Palacký University Olomouc

Faculty of Science

Department of Botany

And

Institute of Experimental Botany

Centre of Region Haná for Biotechnological and Agricultural Research

Olomouc

Miroslava Karafiátová

Cytogenetic analysis of crop genomes

Ph.D. Thesis

Supervisor: Mgr. Jan Bartoš, Ph.D.

Olomouc 2014

Acknowledgements:

First, I would like to thank to the head of laboratory, prof. Ing. Jaroslav Doležel, DrSc., for the great opportunity to work in his famous and successful group. It was my pleasure to be a part of his excellent team for these several years. My cordial thanks also go to my supervisor, Mgr. Jan Bartoš, Ph.D., for his stimulating ideas, kindness and professional guidance. Further, I would like to express my thanks to all my colleagues and friends in Centre of Plant Structural and Functional Genomics for their never-ending encouragements and friendly atmosphere, which make my Ph.D. study unforgettable period of my life. Finally, I would like to thanks to my family for their patience and all their support.

Declaration:

I hereby declare that I elaborated this Ph.D. thesis independently under the supervision of Mgr. Jan Bartoš, Ph.D. using only information sources referred in the Reference chapter.

……

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01), Centre of Region Haná for Biotechnological and Agricultural Research and the grant LO1204 from the National Program of Sustainability I and Internal Grant Agency of Palacký University (grant award No. IGA PrF/2012/001).

BIBLIOGRAPHIC IDENTIFICATION

Author's name: Mgr. Miroslava Karafiátová

Title: Cytogenetic analysis of crop genomes

Type of Thesis: Ph.D. thesis

Department: Botany

Supervisor: Mgr. Jan Bartoš, Ph.D.

The Year of Presentation: 2014

Abstract:

Crops represent highly variable plant group accompanying the human society from beginning of time and having an essential impact also on recent human living. These plants contribute to the standard of living with various aspects, nevertheless the crops used as a food and fodder have influenced the human life most significantly. Wheat, maize, rice, soybean, sugarcane and further barley and chickpea represent the fundamental nutrition source for enormous number of people all over the world. Reduced crop production or their total absence results in serious malnutrition, what substantially stimulate the crop plant research.

Cytogenetics is a versatile tool for plant genome analysis. It provides the number of fundamental characteristics of each species, which are considered to be the essential background for more extended studies. In the first part of the thesis, potential of cytogenetics in plant genome mapping was emphasised on cultivated barley (*Hordeum vulgare* L.). Mapping of short, low-copy full length cDNA clones on barley metaphase chromosomes has proved that even DNA fragments of limited length can be reliably localize using fluorescence *in situ* hybridization (FISH), what provide an alternative approach for physical map construction or verifying the contigs assembly in problematic region like centromere. Further, the second part of the thesis was focused on developing of chromosome painting protocol intended primarily for the plant species with large genome and considerable portion of repetitive sequences. The protocol was applied on barley metaphase chromosomes with the aim to paint barley chromosome 1H. Highly innovative approach employs the procedure of gene-capture and flow cytometry in combination with FISH. Our results represent very important leap forward the study of chromosome organization in interphase nuclei in species with large and complex genetic information.

Key words: crop, cytogenetics, FISH, cytogenetic mapping, probe, single-copy region, chromosome painting

Number of pages/Apendices: 75/XI

Language: English

BIBLIOGRAFICKÁ IDENTIFIKACE

Jméno: Mgr. Miroslava Karafiátová

Název: Cytogenetická analýza genomů hospodářsky významných plodin

Typ práce: Disertační práce

Obor: Botanika

Školitel: Mgr. Jan Bartoš, Ph.D.

Rok obhajoby: 2014

Abstrakt:

Hospodářsky významné plodiny představují velice rozmanitou skupinu rostlin, které jsou od prvopočátku neodmyslitelně spjaty s utvářením lidské civilizace a které ve značné míře ovlivňují i život dnešního člověka. Tyto plodiny přispívají ke kvalitě života různými aspekty, přičemž největší dopad mají plodiny, které slouží jako zdroj potravy buď pro lidstvo samotné, nebo jsou využívány jako krmivo pro hospodářská zvířata. Mezi plodiny tvořící esenciální složku potravy obyvatel v různých částech světa patří hlavně pšenice, kukuřice, rýže, sója, cukrová třtina a v neposlední řadě i ječmen či cizrna. Jejich nedostatek se odráží v podvýživě obyvatel, což významně podněcuje výzkum v této oblasti.

Velmi široké možnosti pro studium rostlinných genomů nabízí cytogenetické techniky. Cytogenetika umožňuje řadu analýz, jež poskytují základní charakteristiku všech druhů a jsou nutným základem pro většinu komplexnějších studií. Potenciál cytogenetiky při mapování rostlinných genomů byl zdůrazněn v první části disertační práce na ječmeni setém (*Hordeum vulgare* L.). Mapování krátkých a nízkokopiových fl-cDNA klonů na metafázní chromozómy ječmene prokázalo, že i DNA fragmenty s omezenou délkou mohou být spolehlivě lokalizovány pomocí fluorescenční *in situ* hybridizace (FISH), což nabízí alternativní variantu při sestavování fyzických map a ověřování správnosti uspořádání jednotlivých kontigů především v problematických oblastech jako jsou např. centromery. Další podstatná část práce byla zaměřena na vytvoření protokolu tzv. malování chromozómů (chromosome painting) určeného především pro druhy s komplexními genomy s vyšším obsahem repetitivních sekvencí. Protokol byl testován na metafázních chromozómech ječmene se sondou pro chromozóm 1H. Vysoce inovativní přístup využívá metodiku tzv. vychytávání genů a průtokové cytometrie v kombinaci s FISH. Dosažené výsledky představují velmi důležitý krok ke studiu uspořádání chromozómů v interfázních jádrech.

Klíčová slova: hospodářsky významné plodiny, cytogenetika, FISH, cytogenetické mapování, sonda, jednokopiové oblasti, „malování“ chromozómů

Počet stránek/ příloh: 75/XI

Jazyk: anglický

CONTENT

1 INTRODUCTION

A crop is any cultivated plant, whose product is harvested by a human at some point of its growth stage. The term „crop“ comprises many species classified in different taxonomic groups across the whole plant kingdom and whose yields are used for various purposes. Human evolution has been accompanied with the domestication of the wildly growing plants since time immemorial. The process of cultivation has highlighted the potential of plants and has strengthened their role in human´s life. Over the years, several of these plant species became the key food source for the majority of the world population. The other crops are effectively used as a fuel, fodder for herbivorous animals or are grown for any other economic purpose and human utilization. Major crops include maize, wheat, rice, sugarcane, pumpkin, cassava, soybeans, potatoes and cotton (FAOSTAT, 2009). Besides the mentioned species, several representatives from the grass family Poaceae (Gramineae) and the legume family Fabaceae (Leguminosae) represent another significant group of crops.

The fact that human existence is highly dependent to the crop harvest, significantly increase the importance of these plants as a scientific objects. Facing the ever-increasing demand for food, the crops improvement is still among the top of the researcher´s aims. The breeding programs are mainly focused on increasing the grain yield, improving the nutritional quality or pathogen resistance. Considering the agronomical importance, crops mostly have well researched scientific background, but despite that there are still a lot of unsatisfactory explored areas, which need the further research.

In this thesis, we aimed to enlarge the knowledge of nuclear genome of four selected crop species (barley, chickpea, meadow fescue and ryegrass) using the cytogenetic approaches. We hope the outcomes could be useful in subsequent analysis in this field and could contribute to increase the general knowledge of these crops.

2 LITERATURE OVERVIEW

2.1 CROPS AND THEIR IMPORTANCE

In this chapter, I am going to present in more detail three crop groups, which include species I focused on in experimental part of my work. I would like to emphasize that taxonomic classifying and grouping of crop plants are not consistent. The representatives from one taxonomic family Leguminosae are assigned into two various crop groups. Legumes, whose are cultivated for grain, are classified as grain crop. On the contrary, some Legumes utilized for the similar purposes like grasses, are detached from grain crops and belong to the forage crops.

2.1.1 Cereal crops

From all crop plants, the human civilization is probably the most affected and influence by cereals. Growing of cereals has accompanied the human societies since the prehistoric times. Nutrients, namely high content of carbohydrates, represent the majority of daily sustenance in developing countries and recently, cereals serve as a food source for more than two third of world population (Jauhar, 2006). In majority of African and Asian countries, cereals represent over 80% of dietary proteins. In these countries, decrease in cereal production leads to severe malnutrition and starving of inhabitants. This serious problem has driven the cereal research forward and while the cereal breeding programmes are enormously supported and preferred, the situation remains critical and continuously increasing world population and its demand is warning. Except the developing nations, cereals are staple food also in developed countries, where cereal consumption is moderate and varied, but still substantial.

Systematically, cereals are grasses belonging to monocots, family Poaceae. Wheat, rice, sorghum, maize and pearl millet are among the top of crop plants at all considering the number of people nutritionally relied on the crop. Other, not less important, species from this family are barley, oat, rye and triticale. Extreme diversity of the cereal species is connected to their diverse origin and significantly contributes to their world importance and allows them to be grown in various parts of the world under different conditions.

Barley (*Hordeum vulgare* L., Figure 1) is one of the most ancient, most extensively studied crop and it is considered to be the fourth in total production among the cereals (Costa and Singh, 2006). Being diploid, it serves as a genetic model for more complicated and complex genomes of Triticeae tribe representatives. Comparing to the other cereals, barley shows a higher tolerance to several abiotic stresses like saline, drought or cold. Barley is used in agriculture for feeding the livestock and in food industry as well, mainly as a material for the alcoholic beverages.

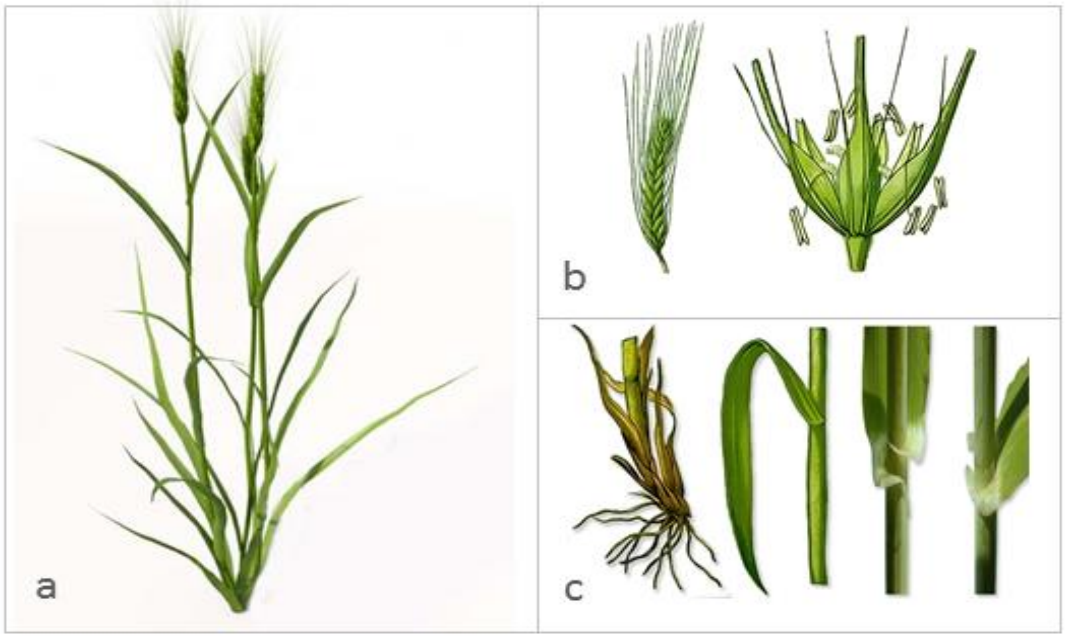


Figure 1: Cereal crop – barley (*Hordeum vulgare* L.). a) complete plant, b) spike and floret detail, c) detail of roots and leaf sheat with auricles. (taken from http://www.turbosquid.com)

2.1.2 Forage crops

Forage crops belong mostly to the grass (Poaceae) and legume (Fabaceae) families. Except their irreplaceable role in maintaining of the landscape by preventing soil erosion and degradation, they exclusively served as livestock feed. Alfalfa, trefoil, birdsfoot and clover from Fabaceae together with ryegrass, wheatgrass, wildrye grasses or bermudagrass from Poaceae are the most important representatives of forage crops (Warnke *et al*, 2009).

Presence of forage legumes in meadows is reflected in enrichment of the soil with nitrogen, which is fixed directly from the atmosphere. Furthermore, legumes

also carry the number of economically important traits. They have a high nutrition value, are significant source of minerals, vitamins, proteins and beginning to be used as a biofuel. Further, because of relatively small genome, some species are considered as the model crops for genome research. For instance, genome of *Medicago truncatula* (close relative of alfalfa) has about 500 Mb and has already been completely sequenced (Yong *et al*., 2011).

Grasses belongs to the most important plants growing all over the world. The grassland area is guessed on 52.5 x 10^6 km^2 (20 %) of the globe vegetation (Reheul *et al*., 2010). Crop grasses mainly produce feed for livestock and are used for forage and turf. Besides, grassland significantly contribute to the forming the surroundings, improves the soil quality, protects the environments and influences climate. Ryegrass (*Lolium L.*), especially Italian ryegrass (*Lolium multiflorum* Lam., Figure 2a) and meadow fescues (*Festuca pratensis* L., Figure 2b), represents the fundamental part of the majority of the grasslands in moderate climate. Both species carry the extremely valuable traits, which in combination contribute to relatively resistant and quality grassland. While meadow fescues show high resistance to abiotic stress, the Italian ryegrass representatives are characterised by the high nutritive index, yield and easy digestibility.

Figure 2: Forage crops a) meadow fescue (*Festuca pratensis* L.) b) ryegrass (*Lolium multiflorum* Lam.) (taken from http://www.lizzieharper.co.uk)

2.1.3 Grain legumes

Legumes are member of the family Fabaceae (also Leguminosae). Many agronomically significant species like common bean, pea, chickpea, soyabean, lentil or mungbean belong to this group. Within the family, legumes could be sorted according to the weather condition preferences into warm weather (*Vigna*, *Glycine*, *Phaseolus*) and cold weather (*Vicia*, *Pisum*) legumes. As cereals, legumes enrich a human diet, but while cereals are significant source of carbohydrates, legumes are rich in proteins. The high protein content (20-50%) is an excellent trait and legumes supply the protein requirement, especially in the countries, where the people suffer from the lack of meat in diet due to the religious or poverty. Recently, legumes become increasingly important also in developed countries, where the vegetarian lifestyle is trendy. However, legumes are stable compound of the diet even in populations with balanced diet and sufficient supply of meat.

Although the grain legumes are comparably relevant for human and animal nourishment as cereals, legumes research is far behind the cereal research. There were many efforts to produce high yield cultivars, but without the breakthrough results.

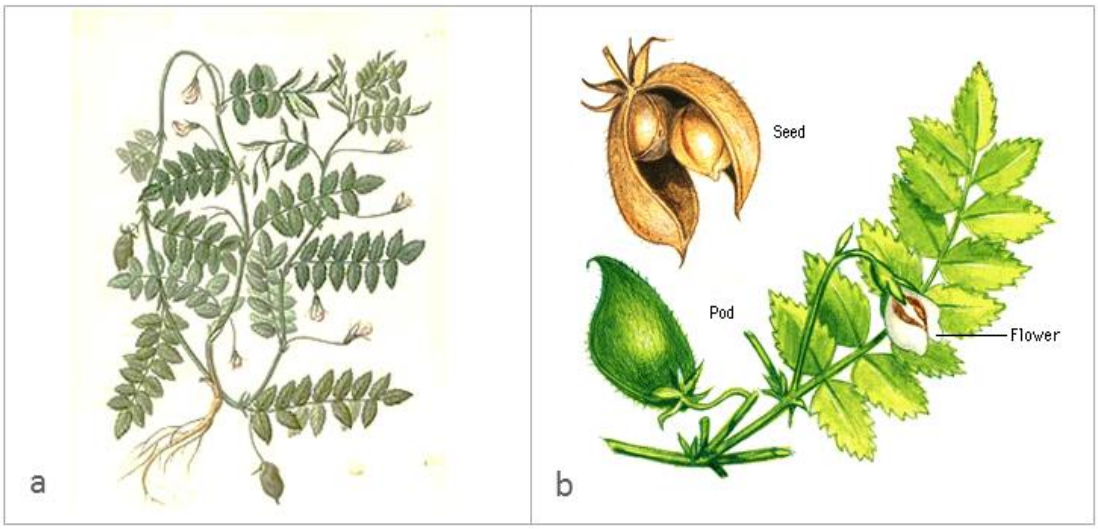


Figure 3: Chickpea (*Cicer arietinum* L.) a) complete plant, b) offshoot detail. (taken from www.plantillustrations.org)

Chickpea (*Cicer arietinum* L., Figure 3) is one is the earliest cultivated legumes. Originally, it is subtropical species used to grown in warm climate. It can be grown in temperate zone as well, but with much lower yield. There are two main types of chickpea; Desi and Kabuli. While Desi is typical Asian cultivar, Kabuli is grown mostly in Europe and Africa. India is world leader in chickpea production for several years and it is far ahead the following Australia (FAO, 2011). Similarly to the other crops, chickpea is grown for two main purposes: food and feed. Its grains have a high nutritive value. Except already mentioned proteins, they are excellent, low fat source of essential nutrients, iron, folate, phosphorus and dietary fibre (http://ndb.nal.usda.gov/ndb/search/list). Due to that it is popular vegetable in worldwide cuisine.

2.2 PLANT NUCLEAR GENOME

Genetic information of the cell is carried in DNA molecule. In eukaryots, main genetic information is located in nucleus and represents the nuclear genome. Except the nucleus, the smaller amount of DNA is stored within semiautonomic organelles like mitochondria and plastids (in plants). Similarly, complete genetic information comprising all mentioned DNA in cell is called cell genome.

2.2.1 Genome size

The plant nuclear genomes are extremely variable in terms of size, ploidy level or chromosome number (Figure 4). Even the species within one family can have a large difference in DNA content. Genome size indicates the amount of nuclear DNA in the individual cell and is expressed using the C-value. This value is defined as DNA content of unreplicated haploid chromosome complement (n) (Bennett and Smith, 1976) and it is quantified in picograms (pg) or in number of base pairs (bp) in the relation 1 pg=0.978×10^9 bp (Doležel *et al*., 2003). Genome size is most frequently measured using the flow cytometry. This robust and reliable method allowed us to analyse the genome size in many different species and revealed very interesting phenomenon. It was found out, that C-value is extremely variable value, which does not correspond to the complexity or evolutionary stage of the organism. The small, simple and taxonomically lower organisms often have larger and more complicated genomes than some higher organisms. This phenomenon is called C-value paradox (Thomas, 1971) and the explanation is connected with the evolutionary processes. During the evolution, the organisms have undergone the processes of crosses, fusions or polyploidizations, which usually resulted in accumulation of genetic information and individuals with bigger genomes. Lower organisms with bigger genomes preserve the huge amount of non-essential repetitive DNA, which numerous higher organisms with smaller genomes did not accumulated or get rid of during the evolution. With 670 Gbp, *Polychaos dubium*, representative of Protozoa, has the largest known genome at all (Parfrey *et al*., 2008) and genome of *Paris Japonica* (150 Gbp) is largest genome in plant kingdom (Pellicer *et al*, 2010). *Genlisea margaretae*, whose genome size is 63.4 Mb, is on the other end of the scale and represents the smallest known plant genome (Greilhuber *et al*., 2006).

In general, the high DNA content is hallmark trait of the widely cultivated Triticeae. Barley genome contains approximately 5 Gbp/1C (Doležel *et al*., 1998) and represents the largest genomes among the species involved in experimental part of the thesis. Comparing to another model plant rice (*Oryza sativa* L.), barley genome is 10 times bigger, but considering the genome size of rye (*Secale cereal* L., 8 Gbp), oat (*Avena sativa* L., 11.3 Gbp) or wheat (*Triticum aestivum* L., 17 Gbp) (Bennett and Smith, 1976), barley yet has the smallest genome among the mentioned species and due to the large genome homology, barley is considered to be a great model species for cereal research.

Genomes of the two involved grass species are smaller. Diploid meadow fescue *F. pratensis* has 3.25 Gbp (Kopecký *et al*., 2010) and DNA content of Italian ryegrass *L. multiflorum* is 2.62 Gbp (Kopecký *et al*., 2010). Finally, genome size value of chickpea is diametrically different. Chickpea has the lowest DNA content of the chosen species. The genome size is roughly 900 Mb and this value is similar for both Kabuli and Desi types (Bennett and Smith, 1976).

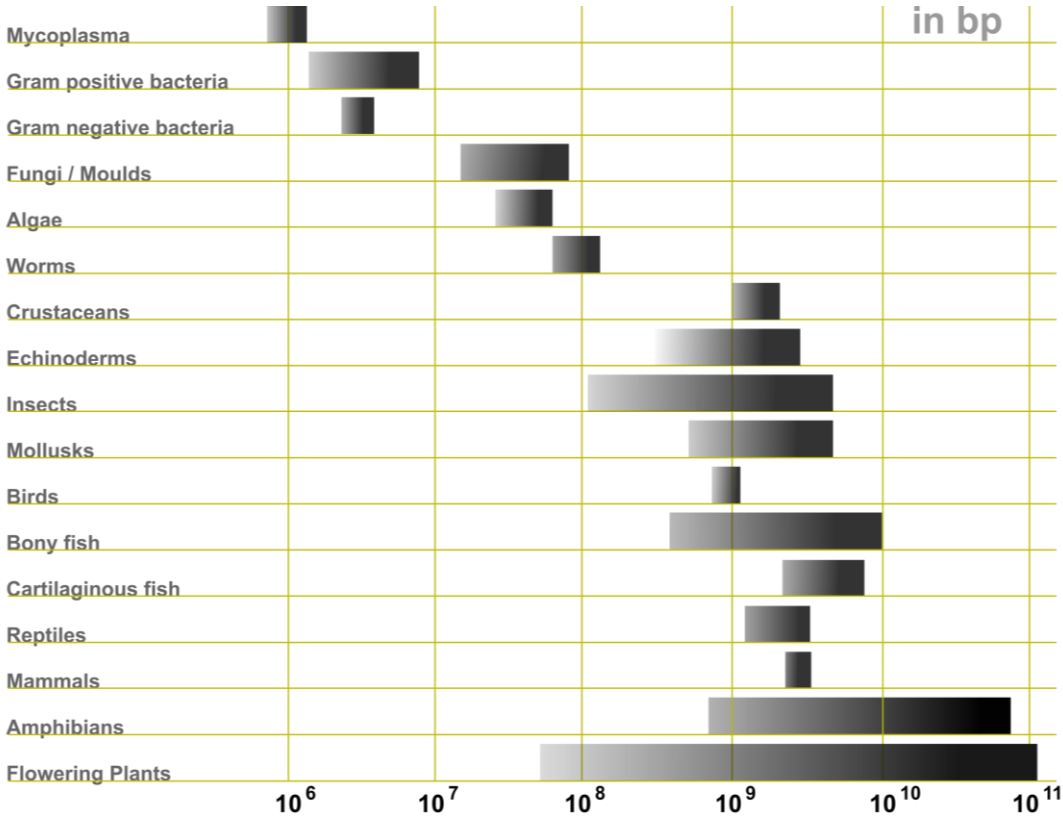


Figure 4: C-value variability among all kingdoms showing slight correlation between genome size and organismal complexity. (Gregory TR, 2014).

2.2.2 Genome structure and organization

Despite that the plant nuclear genomes differ in many aspects, they show a high degree of uniformity in the way of nucleus structure and chromatin organization. Genome consists of DNA molecule and together with associated proteins they form the structure called chromatin. Chromatin appears in nucleus in two distinctive structural variants. During the majority of the cell cycle, chromatin exists in the nucleus as a relaxed fiber, while in dividing cells the chromatin is arranged into highly organized structures – chromosomes. Chromatin formation in individual cell cycle stages is related to the genome function in particular stage.

2.2.2.1 Chromosomes

Nuclear DNA is divided into the highly organized objects called chromosomes. The term chromosome comes from Greek merging two words – *chromo* (colour – due to the property being strongly stained with several dyes) and *soma* (body). The term was firstly introduced by German anatomist Heinrich Wilhelm Gottfried von Waldeyer-Hartz in 1888 in effort to describe the objects in the process of cell division.

Chromosomes appear in its typical X-shaped form, clearly visible in light microscope during the cell division. As visible and easily available distinct objects in majority of plants, they are predetermined to be a strong tool for basic characterization of plant genomes. Using visual observation it is possible to compile the fundamental feature of each genome – karyotype. The karyotype assembling has been mostly based on the clearly distinguishable markers resulting from chromosome morphology. Each chromosome consists of two sister chromatids, which are constricted in the point called centromere (Figure 5). This grip divides the chromatids into smaller subunits named “arms”. The position of centromere (also called primary constriction) serves as one of the marker for elementary chromosome classification and is expressed as centromeric index (CI) - the ratio of shorter chromosome arm to the total chromosome length, originally expressed as a percentage. According to the centromeric index, the chromosomes are classified as metacentric, submetacentric, acrocentric and telocentric.

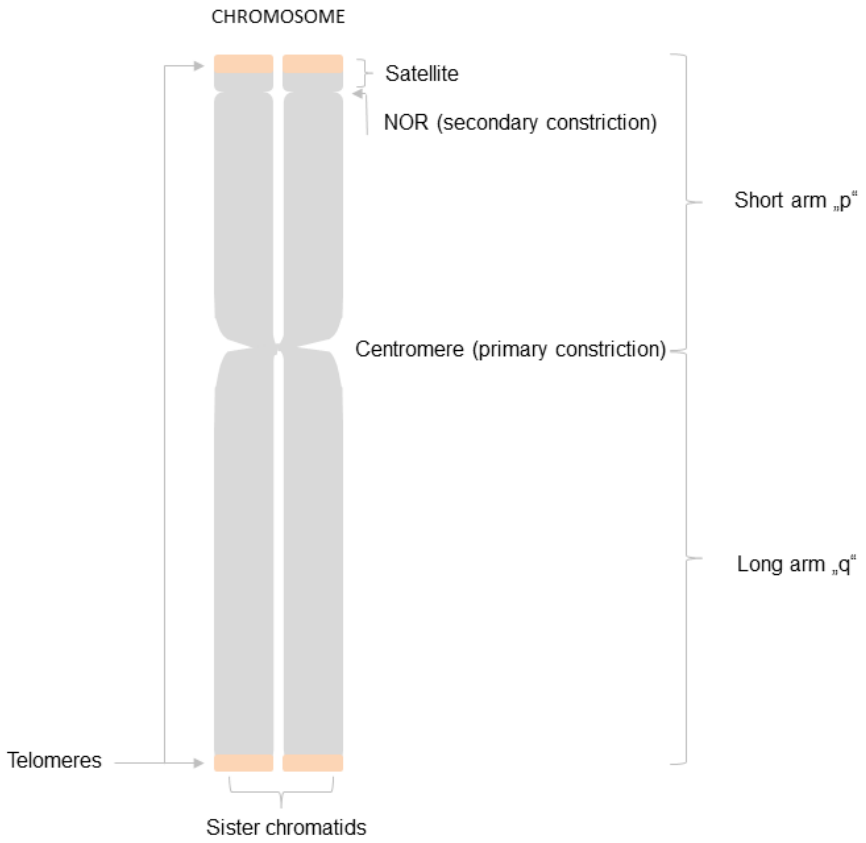


Figure 5: Scheme showing main morphological features of mitotic chromosome

The presence of centromere is essential during the cell division, when microtubules of mitotic spindle are attached to the kinetochore and ensured the equal distribution of homologue chromosomes towards the opposite spindle poles. Usually, chromosomes have only one centromere. The appearance of none or more centromeres is rare, resulting in pathological chromosomal rearrangements and the offspring is mostly not viable. Nevertheless, there exist a few species, which originally have the different centromeric distributions. In plant, *Ophioglossum* carries chromosomes with more than two centromeres – polycentric chromosomes (Khandelwal, 1990). Extreme form of polycentric chromosomes is holocentric chromosome, where whole chromosome length represents functional centromere. This phenomenon is described for instance in flowering plant *Luzula elegans* (Heckmann *et al.*, 2011). Interestingly, it was proved that the centromere substantiality does not depend on DNA sequence, but it is encoded in epigenetic modification of chromatin (Carroll and Straight, 2006) and therefore chromosome, which lost its centromere, is not necessarily convicted to elimination. Due to the

evolutionary pressure, it can establish new centromere (neocentromere) by epigenetic reprogramming and chromosome could be maintained to next generation.

The chromosome stability depends on co-existence of centromere together with telomeres. Telomere represents terminal part of chromosome and protects the chromosome ends from the fusion or degradation. Naturally, each cell has a machinery of the proteins checking DNA integrity. The free DNA ends signalize the DNA damage and these ends are immediately recognized by the machinery, and subsequently repaired. The specific structure of chromatin in these regions prevents the telomeres to be targeted by the nucleases and thus contribute the chromosome integrity and stability. The chromosome stability is further underlined by the local chromatin structure. Both centromeres and telomeres consist of highly condensed heterochromatin with reduced transcriptional activity and rare recombination events, what makes these regions extremely stable.

In diploid organism (2n=2x), chromosomes exist as homologue pairs with the same or very similar DNA sequence. The number of chromosomes is another specific feature characterized all species and distribution of DNA among the chromosomes within one species is steady. Chromosome number is highly variable value in plant kingdom. It varies from 2n=4 in daisy species *Haplopappus gracilit* and *Brachycome lineariloba* to 2n=1440 in fern *Ophioglossum reticulatum* and there is no correlation between the genome size and the number of chromosomes. In the group of chosen crops, except tetraploid *Lolium multifluorum* (2n=4x=28), all representatives show diploid genome constitutions with huge differences in genome size, but similar chromosome numbers (barley 2n=14, chickpea 2n=16, *Festuca pratensis* 2n=14).

2.2.2.2 Interphase nucleus

In nondividing cells, the chromatin is located within the nucleus distinctly bordered by nuclear membrane. Chromosomes became much more relaxed, but still exhibit high level of organization. Far from being static, interphase nucleus is highly dynamic structure allowing the events foregoing to the cell division like replication, transcription or repairing processes. Although the way, how DNA is organized and folded into the chromatin, is intensively studied, the structural and functional aspects of the process are still poorly understood.

The one of the first studies focused on arrangements of the chromosome in interphase nuclei was published in 1885 by Rabl. The experiments were performed on salamander nuclei and showed that telomeres and centromeres are placed at the opposite nucleus periphery in the spots correspond to the relic cell poles in anaphase. This organization was called according to the discoverer "Rabl arrangement" (Figure 6). Nevertheless, this formation has not been observed in all species. Furthermore, its occurrence does not correlate with the genome complexity, but is conditional to the presence of big, metaphase chromosomes. In plants, Rabl configuration is shared by such big genomes as wheat, barley or oat genome and on the contrary, genomes of the species with the lower complexity like *Arabidopsis thaliana, Zea mays* or *Sorghum bicolor* do not show this arrangement (Dong *et* Jiang, 1998). In *A. thaliana*, chromosomes exhibits the rossete-like structure, when the condensed hetechromatin of the chromosome is accumulated in the one chromocenter and euchromatin is formed into the relaxed loops, which comes out from these chromocenters (Francz *et al*, 2002; Figure 6).

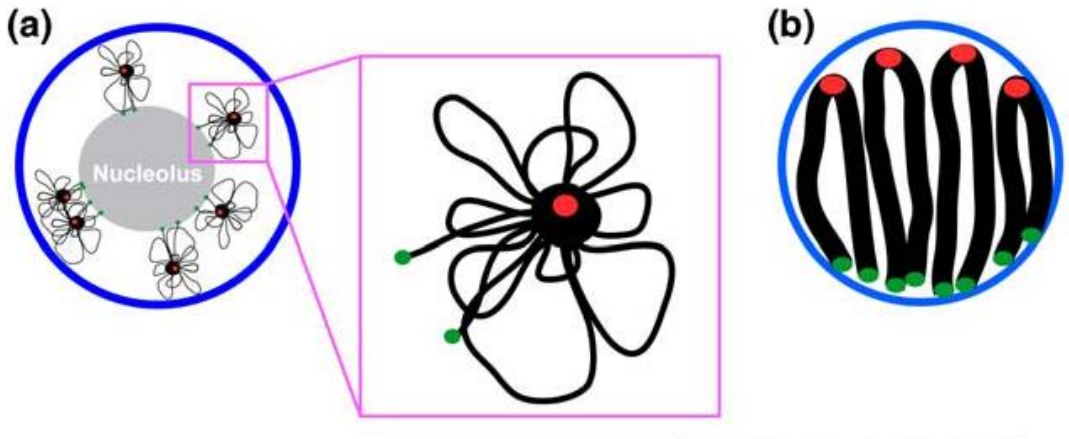


Figure 6: Models for arrangements of chromosome territories in interphase nuclei. a) rossete-like structure typical for *A. thaliana* or maize b) Rabl configuration sheared by wheat or barley (Pawlowski *et al*., 2010).

The outcomes of several subsequent studies, which have been done more than century ago, led to the conclusion that chromosome continuity is maintained throughout the cell cycle and chromatin of individual chromosomes is preserved in

discrete structural units referred to as a chromosome territory (CT; Figure 7) (Boveri *et al*., 1909). This theory was consistent with the previous finding of Laibach (1907), who observed the chromocenters in interphase nuclei of *Arabidopsis* and realized, that the number of chromocenters correspond to the number of chromosomes. Further, Heitz (1928) also support this finding, when he identified the individual chromosomes using the heterochromatin blocks in both interphase and dividing cells. Despite the results of these two studies indicated the existence of chromosomal territories, the territorial organization was clearly confirmed in 1982 by Cremer *et al*. using UV-microirradiation and pulse labelling with ^{3}H-thymidine in interphase nuclei. However, the visualization of the interphase chromosome was limited till the introduction of FISH technique (Schardin *et al*. 1985, Pinkel *et al.* 1986).

