tulpová, Z., Luo, M.C., Toegelová, H., Visendi, P., Hayashi, S., Vojta, P., Paux, E., Kilian, A., Abrouk, M., Bartoš, J., Hajdúch, M., Batley, J., Edwards, D., Doležel, J., Šimková, H., 2019a. Integrated physical map of bread wheat chromosome arm 7DS to facilitate gene cloning and comparative studies. New Biotechnol. 48, 12–19. https://doi.org/10.1016/j.nbt.2018.03.003.

Tulpová, Z., Toegelová, H., Lapitan, N.L.V., Peairs, F.B., Macas, J., Novák, P., Lukaszewski, A.J., Kopecký, D., Mazáčová, M., Vrána, J., Holušová, K., Leroy, P., Doležel, J., Šimková, H., 2019b. Accessing a russian wheat aphid resistance gene in bread wheat by long-read technologies. Plant Genome. 12 https://doi.org/10.3835/plantgenome2018.09.0065.

Valárik, M., Bartoš, J., Kovářová, P., Kubaláková, M., De Jong, J.H., Doležel, J., 2004. High-resolution FISH on super-stretched flow-sorted plant chromosomes. Plant J. 37, 940–950. https://doi.org/10.1111/j.1365-313X.2003.02010.x.

Van Oeveren, J., De Ruiter, M., Jesse, T., Van Der Poel, H., Tang, J., Yalcin, F., Janssen, A., Volpin, H., Stormo, K.E., Bogden, R., Van Eijk, M.J.T., Prins, M., 2011. Sequence-based physical mapping of complex genomes by whole genome profiling. Genome Res. 21, 618–625. https://doi.org/10.1101/gr.112094.110.

Varshney, R.K., Song, C., Saxena, R.K., Azam, S., Yu, S., Sharpe, A.G., Cannon, S., Baek, J., Rosen, B.D., Tar'an, B., Millan, T., Zhang, X., Ramsay, L.D., Iwata, A., Wang, Y., Nelson, W., Farmer, A.D., Gaur, P.M., Soderlund, C., Penmetsa, R.V., Xu, C., Bharti, A.K., He, W., Winter, P., Zhao, S., Hane, J.K., Carrasquilla-Garcia, N., Condie, J.A., Padhyaya, H.D., Luo, M.C., Thudi, M., Gowda, C.L.L., Singh, N.P., Lichtenzveig, J., Gali, K.K., Rubio, J., Nadarajan, N., Dolezel, J., Bansal, K.C., Xu, X., Edwards, D., Zhang, G., Kahl, G., Gil, J., Singh, K.B., Datta, S.K., Jackson, S.A., Wang, J., Cook, D.R., 2013. Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nat. Biotechnol. 31, 240–246. https://doi.org/10.1038/nbt.2491.

Varshney, R.K., Sinha, P., Singh, V.K., Kumar, A., Zhang, Q., Bennetzen, J.L., 2020. 5Gs for crop genetic improvement. Curr. Opin. Plant Biol. 56, 190–196. https://doi.org/10.1016/j.pbi.2019.12.004.

Veuskens, J., Marie, D., Brown, S.C., Jacobs, M., Negrutiu, I., 1995. Flow sorting of the Y sex-chromosome in the dioecious plant *Melandrium album*. Cytometry 21, 363–373. https://doi.org/10.1002/cyto.990210408.

Vitulo, N., Albiero, A., Forcato, C., Campagna, D., Dal Pero, F., Bagnaresi, P., Colaiacovo, M., Faccioli, P., Lamontanara, A., Šimková, H., Kubaláková, M., Perrotta, G., Facella, P., Lopez, L., Pietrella, M., Gianese, G., Doležel, J., Giuliano, G., Cattivelli, L., Valle, G., Stanca, A.M., 2011. First survey of the wheat chromosome 5A composition through a next generation sequencing approach. PLoS One 6. https://doi.org/10.1371/journal.pone.0026421.

Vrána, J., Cápal, P., Šimková, H., Karafiátová, M., Čížková, J., Doležel, J., 2016. Flow analysis and sorting of plant chromosomes. Curr. Protoc. Cytom. 2016, 5.3.1–5.3.43. https://doi.org/10.1002/cpcy.9.

Wang, M.L., Leitch, A.R., Schwarzacher, T., Heslop-Harrison, J.S., Moore, G., 1992. Construction of a chromosome-enriched *Hpa*II library from flow-sorted wheat chromosomes. Nucleic Acids Res. 20, 1897–1901. https://doi.org/10.1093/nar/20.8.1897.

Wang, H., Dai, K., Xiao, J., Yuan, C., Zhao, R., Doležel, J., Wu, Y., Cao, A., Chen, P., Zhang, S., Wang, X., 2017. Development of intron targeting (IT) markers specific for chromosome arm 4VS of *Haynaldia villosa* by chromosome sorting and next-generation sequencing. BMC Genomics 18, 167. https://doi.org/10.1186/s12864-017-3567-z.

Wenzl, P., Suchánková, P., Carling, J., Šimková, H., Huttner, E., Kubaláková, M., Sourdille, P., Paul, E., Feuillet, C., Kilian, A., Doležel, J., 2010. Isolated chromosomes as a new and efficient source of DArT markers for the saturation of genetic maps. Theor. Appl. Genet. 121, 465–474. https://doi.org/10.1007/s00122-010-1323-8.

Wicker, T., Mayer, K.F.X., Gundlach, H., Martis, M., Steuernagel, B., Scholz, U., Šimková, H., Kubaláková, M., Choulet, F., Taudien, S., Platzer, M., Feuillet, C., Fahima, T., Budak, H., Doležel, J., Keller, B., Stein, N., 2011. Frequent gene movement and pseudogene evolution is common to the large and complex genomes of wheat, barley, and their relatives. Plant Cell 23, 1706–1718. https://doi.org/10.1105/tpc.111.086629.