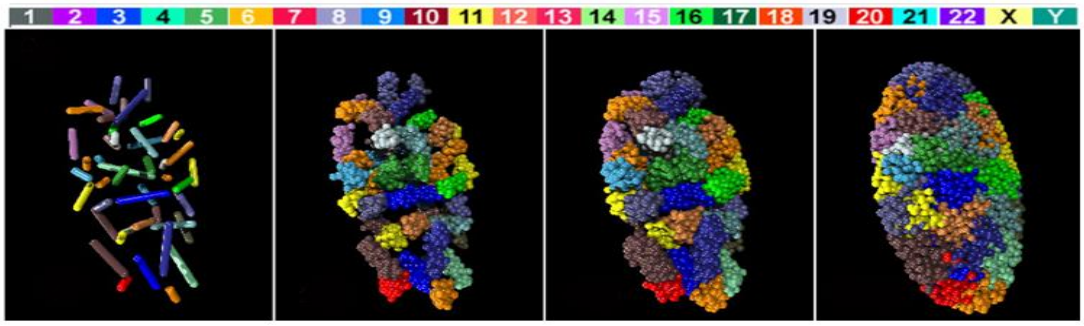


Figure 7: Simulation of complete arrangements of chromosome territories at different de-condensation levels in human fibroblast (Bolzer *et al*., 2005)

Despite the internal organization within the territories is still mostly unclear, high-resolution fluorescent microscopy and electron microscopy elucidated some aspects of the CT structure. The attention was aimed preferably on the processes related to the cell division like replication and gene transcription. It was observed that replication domains make the distinct area of CT and are maintained during the following cell cycles (Visser *et al*, 1999; Zink *et al*., 1999). Similarly, there is the evidence that chromatin conformation itself regulates the transcriptional process, so it is presumed that also higher arrangements into the chromosomal territories could play the key role in transcription. Gene transcription take place in the parts of chromatin called transcription sites, which supposed to be organized as transcription

factories and have been proposed to be located in the periphery of the nucleus. Although this model is common in many species, wheat does not share this arrangement. Regardless to their activity, the transcription sites are distributed throughout the CTs uniformly without any obvious clustering and do not reflect the gradient of gene concentration. It means that genes could migrate to the transcription site within the nucleus in some limited distance (Abranches *et al.* 1998).

2.2.3 Composition of plant nuclear DNA

DNA molecule is a substance that encodes genetic instructions in all living organisms and all characters and features forming the living organism are determined in the specific nucleotide order. The nuclear DNA consists of sequences, which can be simply classified as coding sequences (genes) and non-coding sequences without any obvious function. The era of next generation sequencing technologies triggered off the massive whole genome sequencing in various species and shed more lights on the genome composition. Sequence analysis in all genome projects confirmed the presumption that the coding sequences represent only a part of the nuclear genome and the majority of DNA is made up of regulatory sequences coordinating e.g. gene transcription or different repetitive sequences, which role has not been elucidated yet (TAGI 2000, IRGSP 2004). In barley sequencing project, there were more than 32,000 single or low copy sequences identified as genes, what represents only minor part of whole genome sequence (TIBGSC 2012). The representation of genes in genomes is variable depending on species, neither affect the genome size nor correspond to the organism complexity. The model plant *A. thaliana* with the genome consisting of 157 Mb and roughly 26,000 genes (TAGI, 2000) has more genes than evolutionary one of the most perfect being – human. Human genome has 3.2 Gbp and gene content is approximately 20,000 genes (IHGSC 2001).

Except genes, which occur in the genome generally as unique, single-copy or low-copy sequences, nuclear DNA includes further promoters, introns, gene-associated regulatory sequences and various types of abundant repetitive sequences. The repetitive DNA can make up the major part of entire genome (Heslop-Harrison and Schmidt, 2012) and constitute the large, highly variable group of sequences. Repetitive motifs are hardly classifiable into the classes, because in spite the majority of representatives befit to one of the classes, there are always some intermediate forms. Some of the repeats are positioned just in one or a few spots in the genomes

and the others are dispersed across the whole nuclear genome. The length of repetitive unit can vary from two nucleotides up to the motifs several kilobases long with abundance from hundreds to thousands and millions copies of fundamental repetitive unit.

Although the repetitive DNA is commonly referred to be redundant, it is clear that the repeats play a significant role in the genome. Being highly dynamic, the presence of repetitive sequences is reflected in genome plasticity. Repetitive DNA may change very dramatically in the meaning of the sequence or copy numbers. This way they increase the genome diversity, divergence and finally can lead up to speciation.

Considering the structural significance, repeats constitute the main part of the centromeric and telomeric regions. Although the function of these areas is conserve and stable, not so DNA sequence. Despite there exist several repeats, which are common in centromeres of several species, the sequence spanning these regions show low level of conservation.

The most of repetitive sequences belong into two main groups, tandemely repeated DNA and transposable elements. Tandem repeats were named according to specific organization of their repetitive units. Repetitive motifs are arranged one by each other in tandems counted various numbers of copies. *Satellites and microsatellites* are the most common sub-groups of tandem repeats. Very short repetitive unit (1-5bp) forming blocks up around the 150 bp long are a typical feature of microsatellites, also called simple sequence repeats (SSRs). Repetitive pattern of so small structural units very often cause the slipped-strand mispairing during the replication, what resulted in high variability of microsatellites. Microsatellites have gained considerable importance in plant genetics. Being codominant, hypervariable, reproducible and abundant, they are highly valuable markers in breeding (Parida *et al*, 2009) and robust cytogenetic markers widely used for chromosome identification (Bush *et al*., 1995; Tsijumoto *et al*., 1997).

In contrast to tandem repeats, *transposable elements* (also called mobile DNA elements) are dispersed throughout whole genome. The ability to “jump” from one locus to another and inserted in the middle of any regions is a specific feature of these elements. They are divided into two classes (I and II) depending on the mode of transposition. While class I-related mobile elements (retroelements) are dispersed using the process of reverse transcription and always results in elements

multiplication, class II related elements can either simply jump out of the original loci and reinsert randomly somewhere else without the element multiplication or can be first replicate and subsequently reinsert (Feschotte *et al*, 2002). Retrotransposons, long terminal repeat (LTR) transposons respectively and namely superfamilies *Gypsy* and *Copia*, were identified as extremely abundant class of repeats in barley genome (Table 1). They represent more than 75 % of all identified repeats. The other transposable elements as SINE, LINE or MITE were detected in negligible amount (TIBGSC, 2012).

Table 1: An overview of different types and contribution of repetitive elements identified in whole genome sequence of barley (TIBGSC, 2012).

% of all bp	WGS reads subset of 850 Mb
MDR: unique/low copy (20mer <10x)	30,63
MDR: medium/high copy (>=10 to <1000x)	33,88
MDR: very high copy (20mer >= 1000x)	35,49
Mobile Element	**81,58**
Class I: Retroelement (RXX)	**75,33**
LTR Retrotransposon (RLX)	75,10
Copia (RLC)	15,32
Gypsy (RLG)	22,29
Gypsy/Copia ratio	1,45
unclassified LTR	37,49
non-LTR Retrotransposon	0,22
Class II: DNA Transposon (DXX)	**5,60**
Retro-TE/DNA-TE ratio	13,44
DNA Transposon Superfamily (DTX)	5,39
CACTA superfamily (DTC)	5,16
hAT superfamily (DTA)	0,02
Mutator superfamily (DTM)	0,12
Tc1/Mariner superfamily (DTT)	0,02
PIF/Harbinger (DTH)	0,06
unclassified	0,02
MITE (DXX)	0,18
Helitron (DHH)	0,01
unclassified DNA transposon	0,02
Class I/Class II- ratio	**13,44**
Unclassified Element (XXX)	0,65
Simple Sequence Repeat	0,32
rRNA gene	0,29

Despite all importance and impact of repetitive sequences on the genome constitution, repetitive DNA significantly complicate the genome analysis in all aspects and therefore the scientist make a big effort to eliminate or completely get rid of the repeats in number of analysis.

2.3 CYTOGENETIC ANALYSIS OF PLANT NUCLEAR GENOMES

The phrase "genome analysis" associates all procedures and approaches leading to the deeper knowledge of the genome in the term of structure, organization or function. Recently, there is countless of various techniques allowing characterization of the genomes on different levels. Nevertheless, the study of each species generally starts with basic analysis such as to define the ploidy level and chromosome number, compile the karyotype or measure the genome size. Cytogenetic methods have been the first tool uncovering the mysteries of the cell nucleus, thanks to the availability of the microscopes for early cellular investigations.

2.3.1 Banding

Plant chromosomes were observed for the first time by Carl Wilhelm von Nägeli in 1842, when he studied cell division in pollen. However**,** except of the basic microscopic observations, banding is considered to be the first cytogenetic technique allowing deeper insight into the nuclear genome. The technique got the name according to the pattern – "bands", which appear on the chromosomes after the staining with appropriate dyes. The principle is based on different colourability of euchromatin and heterochromatin regions of chromosomes and the banding pattern is specific for each chromosome. The technique was originally described by Barbara McClintock in 1929, who thanks to simple staining of maize chromosomes in carmine dye in the combination with length measurement and chromosome morphology developed the tool, which allowed her to discriminate individual chromosomes (McClintock, 1929). This carmine-based staining procedure was found useful for chromosome characterization in various species including e.g. sorghum (Magoon, 1961) and rice (Shastry 1960).

Afterwards, several others, advanced banding variations were introduced. In, 1968, Casperson *et al*. stained the chromosomes with the fluorescent dye quinarcine resulting in the patern called Q-banding (Casperson, 1968). Alternatively, C-banding, preferably staining the constitutive heterochromatin located mainly in centomeres (Unakul and Hsu, 1972) permitted the development of karyotype of diploid barley (Figure 8, Linde-Laursen, 1975). Further modification, G-banding, produces the dark bands in late replicated regions with high content of AT nucleotides (Wurster, 1972)

and the last alternation, N-banding, is intended for identification of nucleolar organizing region (NOR), which is visualized by silver staining.

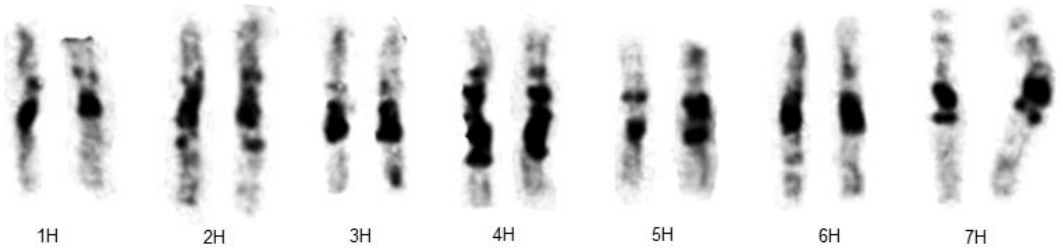


Figure 8: C-banded chromosomes of barley (alternate from Noda and Kasha, 1978)

Additionally to the chromosome pair identification, comparison of banding pattern among the individuals can discover some chromosome aberration (Misra, 1967). Nevertheless, this application is hampered by the extent of the chromosome rearrangements. It uncovers the massive abnormalities, but banding pattern is not so fine to capture slight discrepancies. The banding resolution is influenced by the chromatin condensation. Thus, prometaphase chromosomes provide more bands, but unfortunately, the band interpretation is more difficult.

In spite of the significant progress, which classical cytogenetics have undergone due to banding introduction, the limits of this technique precluded its further upswing. Even under the ideal staining condition, the ability to discriminate particular chromosomes is reduced due to the chromosome length (banana, chickpea; Ruperao *et al.*, 2014) or due to natural chromosome similarity in many species as for instance in bean (Mok *et al*., 1976).

2.3.2 Fluorescence *in situ* hybridization (FISH)

The innovative idea to localize or visualize the specific DNA sequences directly onto the chromosomes meant a huge leap forward to the advanced cytogenetic technique called *in situ* hybridization (ISH), which mixed classical cytogenetic approaches with the novel procedures from molecular biology. ISH was newly performed in 1969, but nevertheless in its original form was far from to be convenient. The probes carried the radioactive tracer and the signal was visualized using the X-rays. In addition to the demands on equipment to prevent the human

health, the technique was also time-consuming, because signal generating lasted several days (Gall *et* Pardue, 1969; John *et al*., 1969).

Although the sensitivity of radioactive labelling was extremely high, the negative aspects moved the procedure towards the alternative labelling with enzymatic or fluorescent tags. Over the decades, *in situ* hybridization with fluorescently labelled probes (FISH) expanded in such an overwhelming way that it overshadowed the other cytogenetics approaches and became a synonym for cytogenetics in general.

The principle of the ISH is based on the ability of DNA sequence of interest (probe) to hybridize with high specificity to the complementary sequence on the target DNA under the controlled conditions (Figure 9). The ion concentration and temperature represent two main factors influencing the precision, with which the probe match to the target sequences. This parameter is called hybridization specificity (stringency) and it can be claimed that the higher stringency is, the more accurate (with less mismatches allowed) probe bind to the target. Nevertheless, the several other aspect like formamide concentration, probe length or probe nucleotide composition directly affect stringency and hybridization success rate. Inappropriate setting of reaction conditions lower the specificity and increase the background of reaction (Schwarzacher and Heslop-Harrison, 2000).

Even under the optimal condition, the high noise:signal ratio is ordinarily observed in the species with large genomes, where the probes mostly contain some repetitive motif, which is dispersed in several copies through all chromosomes and produced the significant background. This unspecific signal coming from repeats can be partly reduced by adding of blocking DNA, what is substantially non-labelled genomic DNA. During the hybridization, unlabelled fragments compete with labelled ones, saturate the complementary regions on the target DNA and prevent probe cross-hybridization. The use of blocking DNA is especially helpful in experiments, where the probes originate from closely related species as exemplary in genomic *in situ* hybridization (GISH) experiments (Kopecký *et al*., 2006; 2010). Alternatively, repetitive fraction of analysed genome, C_0t – DNA, can be added to the hybridization mixture as blocking DNA with the aim to increase the signal specificity. Enrichment of blocking DNA by repetitive fractions improves the competitiveness of unlabelled fragments against dispersed signal (Harson *et al*., 1995; Kim *et al*., 2002).

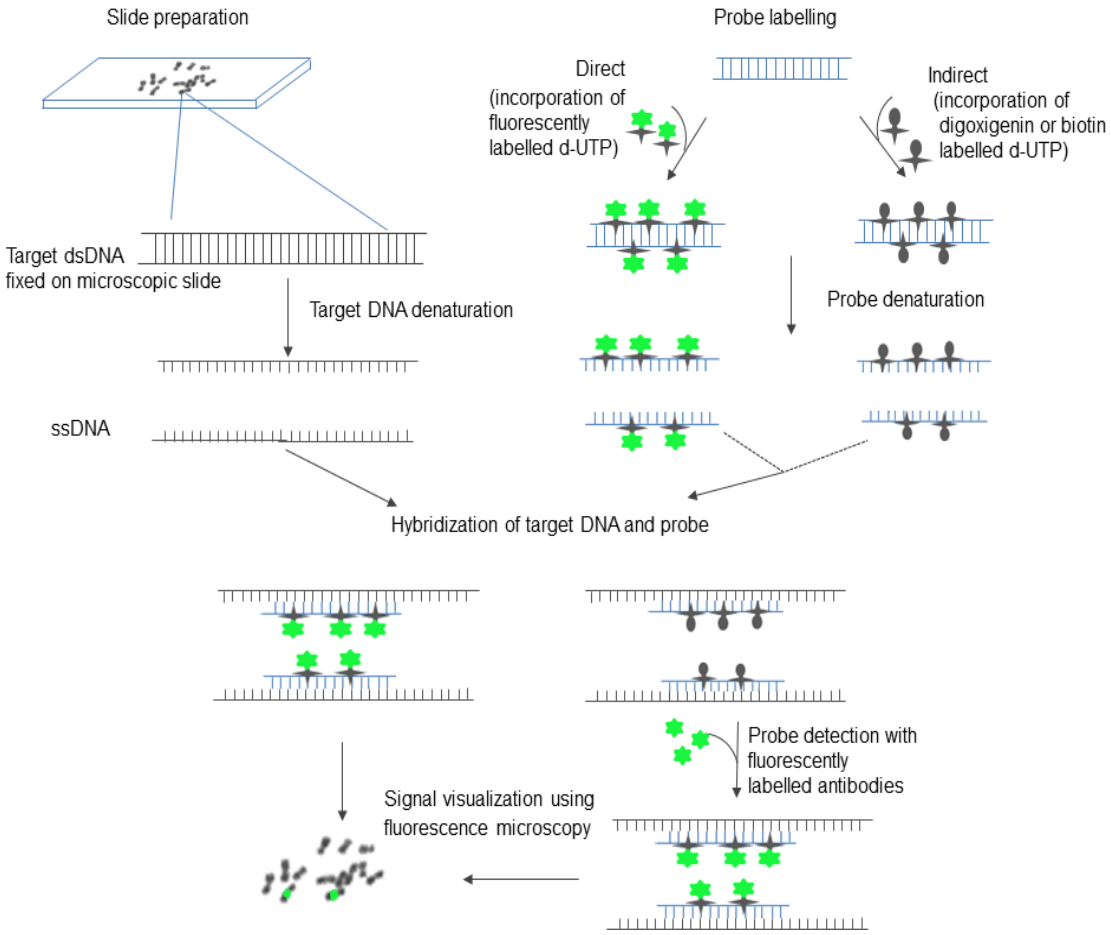


Figure 9: Scheme of fluorescence *in situ* hybridization (FISH)

Naturally, principle of *in situ* hybridization is transferable for detection of RNA molecules as RNA *in situ* hybridization. Recently, the advent of proteomics and analysis of gene expression enormously increase the attractiveness of this modification. Allowing identification of the expression pattern of individual transcripts with an incredible sensitivity, it has started new trend of *in situ* hybridization. Till date, the technique was used mainly in human research (Hamilton-Dutoit JS and Pallesen G, 2007), nevertheless the first studies started to appear also in plants (Wu and Wagner, 2012).

In general, FISH is several-steps procedure and the quality of resulting signal is formed in each of these stages. The probe quality, target DNA, type of fluorochrome, labelling efficiency and not last the detection device could be essentially reflected in signal quality. Anyway, over the 30 years, it has been successfully applied on countless species gaining the miscellaneous scientific aims as

reviewed in Jiang and Gill (2006). The versatility of the method flung out FISH among the top of the researches procedures and highlighted the ability of cytogenetics to complement or even accelerate the plant-genome research.

2.3.2.1 Probe labelling procedures

In recent labelling procedures, modified nucleotides, which include the detectable ligand, are incorporated into the probe. Modified dUTP is most frequently uses nucleotide as analogue to dTTP in process of DNA synthesis. The modified nucleotide can carry the molecule of fluorochrome and generate the signal by itself (direct labelling) or the nucleotides can be modified with hapten moieties (biotin – vitamin H or digoxigenin – steroid from *Digitalis purpurea*), which are further visualised using the fluorescently labelled antibody (indirect labelling, Langer-Safer *et al*., 1982).

Fluorochromes are chemical compounds with one functional group able to absorb energy from the excitation beam, transfer the electrons and emitted the energy excess as signal of particular wavelength. The signal is emitted in full intensity for quite short time to be detected and then gets weaker till burn out definitively. This process of fading is called photobleaching and is considered to be major disadvantage of FISH. The photobleaching effect is most obvious on short probes, where the amount of incorporated fluorochrome is limited. However, while the signal from single copy probe could be bleached after one observation, the life-time of larger probes is much longer and it is possible to observe it several times.

Fluorescein isothiokyanate (FITC) was the first fluorochrome used for FISH (Pinkel *et al*., 1986; Schwarzacher *et al*., 1989) and together with red Cyanine 3 (Cy3), they are the most widely used fluorochromes for FISH. Nowadays, trends lead toward the multicolour FISH employing additional fluorochromes like Texas Red, Cyanine 5 (Cy5), diethylaminocumarine (DEAC) in one reaction, what accelerate the experiments and widened the range of FISH applications (Tang *et al*., 2009). There are number of others fluorochromes with different Ex/Em spectrum available, thus the lacking filtersets discriminating the individual colours in similar spectrum represent the bottleneck of the multicolour applications.

There are several molecular procedures available for DNA labelling, each appropriate for different input DNA. Original labelling method, *nick translation*, was developed almost 40 years ago (Rigby *et al*., 1977) and it is only protocol suitable for

labelling of long DNA fragments like BAC and YAC clones or the whole genomic DNA. The loading of modified nucleotide into the DNA is performed in cooperation of two enzymes. While DNase I creates the gaps called "nicks" in double-stranded DNA (dsDNA), DNA polymerase subsequently replaces the nucleotides in these nicks and re-synthetized the missing part using the mixture of nucleotides (dNTPs) containing modified analogue.

In contrast to nick translation, which requires big amount of input DNA, the other methods can generate probe from several nanograms (ng) of DNA using the principle of PCR. The process of *PCR labelling* is identical to general PCR with only one difference. The reaction mixture is enriched by labelled nucleotide analogue (mostly dUTP) in appropriate ratio to unlabelled nucleotide (dTTP). Another method, *random priming*, works on the principle of DNA amplification from random hexamer primers on single-stranded (ssDNA) template using the part of DNA polymerase called Klenow fragment. The last method, *end labelling*, is suitable for labelling very small DNA fragments (up to 100 bp). Comparing to the previous methods, the end labelling, as the name suggests, provides the fluorescent tag just at the end of the fragment. The label is incorporated with the involvement of terminal deoxynucleotidyl transferase and Klenow fragment, and can be performed on both ends.

2.3.2.2 Probes for fluorescence *in situ* hybridization

Any DNA (RNA) of interest can serve as a probe regarding to the aim of experiment. The length of the probe widely differs, but substantially, each probe is finally chopped into short fragments regardless the cumulative probe length, because the hybridization efficiency was proved to be the highest with the fragments of the length in the range 200 – 600 bp (Trask, 1999).

Cloned DNA sequences, PCR fragments, oligonucleotides or genomic DNA are the most common probes for FISH. Generally, sequences cloned in some kind of vector have become highly popular due to the usage versatility. Labelled whole genomic DNA is an extreme probe example, which is widely used for GISH in evolutionary studies or in analysis of polyploidy species (Humphreys *et al*., 1995; Zwierzykowski *et al.,* 1998). PCR-based probes and oligonucleotides represent the group of probes requiring the sequence knowledge, but since the whole genome or partial sequences of various species are ordinarily available, this aspect no longer

hampers their utilization. These kinds of probes allow us to detect mainly the short motifs, which appear in genome in high-copy number.

Probes can be also classified as single- or low-copy sequence, repetitive sequence or chromosome-specific probe considering the probe specificity. *Single-copy sequence* appears in genome as unique motif in one copy usually representing genes or regulatory sequences. The uniqueness of single-copy probes gives them a notable informative value, because the hybridization of these probes results in just one specific signal, what allows its precise and undoubted localization on particular loci.

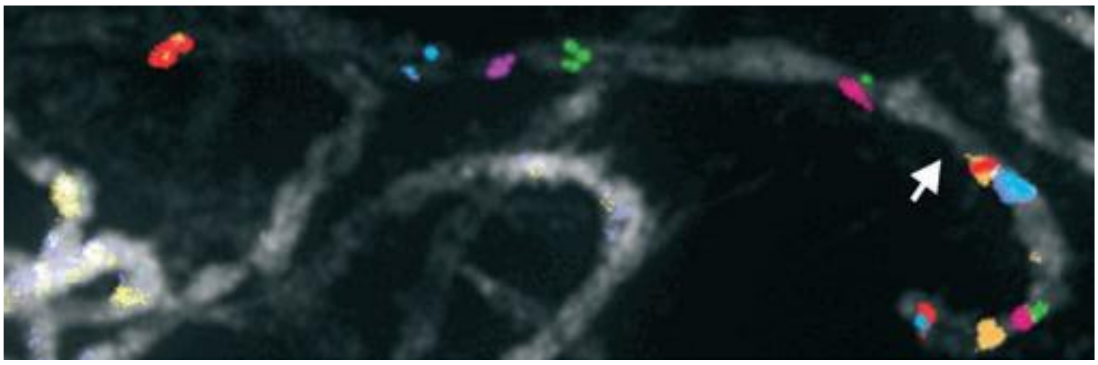


Figure 10: Multi-colour FISH on tomato pachytene chromosome 6 using BAC clones. The white arrow indicates the centromere position (Szinay *et al*., 2008)

One of the first studies using the single-copy probes was done nearly 20 years ago by Jiang *et al*., who mapped selected low-copy BAC clones on metaphase chromosomes of rice (Jiang *et al*., 1995). Except the eventual single-copy BAC clones (Suzuki *et al*. 2012; Szinay *et al*. 2008 -Figure 10), single-copy motifs are generally only a few kilobases (kb) long and the limited probe length hampered the wide utilization of these sequences. The visualization of such a small probes is on the detection limit of the method and makes the application greatly challenging. Anyway in the past years, the advances in microscopic technologies, fluorochrome chemistry and procedure itself have all contributed to improvement of the detection limits, opened the avenue to map almost any given gene-size sequence and allowed entirely use the potential, which short probes provide.

Lowering the detection limit allowed to short the probe to the length 3 kb and less. Repeats and introns-free cDNA are very useful sequence source in this length range especially in species with large genomes like wheat and barley, in which no

single-copy BAC clones have been detected. The partial cDNA libraries for both species are available (Sato *et al*., 2009, Kawaura *et al*., 2009) and can be used for wide range of application as analysing of evolutionary relationships (Danilova *et al*., 2012), chromosome identification (Danilova *et al*., 2013; Figure 11) or physical mapping. The enormous expansion of sequencing technologies and the accessibility to sequencing data mediated an alternative source for single-copy regions, especially in the species, where the cDNA clones are not available or the collections are not complete. The knowledge of physical location of even very short sequences allows us to use the surrounding sequences in whole BAC clone and *in silico* design the probe of sufficient length (Ma *et al*., 2010). This approach was successfully tested in combination with Kmasker tool for predicting of single-copy sequences in barley (Alyieva *et al*., 2013).

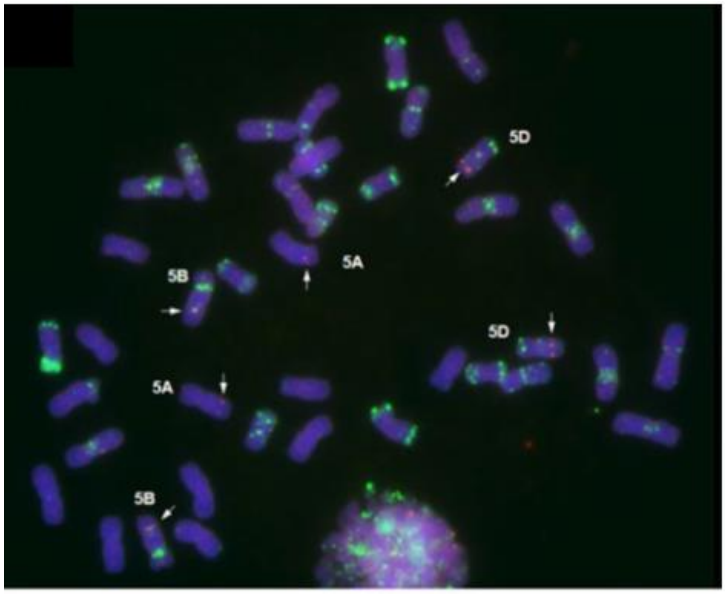


Figure 11: FISH mapping of low-copy cDNA 5L-3 (red) in combination with GAA pAs1 (green) on metaphase chromosomes in wheat. Probe provided signal on all three homeologous chromosome pairs (Danilova *et al*., 2013)

To date, 700 bp long T-DNA sequence localized in *Alium cepa* is probably the smallest probe, which has been ever detected using modified FISH protocol called Tyr-FISH (Khrustaleva *et al*., 2001). However in spite of all improvements, localization of the probes smaller than 1 kb remains very troublesome.

Repetitive sequences represent another large group of potential probes and due to their abundance in plant genomes; they are easily accessible in number of

species and are widely used in various studies. Specifically, the repetitive sequences like tandem repeats or microsatellites, whose repetitive units exist in larger clusters preferably in heterochromatic regions and their distributions show any chromosome specific pattern, can be extremely valuable FISH markers. However, despite the similarities among the plant genomes, the representation of various repetitive sequences and differences in their distribution make some repetitive motifs enormously useful in one species and conversely absolutely futile in another one. For instance, microsatellites and tandem repeats represent highly versatile probes frequently employed in Triticeae cytogenetics, where their use allows chromosome discrimination and identification (Pedersen and Landridge, 1997; Tsujimoto *et al.*, 1997), attestation of evolutionary relationship and phylogenesis (Cuadrado *et al.*, 2008; Carmona *et al*., 2013) or serves as a tool for hybrids characterization (Marín *et al.*, 2008; Szakascz and Molnar-Lang, 2010; Cseh *et al*., 2011).

On the contrary, karyotyping and cytogenetic mapping in fescues and chickpea is lagging behind cereals. Although that both repetitive groups were identified also in chickpea and fescues genomes, their application has not moved the cytogenetic state of art in these species too much ahead. The poor knowledge in this area in fescues is most probably due to the lack of whole genome sequence (WGS) information and thus no available data for repeat mining. This premise is supported by the outcomes of Kopecký *et al.* (2013), where the analysis of sequencing data of chromosome 4F revealed a few tandem repeats, which subsequently have shown chromosome specific patterns (Chmelařová *et al.*, 2014 - unpublished). In contrast, there were several tandem repeats identified in chickpea WGS, but almost all of them were clustered in the same region of two biggest chromosomes.

The probes can be also derived from the individual chromosomes. *Chromosome-specific probes* can be derived from chromosome parts or whole chromosome and they are mainly employed in evolutionary studies or in chromosome rearrangements and hybrid analysis. Due to the defined and homogenous signal, chromosome specific probes are also called “paints” and their application is known as chromosome painting (Pinkel *et al*., 1988; Guan *et al*., 1993; Chang *et al*. 1992)

Recently, it is routinely used in human diagnostic and it also represents helpful tool for comparative cytogenetics and analysis the evolutionary relationships among the animal species, but so far the use in plants is still troublesome and

application is restricted to the species with relatively small, non-complex genomes as *Arabidopsis* (Lysák *et al*., 2003), *Brachypodium* (Idziak *et al*. 2011) or *Cucumis* (Lou *et al*., 2014). Chromosome painting will be described in more details in the chapter "FISH modification and technical developments".

2.3.2.3 Cytological targets for FISH

In general, mapping of nuclear genome is laborious process with the aim to identify the mutual position and distance of individual markers and localize them on the chromosome. According to the used approach, we talk about either genetic or physical mapping. Cytogenetic mapping represents the sub-group of physical mapping and allows visualization of DNA sequences directly on chromosomes. Cytogenetic maps are compiled based on the order of localized sequences. The marker distance is quantified in micrometers (μm) or as a percentage of total chromosome length. The different nature of cytogenetic maps gives them the unique informative value, which after the integration with physical and genetic maps provides more complex information about studied genome.

In contrast to genetic mapping, cytogenetic mapping does not depend on cross-over occurrence and therefore it can dramatically improve the mapping resolution in the regions like centromeres and pericentromeres, where the recombination frequency is enormously limited and thus the genetic maps suffer from poor resolution in those areas. Due to this different nature of mentioned maps, there are huge discrepancies between the distance of two markers anchored on both genetic and physical maps.

Besides, cytogenetic mapping profits from the events naturally occurring in dividing cells. The process of cell division is generally characterized by the significant changes in chromatin structure. Because the transition between individual stages of chromatin condensation in dividing cell is continual, chromatin exists in various intermediate forms, which provides different resolving power and sensitivity. This phenomenon represents the major and highly valuable advantage in contrast to the others mapping approaches and it ensure to cytogenetics a stable and essential role in mapping process.

FISH is mainly performed on squashed or dropped *metaphase chromosomes* with tightly condense chromatin, what significantly reduces FISH resolution. Two sequences has to be theoretically located at least 0.5-1 Mb far apart to be

distinguishable as two separated signals (Trask, 1999). However, Pedersen and Linde-Laursen (1995) published study, where they localized two DNA clones on metaphase chromosomes of barley and they proved that practically, the distance between two clones has to be 10 times bigger (5 - 10 Mb) than was predicted. Nevertheless, limited resolution is compensated with the availability of metaphase chromosomes and therefore they permanently represent the most popular cytological targets.

Chromatin of *prometaphase chromosomes* offers better resolution, but their utilization is laborious and the results interpretation rather complicated. In prometaphase, level of chromatin condensation is not uniform throughout the whole chromosome and final resolution depends on the level of condensation of target sequence (Jiang *et al*., 2006). Cheng *et al*. (2002) performed FISH on rice chromosomes and reached the resolution 2 Mb, nevertheless even this improvement does not increase the attractiveness of prometaphase chromosomes and they remain rarely used object.

Pachytene chromosomes provide even higher resolution than prometaphase chromosomes. Chromatin of pachytene chromosomes can be several times enlarged compared to metaphase chromosomes. For instance, pachytene chromosomes of rice are 20 times larger than chromosomes in metaphase (Cheng *et al*., 2002; Figure 12). Furthermore, there are big differences in chromatin condensation even within the pachytene stage. While early pachytene chromosomes are characterized by more relaxed structure and allow us to distinguish the sequences 40 kb far apart, in later pachytene the resolution is remarkably lower and these two probes co-localize into the same position (Cheng *et al*., 2001). However, the early pachytene chromosomes maintain fibrillar structures more than typical chromosome and chromosome discrimination or identification is troublesome. Similarly to prometaphase chromosomes, pachytene chromatin shows variable level of condensation and resolution corresponding to heterochromatin and euchromatin regions. De Jong *et al*. (1999) described the huge differences in resolution depending on the type of target chromatin of pachytene chromosomes in tomato, where they got the resolution 1.2 Mb in heterochromatin and 120 kb in euchromatin areas. Surprisingly, these outcomes significantly differ from the results they reached in *Arabidopsis thaliana*, where the resolution in heterochromatin was only twice lower than in euchromatin.

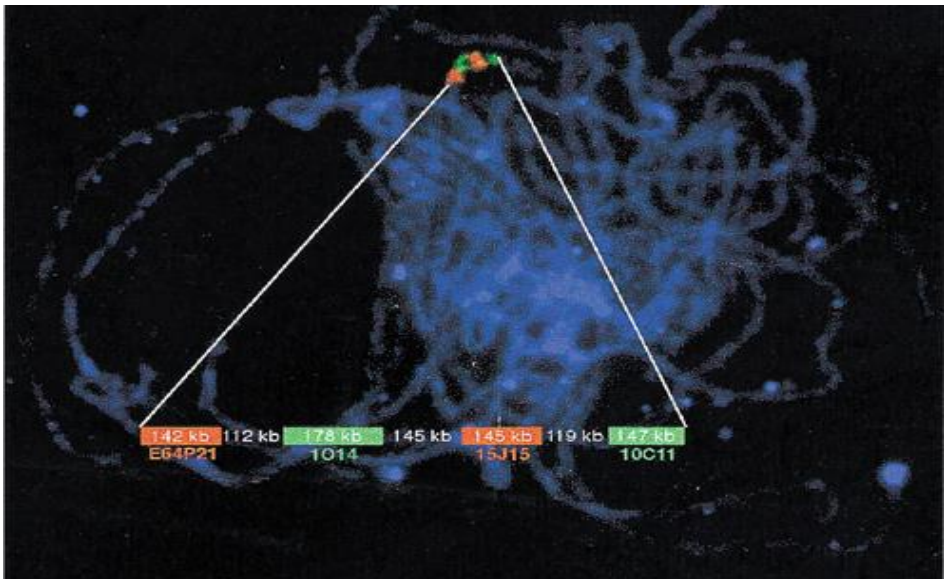


Figure 12: The proof of resolution power of pachytene chromosomes. Discrimination of four BAC clones separated by approximately 100 kb regions on early pachytene chromosomes (Cheng *et al*., 2002).

Higher resolution can be further reached only on artificially stretched chromatin. Mapping on super-stretched chromosomes was performed originally on flow-sorted chromosomes on rye, wheat, barley and chickpea (Valárik *et al*., 2004). Extended chromosomes were up to 100-times longer than in metaphase and allowed them to distinguish two probes 70 kb far apart. The level of chromosomes decondensation is not consistent within one slide, what provides one unique feature of this procedure that mutual position of two probes can be visualized with various resolutions all at once. Furthermore, in contrast to DNA fibers, stretched chromosome still preserved their integrity and preparations remain synoptical. This approach is especially suited for species with complex genomes, but the applications are restricted only to the species, whose chromosomes can be sorted using flow cytometry.

Further, FISH can be performed on *DNA fibers* stretched out from lyzed interphase nuclei (Francz *et al*., 1998; Jackson 1998). The use of DNA fibers leaped the resolution limit of FISH remarkably forward. The procedure has been originally invented by Heng *et al*. (1992), who localized sequences separated from 21 kb to 350 kb. General resolution of fiber-FISH is ranging from 1 kb to 5 kb (Ikeuchi *et al*., 1997) and allowed to detect the probe shorter than 1 kb (Weier, 2001). Due to the great resolution, fiber-FISH is extremely useful tool for fine mapping of various

sequences and sizing the gaps in physical maps (Jackson *et al*., 1998). On contrary, the localization of large probes containing significant amount of repeats is troublesome due to the unspecific signal, which complicates the orientation in fiber network. Nevertheless, this background can be partially supressed using blocking DNA (Jiang *et al,* 2006, Tomita *et al*., 2008 - Figure 13).

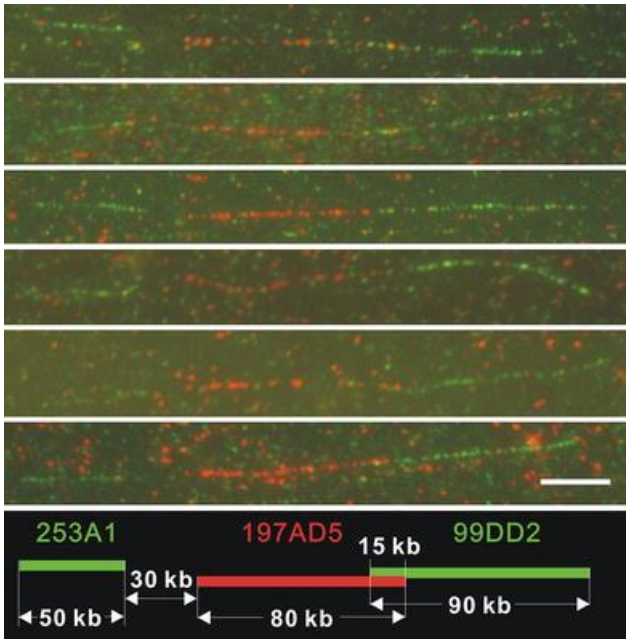


Figure 13: FISH analysis performed on extended DNA fiber of *Capsicum chinense* to size the gap between BAC contigs 253A1 (green), 197AD5 (red) and 99DD2 (green) (Tomita *et al.* 2008).

However, not only chromatin of dividing cells catches the scientist´s attention. Intact, *interphase nuclei* represent highly valuable source of information about organization of chromatin between two cell divisions and elucidate the behaviour of chromosomal domains or territories in relaxed chromatin stage (Sach *et al*., 1995; Yokota *et al*., 1995). FISH on interphase nuclei has enormous potential mainly in human medicine to diagnose various chromosomal aberrations, but the method found a solid place also in plant research (Pecinka *et al*., 2005; Berr *et al*., 2007; Schubert *et al*., 2011).

2.3.2.4 FISH modifications and technical developments

2.3.2.4.1 Chromosome painting

Chromosome painting represents an appealing cytogenetic technique with large scale of application. The method was firstly described in 1988 (Pinkel *et al*.) and the main idea of this method is generating of fluorescently labelled probes for whole chromosomes (WCPs), which can be subsequently visualized in the process of fluorescent *in situ* hybridization on metaphase chromosomes and in interphase nuclei as well. Chromosome specific probes are prepared mainly from microdissected (Guan *et al*., 1993) or flow-sorted chromosomes (Chang *et al*. 1992), but alternatively the painting probe can consist of large number of single copy probes (Lou et *al.* 2014). The technique was originally developed for analyses of human cells with the purpose to make the medicine diagnostic faster and more accurate. Nowadays, it is a robust technique widely used also in animals in phylogenetic and evolutionary studies (Ferguson-Smith 1997, Nie *et al*., 2009, 2012), in the studies of nucleus compartmentalization (Cremer *et al.,* 1993, 2008) or as a tool to analyse structural mutants (Liechty *et al.,* 1995).

Though chromosome painting is well established for human and many animal species with the wide range of genome size, its application in plants is enormously limited. After the undoubted benefits in executed studies, several groups have unsuccessfully tried to apply the technique in plants. The achievement has been confounded most probably by the enormous amounts of repetitive sequences (Schubert *et al.,* 2001), which produced significant non-specific signal background on all chromosomes.

Finally, there were a few successful examples, which were obtained with different strategies. All studies were performed on the plant species with relatively small, non-complex genomes. Shibata *et al*. (1999) based their success on the presence of specific repeats carried on sex chromosome Y of *Rumex acetosa*. Soon after, Lysak *et al.* (2001) published the novel strategy of using the large number of BAC clones in one FISH experiment in *Arabidopsis thaliana*, which finally have resulted in painting of entire chromosome. Nevertheless, the expansion of this approach is restricted only to the genomes with negligible portion of repeats, because BAC clones carrying the fragments from large genomes usually contain repetitive motives, which make BAC-FISH very worrisome.

Subsequently, chromosome painting of B chromosomes of rye and *Brachycome dichromosomatica* (Houben *et al*., 2001) and chromosomes of rice (Shishido *et al*., 2001) were performed based on chromosome specific repetitive sequences. A few years later, Kato *et a*l. (2004) came up with another attempt, where they hybridized cocktail of chromosome specific repeats in one experiment to paint the chromosomes of maize. Most recently, Lou *et al*. (2014) published the next chromosome painting approach based on hybridization of gene repeat-free regions in *Cucumis sativus* (Figure 14). In simultaneous hybridization of gene pools they painted the metaphase and pachytene chromosomes and even also the interphase nuclei. This way have identified the large miss-assembled sequence blocks of chromosome 4 in cucumber.

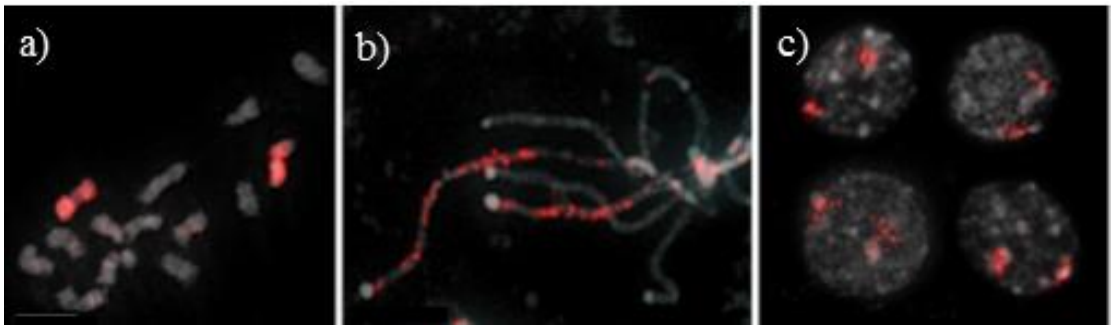


Figure 14: Chromosome painting performed on different level of chromatin condensation (a) metaphase, (b) pachytene (c) interphase of chromosome 4 in cucumber (Lou *et al*. 2014)

2.3.2.4.2 Primed *in situ* labelling (PRINS)

PRINS represents an alternative method for localization of sequences on chromosomes in primed-induced reaction, which is performed *in situ* on microscopic slide (Koch *et al*., 1989). Whole chromosomes or nuclei serve as template DNA and the probe is generated from primer pairs directly on the target loci from nucleotide mixture containing one labelled analogue. Original PRINS method involves only one cycle and in the case of direct labelling, the signal can be observed immediately after the reaction. Compared to classical FISH procedure, PRINS, avoiding the standard washing steps, is gentler and reduce the chromosome damage during the manipulation. In addition, it is versatile and very fast technique due to absence of overnight hybridization step generally performed in FISH. Furthermore, applying directly labelled nucleotides can even speed up the protocol and simultaneously

contributes to the very low unspecific background. This method was precisely worked out in human, where it is employed in physical mapping and in routine diagnostics. In plants, Kubaláková *et al*. (1997, 2000) used PRINS to detect tandem repeats in barley, pea, bean and wheat.

Sequences with lower copy number can be localized using so-called C-PRINS (cycling PRINS), where similarly to PCR the synthesis runs in more reaction cycles (Gosden *et al*, 1991, Kaczmarek *et al*., 2006) and this variant is even suitable for visualizing of unique single copy sequences. Except C-PRINS, there are a few other PRINS modifications like MULTIPRINS (multicolour PRINS; Volpi *et* Baldini, 1993), PRINSES (primed *in situ* labelling *en* suspension, Macas *et al*., 1995) with application in flow cytometry or SPRINS for detection CpG islands (Andersen *et al.* 1993).

2.3.2.4.3 Fluorescence *in situ* hybridization in suspension (FISHIS)

In spite of all methodical developments, the analysis of large size genomes remains challenging and the separation of studied genomes into the smaller units appears to be an only meaningful way leading to successful outcomes. The genome complexity could be reduced by portioning of whole genome into individual chromosome using flow cytometry. Nevertheless until date, the chromosome sorting was plagued by the demands on considerable difference in DNA content of chromosome of interest or need of special cytological stocks allowing discrimination of chromosomes or their arms, which would normally fall in composite peaks.

Innovative idea to perform FISH in suspension instead of on fixed preparation offers an enormous potential and predetermines the labelled particles to be further used for various purposes. The original idea has appeared roughly twenty years ago (Macas *et al*.1995; Pich *et al*., 1995), nevertheless without satisfactory outcomes. While they achieved some successful results, the procedure was burdened with poor reproducibility. Finally, Giorgi *et al*. (2013) developed the reliable protocol termed fluorescence *in situ* hybridization in suspension (FISHIS), which provide the stable outcomes. This protocol was invented regarding to the subsequent application in flow cytometry, where it overcome previously mentioned limitations and extend the range of possible candidate species suitable for chromosome discrimination using flow cytometry.

FISHIS is fast, wash-less procedure with almost immediate feedback in term of the quality of sorted material, because it allows us to check the fraction purity in real time. Chromosome sorting using FISHIS relies on chromosome specific hybridization pattern of the probe, which produce a different signal on individual chromosomes and just the signal intensity allows sorter to detect and separate appropriate chromosome. Giorgi *et al*. demonstrated the protocol in wheat and *Dasypyrum villosum* employing synthetic microsatellite sequences as a probe. In flow karyotype of hexaploid wheat (Chinese Spring), where originally only one peak is formed from single-type chromosome 3B, FISHIS remarkably increase the flow cytometry resolution. In total, twelve different chromosomes can be isolated directly from wild wheat with the purity levels above 90 % for each chromosome.

2.3.2.5 Other FISH variants

Although FISH represents extremely versatile method with a wide spectrum of application, the increasing scientific demands and natural need to overcome the existing limitations have stimulated development of improved variants of FISH. Recently, there are several advanced techniques based on the FISH, which introduce innovative methods of molecular biology approaches and provide the new perspective of plant cytogenetics.

However, some of them were specifically named only according to particular application or probe type. One of the most expanded FISH stream, *GISH (*genomic *in situ* hybridization*)*, uses labelled whole nuclear genome as a probe. This robust procedure is enormously valuable information source in grasses, where it is widely used for studying of hybrids (Schubert *et al*., 1998; Shi and Endo, 1999) or alien chromosome behaviour during the meiosis (Kopecký *et al*., 2009, Figure 15). However, this application fails in closely related species, whose genome homology hampers the unambiguous discrimination of one progenitor. For instance, hexaploid wheat originates from allopolyploidization of three diploid wide species, which are so closely related that the visualization of particular progenitors in wheat genome is almost impossible.

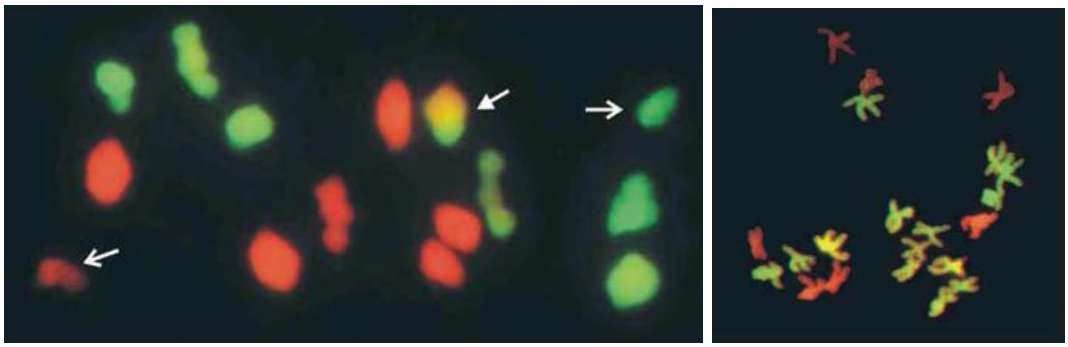


Figure 15: GISH analysis of chromosome pairing in hybrids of *L. multiflorum* (red pseudocolour) x *F. glaucescens* (green genomic probe) during the meiotic metaphase I (a) and diakinesis (b). The arrows indicate bivalent and univalents of homeologous chromosomes (Kopecký *et al*., 2009).

A few other variants aimed mainly on animal and human studies originated from the chromosome painting protocol. In 1995, Scherthan *et al*. have introduced method termed *ZOO-FISH* with high potential in evolutionary studies. ZOO-FISH uses the probes of human chromosomes to paint the chromosomes of distantly related species as whales or rodents. Further, special trend called comparative genomic *in situ* hybridization (CGH) appeared in 1992 (Kallioniemi *et al*.). They have invented the CGH in study with direct application in human diagnostic, when they have pointed and analysed the variable karyotypes in solid tumour. In CGH, both tested and control samples are labelled with different dyes and subsequently, they are hybridized against each other. The signal ratio is evaluated and the unbalanced hybridization suggests the possible changes in examined genome in contrast to the control.

Several following variants are characterized by any innovative approach, which significantly improve the existing FISH protocol. *Tyramide signal amplification (TSA)* is an enzyme-mediated detection procedure, in which first the probes is indirectly detected using the antibody conjugated with horseradish peroxidase (HRP) and subsequently, the signal is generated by adding the fluorescently labelled tyramide as enzyme substrate. The TSA method has been reported to increase the sensitivity by up to 100 - fold comparing to conventional FISH procedure and allow us to detect the probes shorter than 1kb (Krusthaleva *et al*., 2001).

3D-FISH protocol enriched FISH procedure with resolution in the third dimension, what allows more accurate signal location in the frame of intact cell

(Figure 16). The preservation of cell spatial structure is reached by mounting into the polyacrylamide (PAA) gel, in which the cells or chromosomes are afterwards labelled and analysed using the confocal microscopy (Bass *et al*., 1997, 2014; Howe *et al*., 2013).

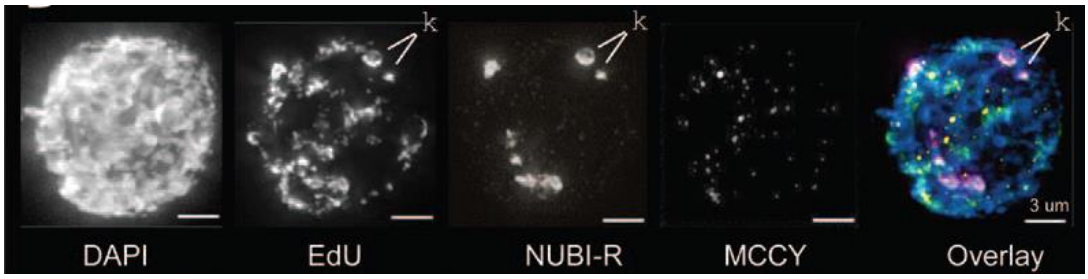


Figure 16: 3D-FISH on maize interphase nuclei visualizing the centromeres (yellow) and replicating knobs (purple) (Bass *et al*., 2014).

Simultaneous visualization of DNA and proteins can be mediated using immuno-FISH, which combine the classical FISH protocol and immunostaining. This approach provides the unique insight into the spatial relations of these two structures *in situ* and dramatically enhances the power of FISH technique. Immuno-FISH is largely employed in centromeric studies to elucidate the association of various DNA sequences and histone modifications (Zhang *et al*., 2013) or reveal the specific DNA-protein interactions (Jin *et al*., 2004)

2.3.2.6 FISH-based karyotyping and chromosome identification

A reliable system for chromosome identification represents the key foundation in cytogenetic research and unfortunately, that sort of method is missing in many species. Several decades ago, chromosome identification relied mostly on observation of chromosome morphology in combination with any banding method. Chromosomes were paired and distinguished according to the shape and the appropriate banding pattern. While this relatively simple approach was applicable on wide range of species, the precise identification required high quality preparation and detectable differences in banding pattern among the chromosomes.

With the advent of molecular techniques and developing of wide range of species-specific probes, FISH has become a robust tool for chromosome identification. Due to the abundance of repetitive elements in plant genomes, they

represent a generous source of possible probes suitable for karyotyping. Mainly signals derived from repeats or probe cocktails produce a chromosome specific pattern, what lead to unlocking the chromosomal identity.

There are the karyotyping systems based on repetitive DNA in many plant species, for instance in cucumber (Koo *et al*., 2005), vicia species (Navratilova *et al*., 2003, Figure 17) or in barley (Bush *et al*., 1995; Tsijumoto *et al*., 1997). When the signal from one probe is not powerful enough to decode all chromosomes, the involvement of more probes can lead to clear chromosome discrimination. Such cocktails have been developed in various species (Francz *et al*., 1998; Sadder *et* Weber, 2001; Lengerova *et al*., 2004 *etc*.) The combination of GAA microsatellite and Afa repeat allows the determination of every single chromosome even in huge genome of hexaploid wheat (Pedersen-Langridge, 1997).

Chromosome-specific markers, mainly BAC clones, represent an alternative way to repeat-derived probes, where appropriate combination of BAC clones generates the signals on individual chromosomes, what allows the chromosome identification (Kim *et al*., 2002). This approach can be valuable in species with smaller genomes, where the amount and representation of repetitive sequences is limited and the possible hybridization pattern is hardly traceable on the small chromosomes. Anyway, sometimes none of above mentioned techniques lead to satisfactory results and thus many species still lack this fundamental characteristic and their karyotyping remains challenging (Hřibová *et al*., 2008).

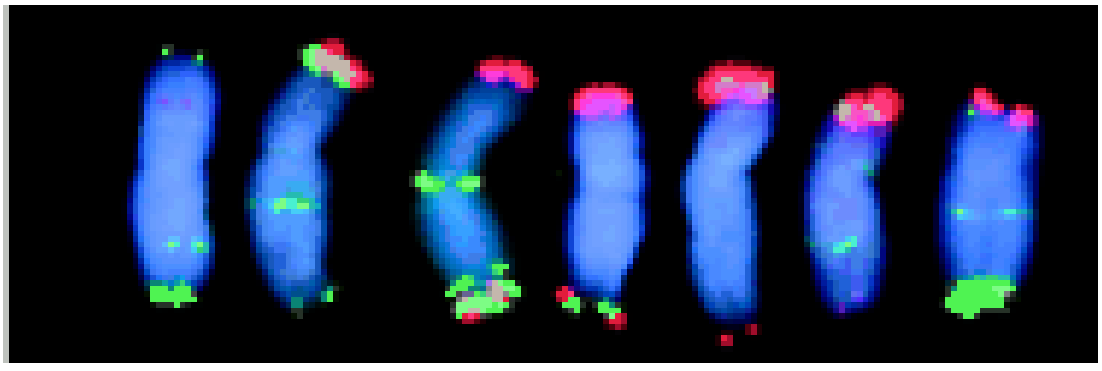


Figure 17: Identification system of individual chromosomes of *Vicia narbonensis* based on repetitive elements (Navratilová *et al*., 2003).

2.4 REFERENCES

Abranches R, Beven AF, Aragon-Alcaide L, Shaw PJ (1998) Transcription sites are not correlated with the chromosome territories in wheat nuclei. J.Cell.Biol. 143:5-12

Aliyeva L, Ma L, Schmutzer T, Stein N, Houben A (2013) Sizing the genetic and physical centromere of barley chromosomes as a prerequisite of understanding genus evolution in genus *Hordeum*. In book of abstract: 19th International chromosome conference. Bologna, Italy. p:19

Andersen CL., Koch J, Kjeldsen E (1998) CpG islands detected by self-primed *in situ* labelling (SPRINS). Chromosoma 107: 260–266

Bass HW, Marshall WF, Sedat JW, Agard DA, Cande WZ (1997) Telomeres cluster *de novo* before initiation of synapsis: a three-dimensional spatial analysis of telomere positions before and during meiotic prophase. J Cell Biol. 137: 5-18

Bass HW, Wear EE, Lee TJ, Hoffman GG, Gumber HK, Allen GC, Thompson WF, Hanley-Bowdoin L (2014) A maize root tip system to study DNA replication programmes in somatic and endocycling nuclei during plant development. J Exp Bot. 21: doi:10.1093/jxb/ert470

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Philosophical Transactions of the Royal Society in London, Series B 274: 227-274

Berr A, Schubert I (2007) Interphase chromosome arrangement in *Arabidopsis thaliana* is similar in differentiated and meristematic tissues and shows a transient mirror symmetry after nuclear division. Genetics. Jun;176(2):853-63

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth Ch, Müller S, Eils R, Cremer Ch, Speicher MR, Cremer T (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. Plos Biol 3(5):e157

Boveri T (1909) Die Blastomerenkerne von Ascaris megalocephala und die Theorie der Chromosomenindividualität. Arch Zellforsch 3: 181-268

Bush W, Martin R, Hermann RG, Hohmann U (1995) Repeated DNA sequences isolated by microdissection. I.Kayrotyping of barley (*Hordeum vulgare* L). Genome 38:1082-1090

Chang KS, Vyas RC, Deaven LL, Trujillo JM, Stass SA, Hittelman WN (1992) PCR amplification of chromosome-specific DNA isolated from flow cytometry-sorted chromosomes. Genomics 12(2):307-12

Cheng Z, Presting GG, Buell CR, Wing RA, Jiang J (2001) High resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the centromere location and the distribution of genetic recombination along chromosome 10 of rice. Genetics 157:1749-1757

Cheng Z, Buell CR, Wing RA, Jiang J (2002) Resolution of fluorescence *in situ* hybridization mapping on rice mitotic prometafase chromosomes, meiotic

pachytene chromosomes and extended DNA fibers. Chromosome Res 10: 379-387

Chmelařová A, Hřibová E, Kopecký D, Doležel J (2014) Tandem repeats as a useful tool for karyotyping in grasses. In book of abstracts:“ Molecular Breeding of Forage and Crops“ Istanbul, Turkey.

Carroll CHW and Straight AF (2006) Centromere formation: from epigenetics to self-assembly. Trends Cell Biol 16:70-78

Caspersson T, Farber S, Foley GE, Kudynowski J, Modest EJ, Simonsson E, Wagh U, Zech L (1968) Chemical differentiation along metaphase chromosomes. Exp Cell Res.49(1):219-222

Costa JM and Singh RJ (2006) Singh J. R. (2006) Chromosome mapping in barley (*Hordeum vulgare* L.). In: Singh JR: Genetic resources, chromosome engineering, and crop improvement: Cereals. CRC Press. New York. Issue 2:257-280

Cremer T, Cremer C, Baumann H, Luedtke EK, Sperling K, Teuber V, Zorn C (1982) Rabl´s model of the interphase chromosomes arrangement tested in Chinese hamster cell by premature chromosome condensation and laser UV microbeam experiments. Hum Genet. 60:46-56

Cremer T, Kurz A, Zirbel R, Dietzel S, Rinke B (1993) Role of chromosome territories in the functional compartmentalization of the cell nucleus. Cold Spring Harb Symp Quant Biol 58: 777-72

Cremer M, Grasser F, Lanctot C, Muller S, Neusser M (2008) Multicolor 3D fluorescence *in situ* hybridization for mapping interphase chromosomes. Methods Mol Biol 463: 205-239

Cseh A, Kruppa K, Molnár I, Rakszegi M, Doležel J, Molnár-Láng M (2011) Characterization of a new 4BS.7HL wheat-barley translocation line using GISH, FISH, and SSR markers and its effect on the β-glucan content of wheat. Genome 54(10):795-804

Cuadrado A, Cardoso M, Jouve N (2008) Physical organisation of simple sequence repeats (SSRs) in Triticeae: structural, functional and evolutionary implications. Cytogenet Genome Res 120(3-4):210-219

Danilova TV, Friebe B, Gill BS (2012) Single-copy gene fluorescence *in situ* hybridization and genome analysis: Acc-2 loci mark evolutionary chromosomal rearrangements in wheat. Chromosoma 121:597–611

Danilova TV, Friebe B., Gill BS (2013) Development of wheat single gene FISH mapping for analyzing homeologous relationship and chromosomal rearrangements within the Triticeae. Theor Appl Genet 127: 715-730

Doležel J, Greilhuber J, Lucretti S, Meister A, Lysák M A, Nardi L, Obermayer R (1998) Plant genome size estimation by flow cytometry: Inter-laboratory comparison. Ann. Bot. 82 (Suppl. A): 17-26

Doležel J, Bartoš J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. Cytometry A 51: 127-128.

Dong F and Jiang J (1998) Non-Rabl patterns of centromere and telomere distribution in interphase nuclei of plant cells. Chromosome Res. 6:551-558

FAO (2011) Production of chickpea by countries. Retrieved: 2013

FAOSTAT, Statistics Division (2009) "Maize, rice and wheat: area harvested, production quantity, yield".

Ferguson-Smith MA (1997) Genetic analysis by chromosome sorting and painting: phylogenetic and diagnostic applications. Eur J Hum Genet 5:253–265

Feschotte C, Jiang N, Wessler SR (2002) Plant transposable elements: where genetics meets genomics. Nat Rev Genet. 3(5):329-341

Francz P, Armstrong S, Alonso-Blanco C, Fischer TC, Torres-Ruiz RA, Jones GH (1998) Cytogenetic for the model system Arabidopsis thaliana. Plant J 13: 867-876

Francz P, de Jong H, Lysak M., Castiglione MR, Schubert I (2002) Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate. Proc Natl Acad Sci USA 99:14584-14589.

Gall JG, Pardue ML (1969) Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc. Natl. Acad. Sci. USA 63: 378-383

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(2):e57994

Gosden J, Hantratty D, Starling J, Fantes J, Mitchell A, Porteous D (1991) Oligonucleotide primed in situ DNA synthesis (PRINS) – A method for chromosome mapping, banding and investigation of sequence organization. Cytogenetic and Cell Genetics 57: 100-104

Gregory RT (2014) Animal Genome Size database. http://www.genomesize.com

Greilhuber J, Borsch T, Muller K, Worberq A, Porembski S, Barthlott W (2006) Smallest angiosperm genomes found in lentibulariaceae, with chromosomes of bacterial size. Plant Biology 8(6): 770-777

Guan XY, Trent JM, Meltzer PS (1993) Generation of band-specific painting probes from a single microdissected chromosome. Hum Mol Genet.2 (8):1117-21

Hamilton-Dutoit SJ and Pallesen G (2007) Detection of Epstain-Barr virus small RNAs in routine paraffin sections using non-isotopic RNA/RNA in situ hybridization. Histopathology 25:101-111

Harson RE, Zwick MS, Choi S, Islam-Faridi MN, McKnight TD, Wing RA, Price HJ, Stelly DM (1995) Fluorescent in situ hybridization of a bacterial artificial chromosome. Genome 38(4):646-651

Heckmann S, Schroeder-Reiter E, Kumke K, Ma L, Nagaki K, Murata M, Wanner G, Houben A (2011) Holocentric chromosomes of *Lusula elegans* are characterized by a longitudinal centromere groove, chromosome bending, and a terminal nucleolus organizer region. Cytogenet Genome Res. 134(3):220-228

Heitz E (1928) Das Heterochromatin der Moose. I. Jahrb Wiss Bot 69:762-818

Heng HHQ, Squire J, Tsui LC (1992) High-resolution mapping of mammalian genes by in situ hybridization to free chromatin. Proc Natl Acad Sci USA 89:9509–9513

Heslop-Harrison JS, Schmidt T (2012) Plant nuclear genome composition. Encyclopedia of Life Sciences eLS: John Wiley & Sons Ltd, Chichester. http://www.els.net [doi: 10.1002/9780470015902???]

Houben A, Field BL, Saunders VA (2001) Microdissection and chromosome painting of plant B chromosomes. Methods in Cell Sci 23: 115-123

Howe ES, Murphy SP, Bass HW (2013) Three-dimensional acrylamide fluorescence *in situ* hybridization for plant cells. Methods Mol Biol. 990:53-66

Hřibová E, Doleželová M, Doležel J (2008) Localization of BAC clones on mitotic chromosomes of *Musa acuminata* using fluorescence in situ hybridization. Biologia Plantarum 52:445-452

Humphreys J, Harper JA, Armstead IP, Humphreys MW (2005) Introgression mapping of genes for drought resistance transferred from *Festuca arundinacea* var. *glaucescens* into *Lolium multiflorum*. Ther. Appl. Genet. 110: 579-587

Ikeuchi T (1997): Extending the capabilities of human chromosome analysis from high-resolution banding to chromatine fiber-FISH. Human Cell 10: 121-134

International Human Genome Sequencing Consortium - IHGSC (2001) "Initial sequencing and analysis of the human genome". Nature 409 (6822): 860–921

International Rice Genome Sequencing Project - IRGSP (2004) The map-based sequence of rice genome. Nature l 436:793-800

Izdiak D, Betekhtin A, Wolny E, Lesniewska K, Wright J, Febrer M, Bevan MW, Jenkins G, Hasterok R (2011) Painting the chromosomes of *Brachypodium* – current status and the future prospects. Chromosoma 120:469-479.