Wricke, G., Wehling, P., 1985. Linkage between an incompatibility locus and a peroxidase isozyme locus (Prx 7) in rye. Theor. Appl. Genet. 71, 289–291. https://doi.org/10.1007/BF00252069.

Xiao, J., Dai, K., Fu, L., Vrána, J., Kubaláková, M., Wan, W., Sun, H., Zhao, J., Yu, C., Wu, Y., Abrouk, M., Wang, H., Doležel, J., Wang, X., 2017. Sequencing flow-sorted short arm of *Haynaldia villosa* chromosome 4V provides insights into its molecular structure and virtual gene order. BMC Genomics 18, 791. https://doi.org/10.1186/s12864-017-4211-7.

Xing, L., Hu, P., Liu, J., Witek, K., Zhou, S., Xu, J., Zhou, W., Gao, L., Huang, Z., Zhang, R., Wang, X., Chen, P., Wang, H., Jones, J.D.G., Karafiátová, M., Vrána, J., Bartoš, J., Doležel, J., Tian, Y., Wu, Y., Cao, A., 2018. Pm21 from *Haynaldia villosa* encodes a CC-NBS-LRR protein conferring powdery mildew resistance in wheat. Mol. Plant 11, 874–878. https://doi.org/10.1016/j.molp.2018.02.013.

Yang, H., Chen, X., Wong, W.H., 2011. Completely phased genome sequencing through chromosome sorting. Proc. Natl. Acad. Sci. U. S. A. 108, 12–17. https://doi.org/10.1073/pnas.1016725108.

Yano, Y., Chiba, T., Asahara, H., 2020. Analysis of the mouse Y chromosome by single-molecule sequencing with Y chromosome enrichment. Front. Genet. 11, 406 https://doi.org/10.3389/fgene.2020.00406.

Zhang, J., Zhang, P., Dodds, P., Lagudah, E., 2020. How target-sequence enrichment and sequencing (TEnSeq) pipelines have catalyzed resistance gene cloning in the wheat-rust pathosystem. Front. Plant Sci. 11, 678. https://doi.org/10.3389/fpls.2020.00678.

APPENDIX III

Comparative analysis of repetitive parts of genomes in eight *Loliinae* representatives

Zwyrtková, J., Kapustová, V., Doležel, J., Hřibová, E.

In: Abstracts of the International Conference “Plant Genome Evolution”

P1.24, Sitges, Spain, 2017

Comparative analysis of repetitive parts of genomes in eight *Loliinae* representatives

Jana Zwyrtková[1], Veronika Kapustová[1], Jaroslav Doležel[1] and Eva Hřibová[1]

[1] *Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic, e-mail: zwyrtkova@ueb.cas.cz*

Introduction

- *Loliinae* subtribe is one of the largest of the *Poeae* tribe. Broad genus of Festuca contains more than 500 species.
- *Festuca* and *Lolium* genera represent closely related complex of species.
- The evolution of grasses was accompanied by frequent and repeated genome size gain and loss. About 70% of species are polyploid.
- Monoploid genome size ranges between 1.58 – 4.03 pg.

Aims of the study

- Complementing the information about the repetitive part of selected fescue and ryegrass species.
- Identify and quantify groups of repeats, assessing their diversity using Repeat Explorer pipeline.
- Studying their contribution to nuclear genome sizes.
- Analysis of centromeric regions.
- Generating bioinformatic resources and cytogenetic markers.

In silico repeat analysis

- Clustering was done on 250 bp long reads, representing ~0.1 % of monoploid genomes.
- Clusters representing at least 0.01% of analyzed data were used for further characterization.
- blastn and blastx program and phylogenetic analysis of coding domains were used to characterize sequences within individual clusters.
- All putative tandem organized repeats were identified by Dotter.

Comparative analysis

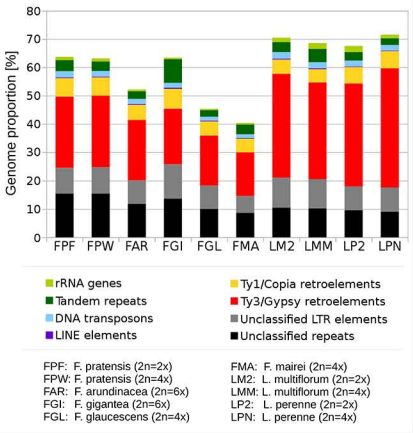
- RepEx pipeline enables to analyze intra- and inter-specific similarities as well as to reconstruct and quantify major repeat families not only within one species but also between several species – process called **comparative analysis**
- Orthologous repeat families from different species are grouped in the same cluster, which facilitates their identification and quantification between the species.
- SeqGrapheR enables to visualize distribution of reads along the graph

Conclusion

- ✔ We provide detailed insights into the genome composition and its variation in fescues and ryegrasses.
- ✔ Within the fescues, *F. mairei* and *F. glaucescens* showed the lowest similarity of DNA repeats as compared to other species of *Festuca*
- ✔ Ty3/*Gypsy* elements were the most frequent repetitive elements.
- ✔ Relatively high number of different tandem repeats was identified in nuclear genomes of all species. None of them is centromere specific.
- ✔ Centromeric region of all analyzed species contains specific Ty3/Gypsy element closely related to Cereba element identified in Hordeum vulgare.
- ✔ Co-localization of CEN-H3 protein and centromeric Ty3/Gypsy element indicates role of this LTR elements for proper function of the centromeres.
- ✔ Databases of repeats were created and are well suited for annotation of genome assemblies.