Jackson SJ, Wang ML, Goodman HM, Jiang J (1998) Application of fiber-FISH in physical mapping of *Arabidopsis thaliana*. Genome 41: 566-572

Jauhar PR (2006) Cytogenetic architecture of cereal crops and their manipulation to fit human needs: Opportunities and challenges. In: Singh RJ and Jauhar PR: Genetic resources, chromosome engineering, and crop improvement: Forage Crops. CRC Press. New York. Issue 2:2-25

Jiang J, Gill B, Wang G-L, Ronald PC, Ward DC (1995) Metaphase and interphase fluorescence *in situ* hybridization mapping of the rice genome with bacterial artificial chromosomes. Proc. Natl. Acad. Sci. 92: 4487-4491

Jiang J, Gill BS (2006) Current status and the future of fluorescence *in situ* hybridization (FISH) in plant genome research. Genome 49:1057–68

Jin WW, Melo JR, Nagaki K *et al* (2004) Maize centromeres: organization and functional adaptation in the genetic background of oat. Plant Cell 16:571-581

John HA, Birnstiel ML, Jones KW (1969) RNA-DNA hybrids at the cytological level. Nature 223:582-587

de Jong JH, Fransz P, Zabel P (1999) High-resolution FISH in plants-techniques and applications. Trends Plant Sci 4:258-263

Kaczmarek A, Naganowska B, Wolko B (2007) PRINS and C-PRINS: promissing tools for physical mapping of the lupin genome. Cell and Mol Biol Letters 12:16-24

Kallioniemi A, Kallioniemi OP, Sudar Da, Rutovitz D, Gray JW, Waldman F, Pinkel D (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258:818-821

Kato A, Lamb JC, Bichler A (2004) Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. PNAS 37: 13554-13559

Kawaura K, Mochida K, Enju A, Totoki Y, Toyoda A *et al* (2009) Assessment of adaptive evolution between wheat and rice as deduced from full-length common wheat cDNA sequence data and expression patterns. BMC Genome 10:271.

Khandelwal S (1990) Chromosome evolution in the genus *Ophioglossum* L. J. Linn. Soc. Bot. 102: 205-217

Khrustaleva LI, Kik CH (2001) Localization of single-copy T-DNA insertion in transgenic shallots (*Allium cepa*) by using ultra-sensitive FISH with tyramide signal amplification. The Plant Journal 25(6): 699-707

Kim JS, Childs KL, Islam-Faridi MN, Menz MA, Klein RR, Klein PE, Price HJ, Mullet JE, Stelly DM (2002) Integrated karyotyping of sorghum by *in situ* hybridization of landed BACs. Genome 45(2):402-412

Koch J, Kolyraa S, Petersen KB, Gregersen N, Bolund L (1989) Oligonucleotide-priming methods for the chromosome-specific labelling of alpha satellite DNA *in situ*. Chromosoma 98:259-265

Koo DH, Choi HW, Cho J, Hur Y, Bang JW (2005) A high-resolution karyotype of cucumber (Cucumis sativus L. „Winter long“) revealed by C-banding, pachytene analysis and RAPD-aided fluorescence in situ hybridization. Genome 48:5345-540

Kopecký D, Loureiro L, Zwierzykowski Z, Ghesquiere M, Doležel J (2006) Genome constitution and evolution in Lolium x Festuca hybrid cultivars (Festulolium). Theor Appl Genet 113: 731-742

Kopecký D, Bartoš J, Zwierzykowski Z, Doležel J (2009) Chromosome pairing of individual genomes in tall fescue (*Festuca arundiancea* Schreb.), its progenitors

and hybrids with Italian ryegrass (*Lolium multiflorum* Lam.) Cytogenet and Genome Res 124: 170-178

Kopecký D, Havránková M., Loureiro J., Castro S, Lukaszewski AJ, Bartoš J, Kopecká J, Doležel J (2010) Physical distribution of homeologous recombination in individual chromosomes of *Festuca pratensis* in *Lolium multiflorum*. Cytogenet and Genome Res.129:162-172

Kopecký D, Martis M, Číhalíková K, Hřibová E, Vrána J, Bartoš J *et al*. (2013) Flow sorting and sequencing meadow fescue chromosome 4F. Plant Physiology 163, 1323-1337

Kubaláková M, Macas J, Doležel J (1997) Mapping of repeated DNA sequences in plant chromosomes by PRINS and C-PRINS. Theor Appl Genet 94: 758–763

Kubaláková M, Lysák MA, Vrána J, Šimková H, Číhalíková J, Doležel J (2000) Rapid identification and determination of purity of flow-sorted plant chromosomes using C-PRINS. Cytometry 41: 102–108

Laibach F (1907) Zur Frage nach der Individualität der Chromosomen in Pflanzereich. Beih Bot Centralbl 22:191-210

Langer-Safer PR, Levine M, Ward DC (1982) Immunocytological method for mapping genes on *Drosophila* polytene chromosomes. Proc. Natl. Acad. Sci. USA 79: 4381-4395

Legnerova M, Kejnosky E, Hobza R, Macas J, Grant SR, Vyskot B (2004) Multicolour FISH mapping of the dioecious plant, *Silene latifolia*. Theor Apl Genet. 108: 1193-1199

Liechty MC, Hall BK, Scalzi JM, Davis LM, Caspary WJ, Hozier JC (1995) Mouse chromosome-specific painting probes generated from microdissected chromosomes. 6(9):592-4

Linde-Laursen I. (1975) Giemsa C-banding of the chromosomes of "Emir" barley. Hereditas 81:285–9.

Lou O, Zhang Y, He Y, Li J, Jia L, Cheng Ch, Guan W, Yang S, Chen J (2014) Single-copy gene-based chromosome painting in cucumber and its application for chromosome rearrangement analysis in Cucumis. The Plant J 78:169-179

Lysak M, Fransz PF, Ali HMA, Schubert I (2001) Chromosome painting in *Arabidopsis thaliana*. The Plant J 28: 689-697

Lysak MA, Pecinka A, Schubert I (2003) Recent progress in chromosome painting of *Arabidopsis* and related species. Chromosome Res. 11: 195-204

Ma L, Vu GTH, Schubert V, Watanabe K, Stein N, Houben A, Schubert I (2010) Synteny between *Brachypodium distachyon* and *Hordeum vulgare* as revealed by FISH. Chromosome Res. 18:841-850

Macas J, Doležel J, Gualberti G, Pich U, Schubert I, Lucretti S (1995) Primed-induced labelling of pea and filed bean chromosomes *in situ* and in suspension. Biotechniques 19:402-408

Magoon ML, Shambulinguppa KG (1961) Karyomorphology of *Sorghum propinquum* and its bearing on origin of 40- chromosome sorghum. Chromosoma 12:460–465

Marín S, Martín A, Barro F. (2008) Comparative FISH mapping of two highly repetitive DNA sequences in *Hordeum chilense* (Roem. et Schult.). Genome. 51(8):580-588

McClintock B. (1929) Chromosome morphology in *Zea mays*. Science 69:629

Misra RN, Shastry SVS (1967) Pachytene analysis in Oryza. VIII. Chromosome morphology and karyotypic variation in *O. sativa*. Indian J Genet & Plant Breed 27:349–68

Mok DWS, Mok MC (1976) A modified Giemsa technique for identifying bean chromosomes. JHered 67:187–8

Navratilova A, Neumann P, Macas J (2003) Karypotype analysis of four vicia species using in situ hybridization with repetitive sequences. Ann.Bot (London) 91:921-926

Nie W, O´Brien PCM, Ng BL, Fu B, Volobouev V, Carter NP, Ferguson-Smith MA, Yang F (2009) Avian comparative genomics: reciprocal chromosome painting between domestic chicken (*Gallus gallus*) and the stone curlew (*Burhinus oedicnemus*, Charadriiformes) – An atypical species with low diploid number. Chromosome Res 17(1):99-113

Nie W,Wang J, SuW,Wang D, Tanomtong A, Perelman PL, Graphodatsky AS, Yang F (2012) Chromosomal rearrangements and karyotype evolution in carnivores revealed by chromosome painting. Heredity 108:17–27

Noda K and Kasha KJ (1978) A proposed barley karyotype revision based on C-band chromosome identification. Crop Science 18: 925-930

Parfrey LW, Lahr DJG, Katz LA (2008). "The Dynamic Nature of Eukaryotic Genomes". Molecular Biology and Evolution 25 (4): 787–94

Parida SK, Dalal V, Singh AK , Singh ND and Mohapatra T (2009) Genic non-coding microsatellites in the rice genome: characterization, marker design and use in assessing genetic and evolutionary relationships among domesticated groups, BMC Genomics 10:140

Pawlowski WP (2010) Chromosome organization and dynamics in plants. Curr Opin Plant Biol 13:645-650

Pecinka A, Kato N, Meister A, Probst AV, Schubert I, Lam E (2005) Tandem repetitive transgenes and fluorescent chromatin tags alter local interphase chromosome arrangement in *Arabidopsis thaliana*. J Cell Sci.118:3751-8

Pedersen C, Linde-Laursen I (1995) The relationship between physical and genetic distances at the Hor 1 and Hor 2 loci of barley estimated by 2-color fluorescent in situ hybridization. Theor Appl Genet 91: 941-946

Pedersen C., Langridge P. (1997) Identification of the entire chromosome complement of bread wheat by two-colour FISH. Genome 40(5): 589-593

Pellicer J, Fay MF, Leitch IJ (2010) The largest eukaryotic genome of them all? Bot J Linean Soc 164: 10-15

Pich U, Meister A, Macas J, Doležel J, Lucretti S, Schubert I (1995) Primed *in situ* hybridization facilitates flow sorting of similar size chromosomes. Plant J 7:1039-1044

Pinkel D, Straume T, Gray JW (1986) Cytogentic analysis using quantitative, high sensitivity fluorescence hybridization. Proc. Natl. Acad. Sci.U.S.A.83: 2934-2938

Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R, Lucas J, Gray J. (1988) Fluorescence *in situ* hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc Natl Acad Sci U S A.85: 9138-42

Rabl C (1885) Über Zelltheilung. Morphol Jahrb 10:214-330

Reheul D, de Cauwer B, Cougnon M (2010) The Role of Forage Crops in Multifunctional Agriculture. In: Boller B, Posselt UK, Veronesi F: The Handbook of Plant Breeding – Fodder Crops and Amenity Grasses, Springer, V. 1-12

Rigby PW, Dieckmann M, Rhodes C, Berg P (1977) Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113 (1): 237–51

Ruperao P, Chan Ch-K K, Azam S, Karafiátová *et al*. (2014) A chromosomal genomics approach to assess and validate the desi and kabuli draft chickpea genome assembly. Plant Biotech J.1-9

Saader T and Weber G (2001) Karyotype of maize (*Zea mays* L.) mitotic metaphase chromosomes as revealed by fluorescence *in situ* hybridization (FISH) with cytogenetic DNA markers. Plant Mol.Biol. Rep. 19: 117-123

Sach RK, van den Engh, Trask B, Yokota H, Hearts JE (1995) A random-walk/giant loop model for interphase chromosomes. Cold Spring Harbor Symp. Quant. Biol.58:755-765

Sato K, Shin-I T, Seki M, Shinozaki K, Yoshida H, Takeda K, Yamanaki Y, Conte M, Kohara Y (2009) Development of 5006 full-length CDNAs in barley: a tool for accessing cereal genomics resources. DNA Res 16:81–89

Schardin M, Cremer T, Hager HD, Lang M (1985) Specific staining of human chromosomes in Chinese hamster x man hybrid cell lines demonstrates interphase chromosome territories. Hum. Genet. 71:281-287

Scherthan H, Cremer T, Arnason U, Weier H-U, Lima de Faria A, Frönicke L (1994) Comparative chromosome painting discloses homologous segments in distantly related mammals. Nature Genetics 6:342-347

Schubert I, Shi F, Fuchs J and E (1998) An efficient screening for terminal deletions and translocations of barley chromosomes added to common wheat. The Plant Journal 14:489–495

Schubert I, Fransz PF, Fuchs J, de Jong H (2001) Chromosome painting in plants. Methods in Cell Science 23:57-69

Schubert V, Meister A, Tsujimoto H, Endo TR, Houben A (2011) Similar rye A and B chromosome organization in meristematic and differentiated interphase nuclei. Chromosome Res. 19(5):645-55. doi: 10.1007/s10577-011-9224-5

Schwarzacher T, Leitch AR, Bennett MD, Heslop-Harrison JS (1989) *In situ* localization on parental genomes in a wide hybrids. Ann. Bot. 64: 315-324

Schwarzacher T and Heslop-Harrison P (2000): Practical *in situ* hybridization. BIOS Scientific Publishers Limited, Oxford

Sedgwick WT (2009) A short history of science. BiblioBazaar. Carolina, US: p. 429.

Shastry SVS, Ranga Rao DR, Misra RN (1960) Pachytene analysis in Oryza. I. Chromosome morphology in Oryza sativa. IndianJ Genet & Plant Breed 20:15–21

Shi F and Endo TR (1999) Genetic induction of structural changes in barley chromosomes added to common wheat by a gametocidal chromosome derived from *Aegilops cylindrica*. Genes Genet. Syst. 74:49-54

Shishido R, Ohmido N, Fukui K (2001) Chromosome painting as a tool for rice genetics and breeding. Methods in Cell Sci 23:125-132

Shibata F, Hizume M, Kuroki Y (1999) Chromosome painting of Y chromosomes and isolation of a Y chromosome-specific repetitive sequences in the dioecious plant *Rumex acetosa*. Chromosoma 108: 266-270

Suzuki G, Ogaki Y, Hokimoto N, Xiao L *et al*. (2012) Random BAC FISH of monocot plants reveals differential distribution of repetitive DNA elements in small and large chromosome species. Plant Cell Rep (2012) 31:621–628

Szakács E, Molnár-Láng M. (2010) Identification of new winter wheat - winter barley addition lines (6HS and 7H) using fluorescence *in situ* hybridization and the stability of the whole 'Martonvásári 9 kr1' - 'Igri' addition set. Genome 53(1):35-44

Szinay D, Chang SB, Khrustaleva L, Peters S *et al* (2008) High-resolution chromosome mapping of BACs using multi-colour FISH and pooled-BAC FISH as a backbone for sequencing tomato chromosome 6. Plant J. 56 (4):627-37

Tang X, de Boer JM, van Eck JH, Bachem Ch, Visser GRF, de Jong H (2009) Assignment of genetic linkage map to diploid *Solanum tuberosum* pachytene chromosomes by BAC-FISH technology. Chromosome Res 17: 899-915

The Arabidopsis Genome Initiative- TAGI (2000) Analysis of the geonome sequence of the flowering plant *Arabidopsis thaliana*. Nature.vol 408:796-815

The International Barley Geonome Sequencing Consorcium – TIBGSC (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature. Vol 491:711-716

Thomas CA (1971) The genetic organization of chromosomes. Annual Review of Genetics 5: 237-256

Tomita R, Murai J, Miura Y, Ishihara H *et al*. (2008) Fine mapping and DNA fiber FISH analysis locates the tobamovirus resistance gene L3 of *Capsicum chinense* in a 400-kb region of R-like genes cluster embedded in highly repetitive sequences. Theor Appl Genet. 2008 Nov;117(7):1107-18

Trask B (1999) Fluorescence In situ Hybridization. In: Birren B, Green ED, Hidžer P, Klapholz S, Myers RM, Riethman H, Roskams J: Genome analysis (a laboratory manual)-Mapping genome, Volume 4,Cold Spring Harbor Laboratory Press, USA, 303-407

Tsujimoto H, Mukai Y, Akagawa K, Nagaki K, Fujigaki J, Yamamoto M, Sasakuma T (1997) Identification of individual barley chromosomes based on repetitive sequences: conservative distribution of Afa-family repetitive sequences on the chromosomes of barley and wheat. Genes Genet Syst.72(5):303-309

Unakul W, Hsu TC (1972) The C- and G-banding patterns of *Rattus norvegicus* chromosomes. J Natl Cancer Inst.49(5):1425-31

Valárik M, Bartoš J, Kovářová P, Kubaláková M, de Jong JH, Doležel J (2004) High resolution FISH on super-stretched flow-sorted chromosomes. The Plant Journal 27: 940-950

Visser AE, Jaunin F, Fakan S, Aten JA (1999) High resolution analysis of interphase chromosome domains. J.Cell.Sci. 113:2585-2593

Volpi EV, Baldini A (1993) MULTIPRINS: a method for multicolour primed *in situ* labelling. Chromosome Res 1: 257–260

Warnke S, Barker R, Jung G (2009) Ryegrass. In: Singh J. R.: Genetic resources, chromosome engineering, and crop improvement: Forage Crops. CRC Press. New York. Issue 5:275-286

Weier H-UG (2001) DNA Fiber Mapping Techniques for the Assembly of Highresolution Physical Maps. J Hystochem Cytochem 49(8): 939-948

Wu MF, Wagner D (2012) RNA *in situ* hybridization in Arabidopsis. Methods Mol Biol.883:75-86

Wurster DH (1972) Mouse chromosomes identified by trypsin-Giemsa (T-G) banding. Cytogenetics.11(5):379-87

Yokota H, van den Engh, Hearst JE, Sachs RK, Trask JB (1995) Evidence for the organisation of chromatin in megabase pair-sized loops arranged along random-walked path in the human G0/G1 interphase nuclei. J.Cell.Biol.130:3226-3232

Yong ND, Debellé F, Oldroyd GE, Geutrs R *et al*. (2011) The medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480:520-524

Zhang WL, Lee HR, Koo DH *et al* (2013) Epigenetic modification of centromeric chromatin: hypomethylation of DNA sequences in the CENH3-associated chromatin in *Arabidopsis thaliana* and maize. Plant Cell 20:25-34

Zink D, Bornfleth H, Visser A, Cremer C, Cremer T (1999) Organization of early and late replicating DNA in human chromosome territories. Exp. Cell. Res. 247:176-188

Zwierzykowski Z., Lukaszewski A.J., Lesniewska A., Naganowska B. (1998) Genomic structure of androgenic progen of pentaploid hybrids, *Festuca arundinacea* x *Lolium multiflorum*. Plant Breeding 117: 457-462

3 AIMS OF THE THESIS

I. Cytogenetic mapping of low-copy sequences on barley metaphase chromosomes

II. Chromosome painting of barley metaphase chromosomes

III. Cytogenetic analysis of forage grasses meadow fescue and ryegrass

IV. Cytogenetic analysis of two chickpea varieties Kabuli and Desi

4 RESULTS AND CONCLUSIONS

4.1 PUBLISHED PAPERS

4.1.1 Mapping nonrecombining regions in barley using multicolor FISH

(Appendix I)

4.1.2 Chromosome painting in barley – the new milestone in cytogenetics of cereals?

(Appendix II)

4.1.3 Physical distribution of homoeologous recombination in individual chromosomes of *Festuca pratensis* in *Lolium multiflorum*

(Appendix III)

4.1.4 A chromosomal genomics approach to assess and validate the *desi* and *kabuli* draft chickpea genome assemblies

(Appendix IV)

4.1.5 Flow sorting and sequencing meadow fescue chromosome 4F

(Appendix V)

4.1.1 Mapping nonrecombining regions in barley using multicolor FISH

Karafiátová M., Bartoš J., Kopecký D., Ma L., Houben A., Stein N.,

Doležel J.

Chromosome Research DOI:10.1007/s10577-013-9380-x, 2013

IF: 2.847

The first major part of the thesis was concentrated on clarification of the sustainability of cytogenetics as a tool for physical mapping. The need of additional approach useful in construction of physical map is highlighted in each study, where the high density genetic maps are not available and any other control for verifying of contigs assembly correctness is missing. In our study, we have tested the potential of fluorescence *in situ* hybridization on the set of 15 full-length cDNA clones previously assigned into the same genetic position on genetic map in centromeric region of barley chromosome 7H. Using the multicolour FISH we have established the position of 13 out of 15 clones on metaphase and prometaphase chromosomes of barley. Except the order of the cDNA clones, FISH results further revealed the actual size of the tested region, which represent negligible part of genetic map but equalled more than 40 % of whole chromosome length. This finding clearly pointed to distorted information coming from the regions of genetic maps plagued by poor resolution, what results in frequently discussed irregularities between genetic and physical maps. Furthermore, the regularity of observed signal was impressive. Even the probes roughly 3 kb long provided reliable signal in more than 90 % examined figures, what allows us to use coding sequences of limited length from assembled BAC or contigs as a probe and check their position and order using FISH. These outcomes provide the clear evidence of cytogenetics power and widen the range of alternative methods for physical mapping if it is needed.

4.1.2 Chromosome painting in barley – the new milestone in cytogenetics of cereals?

Karafiátová M., Mascher M., Bartoš J., Vrána J., Stein N., Doležel J.

Manuscript - prepared for submittion

The second significant part of the presented thesis was gained to develop the chromosome painting protocol applicable in the genome analysis of the species with the high level of sequence homology and relevant repeat abundance. The experiments were performed on barley as representative of the species with the complex genome. The reported protocol couple the gene-capture technology and flow-cytometry, which significantly contribute to our breakthrough results. Highly specific probe and unique blocking DNA prepared from sorted chromosomes represent key features of described protocol. While the probe mainly consists of coding sequences extracted from barley chromosome 1H using microarray designed for exome capture, the blocking DNA was prepared from the fraction of six remaining chromosomes sorted by flow cytometry. Using these two essential compounds we successfully performed chromosome painting of barley chromosome 1H. The probe generated distinct and unambiguous signal on both homologue chromosome and only negligible background on the entire genome. Our experiments bring the first proof that so far unsuccessful chromosome painting in cereals is possible and this study means the first step on the long way to chromosome painting on interphase nuclei, what could provide valuable knowledge about the spatial chromatin organization in non-dividing cells.

4.1.3 Physical distribution of homoeologous recombination in individual chromosomes of *Festuca pratensis* in *Lolium multiflorum*

Kopecký D., Havránková M., Castro J., Lukaszewski A.J., Bartoš J., Kopecká J., Doležel J.

Cytogenetic and Genome Research DOI:10.1159/000313379, 2010

IF: 1.885

Recombination events occurring between the homologous chromosomes represent the natural an essential source of variability in generatively reproducing individuals. Here, we reported on the occurrence of the homoeologous recombination in unique *Festuca* x *Lolium* hybrid system. We analysed the distribution of crossing-overs between the introgressed chromosome of donor (*F. pratensis*) and homoeologous chromosome in the genome of acceptor species (*L. multiflorum*) in complete set of substitution lines. The results revealed that recombination events do not show the balanced distribution along the whole chromosome length. We identified several regions called "hot spots, where the frequency of crossing-over was significantly higher and on the contrary, there were several "cold spots" areas with highly reduced level of recombination. Nevertheless, despite the unequal distribution of recombination, we proved that almost any segment of *F. pratensis* can be introgressed into the genome of *L. multiflorum* and clarification of that principle significantly contributes to introgression breeding in these species.

4.1.4 A chromosomal genomics approach to assess and validate the *desi* and *kabuli* draft chickpea genome assemblies

Ruperao P., Chan Ch.-K.K., Azam S., Karafiátová M., Hayashi S., Čížková J., Saxena R.K., Song Ch., Vrána J., Chitikineni A., Visendi P., Gaur P.M., Milán T., Singh K.B., Taran B., Wang J., Batley J., Doležel J., Varshney R.K., Edwards D.

Plant Biotechnology Journal, DOI:10.1111/pbi.12182, 2014

IF: 5.181

There are two varieties of chickpea (*Cicer arietinum* L.), desi and kabuli, available on the market. According to the previous studies, both types have shown the high level of similarity, nevertheless the recent results disclose the differencies between the kabuli and desi genomes, which are more signifficant than was expected. In our work, we performed the complex comparative study of both genomes. We updated the genome size values and complete the fundamental cytogenetic analysis by the measurement of the length of individual chromosomes in both varieties. Furthermore, we introduce the chromosomal approach for sequencing of chickpea genome. Isolation of chickpea chromosome using flow cytometry and subsequent sequencing of sorted chromosome allowed validate the assembly of the draft of whole genome sequence at the chromosome level. The comparison of our sequencing data and already existed assembled whole genome sequences revealed smaller missassembled regions in the kabuli draft and significant discrepancies in the draft sequence of desi type. This way we clearly proved that reduction of the amount of analyzed information is valuable even in the species with smaller genomes and the coupling of chromosomal genomics with the genome sequencing project improve the quality of whole genome assembly.

4.1.5 Flow sorting and sequencing meadow fescue chromosome 4F

Kopecký D., Martis M., Číhalíková K., Hřibová E., Vrána J., Bartoš J., Kopecká J., Cattonaro F., Stočes Š., Novák P., Neumann P., Macas J., Šinková H., Studer B., Asp T., Baird J.H., Navrátil P., Karafiátová M., Kubaláková M., Šafář J., Mayer K., Doležel J.

Plant Physiology Vol.163, 1323-1337, 2013

IF: 7.084

The genomic analysis of the species with the abundance of repetitive sequences is difficult due to the large number of similar sequences, which can frequently lead to fragmented assembly or even miss-assembly of particular regions. Here, we offer the approach to reduce the genome complexity of *Festuca pratensis* using the flow cytometry, which so far has not been applied in fescues. We developed the protocol for cell cycle synchronization as necessary prerequisite for isolation of chromosome 4F from the fescue genome by flow sorting. As the smallest chromosome with considerable difference in DNA content, chromosome 4F was discriminated directly from the genome of meadow fescue without need of special addition or substitution lines. Subsequently, amplified DNA of sorted chromosomes 4F was sequenced using Illumina technology and the obtained data provide the first real insight into the composition of the genome of *F. pratensis*. The sequencing data allowed identification of several new tandemely organized repeats, which were found as sturdy cytogenetic markers. Further, we estimated the gene content, established gene virtual order and performed the study of collinearity of chromosome 4F with genomes of model species. Using GemoneZipper approach we also detected the conclusive syntheny between chromosome 4 of fescues and barley. Presented study shows the possible way, how to deal with the analysis of complex genomes and we believe that this approach will significantly improve the state of art in fescue research and will lead to the acceleration of the breeding process.

4.2 PUBLISHED ABSTRACTS

4.2.1 Low-copy FISH on barley chromosomes

(Appendix VI)

4.2.2 5H/7H segmental duplication in barley revealed by cDNA-FISH

(Appendix VII)

4.2.3 Mapping of non-recombining region of barley chromosome 7H using multicolor FISH

(Appendix VIII)

4.2.4 Physical mapping of non-recombining genomic regions in barley using FISH

(Appendix IX)

4.2.5 The potential of low-copy FISH in physical mapping

(Appendix X)

4.2.6 Chromosome painting in barley – a new milestone in cytogenetics of cereals

(Appendix XI)

4.2.1 Low-copy FISH on barley chromosomes

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Book of Abstracts "Olomouc Biotech 2011 - Plant Biotechnology: Green for Good". P.49. Olomouc, 2011.

[Poster presentation]

Barley (*Hordeum vulgare* L.) belongs to important cereal crops. Cultivated barley is diploid (2n = 2x = 14) with genome size 5.1 Gbp / 1C. Large genome size and large portion of repetitive DNA (~80%) make genomic studies in this crop difficult. In barley, a number of molecular markers were developed and genetic maps were established so far. However, the position of markers on genetic map does not necessary reflect their physical position on chromosomes. Moreover, it is difficult to map genetic markers in proximal chromosomal regions due lack of recombination. One of the methods, which can be used for localization of specific DNA sequences on chromosomes is the fluorescence in situ hybridization (FISH). Currently it is possible to detect ~3 kb targets. However, localization of smaller targets is more challenging. To increase the sensitivity and resolution of FISH, flow-sorted chromosomes can be used. Chromatin of sorted chromosomes is less condensed and therefore better sensitivity and resolution can be achieved. In this project, we localized 22 out of 25 full-length cDNAs on barley metaphase and flow-sorted chromosomes using FISH. These fl-cDNAs were previously genetically mapped in centromeric region of chromosome 7H. The length of fl-cDNAs ranged from 2 to 3.5 kb. Out of these, 11 cDNAs provided fluorescent signal on chromosome 7H. Most of the clones (9) were localized in centromeric or pericentromeric regions of the chromosome; two clones were localized in the middle of the short arm of chromosome 7. Surprisingly, 19 out of 25 cDNAs were also localized on chromosome 5H. This can be explained by the existence of ancient segmental duplication between chromosomes 5H and 7H. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. CZ.1.05/2.1.00/01.0007).

4.2.2 5H/7H segmental duplication in barley revealed by cDNA-FISH

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Abstracts of the 18th International Chromosome Conference. P. 85. Manchester, 2011.

[Poster presentation]

The evolution of plant species is frequently accompanied by interspecific hybridization and polyploidization, followed by chromosome rearrangements. These changes have played a crucial role in the speciation of polyploid as well as diploid species. Ancient duplications were found in several diploid species, including barley (*Hordeum vulgare* L.). Based on the analysis of DNA sequence data, one of the segmental duplications in barley was suggested to involve chromosomes 5H and 7H. With the aim to prove this duplication, we used fluorescence *in situ* hybridization (FISH) with cDNA clones. The sensitivity and spatial resolution of the method was increased by using flow-sorted chromosome arms 5HS, 5HL, 7HS and 7HL as a target DNA and a set of 25 fl-cDNAs genetically mapped on chromosome 7H as probes. Out of these, 15 clones provided specific signals on chromosome 7H, while 19 clones could be localized on chromosome 5H. These results confirmed the ancient segmental duplication between the two chromosomes and indicated suitability of this approach to verify similar chromosome rearrangements. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. CZ.1.05/2.1.00/01.0007).

4.2.3 Mapping of nonrecombining region of barley chromosome 7H using multicolor FISH

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Abstracts of the 10th International Ph.D. Student Conference on Experimental Plant Biology. – Bulletin of the Czech Society of Experimental Plant Biology and the Physiological Section of the Slovak Botanical Society. P. 109. Brno, 2012.

[Poster presentation]

In this study, we confirmed the key importance of cytogenetic techniques for physical mapping. We selected 15 cDNAs of variable length (2.3-3.5kb) originating from centromeric and pericentromeric regions of barley chromosome 7, which were anchored in identical genetic position. The probes were prepared with nick translation and mapped using fluorescence *in situ* hybridization. Five clones were localized on the short arm and ten clones on the long arm of chromosome 7H. Surprisingly, 9 cDNA clones also provided a specific signal on chromosome 5H. This finding is consistent with a hypothesis of ancient duplication between these two chromosomes. We localized selected cDNAs on the particular chromosome arms and furthermore we determined the mutual position of 13 of the selected clones. According to the FISH results, non-recombining region spanning the centromere represents about 30 % of the chromosome 7H. Besides BAC-FISH, this approach provides alternative method for overcoming the common problem of mapping in non-recombing regions. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, the European Regional Development Fund (Operational Programme Research and Development for Innovations No. CZ.1.05/2.1.00/01.0007) and IGA UP PrF/2012/001.

4.2.4 Physical mapping of nonrecombining genomic regions in barley using FISH

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Abstracts of the International Conference "Plant and Animal Genome XXI". P. 218. Sherago International, Inc., San Diego, 2013

[Poster presentation]

Development of ready to sequence physical maps relies on establishment of contigs of overlapping BAC clones. The contigs are orientated and ordered along chromosomes after integration with genetic maps. Unfortunately, this approach fails in genomic regions with low recombination where genetic maps suffer from poor resolution. Mapping selected BAC clones to chromosomes using FISH is a solution, but it may not be feasible in species with excess of repetitive DNA such as barley. We demonstrate that FISH with single-copy sequences shorter than 3kb such as cDNAs is an attractive alternative. To test this option we choose 15 full length cDNAs originating from centromeric and pericentromeric regions of barley chromosome 7H. We localized five and ten cDNAs on the short and long arm of 7H, respectively. Nine cDNA clones gave FISH signals also on other chromosomes, which confirmed the presence of ancient duplications in barley. We succeeded in establishing relative order for 13 out of 15 cDNA clones and showed that non-recombining region covering the pericentromeric region represents more than 30% of barley chromosome 7H. This finding confirms the constraints of genetic mapping in anchoring physical maps of significant portions of genomes in plants with large genomes and underlines a need for alternative approaches such as FISH. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01 and Internal Grant Agency of Palacky University (Grant award No. IGA PrF/2012/001).

4.2.5 The potential of low-copy FISH in physical mapping

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Book of Abstracts "Olomouc Biotech 2013 - Plant Biotechnology: Green for Good". P.70. Olomouc, 2013.

[Poster presentation]

Recent genomic trends tend towards modern and high throughput techniques like sequencing. During the last decade, sequencing of the different genome portions or whole genomes of various species became more available and experience a boom. But as the every other technique, also sequencing has its weakness. Correct assembly of a genome sequence is difficult, especially in species with large and complex genomes like cereals have. The knowledge of sequence orientation and of the physical position are essential to assembly genome sequence correctly. We introduce fluorescence *in situ* hybridization (FISH) with single-copy probes shorter than 4kb as an attractive alternative and useful approach for physical mapping. We chose the centromeric region of barley chromosome 7H and we concentrated on the ordering of the set of full-length cDNA originating in this non-recombining region using multi-colour FISH. We successfully establish the order 13 out of 15 cDNA clones and show that region with highly reduced recombination frequency representing genetic centromere constitutes more than 30 % of the barley chromosome 7H. This finding underlines the necessity of combining the different mapping approaches to reach high resolution physical map and we proved that FISH could be one of them. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01 and Internal Grant Agency of Palacky University (Grant award No. IGA PrF/2012/001).

4.2.6 Chromosome painting in barley – a new milestone in cytogenetics of cereals

Havránková M., Knauft M., Bartoš J., Vrána J., Kubaláková M., Stein N., Doležel J.

In: Abstracts of the 19th International Chromosome Conference. P. 78. Bologna, 2013.

[Poster presentation]

Chromosome painting describes fluorescence *in situ* hybridization with chromosome-specific, composite probes to label selected chromosomes in metaphase and interphase. This procedure developed into a major tool in clinical and research molecular cytogenetics. It has been successfully applied in many animals and contributed to the identification of structural chromosome changes that accompanied their evolution and speciation. The attempts to utilize chromosome painting with composite chromosome probes in plants failed, mainly due to presence of dispersed repeats in painting probes. Here we describe a novel approach suitable for chromosome painting in plants with large genomes that are composed mainly of repetitive DNA sequences. We used barley chromosome 1H as a model. The method relies on the ability to prepare chromosome painting probe composed mainly from low-copy coding sequences. DNA for the probe preparation was isolated using gene capture from DNA of flow-sorted chromosome 1H. Illumina sequencing confirmed significant enrichment of the captured DNA for sequences specific for 1H. The use of highly specific probe in combination with a large amount of blocking DNA resulted in specific labelling barley chromosome 1H when applied to barley mitotic metaphase spreads. These results open avenues for a wider use of chromosome painting in plants with large genomes and represent the first step towards chromosome painting in interphase nuclei to study their 3D organization and its dynamics. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01 and Internal Grant Agency of Palacky University (Grant award No. IGA PrF/2012/001).

5 GENERAL CONCLUSION

The thesis was focused on analysis of crop genomes using the cytogenetic approaches in the context of modern plant genome research. Original cytogenetics has to transform a lot to keep the rate of scientific demands and uphold the position in genome research. Recent cytogenetics is not isolated discipline anymore, but the improvement and methodical advance interlink the cytogenetics with larger studies, where it allows provide deeper and complex knowledge about the structure and function of plant genomes.

In this work, we have significantly contributed to expand the knowledge about the genome structure in barley, chickpea, *Festuca pratensis* and *Lolium multiflorum* employing the various modification of fluorescence *in situ* hybridization as a robust cytogenetic technique. The versatility of the FISH was showed in five separated studies; each concentrated on different aims and applications. The projects required managing wide range of techniques directly related to cytogenetics and especially the findings reported in original research articles represent significant developments in cytogenetic field.

6 LIST OF ABBREVIATIONS

BAC	bacterial artificial chromosome
bp	base pair
cDNA	complementary DNA
CI	centromeric index
CGH	comparative genomic hybridization
CT	chromosome territory
Cy3	cyanine 3
Cy5	cyanine 5
C-PRINS	cycling primed *in situ* labelling
DEAC	diethylaminocumarine
DNA	deoxyribonucleic acid
dNTPs	deoxynucleoside triphosphates
dTTP	deoxythymidine triphosphate
dUTP	deoxyuridine triphosphate
dsDNA	double-stranded DNA
Em	emission
EST	expressed sequence tag
Ex	excitation
Gbp	gigabase pairs
GISH	genomic *in situ* hybridization
HRP	horseradish peroxidase
FISH	fluorescence *in situ* hybridization
FISHIS	fluorescence *in situ* hybridization in suspension
FITC	fluorescein isothiocyanate
ISH	*in situ* hybridization

kb	kilobase pairs
LINE	long interspersed nuclear element
LTR	long terminal repeat
Mbp	megabase pairs
MITE	miniature inverted repeat transposable element
n	haploid chromosome dosage
NOR	nucleolar organizing region
PCR	polymerase chain reaction
pg	picograms
PRINS	primed *in situ* labelling
PRINSES	primed *in situ* labelling *en* suspension
RNA	ribonucleic acid
SINE	short interspersed element
SPRINS	self-primed *in situ* labelling
ssDNA	single-stranded DNA
SSR	simple sequence repeat
T-DNA	transferred DNA
TSA	tyramide signal amplification
Tyr-FISH	tyramide-fluorescence *in situ* hybridization
UV	ultraviolet
WGS	whole genome sequence
WCP	whole chromosome probe
YAC	yeast artificial chromosome
µm	micrometer

7 LIST OF APENDICES

Published papers:

Appendix I: Mapping nonrecombining regions in barley using multicolor FISH

Appendix II: Chromosome painting in barley – the new milestone in cytogenetics of cereals?

Appendix III: Physical distribution of homoeologous recombination in individual chromosomes of *Festuca pratensis* in *Lolium multiflorum*

Appendix IV: A chromosomal genomics approach to assess and validate the *desi* and *kabuli* draft chickpea genome assemblies

Appendix V: Flow sorting and sequencing meadow fescue chromosome 4F

Poster presentation:

Appendix VI: Low-copy FISH on barley chromosomes

Appendix VII: 5H/7H segmental duplication in barley revealed by cDNA-FISH

Appendix VIII: Mapping of nonrecombining region of barley chromosome 7H using multicolor FISH

Appendix IX: Physical mapping of nonrecombining genomic regions in barley using FISH

Appendix X: The potential of low-copy FISH in physical mapping

Appendix XI: Chromosome painting in barley – a new milestone in cytogenetics of cereals

APPENDIX I

Mapping nonrecombining regions in barley using multicolour FISH

Karafiátová M., Bartoš J., Kopecký D., Ma L., Houben A., Stein N.,

Doležel J.

Chromosome Research DOI:10.1007/s10577-013-9380-x, 2013

IF: 2.847

Chromosome Res (2013) 21:739–751
DOI 10.1007/s10577-013-9380-x

Mapping nonrecombining regions in barley using multicolor FISH

M. Karafiátová · J. Bartoš · D. Kopecký · L. Ma · K. Sato · A. Houben · N. Stein · J. Doležel

Received: 23 June 2013 / Revised: 26 August 2013 / Accepted: 30 August 2013 / Published online: 12 September 2013

Abstract Fluorescence in situ hybridization (FISH) is a widely used method to localize DNA sequences on chromosomes. Out of the many uses, FISH facilitates construction of physical maps by ordering contigs of large-insert DNA clones, typically bacterial artificial chromosome (BAC) and establishing their orientation. This is important in genomic regions with low recombination frequency where genetic maps suffer from poor resolution. While BAC clones can be mapped directly by FISH in plants with small genomes, excess of repetitive DNA hampers this application in species with large genomes. Mapping single-copy sequences such as complementary DNA (cDNA) is an attractive alternative. Unfortunately, localization of single-copy sequences shorter than 10 kb remains a challenging task in plants. Here, we present a highly efficient FISH technique that enables unambiguous localization of single copy genes. We demonstrated its utility by mapping 13 out of 15 full-length cDNAs of variable length (2,127–3,400 bp), which were genetically defined to centromeric and pericentromeric regions of barley chromosome 7H. We showed that a region of 1.2 cM (0.7 %) on genetic map represented more than 40 % of the physical length of the chromosome. Surprisingly, all cDNA probes occasionally revealed hybridization signals on other chromosomes, indicating the presence of partially homologous sequences. We confirmed the order of 10 cDNA clones and suggested a different position for three cDNAs as compared to published genetic order. These results underline the need for alternative approaches such as FISH, which can resolve the order of markers in genomic regions where genetic mapping fails.

Keywords cDNA · multicolor FISH · low-copy FISH · nonrecombining regions · physical mapping · genetic mapping

Responsible Editor: Jiming Jiang

Electronic supplementary material The online version of this article (doi:10.1007/s10577-013-9380-x) contains supplementary material, which is available to authorized users.

M. Karafiátová · J. Bartoš · D. Kopecký · J. Doležel
Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Šlechtitelů 31, 783 71 Olomouc-Holice, Czech Republic

L. Ma · A. Houben · N. Stein
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Seeland, OT Gatersleben, Germany

K. Sato
Institute of Plant Science and Resources, Okayama University, Kurashiki 710-0046, Japan

J. Doležel (✉)
Institute of Experimental Botany, Šlechtitelů 31, 783 71 Olomouc-Holice, Czech Republic
e-mail: dolezel@ueb.cas.cz

Abbreviations

BAC-FISH	Bacterial artificial chromosome-fluorescence in situ hybridization
Cy3, Cy5	cyanine dyes
dUTP	Deoxyuridine triphosphate
DEAC	Diethylaminocoumarin
fl-cDNA	Full-length complementary DNA
FISH	Fluorescence in situ hybridization
FITC	Fluorescein isothiocyanate
HU	Hydroxyurea
rDNA	Ribosomal DNA

Introduction

Barley (*Hordeum vulgare* L.), a member of the Triticeae tribe within *Pooideae* subfamily of *Poaceae*, is one of major crops in temperate zone. Being diploid ($2n=2x=14$) it serves as a model for studies of complex polyploid Triticeae genomes such as bread wheat (*Triticum aestivum* L.). This is possible because of extensively conserved synteny between chromosomes of barley and wheat (Devos and Gale 1993). Recently, a physical map was published together with significant genomic sequence resources that provide access to most of the expressed barley genes—20,000 in a physical/genetic framework (Mayer et al. 2011; The International Barley Genome Sequencing Consortium 2012).

Because of the agronomic importance and suitability as a model for Triticeae genetics, cytogenetics, and more recently also genomics, barley has been a popular subject of studies on genome structure. Early works focused on chromosome identification (Tjio and Hagberg 1951) and the first karyotypes of barley were established following the introduction of Giemsa N-banding (Singh and Tsuchiya 1982) and C-banding (Linde-Laursen 1978) techniques. Introduction of DNA in situ hybridization techniques initiated a new period of molecular cytogenetics. 5S ribosomal DNA (rDNA) and 45S rDNA sequences were among the first localized using fluorescence in situ hybridization (FISH; Leicht and Heslop-Harrison 1992, 1993; Fukui et al. 1994) and offered a new way for chromosome identification. Subsequently, various DNA probes, such as repetitive sequences (Tsujimoto et al. 1997; Kato 2011), transgenes (Harwood et al. 2005), bacterial artificial chromosome (BAC) clones (Ma et al. 2010), and whole genome probes (Cao et al. 2000) were mapped to chromosomes. Recent cytogenetic studies in barley focus on chromosome rearrangements and meiotic behavior. Thus, Farré et al. (2012) characterized reciprocal translocation between chromosomes 1H and 3H in Spanish barley variety Albacete. They used rDNA and microsatellite probes for FISH to localize the translocation breakpoints at high resolution. In order to describe behavior of chromosomal domains, the progression of synapsis and the structure of the synaptonemal complex during early stages of meiosis in barley, Phillips et al. (2012) emplo-yed advanced fluorescence imaging after FISH with single locus centromeric BAC clones derived from *Brachypodium* and telomeric repeats.

Apart from cytogenetics, other methods have been used to understand barley genome organization with a hope of obtaining a more complete picture. Most of them can be grouped under the term “mapping” and their aim is to determine linear order of various landmarks in a genome. In general, there are two main mapping strategies: genetic mapping and physical mapping, to which also cytogenetic mapping belongs. Genetic (linkage) mapping is based on the number of crossover events during meiosis. It is important to note that recombination frequency is not constant along a chromosome and thus genetic distances between markers do not correspond to their physical distances (The International Brachypodium Initiative 2010; The International Barley Genome Sequencing Consortium 2012). In particular, genetic maps of regions with reduced recombination, such as centromeric and pericentromeric regions, suffer from poor resolution (Künzel et al. 2000). In barley, the integration of physical and genetic maps demonstrated that almost 50 % of physical contigs were anchored to regions that equalled less than 5–10 % of the genetic map (The International Barley Genome Sequencing Consortium 2012). As expected, the correlation between genetic and physical distances improved in telomeric regions in all seven barley chromosomes.

Different physical mapping techniques were established to compensate for deficiencies of genetic mapping by providing physical distances between landmarks on a chromosome. Ideally, genetic and physical maps are integrated, i.e., markers from

physically mapped loci are mapped genetically or vice versa. Like genetic mapping, also physical mapping has its own limitations. Development of physical maps from large insert DNA clones (e.g., bacterial artificial chromosome, BAC) is based on identification of overlapping fingerprint patterns of clones (Meyers et al. 2004). The overlapping clones make up contigs, and after ordering, the contigs constitute a clone-based physical map. However, in genomes with high portion of repetitive sequences like barley (The International Barley Genome Sequencing Consortium 2012) and in polyploid genomes like wheat (Paux et al. 2008), it may be difficult to establish a correct order of clones in contig(s). DNA clones from duplicated regions may have similar fingerprints and cause ambiguities in the assembly. Moreover, some parts of a genome may be missing in DNA libraries. Due to the limited resolution of genetic maps in pericentromeric regions, ordering of contigs of DNA clones is difficult in such regions.

An alternative option for physical mapping is optical mapping (Schwartz et al. 1993). The method was used to develop physical maps of human chromosomes 11, 12, and Y (Cai et al. 1998; Giacalone et al. 2000) and to assist in assembling and validating rice sequence data (Zhou et al. 2007). It also facilitated the physical map assembly of the maize genome with repetitive DNA content of more than 85 % (Zhou et al. 2009). In all these cases, the assembly of optical maps was done in species with genomic sequences available. To date, the method has not been applied in species with large nonsequenced genomes.

The gametocidal chromosomes of some species in the genus *Aegilops*, related to wheat, induce chromosome breakage in common wheat (Endo 1990, 2007). This property has been employed to induce deletions and translocations in barley chromosomes added to common wheat and to generate a series of dissection lines of barley chromosomes (Sakai et al. 2009; Joshi et al. 2011). The lines can be used for physical mapping using PCR-based markers and the resolution of the physical mapping depends on the number of deletion lines and sizes of deletion bins.

Radiation hybrid mapping is another approach of physical mapping (Kumar et al. 2012; Michalak de Jimenez et al. 2013), which allows ordering DNA markers in nonrecombining regions. This approach is based on radiation-induced chromosome breakage. A distance between two markers is estimated after screening the radiation hybrid panel using relevant markers and identification of lines, which lost one of the markers. Thus, the principle of this method is similar to genetic mapping, but unlike the meiotic recombination events, radiation breaks are induced randomly along the chromosome with the same probability. To date, a single radiation hybrid panel was prepared for barley (Wardrop et al. 2002). However, in comparison with the panels generated for mammals, the efficiency of the fusion of irradiated barley protoplasts and untreated recipient protoplast was low. Because of the inefficient number of the produced hybrids, the barley radiation hybrid panel did not allow the high resolution mapping.

Thus, another alternative for physical mapping is needed and FISH has been one of the methods considered. Ideally, large insert (BAC) clones from physical maps are localized on chromosomes. As DNA clones with inserts larger than 10 kb provide distinct FISH signals, BAC-FISH has been used in a number of plant species with rather small genomes. Schnabel et al. (2003) used BAC probes to integrate physical, genetic, and cytogenetic maps in *Medicago truncatula*. Tang et al. (2009) developed a set of 60 BAC clones, which identify bivalents in *Solanum tuberosum* using multicolor FISH and assigned genetically anchored BAC clones to pachytene chromosomes. This cytogenetic study also clarified some problems in the genetic map of potato. Cheng et al. (2001) presented a physical map of rice chromosome 10, which was constructed using BAC-FISH on pachytene chromosomes. In the tomato genome sequencing project BAC clone hybridization helped to scaffold the obtained genomic sequence (Tomato Genome Consortium 2012).

Unfortunately, in species with large genomes and prevalence of repetitive DNA, FISH with BAC clones typically results in hybridization signals dispersed on all chromosomes. The addition of unlabeled C_0t-1 fraction of genomic DNA into hybridization mix can reduce nonspecific hybridization (Cheng et al. 2001), but this approach is not efficient in large genomes (Janda et al. 2006). Alternatively, unique genomic sequences of large genomic fragments are in silico identified and used as FISH probes (Ma et al. 2010). Another solution could be to prepare FISH probes from low-copy coding sequences. This, however, requires the ability to localize probes shorter than 5 kb, which is challenging in plants. With a few exceptions (Fuchs and Schubert 1995; Wang et al. 2006; Pérez et al. 2009; Yang et al. 2011; Kato et al. 2005), probes shorter than 5 kb have not been localized routinely.

Here, we describe a highly effective method to localize cDNA probes shorter than 3.5 kb on mitotic metaphase and prometaphase chromosomes of barley. Using this approach, we established relative order of selected cDNA clones from the genetically defined pericentromeric region of barley chromosome 7H. However, while the clones were mapped genetically to a narrow region of about 1.2 cM, representing 0.7 % of 166 cM of 7H genetic map, the clones appear to occupy 40 % of the chromosome physical length as determined by FISH. Our observations confirm that FISH with single-copy cDNA probes could be a valuable tool to order physical map and provide a complementary approach to genetic mapping for chromosome regions with limited or no recombination in large genome species.

Materials and methods

Plant material and chromosome preparations

Seeds of barley (*Hordeum vulgare* L., 2*n*=2*x*=14) cv. Morex were germinated in a Petri dish for 2 days in the dark at 25 °C. Mitotic metaphase chromosomes were prepared from root tips synchronized according to Lysák et al. (1999) to achieve higher frequency of mitoses. Briefly, germinated seedlings were transferred into Hoagland's nutrient solution (Gamborg and Wetter 1975) containing 2 mM hydroxyurea (HU) for 18 h. Then the roots were washed in distilled water and cultured in HU-free Hoagland's solution for 5.5 h. To accumulate cells at metaphase, the roots were treated 2 h with Hoagland's solution containing 2.5 μM amiprophos-methyl. Subsequently, the root tips were fixed in ethanol/acetic acid (3:1) for a week at 37 °C. The fixed roots were stained in 1 % carmine acid solution and mitotic metaphase spreads were prepared according to Masoudi-Nejad et al. (2002). The quality of chromosome spreads was checked under the microscope and the best slides were used for FISH. Postfixation was performed according to Ma et al. (2010).

Probes for FISH

We used full-length cDNAs (fl-cDNAs) and barley 5S rDNA (containing flanking spacer regions of barley; Fukui et al. 1994) as probes for FISH. 5S rDNA was used as a marker for short arm of chromosome 7H. All fl-cDNAs were obtained from a collection of 5,006 full-length cDNA sequences of barley (Sato et al. 2009). We chose a set of 15 fl-cDNAs with a length between 2,127 and 3,400 bp (Table 1), which had been genetically mapped to the centromeric region of barley chromosome 7H. The fl-cDNAs were cloned in pFLCIII-sfi-cDNA in *Xho*i/*Sal*I and *Bam*HI cloning site (Sato et al. 2009). DNA of individual plasmids was isolated according to standard alkaline extraction protocols (Birnbiom 1983). To obtain high quality probes, we amplified cDNA sequences using PCR with T3/T7 primers. Final concentrations in 50 μl PCR mix were as follows: 1 mM PCR buffer containing 1.5 mM $MgCl_2$ (Finnzymes, Vantaa, Finland), 0.2 mM of each dNTPs (Fermentas, Glen Burnie, MD, USA), 1 mM T3/T7primers and 2U/50 μl Dynazyme^{TM} II DNA polymerase (Finnzymes). PCR was performed under following conditions: 5 min at 94 °C followed by 35 cycles of 50 s denaturation at 94 °C, 50 s annealing at 55 °C, 1.5 min extension at 72 °C, and final extension for 5 min at 72 °C. To get enough DNA, we run eight PCR reactions with each clone, mixed the PCR products, precipitated DNA by isopropanol precipitation (Fischer and Favreau 1991) and dissolved in 20 μl TE buffer. Template for 5S rDNA probe was amplified by PCR according to Fukui et al. (1994) and labeled by nick translation as described by Kato et al. (2006).

In order to achieve bright FISH signals with probes from short cDNAs, we used Texas Red-dUTP (Invitrogen, Camarillo, CA, USA) as primary fluorescent label (Kato et al. 2006). Biotin-dUTP and digoxigenin-dUTP (both from Roche, Mannheim, Germany) were used as alternative labels. 5S rDNA was labeled with diethylaminocoumarin (DEAC)-dUTP. Nick-translation was performed as described by Kato et al. (2006). The probe quality was checked on 1.5 % agarose gel. To increase probe concentration and remove unbound nucleotides, 3 μg of the probes were precipitated and purified by adding 60 μg salmon sperm DNA (Sigma-Aldrich, Schnelldorf, Germany), 157 μl 1× TE buffer (pH 7.5), 20 μl NaAc (3 M, pH 5.2) and 500 μl 96 % ethanol. After overnight precipitation at −20 °C the probes were centrifuged for 30 min at 4 °C at 14,000×*g*, rinsed in 70 % ethanol and dissolved in 10 μl 2× SSC at 37 °C overnight.

Table 1 List of clones localized using FISH

Full-length cDNA	Accession number	Length (bp)	FISH signal location	Relative chromosome position[a]
FLbaf140k15	AK248620	2,127	7HS	0.33
FLbaf67j12	AK250219	2,300	7HS	0.37
FLbaf140c21	AK252013	3,400	7HS	0.38
FLbaf104j18	AK251038	2,475	7HS	0.41
FLbaf151b16	AK252317	2,328	7HS	0.46
FLbaf169o18	AK248228	2,488	7HL	0.52
FLbaf125j04	AK251498	3,382	7HL	0.54
FLbaf54a18	AK249749	2,661	7HL	0.56
FLbaf148b24	AK252034	2,568	7HL	0.59
FLbaf24d09	AK249246	3,499	7HL	0.62
FLbaf25l12	AK249387	2,447	7HL	0.63
FLbaf129g09	AK251673	2,295	7HL	0.63
FLbaf89h06	AK250597	3,083	7HL	0.67
FLbaf107j09	AK248217	2,038	7HL	0.71
FLbaf175h04	AK252946	3,101	7HL	0.74

[a] The positions are the average values determined from 10 measurements

Fluorescence in situ hybridization and microscopy

Hybridization mixture consisting of 10 μl formamide, 5 μl 4× buffer (8× SSC, 0.4 M Tris–HCl pH 8, 0.2 M EDTA, 4 μg/μl salmon sperm DNA), 0.3 μg/μl of each probe and water up to 20 μl was applied onto slides, and the cover slides were taped up. The slides and hybridization mixture were denatured together at 80 °C for 3 min and then incubated in a moist chamber at 37 °C overnight. After hybridization, stringent washes followed in 2× SSC for 10 min, 0.1× SSC, 2 mM $MgCl_2$, and 0.1 % *v/v* Triton X-100 for 5 min, 2× SSC for 10 min (all previous washes were carried out at 42 °C) and 4× SSC at RT. The signals of Texas Red labeled probes were observed directly. Digoxigenin-labeled probes were detected using antidigoxigenin fluorescein isothiocyanate (FITC; Roche, Mannheim, Germany) and the signal was amplified using antisheep FITC (Vector Laboratories, Burlingame, USA). Biotin-labeled probes were detected with Cy3-labeled streptavidin (Invitrogen) after two rounds of amplification with biotinylated streptavidin and Cy3 streptavidin. The antibodies were applied in concentrations recommended by the manufacturers. Chromosome preparations were counterstained with 4′,6′-diamidino-2-phenylindole in Vectashield (Vector Laboratories).

The preparations were evaluated using Axio Imager Z.2 Zeiss microscope (Zeiss, Oberkochen, Germany) equipped with Cool Cube 1 (Metasystems, Altlussheim, Germany) camera and appropriate filter sets—for 4′,6′-diamidino-2-phenylindole G365 (excitation), FT395 (beam splitter), and BP445/50 (emission) (filter set 49, Zeiss); for DEAC: BP436/20 (excitation), FT455 (beam splitter), and BP480/40 (emission) (filter set 47, Zeiss); for FITC: HQ535/50 (excitation), Q505LP (beam splitter), and HQ480/40 (emission) (filter set F41-001, AHF Analysentechnik AC, Tuebingen, Germany); for Cy3: D580/30 (excitation), 560DCLP (beam splitter), and D546/10 (emission) (filter set F31-003, AHF); for Texas Red: HQ617/40 (excitation), Q593LP (beam splitter), and HQ581/10 (emission) (filter set F46-674, AHF). The capture of fluorescence signals and merging the layers were performed with ISIS software (Metasystems). The final image adjustment was done in Adobe Photoshop 6.0 and straightening of prometaphase chromosomes was performed with Image J software (Kocsis et al. 1991). Finally, the relative length of the chromosome 7H and the relative position of the particular signals were measured using ISIS software (Metasystem). The positions given are the average values determined from 10 measurements.

BLASTN against the barley genome

We used barley web BLAST (The International Barley Genome Sequencing Consortium 2012; http://webblast.ipk-gatersleben.de/barley/) to confirm the origin of cDNA sequences from chromosome 7H and investigate the distribution of homologous sequences in the barley genome. In order to link sequences to individual chromosomes/arms BLASTN (Altschul et al. 1997) analyses were performed against the databases "assembly_WGSMorex" and "sorted Chromosomes". Default BLASTN parameters were used for the analysis.

Results and discussion

With the ongoing efforts to deliver reference sequences for species with large genomes, ordering of physical map contigs in nonrecombining chromosome regions is becoming of great importance. Proper order of sequence contigs along a chromosome is important to avoid errors in the reference sequence assembly in such regions. Yet, there is no widely accepted and verified method suitable for this task. As genetic maps in these parts of chromosomes suffer from poor resolution (Künzel et al. 2000), alternative approaches have been considered, including optical mapping (Zhou et al. 2009), deletion bin mapping (Sakai et al. 2009), radiation hybrid mapping (Riera-Lizarazu et al. 2008), and cytogenetic mapping (Cheng et al. 2001; Wang et al. 2006). All these methods have already been used to some extent.

In this work, we verified the feasibility and utility of FISH to support the development of physical maps in plant species with large genomes. We chose barley as a model because it has a complex genome with the size exceeding 5 Gbp (Doležel et al. 1998) and with high content of repetitive DNA (~80 %). Moreover, it is an agronomically important species, has been used as a model for Triticeae genomics, and its BAC contig physical map became available recently (The International Barley Genome Sequencing Consortium 2012). As a response to previous difficulties with localizing large insert DNA clones in plant species with complex genomes (Janda et al. 2006) and to avoid repetitive DNA in hybridization probes, we chose to use fl-cDNAs as probes for FISH. All 15 fl-cDNAs used in the present study were selected from the centromeric region of barley chromosome 7H. Out of them, 11 were genetically mapped to six distinct positions within a narrow interval of 1.2 cM and four additional cDNAs were assigned to the same region indirectly based on synteny with model grass genomes (Mayer et al. 2011).

One disadvantage of using cDNAs as probes for FISH is their small size, ranging from 167 to 6,780 bp in barley (Sato et al. 2009). It was therefore important to choose appropriate fluorochromes for probe labeling so that their hybridization signals could be detected. After screening a variety of fluorescent dyes, we choose Texas Red. The same fluorochrome was selected by Kato et al. (2006) who tested brightness of probes labeled with different fluorochromes under various labeling conditions. The superiority of Texas Red could be explained by quantum yield (QY, ϕ), which is defined as ϕ=# photons emitted/photons absorbed. Except for Alexa 488 (QY=0.94), all dyes we tested had lower QY: QY_{FITC}=0.5, QY_{Cy3}=0.15, QY_{Cy5}=0.28 (http://home.earthlink.net/~fluorescentdyes/McNamara2007FluorophoresTable.xls), QY_{TMR}=0.71 (Wang et al. 2006) as compared to Texas Red (QY=0.94). Despite the relative high QY and matching parameters of optical filters and microscope camera, hybridization signals of Alexa-488 labeled probes were not detectable. A reason for this failure is not clear.

In order to establish relative order of cDNAs on a chromosome, more than one fluorochrome is needed for FISH. Tang et al. (2008) showed that currently available flurochromes (Cy3, Cy3.5, Cy5, FITC, and DEAC) permit multicolor FISH with directly labeled probes prepared from BAC clones, which have large inserts. In contrast, our preliminary experiments revealed that simultaneous visualization of more than one short DNA sequence is hampered by the lack of bright fluorochromes. Thus, we were able to localize simultaneously more than one cDNA probe only if one probe was directly labeled by Texas Red and the remaining probes were detected indirectly after signal amplification. However, the amplification resulted in higher non-specific signal background.

As the first step to physically map cDNA clones, which were genetically mapped to the centromeric region of barley chromosome 7H, we determined on which chromosome arm each cDNA was located. To do this we employed two-color FISH with a DEAC-labeled probe for 5S rDNA and Texas Red-labeled probe for cDNA (Fig. 1). FISH with 5S rDNA was used to identify the short arm of 7H (7HS). Despite the

short length of the cDNAs, their hybridization signals were observed in almost 90 % metaphase plates and the hybridization success did not correlate with the probe length. In many cases, the signal of a probe prepared from a cDNA clone with smaller insert was clearer than the signal observed for the longer one. These experiments revealed that five cDNA clones localized on 7HS and 10 clones on the long arm of chromosome

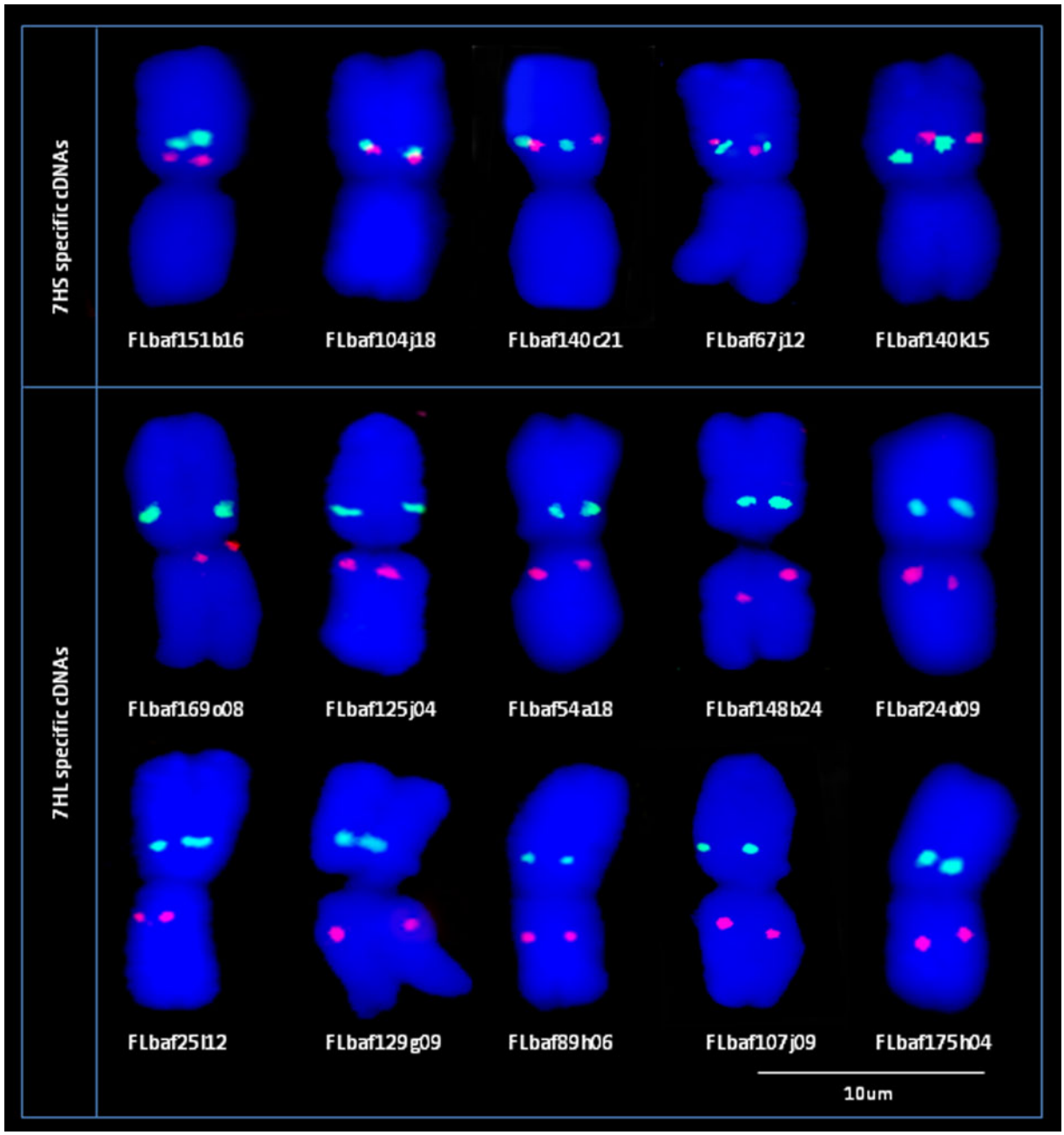


Fig. 1 Localization of fl-cDNAs by fluorescence in situ hybridization (FISH) on barley chromosome 7H. Hybridization signals of Texas Red-labeled cDNA clones are shown in *purple* and signals of DEAC-labeled 5S rDNA are shown in *green*. Chromosomes were counterstained with DAPI (*blue*). The clones are aligned according to the final order on particular arm in direction from proximal to distal part of the chromosome arm. The *upper panel* shows the signals of five clones specific for short arm of 7H (7HS) and the hybridization results of 10 fl-cDNAs specific for long arm of 7H (7HL) are illustrated in the lower panel

7H (7HL). As expected, hybridization signals of most cDNA clones were observed in centromeric and pericentromeric areas. To confirm the localization of particular fl-cDNA on chromosome 7H, we used BLASTN to compare the fl-cDNA sequences with the whole genome sequence data of all barley chromosome arms (The International Barley Genome Sequencing Consortium 2012; http://webblast.ipk-gatersleben.de/barley/). The BLASTN results are summarized in Table 2. Out of the 15 cDNAs, 11 clones matched sequences anchored to chromosome 7H. The remaining four cDNAs were linked to sequence contigs which could not be anchored to any of the seven barley chromosomes. Thus, the in silico sequence comparison supported our localization of 11 cDNA by FISH.

Rather unexpectedly, we found that the cDNAs were not localized only on chromosome 7H. Additional FISH signals were observed for all 15 cDNA clones. The additional signals were irregularly distributed on different chromosomes, they were weaker and hybridization efficiency was lower as compared to primary loci on chromosome 7H (see Electronic supplementary material 1). The occurrence of loci exhibiting additional FISH signals could be due to the presence of gene paralogs or gene fragments, which seem to be frequent in the genomes of Triticeae, including barley (Wicker et al. 2011). To support this hypothesis, we checked homology of cDNAs to sequences from other chromosomes of barley. We found significant BLASTN hits for all but two cDNA clones. All exonic regions showed reduced homology compared to 7H ranging from 70 to 95 % over shorter region (data not shown), explaining weak and irregular FISH signals of the additional loci.