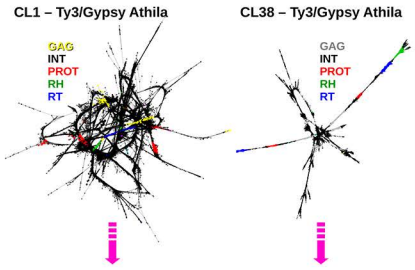
Results

Genome proportion of different types of repeats in nuclear genomes of analyzed accessions

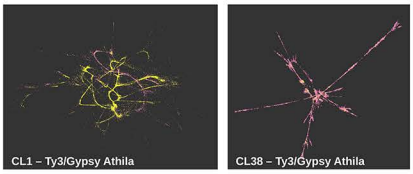


FPF: F. pratensis (2n=2x)
FPW: F. pratensis (2n=4x)
FAR: F. arundinacea (2n=6x)
FGI: F. gigantea (2n=6x)
FGL: F. glaucescens (2n=4x)
FMA: F. mairei (2n=4x)
LM2: L. multiflorum (2n=2x)
LMM: L. multiflorum (2n=4x)
LP2: L. perenne (2n=2x)
LPN: L. perenne (2n=4x)

Graph examples – comparative analysis

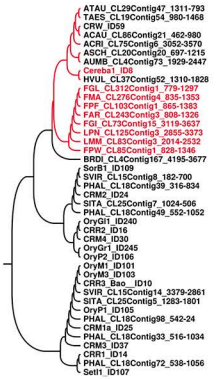


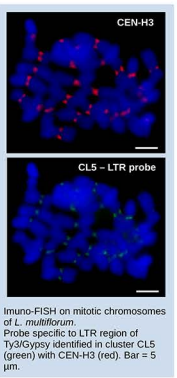
Visualization of reads distribution along the graph using SeqGrapheR – fescue reads in pink, ryegrass reads in yellow



Centromere specific repeats

- Centromeric regions of fescues and ryegrasses do not contain typical satellite repeat
- Fescue and ryegrass centromers contain LTR element from Ty3/Gypsy Chromoviridae family which co-localize with CEN-H3 protein

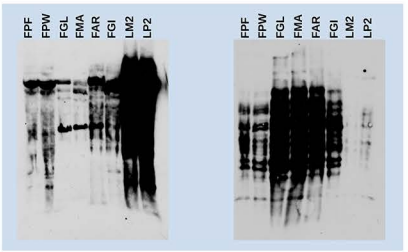




Imuno-FISH on mitotic chromosomes of *L. multiflorum*. Probe specific to LTR region of Ty3/Gypsy identified in cluster CL5 (green) with CEN-H3 (red). Bar = 5 µm.

Experimental verification by Southern hybridization

- Species-specific variants of selected repetitive elements between broad-leaved fescues identified using detailed analysis of individual clusters in SeqGrapheR program were used to design probes for Southern hybridization.



✔ Results of Southern hybridization confirmed *in silico* data.

This work has been supported by the National Program of Sustainability I (grant award LO1204). The computing was supported by the National Grid Infrastructure MetaCentrum (grant No. LM2010005 under the program Projects of Large Infrastructure for Research, Development, and Innovations).

APPENDIX IV

Comparative analysis of repetitive DNA in eight representatives of fescues and ryegrasses

Hřibová, E., Zwyrtková, J., Holušová, K., Čížková, J., Kopecký, D., Doležel, J.

In: Book of Abstracts of the International Conference on Polyploidy, Hybridization and Biodiversity

P. 70. Rovinj, Croatia, 2016

COMPARATIVE ANALYSIS OF REPETITIVE DNA IN EIGHT REPRESENTATIVES OF FESCUES AND RYEGRASSES

Eva Hřibová, Jana Zwyrtková, Kateřina Holušová, Jana Čížková, David Kopecký and Jaroslav Doležel

Introduction

- Broad genus of *Festuca* contains more than 500 species.
- Distributed in holarctic region and in temperate zones of Southern hemisphere.
- *Festuca* is a complex genus divided into several sub-genera and sections.
- The evolution of grasses was accompanied by frequent and repeated genome size gain and loss.
- Monoploid genome size range from 1.58 – 4.03 pg.
- Hybridization and polyploidy play important role in their evolution.
- About 70% of species are polyploid.

Objectives of the study

- Complementing the information about the repetitive part of seven fescue and ryegrass species (*F. arundinaceae*, *F. gigantea*, *F. glaucescens*, *F. mairei*, two cultivars of *F. pratensis*, and *L. perenne* and *L. multiflorum*).
- Identifying and quantifying major groups of repeats, assessing sequence diversity of repeats between and within selected species using graph-based clustering done by RepeatExplorer pipeline.
- Studying the contribution of repetitive DNA to nuclear genome size.
- Generating bioinformatic resources for development of markers and for repeat identification for future genome assembly projects.

Repeat analysis – methods and results

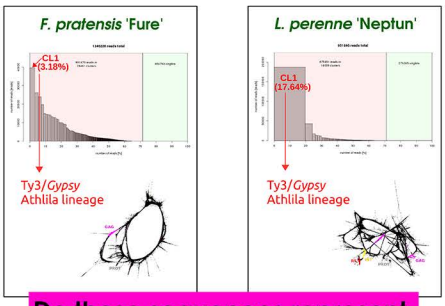
Global repeat identification and characterization using RepeatExplorer

- Clustering was done on 250 bp long reads, representing ~ 0.1 % of nuclear genome.
- Clusters representing at least 0.01% of analyzed data were used for further characterization.
- blastn and blastx program was used to characterize sequences within individual clusters.
- Phylogenetic analysis of coding domains was used for detailed characterization of transposable elements.
- All putative tandem organized repeats were identified by Dotter.

Comparative analysis of repetitive DNA using RepeatExplorer

- RepEx pipeline enables to analyze intra- and inter-specific similarities as well as to reconstruct and quantify major repeat families not only within one species but also between several species – process called **comparative analysis**
- Orthologous repeat families from different species are grouped in the same cluster, which facilitates their identification and quantification between the species.
- SeqGrapheR enable to visualize distribution of reads along the graph

Examples of graph-based clustering



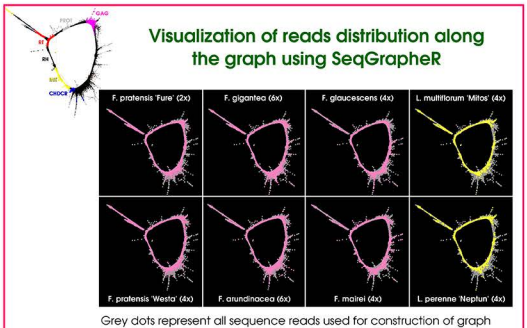
Do these sequences represent one particular element???