After verifying the location of the 15 cDNAs on chromosome 7H, we focused on establishing their linear order by multicolor FISH. Ordering sequences originating from the same genomic region by FISH requires a template, which allows for sufficient resolution. Initially, we used mitotic metaphase stage chromosomes, but the spatial resolution was too low. This was clearly due to the highly condensed chromatin of metaphase chromosomes, which limits the resolution of FISH to only about 3 Mbp (Heng and Tsui 1998). In order to achieve higher spatial resolution, we performed FISH on prometaphase chromosomes. The highest number of simultaneously detected probes in our work was four as we used Texas Red-dUTP, biotin-dUTP, digoxigenin-dUTP for cDNA labeling, and DEAC-dUTP to label 5S rDNA probe. In case of 7HS, we succeeded in determining the positions of all five cDNAs simultaneously using a combination of three fluorochromes (Fig. 2a). According to the approximate clone position from the previous results,

Table 2 BLASTN results of homology search against whole genome shotgun sequence (Assembly3_WGSMorex; The International Barley Genome Sequencing Consortium 2012). Chromosome locations, positions on genetic map and contigs of clone-based physical map are given for particular whole genome shotgun (WGS) contig whenever this information is available

Query	Hit name	Chromosome ID	Genetic position (cM)	FPC_contig
FLbaf140k15	morex_contig_55243	7H	70.68	–
FLbaf67j12	morex_contig_135315	7H	70.68	–
FLbaf140c21	morex_contig_160232	–	–	–
FLbaf104j18	morex_contig_2557171	7H	70.43	–
FLbaf151b16	morex_contig_2548891	7H	70.68	–
FLbaf169o18	morex_contig_123002	7H	71.60	–
FLbaf111n05	morex_contig_125322	7H	70.82	contig_2729
FLbaf54a18	morex_contig_359593	–	–	–
FLbaf148b24	morex_contig_1564111	–	–	–
FLbaf24d09	morex_contig_37438	7H	70.96	contig_6301
FLbaf25l12	morex_contig_45777	7H	70.96	contig_714
FLbaf129g09	morex_contig_45746	7H	70.96	contig_714
FLbaf89h06	morex_contig_6028	–	–	–
FLbaf107j09	morex_contig_2547754	7H	71.07	–
FLbaf175h04	morex_contig_44785	7H	70.96	contig_45985

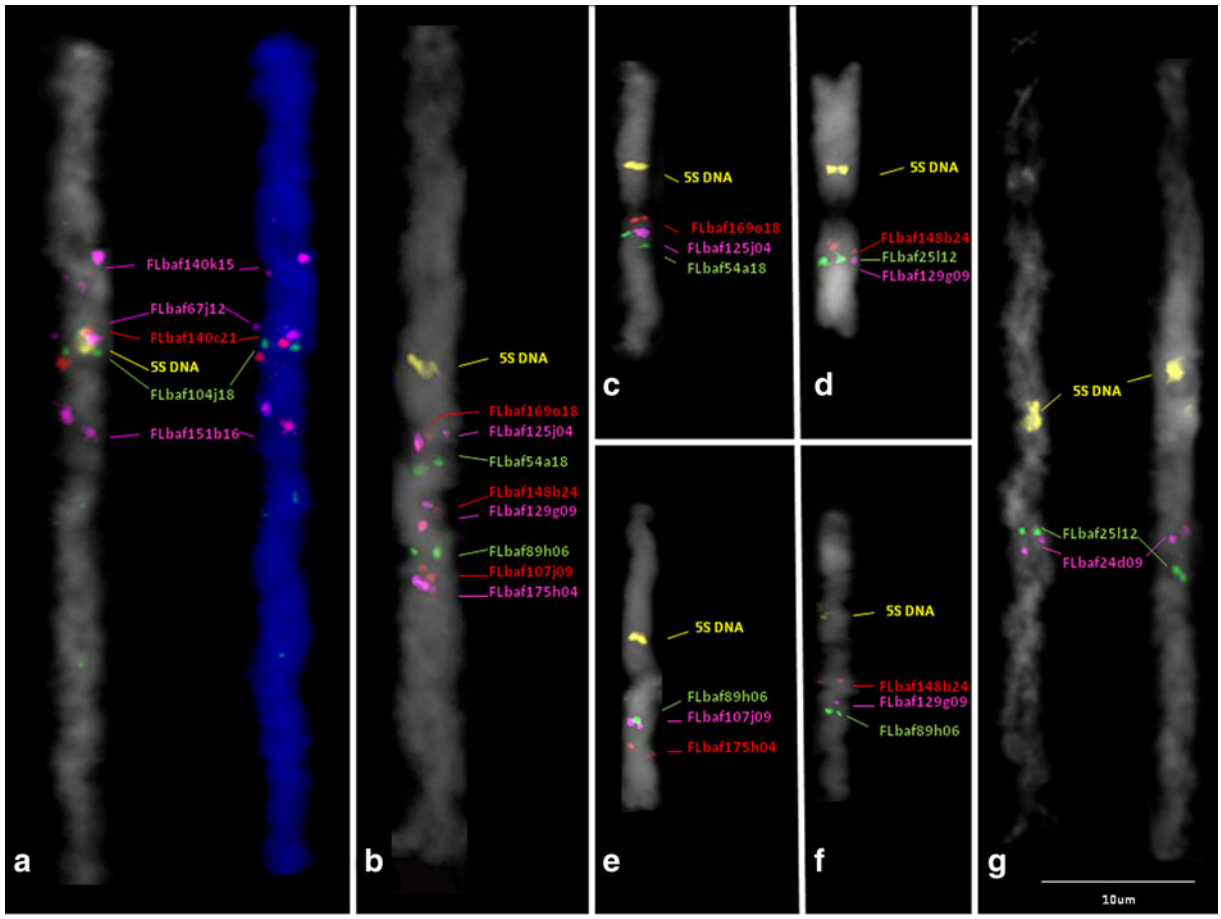


Fig. 2 Multicolor FISH with fl-cDNA probes on barley chromosome 7H at prometaphase stage. The short arm is identified using 5S rDNA (*yellow*) in all images. **a** FISH mapping of five cDNA clones on short arm of 7H (7HS). Probes (*top to bottom*): FLbaf140k15 (*purple*), FLbaf67j12 (*purple*), FLbaf140c21 (*red*), FLbaf104j18 (*green*), and FLbaf151b16 (*purple*). **b**–**g** Multicolor FISH on long arm of chromosome 7H (7HL). **b** Mapping of eight cDNA clones. *Red signals* FLbaf169o18, FLbaf148b24, and FLbaf107j09; *green signals* FLbaf54a18, FLbaf89h06; *purple signals* FLbaf125j04, FLbaf175h04, FLbaf129g09. **c** FLbaf169o18 (*red*), FLbaf125j04 (*purple*), FLbaf54a18 (*green*). **d** FLbaf148b24 (*red*), FLbaf129g09 (*purple*), FLbaf25l12 (*green*). **e** FLbaf175h04 (*red*), FLbaf107j09 (*purple*), FLbaf89h06 (*green*). **f** FLbaf89h06 (*green*), FLbaf129g09 (*purple*), FLbaf148b24 (*red*). **g** FLbaf24d09 (*purple*), FLbaf25l12 (*green*)

we selected three best distinguishable cDNAs and labeled them by Texas Red. The remaining two probes were labeled by biotin-dUTP and digoxigenin-dUTP and detected indirectly. The clone order shown in Fig. 2a was verified in three independent hybridization experiments.

The same approach could not be used to order cDNAs on 7HL as simultaneous hybridization of 10 probes was complicated and the results were ambiguous. In a single hybridization experiment, we were able to observe signals from not more than eight probes (Fig. 2b). Therefore, we established the order of cDNAs on 7HL in several independent hybridizations. To do this, we divided 10 cDNAs into three groups and determined their order following the approach used with 7HS (Fig. 2c–g). Subsequently, we ordered the three groups using the outer clones. Among the 10 clones, three clones (FLbaf24d09, FLbaf25l12, and FLbaf129g09) did not provide unequivocal FISH results and we were not able to order two of them. Signals of clones FLbaf25l12 and FLbaf129g09 overlapped (Fig. 2d). In order to explain this observation, we searched for BAC contigs containing sequences corresponding to fl-cDNA clones in the

physical map of barley (The International Barley Genome Sequencing Consortium 2012). We found that coding sequences for the clones FLbaf25l12 and FLbaf129g09, which could not be resolved by FISH, were contained in a single contig. Based on this finding, we estimated the distance between the two clones not to be more than 200 kb.

Ordering closely spaced cDNAs would require even less condensed targets than mitotic prometaphase chromosomes, such as meiotic pachytene chromosomes, which may be 7–50 times longer than their mitotic counterparts and give a spatial resolution as high as 100 kb (de Jong et al. 1999; Peterson et al. 1999). Pachytene chromosomes were used to order BAC clones in several plant species (Cheng et al. 2001; Tang et al. 2008, 2009). Alternatively, longitudinally stretched chromosomes as described by Valárik et al. (2004), which provide a spatial resolution as high as 70 kb, could be used.

cDNA clones FLbaf24d09 and FLbaf25l12 provided distinct signals, but their order varied between the experiments (Fig. 2g). This could be due to nonhomogenous chromatin condensation and irregularities in the arrangement of chromatin loops, leading occasionally to reverse orientation of FISH signals. The final order of the two clones was established as average based on multiple observations. The average values blurred the differences and resulted in order with the clone FLbaf24d09 slightly closer to the centromere. Finally, 10 cDNA clones which localized to the long arm of 7H could be localized to nine distinct positions.

In the present work, we established relative positions of FISH signals based on the average determined

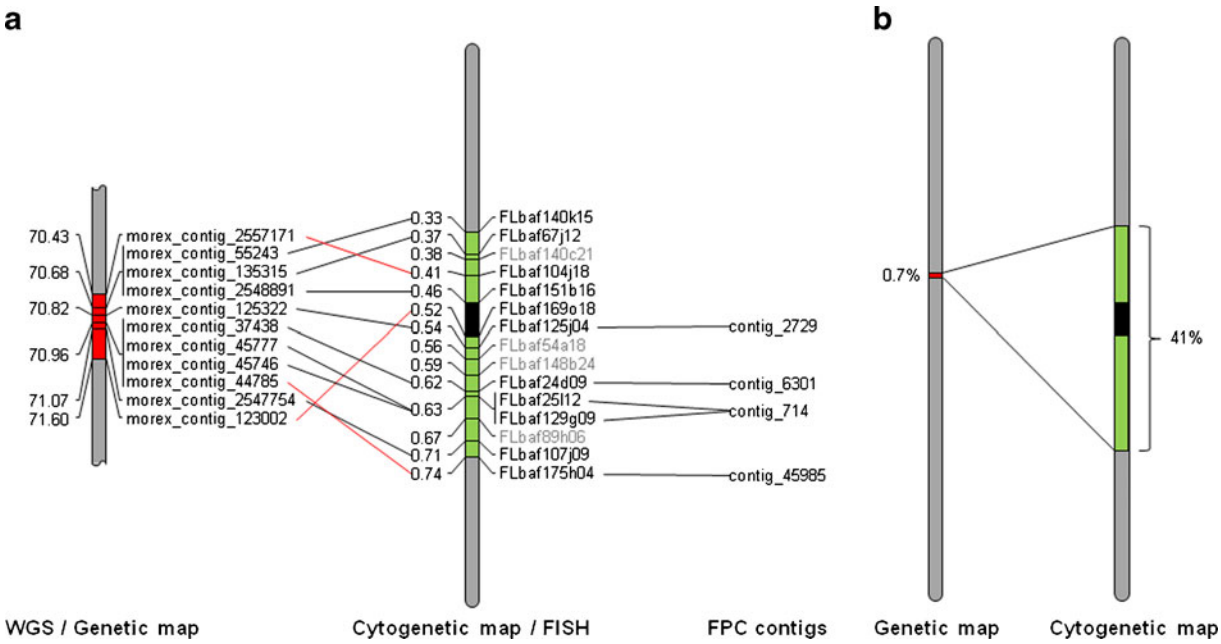


Fig. 3 Comparative map of barley chromosome 7H. **a** Genetic map (in centimorgan; a small segment shown only) displays position of 11 whole genome shotgun (WGS) contigs containing particular fl-cDNAs. The 11 markers are distributed over six distinct positions representing narrow region of only about 1.2 cM of genetic distance. Cytogenetic map represents physical distribution of fl-cDNA clones over region representing about 40 % of physical length of prometaphase chromosome 7H (the numbers represent fraction length determined as average of 10 measurements). Clones in *gray* were selected for FISH based on position in GenomeZipper (Mayer et al. 2011). The link between each pair of WGS contig (genetic map) and particular fl-cDNA (FISH map) is displayed as *black* or *red line*. *Red lines* indicate clones/markers whose position was corrected using FISH mapping on prometaphase chromosomes. FPC_contigs were linked to fl-cDNA clones whenever the position of fl-cDNA in BAC clone-based physical map is known. *Black box* represents tentative position of centromere based on localization of most proximal fl-cDNA clones on short and long arm. **b** Comparison of chromosome regions occupied by selected fl-cDNA clones in genetic (0.7 %) and cytogenetic (41 %) maps. Note several-fold expansion of a region covered with fl-cDNAs and hence resolution on cytogenetic map. *Black box* represents tentative position of centromere based on localization of most proximal fl-cDNA clones on short and long chromosome arms

from 10 independent measurements (Table 1). The final order of cDNA clones is shown in (Fig. 3a). The position of three clones did not correspond with the genetic map. However, the position and order of molecular markers could be incorrect on genetic map due to absent or false genotyping data for some individuals. Our results clearly demonstrate that the selected cDNAs, which were assigned to a narrow region of 1.2 cM (0.7 %) on genetic map, are distributed over 40 % of the physical length of chromosome 7H (Fig. 3b). The multicolor FISH on prometaphase chromosomes facilitated physical ordering of 13 out of the 15 cDNA clones on barley chromosome 7H, providing encouraging results for the use of FISH for physical mapping genome regions with limited recombination frequency.

Acknowledgements We appreciate technical advice from our colleague Marie Kubaláková. This work was supported jointly by the Czech Academy of Sciences and the German Academic Exchange Service–DAAD (grant award no. CZ07-DE12/2013-2014) by Internal Grant Agency of Palacky University in Olomouc (grant award no. IGA PrF/2012/001) and by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01).

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Birnbiom HC (1983) A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol 100:243–255

Cai W, Ling J, Irvin B, Ohler L, Rose E et al (1998) High-resolution restriction maps of bacterial artificial chromosomes constructed by optical mapping. Proc Natl Acad Sci U S A 95:3390–3395

Cao M, Sleper DA, Dong F, Jiang J (2000) Genomic in situ hybridization (GISH) reveals high chromosome pairing affinity between *Lolium perenne* and *Festuca mairei*. Genome 43:398–403

Cheng Z, Presting GG, Buell CR, Wing RA, Jiang J (2001) High-resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the centromere location and the distribution of genetic recombination along chromosomes 10 of rice. Genetics 157:1749–1757

de Jong JH, Fransz P, Zabel P (1999) High resolution FISH in plants—techniques and applications. Trends Plant Sci 4:258–263

Devos KM, Gale MD (1993) Extended genetic maps of the homeologous group-3 chromosomes of wheat, rye and barley. Theor Appl Genet 85:649–652

Doležel J, Greilhuber J, Lucretti S, Meister A, Lysak MA, Nardi L, Obermayer R (1998) Plant genome size estimation by flow-cytometry: inter-laboratory comparison. Ann Bot 82:17–26

Endo TR (1990) Gametocidal chromosomes and their induction of chromosome mutation in wheat. Jpn J Genet 65:135–152

Endo TR (2007) The gametocidal chromosomes as a tool for chromosome manipulation in wheat. Chromosome Res 15:67–75

Farré A, Cuadrado A, Lacasa-Benito I, Cistué L, Schubert I, Comadran J, Jansen J, Romagosa I (2012) Genetic characterization of a reciprocal translocation present in a widely grown barley variety. Mol Breed 30:1109–1119

Fischer JA, Favreau MB (1991) Plasmid purification by phenol extraction from guanidinium thiocyanate solution: development of an automated protocol. Anal Biochem 194:309–315

Fuchs J, Schubert I (1995) Localization of seed protein genes on metaphase chromosomes of *Vicia faba* via fluorescence in situ hybridization. Chromosome Res 3:94–100

Fukui K, Kamisugi Y, Sakai F (1994) Physical mapping of 5S rDNA loci by direct-cloned biotinylated probes in barley chromosomes. Genome 37:105–117

Gamborg OL, Wetter LR (1975) Plant tissue culture methods. National Research Council of Canada, Saskatoon

Giacalone J, Delobette V, Gibaja L, Ni L, Skidas Y et al (2000) Optical mapping of BAC clones from the human Y chromosome DAZ locus. Genome Res 10:1421–1429

Harwood WA, Bilham LJ, Travella S, Salvo-Garrido H, Snape JW (2005) Fluorescence in situ hybridization to localize transgenes in plant chromosomes. Methods Mol Biol 286:327–340

Heng HH, Tsui LC (1998) High resolution free chromatin/DNA fiber fluorescent in situ hybridization. J Chromatogr 86:219–229

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: BAC library specific for the short arm of wheat chromosome 1B. Plant J 47:977–986

Joshi GP, Nasuda S, Endo TR (2011) Dissection and cytological mapping of barley chromosome 2H in the genetic background of common wheat. Genes Genet Syst 86:231–248

Kato A (2011) High-density fluorescence in situ hybridization signal detection on barley (*Hordeum vulgare* L.) chromosomes with improved probe screening and reprobing procedures. Genome 54:151–159

Kato A, Vega JM, Han FP, Lamb JC, Birchler JA (2005) Advances in plant chromosome identification and cytogenetic techniques. Curr Opin Plant Biol 8:148–154

Kato A, Albert PS, Vega JM, Bichler JA (2006) Sensitive fluorescence in situ hybridization signal detection in maize using directly labelled probes produced by high concentration DNA polymerase nick translation. Biotech Histochem 81:71–78

Kocsis E, Trus BL, Steer CJ, Bisher ME, Steven AC (1991) Image averaging of flexible fibrous macromolecules: the

clathrin triskelion has an elastic proximal segment. J Struct Biol 107:6–14

Kumar A, Simons K, Iqbal MJ, de Jiménez MM, Bassi FM, Ghavami F, Al-Azzam O, Drader T, Wang Y, Luo MC, Gu YQ, Denton A, Lazo GR, Xu SS, Dvorak J, Kianian PM, Kianian SF (2012) Physical mapping resources for large plant genomes: radiation hybrids for wheat D-genome progenitor *Aegilops tauschii*. BMC Genomics 13:597

Künzel G, Korzun L, Meister A (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics 154:397–412

Leicht IJ, Heslop-Harrison JS (1992) Physical mapping of 18S-5,8S-26S rRNA genes in barley by in situ hybridization. Genome 35:1013–1018

Leicht IJ, Heslop-Harrison JS (1993) Physical mapping of 4 sites of 5S rDNA sequences and one site of the aplha-amylase-2 gene in barley (*Hordeum vulgare*). Genome 36:517–523

Linde-Laursen IB (1978) Giemsa C-banding of barley chromosomes; II. Banding patterns of trisomics and telotrisomics. Hereditas 89:37–41

Lysák MA, Číhalíková J, Kubaláková M, Šimková H, Künzel G, Doležel J (1999) Flow karyotyping and sorting of mitotic chromosomes of barley (*Hordeum vulgare* L.). Chromosome Res 7:431–444

Ma L, Vu GTH, Schubert V, Watanabe K, Stein N, Houben A, Schubert I (2010) Synteny between *Brachypodium distachyon* and *Hordeum vulgare* as revealed by FISH. Chromosome Res 18:841–850

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

Mayer KFX, Martis M, Hedley PE, Šimková H, Liu H, Morris JA, Steurnagel B, Taudien S, Roessner S, Gundlach H, Kubaláková M, Suchánková P, Murat F, Felder M, Nussbaumer T, Graner A, Slase 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

Meyers BC, Scalabrin S, Morgante M (2004) Mapping and sequencing complex genomes: let's get physical! Nat Rev Genet 5:578–588

Michalak de Jimenez MK, Bassi FM, Ghavami F, Simons K, Dizon R, Seetan RI, Alnemer LM, Denton AM, Doğramacı M, Šimková H, Doležel J, Seth K, Luo MC, Dvorak J, Gu YQ, Kianian SF (2013) A radiation hybrid map of chromosome 1D reveals synteny conservation at a wheat speciation locus. Funct Integr Genomics 13:19–32

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, Safar J, Simkova H, Dolezel J, Bernard M, Feuillet C (2008) A physical map of the 1-gigabase bread wheat chromosome 3B. Science 322:101–104

Pérez R, de Bustos A, Cuadrado Á (2009) Localization of Rad50, a single-copy gene, on group 5 chromosomes of beat, using a FISH protocol employing tyramide for signal amplification (Tyr-FISH). Cytogenet Genome Res 125:321–328

Peterson DG, Lapitan NLV, Stack SM (1999) Localization of single- and low-copy sequences on tomato synaptonemal complex spreads using fluorescence in situ hybridization (FISH). Genetics 152:427–439

Phillips D, Nibau C, Wnetrzak J, Jenkins G (2012) High resolution analysis of meiotic chromosome structure and behaviour in barley (*Hordeum vulgare* L.). PLoS One 7:e39539

Riera-Lizarazu O, Vales MI, Kianian SF (2008) Radiation hybrid (PH) and happy mapping in plants. Cytogenet Genome Res 120:233–240

Sakai K, Nasuda S, Sato K, Endo TR (2009) Dissection of barley chromosome 3H in common wheat and comparison of 3H physical and genetic maps. Genes Genet Syst 84:25–34

Sato K, Shin-I T, Seki M, Shinozaki K, Yoshida H, Takeda K, Yamanaki Y, Conte M, Kohara Y (2009) Development of 5006 full-length CDNAs in barley: a tool for accessing cereal genomics resources. DNA Res 16:81–89

Schnabel E, Kulikova O, Penmetsa RV, Bisseling T, Cook DR, Frugoli J (2003) An integrated physical, genetic and cytogenetic map around the sunn locus of *Medicago truncatula*. Genome 46:665–672

Schwartz DC, Li X, Hernandez LI, Ramnarain SP, Huff EJ, Wang YK (1993) Ordered restriction maps of *Saccharomyces cerevisiae* chromosomes constructed by optical mapping. Science 262:110–114

Singh RJ, Tsuchiya T (1982) An improved Giemsa N-banding technique for the identification of barley chromosomes. J Hered 73:227–229

Tang X, Szinay D, Lang C, Ramanna MS, van der Vossen EAG, Datema E, Lankhorst RK, de Boer J, Peters SA, Bachem C, Stiekema W, VisserR GF, de Jong H, Bai Y (2008) Cross-species bacterial artificial chromosome-fluorescence in situ hybridization painting of the tomato and potato chromosome 6 reveals undescribed chromosomal rearrangements. Genetics 180:1319–1328

Tang X, de Boer JM, van Eck HJ, Bachem CH, Visser RGF, de Jong H (2009) Assignment of genetic linkage maps to diploid *Solanum tuberosum* pachytene chromosomes by BAC-FISH technology. Chromosome Res 17:399–915

The International Barley Genome Sequencing Consortium (2012) A physical, genetical and functional sequence assembly of the barley genome. Nature 491:711–716

The International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463:763–768

Tjio JH, Hagberg A (1951) Cytological studies on some X-ray mutants of barley. Ann Aula Dei 2:149–167

Tomato Genome Consortium (2012) The tomato genome sequence provides the insight into fleshy fruit evolution. Nature 485:635–641

Tsujimoto H, Mukai Y, Keisuke A, Nagaki K, Fujigaki J, Yamamoto M, Sasakuma T (1997) Identification of individual barley chromosomes based on repetitive sequences: conservative distribution of Afa-family repetitive sequences on the chromosome of barley and wheat. Genes Genet Syst 72:303–309

Valárik M, Bartoš J, Kovářová P, Kubaláková M, de Jong JH, Doležel J (2004) High-resolution FISH on super-stretched flow-sorted plant chromosomes. Plant J 37:940–950

Wang CH-JR, Harper L, Cande WZ (2006) High-resolution single-copy gene fluorescence in situ hybridization and its

use in the construction of a cytogenetic map of maize chromosome 9. Plant Cell 18:529–544

Wardrop J, Sape J, Powell W, Machray GC (2002) Constructing plant radiation hybrids panel. Plant J 31:223–228

Wicker T, Mayer KFX, 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

Yang K, Zhang H, Converse R, Wang Y, Rong X, Wu Z, Luo B, Yue L, Jian L, Zhu L, Wang X (2011) Fluorescence in situ hybridization on plant extended chromatin DNA fibres for single-copy and repetitive DNA sequences. Plant Cell Rep 30:1779–1786

Zhou S, Bechner MC, Place M, Churas CP, Pape L, Leong SA, Runnheim R, Forrest D, Goldstein S, Livny M, Schwartz DC (2007) Validation of rice genome sequence by optical mapping. BMC Genomics 8:278

Zhou S, Wei F, Nguyen J, Bechner M, Potamousis K, Goldstain S, Pape L, Mehan MR, Churas C, Pasternak S, Forrest DK, Wise R, Ware D, Wing RA, Waterman MS, Livny M, Schwartz DC (2009) A single molecule scaffold for the maize genome. PLoS Genet 5:e1000711

APPENDIX II

Chromosome painting in barley – the new milestone in cytogenetics of cereals

Karafiátová M., Mascher M., Bartoš J., Vrána J., Stein N., Doležel J.

Manuscript - prepared for submission

CHROMOSOME PAINTING IN BARLEY – THE NEW MILSTONE IN CYTOGENETICS OF CEREALS

Karafiátová M.[1], Mascher M.[2], Bartoš J.[1], Vrána J.[1], Stein N.[2], Doležel J.[1]

1) Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Šlechtitelů 31, 783 71 Olomouc, Czech Republic
2) Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Seeland OT Gatersleben, Germany

*Corresponding author:
Jaroslav Doležel
Institute of Experimental Botany
Šlechtitelů 31, 783 71 Olomouc
Email: dolezel@ueb.cas.cz
Fax: (+420) 585 238 703

Key words: chromosome painting • whole-chromosome probe • gene capture • microarray • blocking DNA

ABSTRACT

The principle of chromosome painting relies on chromosome-specific, composite probes, which allows visualization of selected chromosomes in metaphase and interphase using fluorescence *in situ* hybridization. This procedure triggered off the new cytogenetic stream applied mainly in human cytogenetics as a robust and valuable clinical tool. Subsequently, chromosome painting has been successfully applied in many animals, nevertheless almost all attempts to utilize chromosome painting in plants failed. The failure is linked with the presence of dispersed repeats in painting probes. Here we reported a novel approach allowing paint the chromosomes in species with large genomes that are composed mainly of repetitive DNA sequences. We chose barley chromosome 1H as a model. The method is based on the ability to prepare highly specific chromosome painting probe composed mainly from low-copy coding sequences. DNA for the probe preparation was isolated using gene capture from DNA of flow-sorted chromosome 1H. The use of this probe in combination with a large amount of blocking DNA containing remaining barley chromosomes resulted in specific labelling barley chromosome 1H when applied to barley mitotic metaphase spreads. These results offer the way for exploration of complex plants genomes and represent the important leap towards chromosome painting in interphase nuclei to study their spatial organization and dynamics.

INTRODUCTION

Chromosome painting is an appealing cytogenetic technique, which significantly increased the attractiveness of cytogenetics, because it is not just a technique of the basic research, but it is mainly a tool of clinical cytogenetics, which helps uncover the serious diseases and saves human lives. The main idea of this method is generating of fluorescently labelled probes for whole chromosomes (WCPs), which can be subsequently visualized in the process of fluorescent *in situ* hybridization on chromosomes or interphase nuclei. The technique was originally developed for analysis of human cells with the purpose to make the medicine diagnostic faster and more accurate. The history of this technique covers several decades, has gone through many changes, diversified a lot and was adapted to the current demands.

The first report about chromosome painting appeared in 80´s of the last century, when Pinkel *et al.* (1988) developed the new technique with the aim to diagnose Down´s syndrome and translocation involving human chromosome 4, and they named this general approach "chromosome painting". The technique was based on the use of human chromosome specific phage libraries derived from flow-sorted chromosomes (Van Dilla *et al.*, 1986). This work initiated a new stream of clinical cytogenetics and was followed by several similar studies focused on human diagnostic (Cremer *et al.*, 1988, Callen *et al.*, 1990). In the following years, the technique was distinctly improved and mainly the probe preparation was modified a lot.

In 1992, Chang *et al*. established the protocol for PCR amplification of chromosome-specific DNA isolated from human flow-sorted chromosomes. This method opened the way to generate whole chromosome probes from relatively small amount of starting DNA from flow-sorted and microdissected chromosomes (Guan *et al*., 1994). Furthermore, micromanipulation of the immobilized chromosomes brought another technique extending. Microdissection allowed to cut smaller portion of chromosomes out of the specimen and generated chromosome specific fragments usable for the diagnostic purposes (Guan *et al*., 1993).

The simultaneous identification of all 24 human chromosomes in one experiment using multicolor FISH was the next significant advance in chromosome painting.

Color karyotyping and visualization of human chromosomes was performed by applying of the two similar approaches called mFISH (Speicher *et al*., 1996) or SKY-FISH (Schröck *et al*., 1996) and the enormous clinical potential catapulted both procedures to the top of the diagnostic methods.

After the undoubted success of chromosome painting in human medicine, there were numerous efforts to apply the method in animal and plant species. Recently, chromosome painting is feasible in animals, where it represents the valuable tool of comparative cytogenetics (Jauch *et al*., 1992; Scherthan *et al*., 1994; Yang *et al*., 1995) and animal painting probes are widely used in phylogenic and evolutionary analysis (Ferguson-Smith, 1997, Nie *et al*., 2009, 2012), in the studies of nucleus compartmentalization (Cremer *et al*., 1993, 2008) or as a tool to analyze structural mutants (Liechty *et al*., 1995).

Although chromosome painting procedure is deeply studied and well established for human and many animal species with the wide range of genome size, its application in plants is very restricted. Because of unquestionable benefits for plant research, there were several attempts to work out the procedure in plants, but the vast majority of them failed. The failure is attributed to the huge amounts of repetitive sequences (Schubert *et al*., 2001), which are scattered over all chromosomes and produce perceptible non-specific signal. Due to the specific and non-specific signal ratio, the outcomes were unclear and ambiguous.

Nevertheless, there exist a couple of successful experiments, which were done more than decades ago. All studies were focused on the plant species with relatively small, non-complex genomes. The one of the first painting studies in plants was published in 1999. Shibata *et al*. (1999) painted the microdissected sex chromosome Y of *Rumex acetosa* and hybridized the probe on the chromosomes as well as on nuclei. In this study, they proved and underlined the key role of the presence of unlabelled competitive DNA in the painting experiments. A few years later, Lysak *et al*. (2001) have come up with an idea of using the large number of BAC clones in one FISH experiment in *Arabidopsis thaliana*, which finally resulted in painting of entire chromosome. This alternative approach, adopted later also by Idziak *et al*. (2011) to paint the chromosomes of *Brachypodium*, allows discrimination any chromosome

part of interest, but unfortunately it is suitable only for species with limited genome size, because BAC FISH in species with large genomes is hampered with high background due to the enormous repeat content.

Kato *et al*. (2004) followed the similar idea, but instead of BAC clones, they used cocktail of chromosome-specific repeats in one experiment to paint the chromosomes of maize. Subsequently, chromosome painting of B chromosomes of rye and *Brachycome dichromosomatica* (Houben *et al*., 1997, 2001), and chromosomes of rice (Shishido *et al*., 2001) were performed, however the success of all these attempts were based more on the chromosome specific repeat enrichment rather than on single-copy sequences. Recently, Lou *et al*. (2014) published the single-copy chromosome painting approach based on hybridization of single-copy genes in *Cucumis sativus*. In simultaneous hybridization of gene pools they painted the interphase nuclei, metaphase and pachytene chromosomes and this way they have identified the large miss-assembled sequence blocks of chromosome 4 of cucumber.

In this study, we report on innovative approach for chromosome painting, which overcomes the troubles with enormous repeat content and is suitable for species with complex genomes. We performed chromosome painting in barley despite the fact that repetitive sequences represent more than 80% of its 5Gbp genome. We used chromosome-specific painting probe for barley chromosome 1H, which resulted in unambiguous painting of entire chromosome in metaphase. The presented method relies on gene-capture technology together with flow cytometry and emphasizes the requirements of highly specific probe composed mainly from single copy sequences in combination with very effective blocking DNA. Our outcomes open a way for a wider use of chromosome painting in plants with large genomes and represent the first step towards chromosome painting in interphase nuclei to study their 3D organization and dynamics.

MATERIAL AND METHODS

Mitotic metaphase preparations

Mitotic metaphase chromosomes were isolated from root tips of barley (*Hordeum vulgare* L., 2n=2x=14) cv. Barke. To increase the number of dividing cells (mitotic

index), two days old seedlings were synchronized using 2mM hydroxyurea (HU) and 2.5μM amiprophos-methyl (APM) according to (Lysák *et al.,* 1999). The root tips were fixed in ethanol: acetic acid (3:1) for a week at 37°C and afterwards were stored at 20°C. The fixed roots were stained in 2% carmine acid solution and mitotic metaphase spreads were prepared using the standard squashing technique. Briefly, the root tips were cut off the roots into the drop of 45% acetic acid, cover with cover slip and smashed by tapping the needle. Preparations were carefully flamed for a few seconds, squashed, freeze on dry ice for an hour and finally washed 3 minutes in preheated 45% acetic acid at 50°C. The selected slides were washed in 2xSSC, fixed in 4% formaldehyde in 2xSSC and dehydrated in ethanol series as described in Ma *et al.* (2010).