Comparative analysis

Cluster (repeat type)	Genome representation of repeats in individual species							
	FPF	FPW	FAR	FGI	FGL	FMA	LMM	LPN
CL1 (Athila)	5.16%	5.60%	3.62%	3.67%	0.57%	0.61%	21.46%	22.18%
CL2 (Angela)	4.11%	4.19%	2.76%	4.44%	2.24%	1.99%	2.92%	3.82%
CL3 (Tandem)	1.64%	2.24%	0.70%	6.40%	0.21%	0.21%	5.26%	1.95%
CL4 (Chromoviridea)	3.22%	3.26%	2.42%	1.96%	1.53%	1.71%	1.98%	2.86%
CL5 (Chromoviridea)	2.42%	2.20%	1.66%	1.30%	1.78%	1.41%	0.67%	0.65%

Table: Examples of five most abundant repeats (clusters) obtained after comparative analysis. Individual species are marked by specific prefixes: *F. pratensis* 'Fure' (FPF); *F. pratensis* 'Westa' (FPW); *F. arundinacea* (FAR); *F. gigantea* (FGI); *F. glaucescens* (FGL); *F. mairei* (FMA); *L. multiflorum* 'Mitos' (LMM) and *L. perenne* 'Neptun' (LPN).

Visualization of reads distribution along the graph using SeqGrapheR



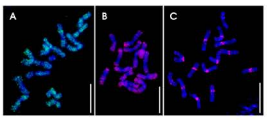
Grey dots represent all sequence reads used for construction of graph (all species), pink dots represent reads specific to different fescue species and yellow dost represent reads specific to Lolium species.

Conclusions

- We provide a detailed characterization of repeat composition of genomes representing the group of 'broad-leaved' *Festuca*.
- Ty3/*Gypsy* elements were the most frequent (more than 3.5 times more abundant than Ty1/*Copia* in fescues and 6 times more abundant than Ty1/*Copia* in ryegrasses).
- Relatively high number of different tandem repeats was identified in nuclear genomes of all species.
- We provide detailed insights into the genome composition and its variation in fescues and ryegrasses.
- In general, high number of similarity hits was found between the species of *Festuca* and *Lolium*.
- Within the fescues, *F. mairei* and *F. glaucescens* showed the lowest similarity of DNA repeats as compared to other species of *Festuca*
- Databases of repeats were created and are well suited for annotation of genome assemblies.

Cytogenetic mapping of repetitive DNA

- Most of all transposable elements resulted in typical dispersed signals on all chromosomes (A).
- New DNA transposon from a *CACTA* family was preferentially localized in subtelomeric region and distal part of all chromosomes (B).
- LTR elements from Chromoviridea lineage of Ty3/*Gypsy* family was preferentially localized to centromeric regions of all chromosomes (C).



- **Tandem organized repeats are one of the most popular cytogenetic markers**
- **They are organized in tandem clusters on the chromosomes and gave good visible signals after FISH**
- **Tandem repeats are often species-specific or chromosome-specific**

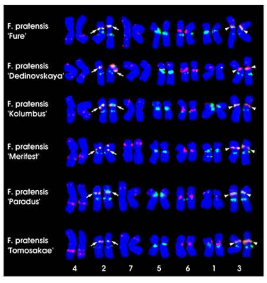


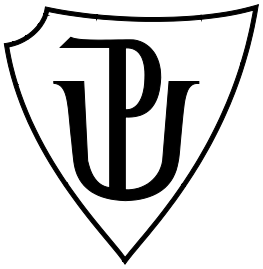
Figure: Localization of fpTR7 , fpTR15 and rRNA genes - 5S rDNA (arrows) and 45S rDNA (arrowheads) on metaphase chromosomes.

This work has been supported by the National Program of Sustainability I (grant award LO1204). We thank V. Černoch (DLF Seeds, Hladké Životice, Czech Republic) for providing seeds of F. pratensis and L. perenne cv. Neptun. The computing was supported by the National Grid Infrastructure MetaCentrum (grant No. LM2010005 under the program Projects of Large Infrastructure for Research, Development, and Innovations).

Palacký University Olomouc
Faculty of Science
Department of Cell Biology and Genetics

and

Institute of Experimental Botany of the Czech Academy of Sciences
Centre of Plant Structural and Functional Genomics



Genome analysis in selected representatives of *Poaceae*

Jana Zwyrtková

Summary of the Ph.D. Thesis

P1527 – Molecular and Cellular Biology

Olomouc 2021

Ph.D. Thesis was carried out at the Department of Cellular Biology and Genetics, Faculty of Science, Palacký University Olomouc, between years 2016–2021.

Candidate: **Mgr. Jana Zwyrtková**

Supervisor: **Prof. Ing. Jaroslav Doležel, DrSc.**

Reviewers: **Prof. Ing. Vladislav Čurn, Ph.D.**
University of South Bohemia in České Budějovice

Doc. Dr. Ing. Pavel Vejl
Czech University of Life Sciences Prague

The evaluation of this Ph.D. Thesis was written by .....
.....

The summary of the Ph.D. Thesis was sent for distribution on .....

The oral defense will take place on .....
in front of the commission for the Ph.D. study of the study program Molecular and Cellular Biology in the conference room at the Institute of Experimental Botany ASCR, Šlechtitelů 31, 779 00 Olomouc – Holice.

The Ph.D. Thesis is available in the Library of the Biological Departments of the Faculty of Science at the Palacký University Olomouc, Šlechtitelů 27, 783 71 Olomouc – Holice.