Flow cytometry

Chromosomes 1H and remaining six chromosomes 2H-7H were sorted from the barley cv. Barke using flow cytometry according to Lysák *et al*. (1999). Briefly, germinated seeds were treated subsequently with hydroxyurea and aminoprophos-methyl to increase the metaphase content. Synchronized seedlings were fixed by formaldehyde, mechanically homogenized and chromosome suspension was analysed using a FACSAria flow cytometer (BD Biosciences, San Jose). Batches of 10,000 chromosomes were sorted into 50 µl deionized water in PCR tubes. Sorted fraction purity was checked using FISH with GAA probe and subsequently, sorted chromosomes were amplified as described in Šimková *et al*. (2008).

Gene capture

We utilized existing exome capture platform for barley developed by Mascher *et al*. (2013). Solution-phase capture pool (liquid array) consisted of more than 4 million probes of variable length (50-100bp) including predicted exons, full-length cDNA and transcript contigs from RNA-seq. Amplified DNA of sorted chromosome 1H of barley cv.” Barke” was used as input DNA. In total, 23 µg of amplified 1H DNA was needed to implement one gene capture run. The exome capture experiment was performed using NimbleGen Plant Sequence Capture (Roche, Madison, WI, USA) according to manufacturer´s instruction. Briefly, sample library was constructed from amplified DNA of chromosome 1H. Afterwards, adapters were ligated to the

fragmented DNA and ligation-mediated (LM) amplification of pre-captured library was performed. Subsequently, mixture containing sample library, enhancer, buffer and oligos was added to the exome library (liquid array) and after denaturation, hybridization was carried out at 47°C for 72 hours. Captured DNA was washed in the row of stringent buffers contained in kit, recovered using Streptavidin Dynabeads and amplified in LM-PCR. Finally, the samples characteristics were checked by BioAnalyzer DNA 1000 chip and NanoDrop spectrophotometer. The final product should exhibit the fragment length ranging 200-400 bp, A260/A280 ration 1.7-2.0 and LM-PCR yield more than 500 ng.

Probe labelling and preparing of blocking DNAs

Probe: First, DNA isolated from the array was amplified using primers TS-Oligos F/R complementary to the existing ligated adaptors. Final concentrations in 50 µl PCR mix were as follows: 1mM PCR buffer containing 1.5mM $MgCl_2$ (Finnzymes, Vantaa, Finland), 0.2mM of each dNTPs (Fermentas, Glen Burnie, Maryland, USA), 1mM TS-Oligos primers and 2U/50 µl Dynazyme ^{TM} II DNA polymerase (Finnzymes). PCR was performed under following conditions: 30 sec at 94°C followed by 35 cycles of 30 sec denaturation at 94°C, 30 sec annealing at 60°C, 30 sec extension at 72°C, and final extension for 5 min at 72°C. Amplified product was directly labelled with Texas Red in following PCR reaction under the same conditions. dNTP mixture in labelling reaction contained dTTP and labelled dUTP-TR in ratio 2:1. PCR product was checked on 1.5% agarose gel in 0.5xTBE. To reach the high concentration of the probe, the products of eight amplifications were merged, DNA was precipitated by isopropanol (Fisher *et al*., 1991) and dissolved in 10 µl 2xSSC buffer in 37°C.

C_0t100: Blocking DNA was prepared according to (Hřibová *et al*., 2007). Concisely, 100 µg of genomic DNA of barley was isolated from the lyophilized leaves using Plant Invisorb MiniKit (Invitek, Berlin,Germany). DNA was cleaved three hours at 37°C in the mixture of restriction enzymes at final concentration in 100 µl as followed: *FOK*I (0.04U/µl), *Hae*III (0.1U/µl) and *Mse*I (0.4U/µl) in 1x BSA and 1x NEB buffer (all produced by New England BioLab, Ipswitch, Massachusetts, USA). The restricted fragments were analysed on 1.5% agarose gel in 0.5xTBE. Sheared

DNA was purified on columns (Microcon 30, Millipore, Billerica, MA, USA) and the concentration of the eluted DNA was measured using fluorimeter. DNA was denatured in boiling water for 10 min and reassotiation was performed in 0.13M phosphate buffer for 94 hours at 60°C. Non-reassotiated, ssDNA fragments were cleaved with 200U of nuclease S1 (Fermentas, Glen Burnie, Maryland, USA) at 37°C and the cleavage was inhibit in 30 min by adding 30 µl of 0.5M EDTA and 10 min incubation at 70°C. Finally, reassociated DNA fraction was evaporated in SpeedVac to reduce the volume to approximately 20 µl. Final concentration of C_0t100 DNA was roughly 1.6 µg/µl.

2H-7H blocking DNA: Blocking DNA was prepared from six remaining barley chromosomes 2-7 sorted using flow cytometry. 200 µg of amplified sorted chromosome 2H-7H was concentrated by evaporation to reach concentration 1 µg/µl. Chromosomes were sheared in addition of NaOH to final concentration 0.4M in boiling water for 90 min. The ideal fragments length ranged from 50bp to 500bp. Sheared DNA were precipitated by adding half volume of 3M sodium acetate and one volume of 96% ethanol. After the precipitation at -20°C, the sample was centrifuged, pellet washed in 70% ethanol, dried and dissolved in 20 µl of TE buffer.

Chromosome painting of barley chromosome 1H

Fluorescence *in situ* hybridization was performed in the volume of 25 µl. Hybridization mixture consisting of 40% formamide, 2xSSC, 10% dextran sulphate, 0.625 µg/µl calf thymus, 1 µg 2H-7H blocking DNA, 1.6 µg C_0t100 DNA, 150 ng/µl 1H probe and water up to 25 µl was dropped on a slide and coverslip was glued around. After three-minute denaturation at 80°C, chromosomes were treated with gradually decreasing temperature. Temperature went down from 80°C by one degree to 41°C for two minutes in each step. The probe was hybridized in a humid chamber at 37°C overnight. After the hybridization, the slides were washed in 2×SSC for 10 min, 0.1×SSC, 2 mM $MgCl_2$, and 0.1% v/v Triton X-100 for 5 min at 42°C, twice in 2×SSC for 10 min at RT and 4×SSC at RT. Chromosomes were mounted in Vectashield containing 4´,6´-diamidino-2-phenylindole (Vector Laboratories).

Microscopy and image processing

The signals were observed using Axio Imager Z.2 Zeiss microscope (Zeiss, Oberkochen, Germany) equipped with Cool Cube 1 (Metasystems, Altlussheim, Germany) camera and filter set for Texas Red: HQ617/40 (excitation), Q593LP (beam splitter), HQ581/10 (emission) (filter set F46-674, AHF) and DAPI: G365 (excitation), FT395 (beam splitter) and BP445/50 (emission) (filter set 49, Zeiss). The images were taken in ISIS software (Metasystems) and the image adjustment was done in Adobe Photoshop 6.0.

RESULTS AND DISSCUSSION

Herein we proposed the new chromosome painting protocol useful for the species with significant portion of repetitive sequences in their genomes. We succeeded in painting barley chromosome 1H after combining two approaches - flow cytometry and exome capture, which showed to be essential and significantly contributed to the breakthrough results.

First, we sorted two different chromosome fractions, 1H and 2H-7H, from barley cv. Barke using flow cytometry. DNA of barley chromosome 1H was amplified, hybridized against the microarray and subsequently used as a DNA source for painting probe. Fraction consisting of the six remaining barley chromosomes served as blocking DNA. Due to significant difference in DNA content between chromosome 1H and other barley chromosomes, 1H could be easily discriminated and flow-sorted (Fig.1). 1H fraction contained barley chromosome 1H with the purity higher than 95% (95.5% ± 0.7%; mean ± SD) and in total, roughly 120,000 chromosomes (corresponding to the 150 ng DNA) were sorted and subsequently amplified in 9 reactions up to amount 23 μg DNA for one gene capture experiment. Similarly, 590,000 chromosomes 2H-7H (850 ng) amplified in 51 amplifications up to 200 μg were used for preparing one stock of blocking DNA.

The highly specific probe represents one of the essential components of described protocol. In order to obtain this kind of probe we hybridized 1H amplified DNA to solution-phase platform originally designed for exome sequencing of barley

(Mascher *et al*., 2013). From one gene-capture experiments we got approximately 280 ng 1H specific DNA, which supposed to be presented in genome in single or low-copy numbers and this DNA was further used for probe preparation. This way we significantly decreased the probability of unspecific signal coming from dispersed repetitive motifs, what represents the critical step in all existing painting protocols so far applied in species with more complex genomes.

In spite the probe specificity, significant background was observed on all chromosomes after the hybridization (Fig.2) and this background had to be supressed by adding blocking DNA and setting the strict hybridization conditions. The possible source of unspecific signal was revealed in analysis of captured DNA fragments. While Illumina sequencing of captured DNA confirmed the significant enrichment of captured DNA specific for 1H chromosome (Mascher *et al*., 2013), we detected also small content of repetitive elements as well. Notwithstanding this very low repeat portion is negligible in comparison to original repeat content in barley genome, the high number of failed experiments proved that even significant elimination of repetitive sequences in probe was not sufficient enough to reduce the unspecific signal and underline the crucial role of blocking DNA. Similar findings were reported in the last chromosome painting study of Lou *et al*. (2014), who performed chromosome painting in *Cucumis*. They used an alternative way to isolate the single copy sequences from *Cucumis* genome. Instead of microarray, they designed hundreds primer pairs for gene regions with the aim to eliminate repetitive elements, but they finally detected the presence Ty1-copia and Ty3-gypsy elements in single copy pools. In our experiment, captured DNA might contain beside repetitive sequences also sequences homologous to duplicated genes or pseudogenes located on the other barley chromosomes (Mayer *et al*., 2011).

The presence of residual repeats and sequences homologous to genes or pseudogenes on chromosomes 2H-7H in our probe was also confirmed experimentally. In our study, two types of blocking DNA were employed. When we used them separately in individual experiments, none of them was able to suppress unspecific signal by itself, but in combination they allowed to obtain distinct signal on chromosome 1H. While blocking DNA from barley chromosomes 2H-7H prevented the probe penetrating to

the loci with high homology to 1H sequences, C_0t100 blocking DNA inhibited the signal from eventual repeats contained in captured DNA.

Using both types of blocking DNA in combination with very stringent conditions we observed unambiguous signal on both homologous chromosomes 1H with irrelevant background on the other chromosomes in metaphase (Fig.3). As expected, the detected signal was not homogenous throughout whole chromosome length. Considering the probe composition, the clear signal was observed in central parts of chromosome arms and on the contrary, the probe provided lower signal in centromeric and telomeric areas, which supposed to be highly repetitive with lower frequency of genes.

Surprisingly, hybridization efficiency in successful experiment was very high. We observed distinct signal on more than 80 % metaphases in successful experiment and the clear signal was always detected just on chromosome 1H and never on any of six other chromosomes. In spite undisputed results, the protocol has not been satisfactory reproducible. Total reproducibility of described protocol was quite low. We succeeded in about 10 % performed experiments, even though we eliminated all possible effects, which could cause the differences among the particular experiments like same stocks of all compounds used in experiment.

Failed experiments together with missing signal on 20 % figures in successful experiments suggest that the accurate concentration in the mixture itself is not able to secure the clear signal, but that the level of chromatin condensation could have a decisive impact on final success. This claim was also supported by the fact that unambiguous signal has never been observed on barley prometaphase chromosome or interphase nuclei even in the successful experiments with clear signal on metaphase chromosomes. These results point on the possible key role of chromatin structure in panting experiments in the species with complex genomes.

The probe nature can also contribute to signal absence on relaxed chromatin structures. The total amount of fluorochrome integrated in probe is due to the chromatin condensation amassed on much smaller area on mitotic metaphase chromosome than on prometaphase chromosomes or interphase nuclei. Furthermore,

visualized single gene regions are originally separated by blocks of sequences not presented in probe, which after the hybridization correspond to the "gaps" among the labelled loci. These gaps started to have the relevant impact on signal detection with the chromatin elongation, when the local density of fluorochrome decrease and the final signal intensity could be below the camera detection sensitivity. This explanation is consistent with results reached on *Cucumis*, where portion of repeats is negligible in comparison to barley and therefore the labelled genes are ordered with higher density and signal is not so dispersed. Whereas in *Cucumis*, painting protocol were successfully adapted to pachytene chromosomes and interphase nuclei as well (Lou *et al*., 2014).

Very low success rate of the protocol is undoubtedly connected to barley genome size and its composition. In comparison to *Cucumis*, barley has almost 20 times larger genome, what is naturally reflected in genome structure. *Cucumis* carries in its roughly 350 Mbp-large genome around 24 % repetitive sequences and also the density of homologous pseudogenes among the chromosomes is remarkably lower (Huang *et al*., 2009).

The presented protocol relies on sophisticated methods and could seem to have the large demands on equipment and costs, however recently, many of needed devices, methods and data are available or easily accessible in various species. In postgenomic era, when next generation sequencing technologies go through massive expansion, there are number of designed platforms (Zhou *et al*., 2012; Neves *et al*., 2013; Tennessen *et al*., 2013) and sequencing data, which can be used also for exome capture design. Similarly, the protocols for sorting of plant chromosomes are worked out for various species and advanced flow sorters in combination with novel methods as FISHIS (Giorgi *et al*., 2013) allow us to overcome recent restrictions of the method.

Using described protocol, we performed so far second chromosome painting on monocot plant, but comparing to *Brachypodium*, on the species with significantly larger and more complex genome. Our promising results clearly show that despite of existing prejudices chromosome painting in cereals is possible. Although the protocol cannot be considered as robust and does not always lead to the satisfactory

outcomes, the uniqueness of reached results allows us to report the new approach, which can be further developed and help to move ahead the research in this field in various species, in which it was so far unthinkable.

AKNOWLEDGEMENTS

We appreciate the technical assistance of Zdeňka Dubská and Romana Šperková. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01 and Internal Grant Agency of Palacky University (Grant award No. IGA PrF-2013-003 and IGA_PrF_2014001).

REFERENCES

Callen DF, Baker E, Eyre HJ, Chernos JE, Bell JA, Sutherland GR. (1990) Reassessment of two apparent deletions of chromosome 16p to an ins(11;16) and at(1;16) by chromosome painting. Ann Genet. 33(4):219-2

Chang KS, Vyas RC, Deaven LL, Trujillo JM, Stass SA, Hittelman WN (1992) PCR amplification of chromosome-specific DNA isolated from flow cytometry-sorted chromosomes. Genomics 12(2):307-12

Cremer T, Richter P, Borcem J, Ward DC, Manuelidis L. (1988) Detection of chromosome aberrations in metaphase and interphase tumor cells by *in situ* hybridization using chromosome-specific library probes. Hum Genet 80:235-246

Cremer T, Kurz A, Zirbel R, Dietzel S, Rinke B (1993) Role of chromosome territories in the functional compartmentalization of the cell nucleus. Cold Spring Harb Symp Quant Biol 58: 777-72

Cremer M, Grasser F, Lanctot C, Muller S, Neusser M (2008) Multicolor 3D fluorescence *in situ* hybridization for mapping interphase chromosomes. Methods Mol Biol 463: 205-239

Ferguson-Smith MA (1997) Genetic analysis by chromosome sorting and painting: phylogenetic and diagnostic applications. Eur J Hum Genet 5:253–265

Fisher JA, Favreau MB (1991) Plasmid purification by phenol extraction from guanidinium thiocyanate solution: development of an automated protocol. Anal Biochem 194(2):309-15

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

Guan XY, Trent JM, Meltzer PS (1993) Generation of band-specific painting probes from a single microdissected chromosome. Hum Mol Genet. 2(8):1117-21

Guan, XY, Meltzer, PS and Trent JM (1994) Rapid generation of whole chromosome painting probes (WCPs) by chromosome microdissection. Genomics 22:101–107

Houben A, Leach CR, Verlin D, Rofe R, Timmis JN (1997) A repetitive DNA sequence common to the different B chromosomes of the genus Brachycome. Chromosoma 106: 513–519

Houben A, Field BL, Saunders VA (2001) Microdissection and chromosome painting of plant B chromosomes. Methods in Cell Sci 23:115-123

Hřibová E, Doleželová M, Tlen CD, Macas J, Doležel J (2007) Isolation and characterization of the highly repeated fraction of the banana geonome. Cytogenet Genome Res 119:268–274

Huang S, Li R., Zhang Z, Li L et al. (2009) The genome of cucumber, *Cucumis sativus* L. Nature Genet 41:1275-1284

Idziak D, Betekhtin A, Wolny E, Lesniewska K *et al.* (2011) Painting the chromosomes of Brachypodium – current status and future prospects. Chromosoma 120:469-479

Jauch A, Wienberg J., Stanyon R., Arnold N., Tofanelli S, Ishida T, Cremer T (1992) Reconstruction of genomic rearrangements in great apes and gibbons by chromosome painting. Proc. Natl. Acad. Aci. USA 89: 8611-8615

Kato A, Lamb JC, Bichler A (2004) Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. PNAS 37: 13554-13559

Liechty MC, Hall BK, Scalzi JM, Davis LM, Caspary WJ, Hozier JC (1995) Mouse chromosome-specific painting probes generated from microdissected chromosomes. Mamm Genome 6(9):592-594

Lou O, Zhang Y, He Y, Li J, Jia L, Cheng Ch, Guan W, Yang S, Chen J (2014) Single-copy gene-based chromosome painting in cucumber and its application for chromosome rearrangement analysis in Cucumis. The Plant J 78:169-179

Lysák MA, Číhalíková J, Kubaláková M, Šimková H, Künzel G, Doležel J (1999) Flow karyotyping and sorting of mitotic chromosomes of barley (*Hordeum vulgare* L.). – Chromosome Res 7: 431-444

Lysák M, Fransz PF, Ali HMA, Schubert I (2001) Chromosome painting in *Arabidopsis thaliana.* The Plant J 28: 689-697

Ma L, Vu GTH, Schubert V, Watanabe K, Stein N, Houben A, Schubert I (2010) Synteny between *Brachypodium distachyon* and *Hordeum vulgare* as revealed by FISH. Chromosome Res 18: 841-850

Mascher M1, Richmond TA, Gerhardt DJ, Himmelbach A *et al.* (2013) Barley whole exome capture: a tool for genomic research in the genus Hordeum and beyond. Plant J 76(3):494-505

Mayer KFX, Martis M, Hedley PE, Šimková *et al* (2011) Unlocking the barley genome by chromosomal and comparative genomics. The Plant Cell 23:1349-1263

Neves LG, Davis JM, Barbazuk WB, Kirst M. (2013) Whole-exome targeted sequencing of the uncharacterized pine genome. Plant J 75(1):146-156

Nie W, O´Brien PCM, Ng BL, Fu B, Volobouev V, Carter NP, Ferguson-Smith MA, Yang F (2009) Avian comparative genomics: reciprocal chromosome painting between domestic chicken (*Gallus Gallus*) and the stone curlew (Burhinus oedicnemus, Charadriiformes) – An atypical species with low diploid number. Chromosome Res 17(1):99-113

Nie W, Wang J, SuW, Wang D, Tanomtong A, Perelman PL, Graphodatsky AS, Yang F (2012) Chromosomal rearrangements and karyotype evolution in carnivores revealed by chromosome painting. Heredity 108:17–27

Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R, Lucas J, Gray J (1988) Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc Natl Acad Sci U S A.85: 9138-42

Scherthan H, Cremer T, Arnason U, Weier H-U, Lima de Faria A, Frönicke L (1994) Comparative chromosome painting disloses homologous segments in distantly related mammals. Nature Genetics 6:342-347

Schröck E, du Manoir S, Veldman T, Schoell B, Wienberg J, Ferguson-Smith MA, Ning Y, Ledbetter DH, Bar-Am I, Soenksen D, Garini Y and Ried T (1996) Multicolor spectral karyotyping of human chromosomes. Science 273: 494–497

Schubert I, Fransz PF, Füchs J, de Jong H (2001) Chromosome painting in plants. Methods in Cell Science 23:57-69

Shibata F, Hizume M, Kuroki Y (1999) Chromosome painting of Y chromosomes and isolation of a Y chromosome-specific repetitive sequences in the dioecious plant *Rumex acetosa*. Chromosoma 108: 266-270

Shishido R, Ohmido N, Fukui K (2001) Chromosome painting as a tool for rice genetics and breeding. Methods in Cell Sci 23:125-132

Speicher MR, Ballard SC, Ward DC (1996) Karyotyping human chromosomes by combinatorial multi-fluor FISH. Nature Genet 12:368–375

Šimková H, Svensson JT, Condamine P, Hřibová E, Suchánková P, Bhat PR, Bartoš J, Šafář J, Close TJ, Doležel J (2008) Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genomics 9: 294-302

Tennessen JA, Govindarajulu R, Liston A, Ashman TL (2013) Targeted sequence capture provides insight into genome structure and genetics of male sterility in a gynodioecious diploid strawberry, *Fragaria vesca* ssp. bracteata (Rosaceae). G3 7;3(8):1341-51

Van Dilla M A, Deaven LL, Albright KL et al. (1986) Human Chromosome–specific DNA Libraries: Construction and Availability. Nature biotechnology 4:537-552

Yang F, Carter NP, Shi L, Ferguson-Smith MA (1995) A comparative study of karyotypes of muntjacs by chromosome painting. Chromosoma 103(9):642-52

Zhou L, Holliday JA (2012) Targeted enrichment of the black cottonwood (*Populus trichocarpa*) gene space using sequence capture. BMC Genomics 14;13:703

Legends to figures:

Fig.1: Flow karyotype of barley cv. Barke. Two peaks are clearly resolved on flow karyotype. First peak correspond to the smallest barley chromosome 1H and the second composed peak contain all six remaining barley chromosomes 2H-7H.

Fig.2: 1H probe hybridization in experiment without employing blocking DNA. Even highly specific probe produce significant unspecific signal (pink) on all barley chromosomes. This experiment proofed essential role of blocking DNA in chromosome painting. Scale bar: 10 μm.

Fig.3: Chromosome painting of barley chromosome 1H. Directly labelled whole-chromosome probe in combination with blocking DNAs (C_0t100 and 2H-7H DNA) generate the unambiguous signal (pink) on target chromosome pair. The signal decrease in centromeric and telomeric regions is clearly visible. Scale bar: 10 μm.

Figure 1:

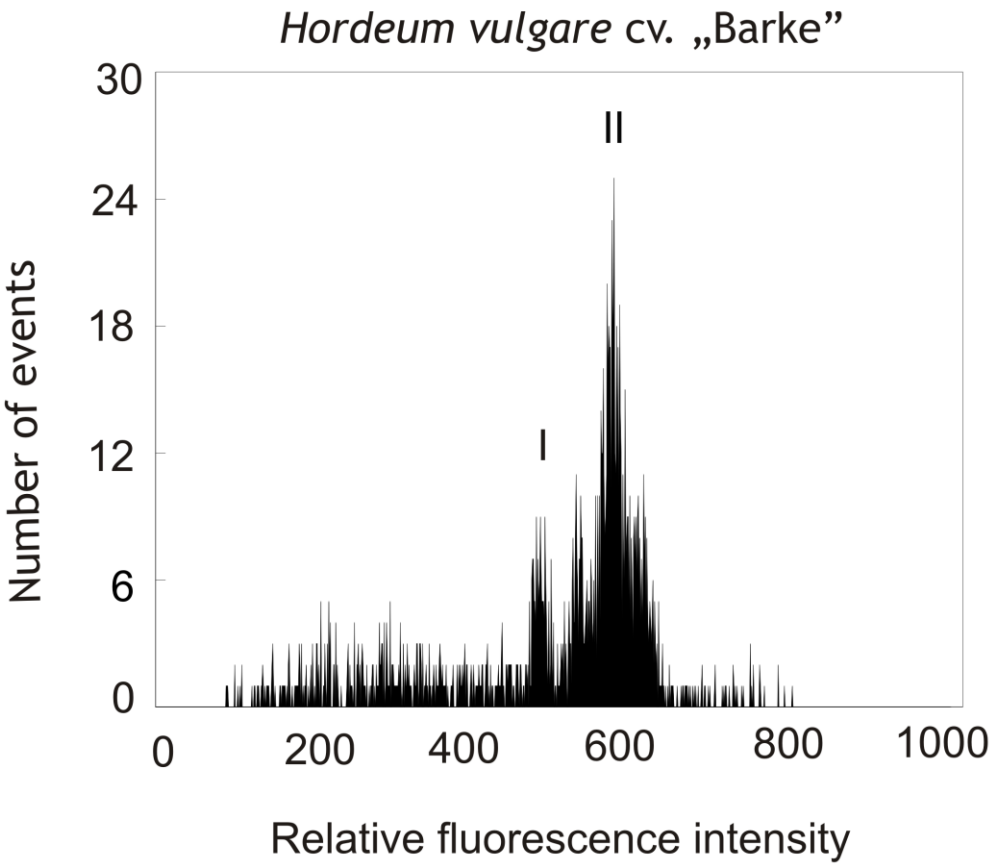

Hordeum vulgare cv. „Barke”
Number of events
Relative fluorescence intensity
I
II
0
6
12
18
24
30
0
200
400
600
800
1000

Figure 2:

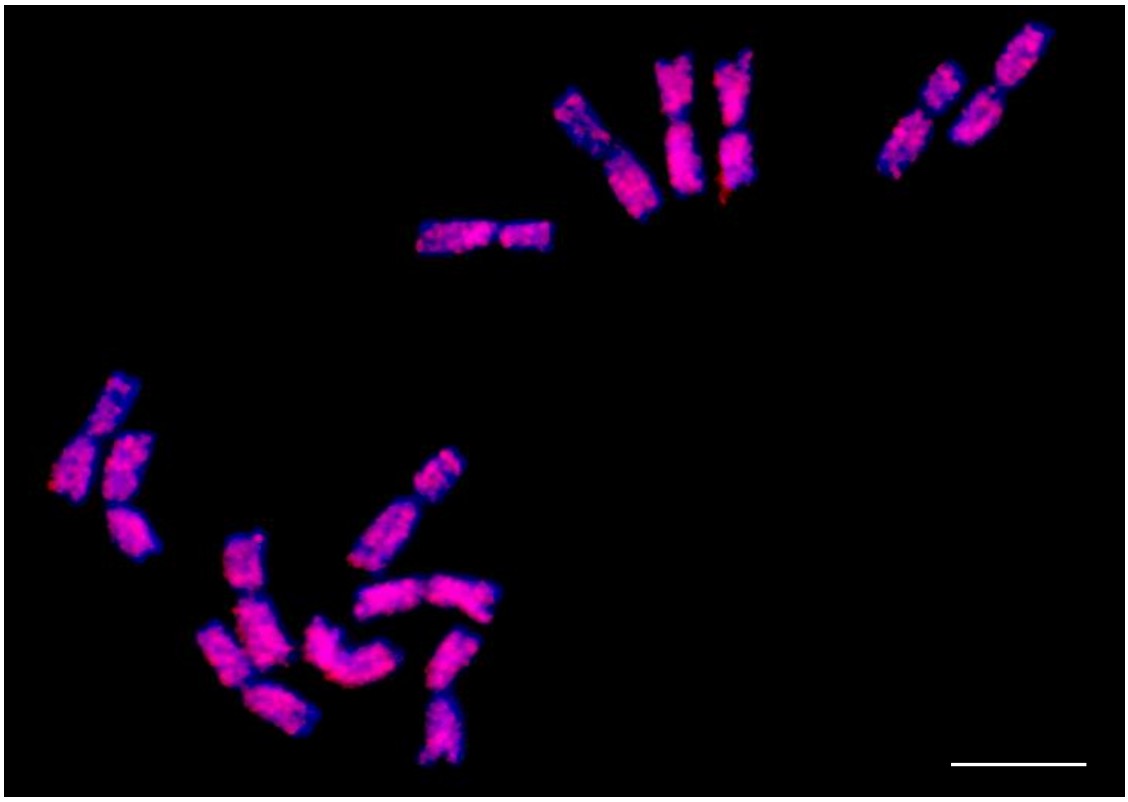
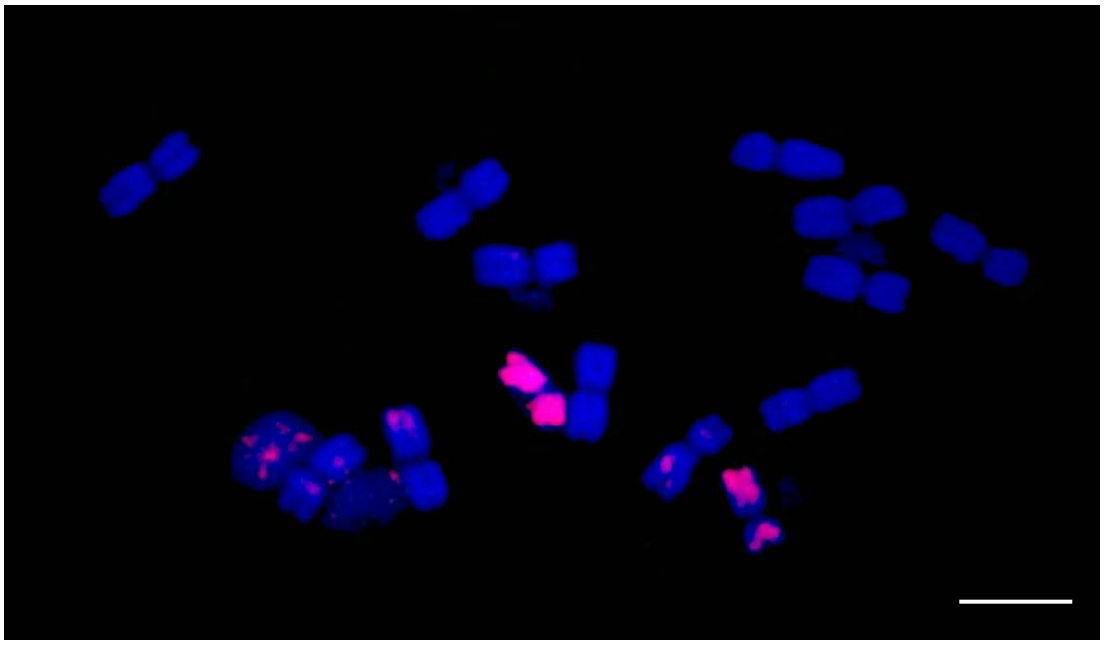


Figure 3:



APPENDIX III

Physical distribution of homoeologous recombination in individual chromosomes of *Festuca pratensis* in *Lolium multiflorum*

Kopecký D., Havránková M., Castro J., Lukaszewski A.J., Bartoš J., Kopecká J., Doležel J.

Cytogenetic and Genome Research DOI:10.1159/000313379, 2010

IF: 1.885

Cytogenetic and Genome Research

Cytogenet Genome Res
DOI: 10.1159/000313379

Published online: May 26, 2010

Physical Distribution of Homoeologous Recombination in Individual Chromosomes of *Festuca pratensis* in *Lolium multiflorum*

D. Kopecký[a] M. Havránková[a] J. Loureiro[b] S. Castro[b] A.J. Lukaszewski[c]
J. Bartoš[a] J. Kopecká[d] J. Doležel[a]

[a]Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Olomouc, Czech Republic; [b]Centre for Functional Ecology, Department of Life Sciences, Faculty of Science and Technology, University of Coimbra, Coimbra, Portugal; [c]Department of Botany and Plant Sciences, University of California, Riverside, Calif., USA; [d]Department of Biology, Pedagogical Faculty, Palacky University, Olomouc, Czech Republic

Key Words
Chromosome introgression · *Festuca* · Flow cytometry · Genome size estimation · Homoeologous recombination · Karyotype · *Lolium*

Abstract
Crossing over-based recombination is a powerful tool for generating new allelic combinations during sexual reproduction. It usually occurs between homologous chromosomes. However, under some conditions, homoeologues may also be capable of crossing over. Whether homologous and homoeologous crossovers are equivalent and governed by the same rules has never been established. Here we report on chromosome distribution of homoeologous crossovers in a unique system of *Festuca* × *Lolium* hybrids. Unlike in most other hybrids, in these intergeneric hybrids, homoeologous chromosomes are capable of pairing and crossing over with frequencies approaching that of homologues. At the same time, genome divergence makes cytological detection of chromosome recombination feasible. We analyzed the distribution of homoeologous recombination along individual chromosomes in a complete set of intergeneric single chromosome substitutions from *F. pratensis* into tetraploid *L. multiflorum*. Homoeologous recombination sites were not evenly distributed along the chromosomes, being concentrated in intercalary regions of the arms and reduced in proximal and distal regions. Several recombination hotspots and cold spots were found along individual chromosomes and the recombination was not affected by the presence of a secondary constriction. Our results indicate that despite the uneven distribution of homoeologous recombination, introgression of any part of the *F. pratensis* genome into *L. multiflorum* is feasible.

Perennial ryegrass (*Lolium perenne* L.) and Italian ryegrass (*L. multiflorum* Lam.) are among the most agriculturally important grass species in the temperate regions. They are widely used for forage and turf. Both species are cultivated in monocultures as well as in mixtures with other species [Fojtík, 1994; Černoch et al., 2003]. Possible changes in the global climate call for broadening of the gene pool of both species as a buffer against potentially harsher growing conditions. Such broadening of the gene

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1424–8581/10/0000–0000$26.00/0

Accessible online at:
www.karger.com/cgr

D. Kopecký
Institute of Experimental Botany
Sokolovská 6
CZ–77200 Olomouc (Czech Republic)
Tel. +420 585 205 857, Fax +420 585 205 853, E-Mail kopecky@ueb.cas.cz

pool can be achieved by intra- and interspecific hybridization. Because of the generally poor survival rate of both ryegrass species under abiotic stress conditions, the latter strategy appears to be more appropriate [Jauhar, 1993]. Ideal donor species of tolerance to abiotic stresses belong to *Festuca* subgen. *Schedonorus*. Species of this subgenus are closely related to ryegrasses [Catalán et al., 2004], and especially *F. arundinacea* Schreb. and *F. pratensis* Huds. exhibit high tolerance to abiotic stresses, such as drought and freezing [Jauhar, 1993].