Prof. RNDr. Zdeněk Dvořák, DrSc. et Ph.D.

Chairman of the Commision for the Ph.D. Thesis of the Study Program Molecular and Cellular Biology
Department of Cellular Biology and Genetics, Faculty of Science, Palacký University Olomouc

Content

1 Introduction

The growth of the world's population places great demands on producing enough food, including the efficient use of arable land. Representatives of the *Poaceae* family, including the important cereals such as wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), rye (*Secale cereale*), and oat (*Avena sativa*) play a significant role in human nutrition. This plant family also includes grass species, which are the main component of pastures and meadows. In addition to providing feed for livestock, they have an important anti-erosion role, help to retain water in the landscape, and perform other functions in the ecosystem. Many types of grasses are grown in decorative gardens, in sports fields, and have many more uses. This Thesis focuses on two groups of grasses, cereals, and fodder crops/amenity grasses.

Meadows, pastures, and lawns are one of the largest ecosystems worldwide. They are sustainable in various geographic and climatic variation (Reheul et al., 2010) and are important from the ecological, economical, and aesthetical points of view (Kopecký and Studer, 2014). Permanent lawns are suitable for biomass production (Searchinger et al., 2008). *Lolium perenne* L. has a high water-soluble carbohydrate content as a polymeric fructan that is easy to extract, split to the separate fructose units, and fermented to bio-ethanol (Farrar et al., 2012). Grasses serve as a readily available source of feed for ruminants that is an economic-friendly way for animals important in agriculture, such as cattle and other farm animals (Reheul et al., 2010). However, in the last decades, meadows are modified by farmers and declining alarmingly due to frequent conversion to arable land (Obermeier et al., 2020). Since the grasses are an important source of feed for farm animals, the need for understanding their genomes to support breeding of improved cultivars is essential. The aesthetical function of grasses is important in decorative areas in parks, around houses, as a recultivation, or in sports ground as a golf course or football field (Kopecký and Studer, 2014).

Growing demands on the performance of new varieties of agricultural crops, including the resistance to biotic and abiotic stresses, are compromised by limited genetic variation in current cultivars, resulting in part from the loss of potentially important genes

during the domestication and breeding. In contrast, wild relatives of crops, which are characterized by larger genetic variation, may serve as an important source of genes and their alleles for breeding. However, the use of this variability is hampered by limited knowledge of their nuclear genomes. Obtaining the missing information is thus necessary for a wider use in the so-called introgressive breeding, which has the potential to contribute to breeding of new varieties that will have higher and stable yields under environmentally sustainable conditions.

Agropyron cristatum (crested wheatgrass) belongs to the family *Poaceae*. It is a perennial low maintenance grass, which can be grown permanently and needs a little additional treatment or special care, such as the application of fertilizers. It has been used for pasture, rangeland, and hayland. Therefore, it has economic importance as a forage, for example, for livestock (Han et al., 2019). It is grown worldwide, but mainly cultivated in arid regions of the United States (Asay et al., 2003; Asay and Jensen, 1996). Wheatgrass has not been domesticated or bred, exhibits large genetic variation, and thus appears a potential source of new genes for wheat improvement (Han et al., 2014). Introgression of its genes to wheat can improve yields, increase resistance to diseases, and improve tolerance to drought (Said et al., 2018).

2 Aims of the Thesis

With the aim to improve the knowledge of genome organization in selected representatives of *Poaceae* and develop molecular resources to support the breeding of crops, this Thesis focuses on three activities:

I Comparative analyses of repetitive DNA elements and identification of centromeric element in representatives of fescues and ryegrasses

The first aim of the Thesis was to characterize repetitive DNA elements in the nuclear genomes in ten representatives of *Festuca* and *Lolium* and provide more insights into DNA sequences in their centromeres using molecular and cytogenetic approaches.

II Genome organization and molecular chromosome structure of *Agropyron cristatum*

The second aim was to characterize repetitive and genic DNA sequences in *Agroypron cristatum*, a wild relative of wheat (*Triticum aestivum*), develop chromosome- and chromosome-arm specific microsatellite (SSR) markers from *A. cristatum*, and perform a comparative genome analysis with the orthologous subgenomes of wheat.

III Comparison of chromosome genomic approaches

The third aim was to review the past and current methods for genome complexity reduction with a focus on flow cytometric chromosome sorting and the use of this method in plant molecular biology and genomics.

3 Materials and Methods

Plant material

Considering the aims of the Thesis a range of plant materials was used. Seeds of fescues and ryegrasses were kindly provided by the Leibniz Institute of Plant Genetics and Crop Plant Research (Gatersleben, Germany), Dr. Arild Larson (Graminor, Norway), Dr. Vladimír Černoch (DLF Seeds, Czech Republic) and Seed Bank Pullman (Washington, USA). Seeds of barley cv. Morex, rye cv. Dánkowskie Diament, and oat cv. Atego were obtained from the Leibniz Institute of Plant Genetics and Crop Plant Research (Gatersleben, Germany). Seeds of wheat cv. Chinese Spring were obtained from Prof. Takashi Ryu Endo (Kyoto University, Japan) and Dr. Pierre Sourdille (INRA, Clermont-Ferrand, France), and seeds of *Aegilops tauschii* were provided by Dr. Miroslav Valárik (Institute of Experimental Botany, Olomouc, Czech Republic).

Seeds of *Agropyron cristatum* cv. Parkway were provided by Dr. Joseph Robins (ARS Forage and Range Research Laboratory, USDA, Logan, USA). Seeds of wheat (cv. Chinese Spring)-*A. cristatum* chromosome disomic addition lines 1P, 2P, 3P, 4P, 5P, 6P, wheat (cv. Chinese Spring)-*A. cristatum* chromosome telosomic arm lines 2PS, 2PL, 4PS, 5PL, 6PS, 6PL, and wheat (cv. Chinese Spring)-*A. cristatum* Robertsonian translocation line TH4 comprising the long arm of wheat chromosome 1B and the short arm of chromosome 1PS of tetraploid *A. cristatum* were provided by Dr. Adoración Cabrera (University of Córdoba, Spain). Seeds of wheat (cv. Chinese Spring)-*A. cristatum* chromosome telocentric arm line 3PS were provided by Dr. Mahmoud Said (Institute of Experimental Botany, Olomouc, Czech Republic).

Chromosome sorting and amplification of DNA

Intact mitotic metaphase chromosome suspensions were prepared from root tips of young seedlings. Three batches of 100,000 chromosomes were flow-sorted following the method of Doležel et al. (1992) and the identity of flow-sorted chromosomes and contamination of the sorted fractions by other chromosomes were checked using fluorescence *in situ* hybridization (FISH, Said et al., 2018) and genome *in situ*

hybridization (GISH, Cabrera et al., 2002). Chromosomal DNA was amplified using standard protocols (Cápal et al., 2015; Šimková et al., 2008).

Illumina sequencing and data analyses

Samples of genomic DNA and DNA from flow-sorted chromosomes were used to prepare Illumina libraries. Briefly, DNA was fragmented, purified, end-repaired, adenylated, size-selected, and ligated with adapters according to the Illumina protocol. DNA sequencing was performed on Illumina MiSeq and HiSeq sequencing platforms (Illumina, San Diego, USA). The sequence reads were processed and used for further bioinformatic analyses, including genome content studies at different levels and SSR markers development.

SSR marker development

Simple sequence repeat (SSR) markers were identified using Microsatellite identification tool (Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany) and appropriate primers were designed. The verification of putative markers was performed by polymerase chain reaction.

Southern hybridization

Genomic DNA was digested by *Hae*III enzyme, size-fractionated by electrophoresis in 1.2% agarose, and transferred onto a nylon membrane. Probes for selected sequences were prepared by polymerase chain reaction with the appropriate template, specific primers, and biotin-labeled dUTP. Southern hybridization was performed overnight at 68 °C and detected with a chemiluminescent approach and 90% stringency.

Immunostaining of chromosomal centromeric regions

To prepare quality chromosome spreads, root tips underwent a preparative process - they were washed in ice water, lysed in LB01 buffer (Doležel et al., 1989), fixed in formaldehyde, and digested in enzymes. Root tips were squashed onto slides and washed in PBS with Triton X-100 and Tween 20. Incubation of chromosome spreads with anti-rabbit primary antibody CENH3 (Nagaki et al., 2004), which was used for centromeric regions labeling, were performed overnight at 4 °C. Detection of signals was done with anti-rabbit Alexa Fluor 488/546 secondary antibody for 1 h at room temperature. Signals were stabilized with ethanol:acetic acid fixative and formaldehyde. Stabilized signals were followed by fluorescent *in situ* hybridization *(*FISH). Probes

for targeted loci or sequence were prepared by polymerase chain reaction with the appropriate template, specific primers, and biotin or digoxigenin-labeled dUTP. Chromosomes were counterstained with 4',6-diamidino-2-phenylindole (DAPI), mounted in Vectashield, and examined with a fluorescent microscope.

4 Summary of results

I Comparative analyses of repetitive elements and identification of centromeric element in representatives of fescues and ryegrasses

Repetitive sequences content of ten *Festuca* and *Lolium* species were compared. The most abundant repeat in *Lolium*, Athila element of the Ty3/gypsy family, was about five times more frequent than in *Festuca* species. A novel centromeric element called Fesreba has been observed and confirmed with the help of molecular methods, mainly by a combination of immunostaining and fluorescence *in situ* hybridization. These comparative analyses confirmed the close relationships between fescues and ryegrasses. The work has been published in BMC Plant Biology (Zwyrtková et al., 2020).

II Genome organization and molecular chromosome structure of *Agropyron cristatum*

The genome of *Agropyron cristatum* has been dissected by flow-sorting into individual chromosomes, which were sequenced and assembled. The resulting draft sequences were used to develop chromosome and chromosome-arm specific SSR markers, which were successfully verified by molecular methods. Comparative analyses of individual *Agropyron* chromosomes with wheat subgenomes were performed to study their orthologous relationships. Repetitive and genic sequence content was characterized and some agronomically important genes were identified. The results of the study will support the use of *A. cristatum* in wheat introgression breeding and the manuscript will be submitted to a journal with an impact factor.

III Comparison of chromosome genomic approaches

The review (Zwyrtková et al., 2021) focuses on experimental approaches to genome complexity reduction with the aim to simplify genome sequencing and gene cloning in plants with large genomes. The potential of flow cytometric chromosome sorting is discussed and overview of the range of applications in molecular biology and genomics is provided. The review has been published in Biotechnology Advances.

5 Summary

The Thesis aimed to improve the knowledge of genome organization in selected representatives of the family *Poaceae*. Their genomes were compared, both at the level of genera, species and cultivars. Important results were obtained using advanced molecular biology methods, including flow cytometry and next-generation sequencing.

In order to compare repetitively arranged DNA sequences and identify a novel centromeric element in the representatives of the *Festuca* and *Lolium* genera, various molecular and genomics approaches were employed. For comparative analyses, sequencing of their genomes by Illumina technology with a small depth of coverage and subsequent *in silico* analyses of the sequence data were used. The analyses showed that retrotransposons were the most abundant type of repetitive DNA sequences in the genomes of fescues and ryegrasses. The biggest difference in the representation of repetitive sequences between genomes was observed for the Athila element, which belongs to the family of Ty3/gypsy retroelements, when the occurrence of this element in ryegrasses was up to five-fold more frequent than in the fescues. The Ty3/gypsy family includes the Cereba element, which typically localizes to (peri)centromeric chromosome regions, as shown for example in barley or wheat. In this work, the long terminal repeat (LTR) element Fesreba was identified, whose relationship with the Cereba element was confirmed by phylogenetic analysis and further confirmed by colocalization with the histone variant CENH3 in cytogenetic experiments. The location of the Cereba element in the centromeric and pericentromeric chromosome regions suggested a possible role of the Fesreba element in these chromosome domains in the studied grass species.

With the aim to facilitate the use of *Agropyron cristatum*, a wild relative of wheat, in introgressive wheat breeding, highly polymorphic microsatellite (SSR) markers specific for individual chromosomes and chromosome arms of this species were developed. The first step in the experimental procedure was the sorting of chromosomes and chromosome arms by flow cytometry. This approach allowed targeted development of SSR markers. The chromosomal DNA was sequenced by Illumina technology and the data were analyzed using appropriate bioinformatics procedures. This led

to *in silico* identification of 250 SSR markers that were verified using PCR. Out of these, 72 were specific for a particular chromosome, or chromosomal arm. The newly developed SSR markers will enable rapid identification of introduced segments of the *A. cristatum* genome into the wheat genome and subsequent monitoring of their transfer to future generations during the breeding process.

There is an increasing need to understand genome organization and function in important crops, many of them with large genomes. The last part of this work reviewed experimental approaches to facilitate the analysis of complex genomes. The focus was on flow cytometric chromosome sorting and on coupling this technology with the methods of molecular biology and genomics. The successful applications include targeted development of DNA markers, physical mapping, chromosome sequencing, and gene cloning.

In summary, the presented Thesis contributed to the improvement of the knowledge on the genome organization in important crops. The results obtained will facilitate the application of molecular and genomic techniques in breeding of new varieties of agricultural crops.

6 References

Asay, K.H., Chatterton, N.J., Jensen, K.B., Jones, T.A., Waldron, B.L., Horton, W.H., 2003. Breeding Improved Grasses for Semiarid Rangelands. Arid L. Res. Manag. 17, 469–478. https://doi.org/10.1080/713936115

Asay, K.H., Jensen, K.B., 1996. Wheatgrasses. Cool. Forage Grasses, Agronomy Monographs. https://doi.org/doi:10.2134/agronmonogr34.c22

Cabrera, A., Martín, A., Barro, F., 2002. *In-situ* comparative mapping (ISCM) of Glu-1 loci in *Triticum* and *Hordeum*, in: Chromosome Research. Springer, pp. 49–54. https://doi.org/10.1023/A:1014270227360

Cápal, P., Blavet, N., Vrána, J., Kubaláková, M., Doležel, J., 2015. Multiple displacement amplification of the DNA from single flow-sorted plant chromosome. Plant J. 84, 838–844. https://doi.org/10.1111/tpj.13035

Doležel, J., Binarová, P., Lucretti, S., 1989. Analysis of Nuclear DNA content in plant cells by flow cytometry. Biol. Plant. 31, 113–120. https://doi.org/10.1007/BF02907241

Doležel, J., Číhalíková, J., Lucretti, S., 1992. A high-yield procedure for isolation of metaphase chromosomes from root tips of *Vicia faba* L. Planta 188, 93–98. https://doi.org/10.1007/BF00198944

Farrar, K., Bryant, D.N., Turner, L., Gallagher, J.A., Thomas, A., Farrell, M., Humphreys, M.O., Donnison, I.S., 2012. Breeding for Bio-ethanol Production in *Lolium perenne* L.: Association of Allelic Variation with High Water-Soluble Carbohydrate Content. Bioenergy Res. 5, 149–157. https://doi.org/10.1007/s12155-011-9156-0

Han, H., Bai, L., Su, J., Zhang, J., Song, L., Gao, A., Yang, X., Li, X., Liu, W., Li, L., 2014. Genetic rearrangements of six wheat-*Agropyron cristatum* 6P addition lines revealed by molecular markers. PLoS One 9. https://doi.org/10.1371/journal.pone.0091066

Han, H., Liu, W., Zhang, J., Zhou, S., Yang, X., Li, X., Li, L., 2019. Identification of P genome chromosomes in *Agropyron cristatum* and wheat-*A. cristatum* derivative lines by FISH. Sci. Rep. 9. https://doi.org/10.1038/s41598-019-46197-6

Kopecký, D., Studer, B., 2014. Emerging technologies advancing forage and turf grass genomics. Biotechnol. Adv. https://doi.org/10.1016/j.biotechadv.2013.11.010

Nagaki, K., Cheng, Z., Ouyang, S., Talbert, P.B., Kim, M., Jones, K.M., Henikoff, S., Buell, C.R., Jiang, J., 2004. Sequencing of a rice centromere uncovers active genes. Nat. Genet. 36, 138–145. https://doi.org/10.1038/ng1289

Obermeier, M.M., Gnädinger, F., Durai Raj, A.C., Obermeier, W.A., Schmid, C.A.O., Balàzs, H., Schröder, P., 2020. Under temperate climate, the conversion of grassland to arable land affects soil nutrient stocks and bacteria in a short term. Sci. Total Environ. 703, 135494. https://doi.org/10.1016/j.scitotenv.2019.135494

Reheul, D., De Cauwer, B., Cougnon, M., 2010. The Role of Forage Crops in Multifunctional Agriculture, in: Fodder Crops and Amenity Grasses. Springer New York, pp. 1–12. https://doi.org/10.1007/978-1-4419-0760-8_1

Said, M., Hřibová, E., Danilova, T. V., Karafiátová, M., Čížková, J., Friebe, B., Doležel, J., Gill, B.S., 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. Theor. Appl. Genet. 131, 2213–2227. https://doi.org/10.1007/s00122-018-3148-9

Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D., Yu, T.H., 2008. Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. Science (80). 319, 1238–1240. https://doi.org/10.1126/science.1151861

Šimková, H., Svensson, J.T., Condamine, P., Hřibová, E., Suchánková, P., Bhat, P.R., Bartoš, J., Šafář, J., Close, T.J., Doležel, J., 2008. Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genomics 9, 294. https://doi.org/10.1186/1471-2164-9-294

Zwyrtková, J., Němečková, A., Čížková, J., Holušová, K., Kapustová, V., Svačina, R., Kopecký, D., Till, B.J., Doležel, J., Hřibová, E., 2020. Comparative analyses of DNA repeats and identification of a novel Fesreba centromeric element in fescues and ryegrasses. BMC Plant Biol. 20, 280. https://doi.org/10.1186/s12870-020-02495-0

Zwyrtková, J., Šimková, H., Doležel, J., 2021. Chromosome genomics uncovers plant genome organization and function. Biotechnol. Adv. 46, 107659. https://doi.org/10.1016/j.biotechadv.2020.107659

7 List of author's publications

7.1 Original papers

Zwyrtková, J., Němečková, A., Čížková, J., Holušová, K., Kapustová, V., Svačina, R., Kopecký, D., Till, B.J., Doležel, J., Hřibová, E., 2020. Comparative analyses of DNA repeats and identification of a novel Fesreba centromeric element in fescues and ryegrasses. BMC Plant Biol. 20, 280. https://doi.org/10.1186/s12870-020-02495-0
IF_{2019}: 3.497

Zwyrtková, J., Šimková, H., Doležel, J., 2021. Chromosome genomics uncovers plant genome organization and function. Biotechnol. Adv. 46, 107659. https://doi.org/10.1016/j.biotechadv.2020.107659
IF_{2019}: 10.744

7.2 Published abstracts – poster presentation

Zwyrtková, J., Kapustová, V., Doležel, J., Hřibová, E.: Comparative analysis of repetitive parts of genomes in eight *Loliinae* representatives. In: Abstracts of the International Conference "Plant Genome Evolution". P1.24, Sitges, Spain, 2017.

Hřibová, E., Zwyrtková, J., Holušová, K., Čížková, J., Kopecký, D., Doležel, J.: Comparative analysis of repetitive DNA in eight representatives of fescues and ryegrasses. In: Book of Abstracts of the International Conference on Polyploidy, Hybridization and Biodiversity. P. 70. Rovinj, Croatia, 2016.

8 Souhrn (Summary in Czech)

Název práce: Analýza genomů vybraných zástupců *Poaceae*

Cílem disertační práce bylo prohloubit znalost struktury genomů vybraných zástupců čeledi *Poaceae*. Jejich genomy byly porovnávány jak na úrovni rodů, tak na úrovni druhů a kultivarů. Pomocí metod molekulární biologie, včetně průtokové cytometrie a metod sekvenování nové generace, byly získány významné výsledky.

Různé molekulární a genomické přístupy byly použity za účelem srovnání zastoupení repetitivně uspořádaných sekvencí DNA a identifikace nového centromerického elementu u zástupců rodů kostřava (*Festuca*) a jílek (*Lolium*). Pro komparativní analýzy bylo využito sekvenování jejich genomů technologií Illumina s malou hloubkou pokrytí a následné *in silico* analýzy sekvenačních dat. Tyto analýzy ukázaly, že retrotranspozóny jsou nejčetnějším typem repetitivních DNA sekvencí v genomech kostřav a jílků. Největší rozdíl v zastoupení repetitivních sekvencí mezi genomy byl pozorován v případě elementu Athila, který patří do rodiny Ty3/gypsy retroelementů, kdy výskyt tohoto elementu u jílků byl až pětkrát četnější než u kostřav. Do rodiny Ty3/gypsy patří Cereba element, který se typicky vyskytuje v (peri)centromerických oblastech, například u ječmene či pšenice. V této práci byla identifikována dlouhá terminální repetice (LTR) element Fesreba, jehož příbuznost s Cereba elementem byla potvrzena fylogenetickou analýzou a také kolokalizací s histonovou variantou CENH3 pomocí cytogenetických experimentů. Výskyt Cereba elementu byl potvrzen v centromerických a pericentromerických oblastech, což naznačuje jeho možnou roli v těchto chromozomálních doménách u studovaných druhů trav.

S cílem prohloubit znalosti o potenciálních zdrojích důležitých genů pro introgresní šlechtění pšenice byl studován její planý příbuzný druh *Agropyron cristatum*. V průběhu disertační práce byly vyvinuty vysoce polymorfní mikrosatelitní (SSR) markery specifické pro jednotlivé chromozómy a ramena chromozómů tohoto druhu. Výchozím experimentálním postupem bylo třídění chromozómů

a chromozómových ramen pomocí průtokové cytometrie. Tento přístup umožnil zacílený vývoj SSR markerů. DNA chromozómů byla sekvenována technologií Illumina a data byla zpracována odpovídajícími bioinformatickými postupy. To umožnilo *in silico* identifikaci 250 SSR markerů, které byly ověřeny pomocí PCR. Z nich 72 bylo specifických pro určitý chromozóm či rameno chromozómu. Nově vyvinuté SSR markery umožní rychlou identifikaci vnesených segmentů genomu *A. cristatum* do genomu pšenice a sledování jejich přenosu do dalších generací v průběhu šlechtění.

S ohledem na rostoucí potřebu znalosti organizace a funkce genomu důležitých zemědělských plodin, které mají obvykle velké genomy, byly v přehledném článku shrnuty experimentální přístupy pro usnadnění jejich studia. Byla popsána kombinace metod třídění chromozómů pomocí průtokové cytometrie a metod molekulární biologie a genomiky. Úspěšné aplikace zahrnují cílený vývoj DNA markerů, fyzické mapování, sekvenování na chromozómové úrovni a klonování genů.

Ve svém souhrnu předkládaná disertační práce přispěla k prohloubení znalosti organizace genomů důležitých rostlin a vytvořila předpoklady pro snadnější aplikaci molekulárních a genomických technik ve šlechtění nových odrůd zemědělských plodin.