The potential of interspecific hybridization for gene transfer via introgression breeding is generally limited by the inability of parental chromosomes to pair and recombine. This is true also in allopolyploid species, which often possess genetic systems preventing homoeologous pairing such as *Ph1* genes in wheat [Riley and Chapman, 1958; Sears and Okamoto, 1958] and *PrBn* in *Brassica napus* [Jenczewski et al., 2003]. However, *Festuca* and *Lolium* species are unique, because their chromosomes pair with each other and recombine freely [Thomas et al., 1994; Kopecký et al., 2008, 2009a].

There are 2 strategies of interspecific breeding. One is the development of amphiploids. A good example of the use of man-made amphiploids is hexaploid triticale (×*Triticosecale* Wittmack), a wheat-rye hybrid. Perhaps the main reason why this amphiploid is agronomically successful is that both parents are cultivated species. When one of the parent species is undomesticated, it tends to contribute to the amphiploid not only desirable but also weedy characteristics and these may require considerable long-term breeding efforts to mitigate. Hence the introgression breeding represents a strategy to introgress only the desirable characteristics into a cultivated species. In this approach, the initial cross is followed by recurrent backcrosses to the recipient species, during which the introgressed region is narrowed down, eliminating the undesired genetic ballast while retaining the specific traits of interest. Clearly, the success of this strategy is dependent on the frequency and distribution of recombination across parental genomes and it is known that the distribution of crossovers along the chromosomes is not random.

In large genomes, such as wheat and barley, the frequency of crossovers increases from centromeres to telomeres [Lukaszewski and Curtis, 1993; Künzel et al., 2000]. However, the increase is not gradual and crossovers preferentially occur in certain regions of chromosomes called 'recombination hotspots'. In contrast to hotspots, there are regions where almost no crossovers occur – the 'recombination cold spots' [Mézard, 2006]. The total number of hotspots in humans was estimated at ~50,000 [Buard and de Massy, 2007]. Meiotic recombination hotspots have also been identified in several plant species, such as *Arabidopsis thaliana* [Mézard, 2006], rice [Kurata et al., 1994], wheat [Gill et al., 1996] and *Lolium/Festuca* [King et al., 2002]. In barley, hotspots were preferentially localized to distal regions of chromosomes, but they were also found in interstitial positions [Künzel et al., 2000]. In *L. perenne*/*F. pratensis* introgression lines, the hotspots were also localized in the distal regions [King et al., 2002], but not in the most distal 10% of the physical chromosome length as in barley [Künzel et al., 2000].

Species-specific distribution of recombination hotspots was described by Winckler et al. [2005]. The authors compared humans and chimpanzees and found that only 8% of recombination hotspots were common to both species. Interestingly, the distribution of recombination hotspots can change during evolution. Several sites with high recombination rates in humans show average or low recombination rates in chimpanzee, and vice versa [Jeffreys et al., 2005]. The size of hotspots can be relatively small; for example in human sperm cells, hotspots of ~2 kb length were identified [Buard and de Massy, 2007]. In plants, regions of extremely high recombination rates may represent only a few kilobases [Dooner and Martinez-Ferez, 1997; Okagaki and Weil, 1997].

The frequency and distribution of recombination events is influenced by intrinsic and extrinsic factors, such as sex [Lenormand and Dutheil, 2005], temperature [Francis et al., 2007], chemical agents and physical stress [Sinha and Helgason, 1969]. Differences in crossover rates in specific regions of the genome were detected among cultivars of many crop species, including barley, suggesting the effect of the genotype [Sall et al., 1990]. Crossover frequency can also be changed by chromosome rearrangements. Thus, even a short terminal deletion of a chromosome arm in *L. perenne* significantly reduced the frequency of crossovers [Jones ES et al., 2002]. In wheat and in rye, crossover distribution can be manipulated by structural aberrations [Sybenga, 1975; Jones LE et al., 2002; Qi et al., 2002], dramatically increasing the crossover saturation rate in designated chromosome segments when they are placed in terminal positions. On the other hand, a whole-arm inversion in rye inverted the pattern of chiasma distribution suggesting that the typical pattern of recombination in cereals, with its concentration at the telomere and absence at the centromere may in fact reflect an innate ability of specific chromosome regions to form crossovers, along the line of the zygomere

concept of Sybenga [1969] and 'pairing centers' of Maguire [1986].

As mentioned earlier, the possibility to introgress just one or a few attractive traits from one species to another (usually an agronomically important crop) is of great importance. For this, an association between the distribution of recombination events and the location of gene-rich regions plays a key role in the success or failure of introgression breeding. In wheat and maize, a high correlation was observed between recombination frequency and the distribution of gene-rich regions [Akhunov et al., 2003a, b; Shah and Hassan, 2005; Anderson et al., 2006]. Similarly in barley, the distribution of recombination hotspots was positively correlated with that of DNA markers assumed to be gene-derived [Künzel et al., 2000]. However, such examples may not represent a general rule. In *Arabidopsis*, most of the recombination events in hotspots were localized to non-coding regions [Drouaud et al., 2006; Mézard, 2006]. Similarly, King et al. [2007] found that most of the functionally annotated genes were located in regions with low recombination rates in *Lolium*/*Festuca* introgression lines.

In this study we performed a detailed analysis of the distribution of homoeologous crossovers at the chromosomal level in *L. multiflorum*/*F. pratensis* substitution lines. As the first step, we established the karyotypes, estimated nuclear genome sizes and determined molecular chromosome sizes in both parental species. Subsequently, we compared the relative lengths of individual homoeologous chromosomes and measured the positions of homoeologous recombination events in each of the 7 chromosomes of *F. pratensis* in *L. multiflorum*/*F. pratensis* substitution lines.

Material and Methods

Plant Material

Nuclear DNA content was estimated in diploid cultivar 'Kolumbus' (2n = 2x = 14) and tetraploid cultivar 'Patra' (2n = 4x = 28) of *F. pratensis*, and in diploid cultivar 'Prolog' (2n = 2x = 14) and tetraploid cultivar 'Lubina' (2n = 4x = 28) of *L. multiflorum*. Moreover, we also estimated nuclear DNA amounts in diploid cultivar 'Handicap' (2n = 2x = 14) and tetraploid cultivar 'Korok' of *L. perenne* and hexaploid cultivar 'Kora' (2n = 6x = 42) of *F. arundinacea*. Five plants of each species were used. With the aim of analyzing the distribution of homoeologous crossovers at the chromosomal level, single chromosome substitution lines of all 7 chromosomes of *F. pratensis* introgressed into the tetraploid *L. multiflorum* were developed [Kopecký et al., 2008]. Plants of individual chromosome substitutions were intercrossed and their progeny was analyzed.

Estimation of Nuclear Genome Size

Seeds were germinated in 12-cm pots filled with substrate in a greenhouse. After 3–4 weeks, we estimated the genome size of 5 plants of each cultivar by flow cytometry following Galbraith et al. [1983]. Nuclear suspensions were prepared by simultaneous chopping of 50 mg leaf tissue of an unknown sample and *Pisum sativum* 'Ctirad', which was used as an internal reference standard with 2C = 9.09 pg [Doležel et al., 1998]. The tissues were homogenized with a razor blade in a Petri dish containing 0.5 ml of Otto I solution (0.1 M citric acid, 0.5% Tween 20 [Otto, 1990]). The nuclear suspension was filtered through 42-μm nylon mesh and stained with 1 ml Otto II solution (0.4 M $Na_2HPO_412H_2O$) containing 2 mg/ml ß-mercaptoethanol, 50 μl/ml propidium iodide and 50 μl/ml RNase IIA. Samples were stained for 10 min at room temperature and analyzed using a CyFlow flow cytometer (Partec GmbH, Münster, Germany) equipped with a green solid-state laser (532 nm). At least 5,000 events were analyzed per sample and only histograms with a coefficient of variation of G_0/G_1 peaks lower than 4.0% were accepted. The analysis of each taxon was repeated on 3 different days. A between-day threshold of 2% was set and samples exhibiting larger variation were excluded and a new sample was analyzed.

Karyotyping of F. pratensis *and* L. multiflorum

To establish karyotypes of *F. pratensis* (cv. 'Kolumbus'; 2n = 2x = 14) and *L. multiflorum* (cv. 'Lubina'; 2n = 4x = 28), we measured chromosome lengths and determined centromeric indices in both species. Individual chromosomes were identified by fluorescence in situ hybridization (FISH) according to Kopecký et al. [2008] with labeled probes for 5S and 45S rDNA and the *F. pratensis* BAC clone 1G18. Chromosomes were numbered following the Triticeae Chromosome Numbering System [Kopecký et al., 2009b]. Chromosomes were counterstained with 1.5 μg/ml 4′,6-diamidino-2-phenylindole (DAPI) prepared in Vectashield antifade solution (Vector Laboratories, Burlingame, Calif., USA). Observations were made using an Olympus AX70 microscope equipped with epifluorescence and SensiCam B/W camera. Chromosomes were measured in 10 metaphase plates in each species using ScionImage software, and mean values were calculated.

Distribution of Homoeologous Crossovers at the Chromosomal Level

Chromosome Preparations and Genomic in situ Hybridization. Single chromosome substitution lines *L. multiflorum*/*F. pratensis* were propagated by intermating within small populations of plants with the same chromosome constitutions, in isolation from all other sources of pollen. Seeds of the progeny thus obtained were germinated in Petri dishes on wet filter paper and seedlings were planted in 30-mm pots in the greenhouse. After 3–4 weeks, plantlets were transferred to a hydroponic culture with aerated solution of Hydroponex at 0.9 g/l (Hu-Ben, Čerčany, Czech Republic). Mitotic metaphase spreads were prepared from root tips according to Masoudi-Nejad et al. [2002]. GISH was performed according to Kopecký et al. [2005]. Total genomic DNA of *F. pratensis* was labeled with digoxigenin using the DIG-Nick Translation Kit (Roche Applied Science, Indianapolis, Ind., USA) and used as a probe. Genomic DNA of *L. multiflorum* was sheared to 200–500-bp fragments by boiling for 45 min and used as blocking DNA. The probe to block ratio was 1:150 with minor variation. Sites of probe hybridization were detected by anti-DIG-FITC con-

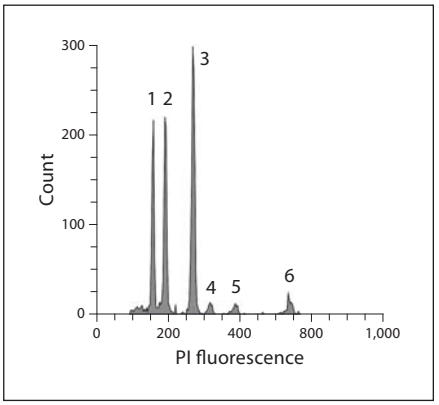


Fig. 1. Histogram of relative fluorescence intensity obtained after simultaneous analysis of nuclei isolated from *Lolium multiflorum* cultivar 'Prolog' (G_0/G_1 nuclei, peak 1; G_2 nuclei, peak 4), *Festuca pratensis* cultivar 'Kolumbus' (G_0/G_1 nuclei, peak 2; G_2 nuclei, peak 5) and *Pisum sativum* cultivar 'Ctirad' (2C = 9.09 pg, as an internal reference standard; G_0/G_1 nuclei, peak 3; G_2 nuclei, peak 6).

jugate (Roche). Fluorescence microscopy was performed as described above and ScionImage and Adobe Photoshop software were used for processing of color pictures.

Distribution of Homoeologous Recombination. To determine the positions of intergeneric crossover points, we measured the length of the introgressed segment(s) and the length of both arms of recombined chromosomes in 5 metaphase plates per plant. Because the absolute lengths of individual pairs of homologues from *Festuca* and *Lolium* are different, a direct calculation of the positions of crossover points along the chromosome length was not possible. Thus, we transformed the measured lengths of the introgressed *Festuca* segments using a calibration factor specific for each chromosome (table 1). The factor was calculated based on the DNA content and relative lengths of homoeologous chromosomes of both species:

Calibration factor = 1CxL × RLF/1CxF × RLL
1CxL = Monoploid genome size of *L. multiflorum*
1CxF = Monoploid genome size of *F. pratensis*
RLL = Relative length of individual *L. multiflorum* chromosome
RLF = Relative length of individual *F. pratensis* chromosome

For each chromosome, the position of over 100 independent homoeologous recombination events was recorded.

Results and Discussion

Estimation of Nuclear Genome Size

Nuclear DNA content of diploid and tetraploid cultivars of *L. multiflorum*, *L. perenne*, *F. pratensis* and of a hexaploid cultivar of *F. arundinacea* was estimated by flow cytometry (table 2). In agreement with the previous study [Loureiro et al., 2007], histograms of relative DNA content contained well-resolved peaks representing nuclei in G_0/G_1 and G_2 phases of the cell cycle (fig. 1). In 91.0% of the assays, the coefficient of variation (CV) of G_0/G_1 peaks was below 3.0% and average CV ranged from 1.96–2.96%. The analyses revealed that monoploid genome sizes (1Cx) [Greilhuber et al., 2005] of *F. pratensis* and *L. multiflorum* were significantly different. While both cultivars of *F. pratensis* measured here had 1Cx = 3.25 pg, diploid *L. multiflorum* cultivar 'Prolog' had 1Cx = 2.62 pg and the tetraploid cultivar 'Lubina' had 1Cx = 2.75 pg. The difference between the 2 *L. multiflorum* cultivars was statistically significant ($p < 0.05$). These observations suggest that, on average, the molecular size of an *F. pratensis* chromosome is 18% larger as compared to an *L. multiflorum* chromosome. This estimate was confirmed independently by the measurement of chromosome lengths (table 3). Estimation of genome size and chromosome lengths in both species was essential to obtain the calibration factor for the analyses of the physical distribution of homoeologous recombination.

With the aim of providing additional data on 1Cx values in *Festuca* and *Lolium* and average chromosome molecular size in the 2 genera, we estimated genome sizes in 1 cultivar of hexaploid *F. arundinacea* and in 2 cultivars of *L. perenne* of different ploidy. The 1Cx value for *F. arundinacea* was 11% lower as compared to that of *F. pratensis* (table 2). This difference could be due to the reduction in genome size known to take place following polyploidization [Leitch and Bennett, 2004], and which was also observed in *Festuca* [Seal, 1983; Loureiro et al., 2007]. On the other hand, the 1Cx values for both cultivars of *L. perenne* (table 2) were statistically not different from the estimates for *L. multiflorum* with the same ploidy level.

Table 1. Calibration factor for the measurements of individual homoeologous chromosomes of *F. pratensis* and *L. multiflorum*

Chromosome	Calibration factor[a]
1	0.792
2	0.863
3	0.877
4	0.862
5	0.801
6	0.767
7	0.914

[a] Molecular size of *L. multiflorum* chromosome/molecular size of *F. pratensis* chromosome. For calculation, see Material and Methods.

Table 2. Estimation of nuclear genome size in *Lolium perenne*, *L. multiflorum*, *Festuca arundinacea* and *F. pratensis*

Species	Cultivar	2C Nuclear DNA amount, pg		1Cx, pg DNA	1Cx, Mb
		mean	SD (n = 5)		
L. multiflorum (2n = 2x = 14)	Prolog	5.25	0.042	2.62[a]	2,567
L. perenne (2n = 2x = 14)	Handicap	5.36	0.047	2.68[a, b]	2,623
L. multiflorum (2n = 4x = 28)	Lubina	10.99	0.092	2.75[b, c]	2,687
L. perenne (2n = 4x = 28)	Korok	11.19	0.057	2.80[c]	2,735
F. arundinacea (2n = 6x = 42)	Kora	17.45	0.078	2.91[d]	2,845
F. pratensis (2n = 2x = 14)	Kolumbus	6.49	0.078	3.25[e]	3,175
F. pratensis (2n = 4x = 28)	Patra	13.01	0.050	3.25[e]	3,181

1Cx = Monoploid genome size. 1 pg DNA = 978 Mb [Doležel et al., 2003].

Small letters label groups with no significant difference at $p < 0.05$ level according to a Bonferroni all pairwise multiple comparisons test.

Table 3. Length, DNA content and molecular size of chromosomes and chromosome arms of *F. pratensis* (cv. 'Kolumbus', 2n = 2x = 14) and *L. multiflorum* (cv. 'Lubina', 2n = 4x = 28)

Chr.	Length μm	% of total genome	1C pg	1C Mb[a]	Chr. arm	Length μm	% of total genome	1C pg	1C Mb[a]	Chr. arm	Length μm	% of total genome	1C pg	1C Mb[a]	CI[b]
F. pratensis Huds.															
1	4.67	11.74	0.382	373	**1S**	1.71	4.30	0.140	137	**1L**	2.96	7.44	0.242	237	37
2	6.07	15.24	0.496	485	**2S**	2.42	6.09	0.198	194	**2L**	3.64	9.15	0.298	291	40
3[c]	6.25	15.70	0.511	499	**3S**	2.92	7.34	0.239	233	**3L**	3.33	8.36	0.272	266	47
4	6.79	17.07	0.555	543	**4S**	3.18	7.99	0.260	254	**4L**	3.61	9.08	0.295	289	47
5	5.04	12.68	0.412	403	**5S**	1.76	4.41	0.143	140	**5L**	3.29	8.27	0.269	263	35
6	4.93	12.38	0.403	394	**6S**	1.97	4.95	0.161	158	**6L**	2.95	7.42	0.241	236	40
7	6.05	15.21	0.495	484	**7S**	2.90	7.28	0.237	231	**7L**	3.16	7.93	0.258	252	48
L. multiflorum Lam.															
1	3.37	11.01	0.303	296	**1S**	1.28	4.18	0.115	112	**1L**	2.09	6.83	0.188	183	38
2[c]	4.77	15.58	0.428	419	**2S**	1.66	5.42	0.149	146	**2L**	3.11	10.16	0.279	273	35
3[c]	4.99	16.30	0.448	438	**3S**	2.37	7.74	0.213	208	**3L**	2.62	8.56	0.235	230	47
4	5.33	17.41	0.479	468	**4S**	2.54	8.30	0.228	223	**4L**	2.79	9.11	0.251	245	48
5	3.68	12.02	0.331	323	**5S**	1.20	3.92	0.108	105	**5L**	2.48	8.10	0.223	218	33
6	3.44	11.23	0.309	302	**6S**	1.47	4.80	0.132	129	**6L**	1.97	6.43	0.177	173	43
7[c]	5.04	16.46	0.453	442	**7S**	2.31	7.54	0.207	203	**7L**	2.73	8.92	0.245	240	46

[a] 1 pg DNA = 978 Mb [Doležel et al., 2003].
[b] CI = Centromeric index (length of the short arm/total chromosome length × 100).
[c] Chromosomes with NOR.

This suggests that the average chromosome molecular size in *L. multiflorum* and *L. perenne* are similar.

The Karyotypes of F. pratensis *and* L. multiflorum

Measurements of chromosome length on 10 metaphase plates in each species established standard karyotypes for the 2 parental species (table 3). Diploid *F. pratensis* has 7 pairs of chromosomes of which 3 are metacentric and 4 are submetacentric; of the metacentric chromosomes, 1 pair (chromosome 3) has a secondary constriction. The length of the chromosomes ranges from 4.67 μm (chromosome 1) to 6.79 μm (chromosome 4) and the total length of the chromosome complement was 39.79 μm. Based on the estimated 1Cx value in *F. pra-*

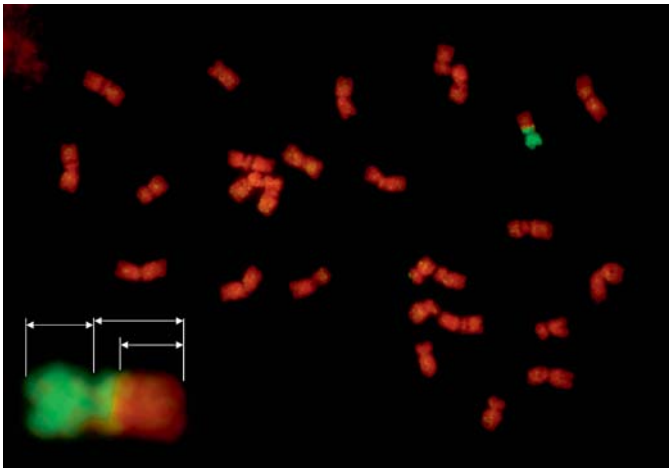


Fig. 2. Molecular cytogenetics analysis (GISH) of chromosomes in recombinant line Bdel21 (2n = 4x = 27), where a segment of *F. pratensis* chromosome 2 has been introgressed into tetraploid *L. multiflorum*. Total genomic DNA of *F. pratensis* was labeled with FITC and used as a probe (green color); unlabeled total genomic DNA of *L. multiflorum* was used as blocking DNA. Chromosomes were counterstained with DAPI (shown in red pseudocolor). Inset shows the chromosome measurement in detail.

tensis, the molecular size of individual chromosomes ranges from approximately 373–543 Mb. Diploid *L. multiflorum* also has 7 pairs of chromosomes; 3 pairs are metacentric and 4 pairs are submetacentric. Three pairs of chromosomes (chromosomes 2, 3 and 7) carry 1 secondary constriction each. The length of individual chromosomes ranges from 3.37 μm in chromosome 1 to 5.33 μm in chromosome 4, and molecular sizes of chromosomes range from 296 Mb (chromosome 1) to 468 Mb (chromosome 4). The length of the entire chromosome complement was 30.62 μm.

Distribution of Homoeologous Recombination Points

To establish the distribution of recombination points along individual chromosomes, we divided each chromosome into 20 segments of equal size. Each segment corresponded to approximately 20 Mb of DNA and its size was chosen considering the spatial resolution of our microscopy-camera system, where 1 pixel represents 0.98–1.43% of a chromosome length (fig. 2). Measurements of recombined chromosomes in 5 different metaphase plates showed that variation in length of up to 5% was not uncommon (data not shown). This variation reflected differences in chromatin condensation among different cells, differences in the labeling efficiency of FISH, and spatial resolution of our optical detection system. Interestingly, King et al. [2002] were able to dissect a chromosome into segments smaller than 0.4% of its entire physical length.

The frequency of recombination events was established for each of the 20 segments in each single chromosome substitution line. The results are graphically displayed in figure 3 in which the chromosome lengths are given in relative values, so that the positions of recombination events are valid for both species. In most of the chromosome arms, the distribution of recombination events was fairly even with a slow gradient of increasing recombination rates towards the telomeres. However, there was a dramatic decrease in recombination at the subtelomeric and telomeric regions. In each chromosome, the lowest recombination rates were in the pericentromeric and centromeric regions. Absence of recombination in the centromeric/pericentromeric and telomeric/subtelomeric regions was not entirely surprising as it has been reported in triploid hybrids *L. multiflorum* × *F. pratensis* and in the pentaploid hybrids *F. arundinacea* × *L. multiflorum* [Zwierzykowski et al., 1998, 1999].

Gradients of recombination along chromosomes were found in other grass species, including wheat [Dvořák and Chen, 1984; Curtis and Lukaszewski, 1991; Erayman et al., 2004], rye [Lukaszewski, 1992], barley [Leitch and Heslop-Harrison, 1993; Künzel et al., 2000] and ryegrass [Hayward et al., 1998]. In barley, Künzel et al. [2000] found that 45% of the physical length of chromosome 3H spanning the centromere showed no recombination at all, while in wheat, Erayman et al. [2004] observed less than 1% of overall chromosome recombination in the 25% around the centromere. Some of the decrease in recombination rates in the vicinity of the telomere may be attributed to detection problems. Given the resolution level of FISH [Lukaszewski et al., 2005] it is possible that very short segments of *F. pratensis* chromatin on *L. multiflorum* chromosomes did not produce bright enough signals to be observed while short terminal segments of *L. multiflorum* on *F. pratensis* chromosomes might have been obscured by the halo of the probe. There are no conceivable technical problems in scoring all remaining intercalary crossover events, including those in the vicinity of the centromere.

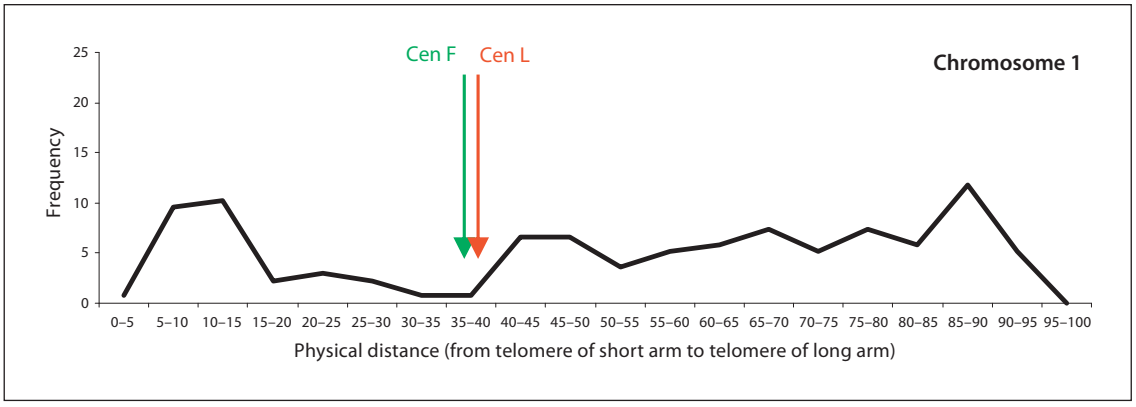
A reduced frequency of recombination was observed to correlate with the presence of the nucleolus organizing regions (NOR) and was observed in NOR-bearing arms such as wheat chromosome 1BS [Dvořák and Chen, 1984], rye chromosome 1RS [Lawrence and Appels, 1986] and chromosome 3 of the *L. perenne*/*F. pratensis* substitution line [King et al., 2002]. The present study indicates that

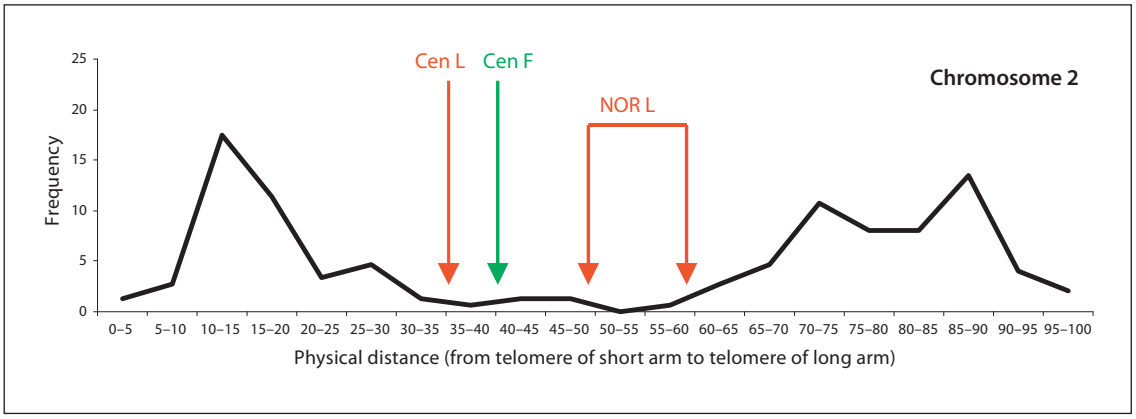
the relationship between the presence of a NOR and recombination may be more complex. While we did not observe any reduction of homoeologous recombination in the NOR region of chromosome 3, chromosome 7 had lower frequency of recombination in the NOR region, and chromosome 2 showed an absence of recombination in the vicinity of the NOR. In the case of chromosome 3 of *F. pratensis*, the differences observed between our results and those of King et al. [2002] might have resulted from the lower sequence homology between *F. pratensis* and *L. perenne* [King et al., 2002] as compared to *F. pratensis* and *L. multiflorum* (this study). In fact, our previous study indicated a closer relationship between the latter 2 species [Kopecký et al., 2009b]. The sequence homology could also explain the differences in recombination observed in chromosome 3 and chromosomes 2 and 7. While chromosome 3 of both parental species *(L. multiflorum* and *F. pratensis)* bears NORs, chromosomes 2 and 7 have NORs only in *L. multiflorum*. Thus, the absence of recombination could be due to structural differences rather than the presence of the NOR per se. This issue clearly deserves further study.

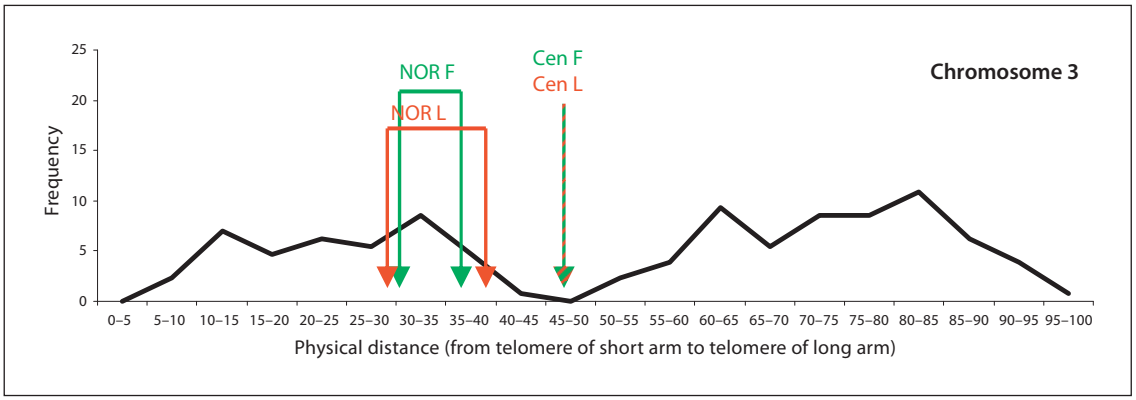
Both *F. pratensis* and *L. multiflorum* are species with typical distal chiasmata producing distinct ring bivalents and end-to-end paired rods [Rees and Dale, 1974; Karp and Jones, 1983]. Thus, the low frequency of homoeologous recombination in the telomeric/subtelomeric regions of *Festuca* × *Lolium* hybrids is difficult to explain. Some part of it could be related to detection problems, as explained earlier. However, the observations of Zwierzykowski et al. [1998, 1999] and King et al. [2002] confirm that this phenomenon is typical for the *Festuca* × *Lolium* hybrids. The decrease in frequency of recombination was also observed at the very distal end of the short arm of wheat chromosome 3B [Saintenac et al., 2009]. As in other chromosomal regions mentioned above, the explanation could be due to lower sequence homology. In hexaploid wheat, Akhunov et al. [2003b] observed a gradient in gene density and recombination rate from the centromere to the telomere of all chromosomes. Higher frequency of recombination may explain a more rapid gene evolution and faster differentiation of homoeologous chromosomes from each other in distal regions [Akhunov et al., 2003a]. As homoeologous chromosomes of cross-pollinating species lose synteny faster than do homoeologues of self-pollinating species [Akhunov et al., 2003a], it is reasonable to expect reduced sequence homology between the distal parts of homoeologous chromosomes of *F. pratensis* and *L. multiflorum*. Moreover, the presence of telomere-associated repeats, which are presumably under low selection pressure, provides another source of sequence variation. Telomeric heterochromatin was suggested as a possible reason for the decrease of recombination in the distal part of the short arm of wheat chromosome 3B [Saintenac et al., 2009]. On the other hand, if the hypothesis of zygomeres or pairing centers is ever confirmed [Sybenga, 1969], many aspects of the particular crossover patterns would be easier to understand and explain.

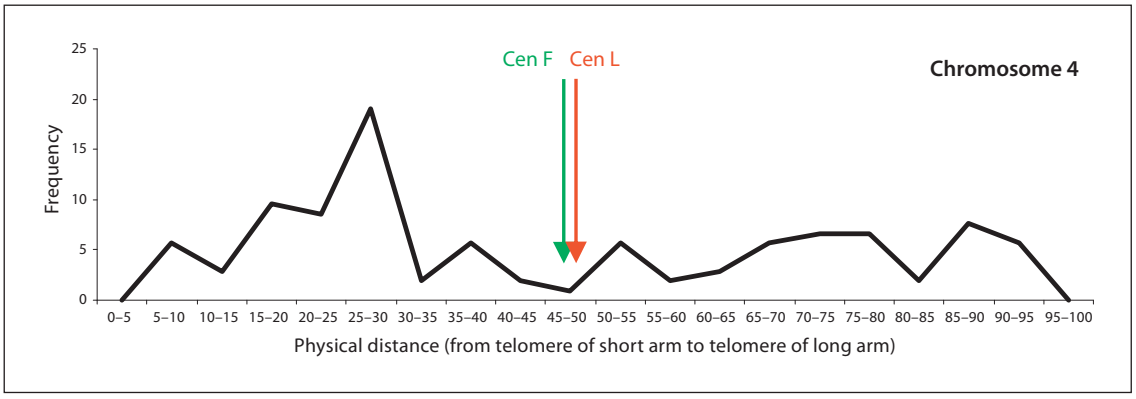
Uneven distribution of recombination along chromosomes was described in a number of studies [Mézard, 2006; Buard and de Massy, 2007]. The molecular nature of the regions with an unusually high or low frequency of recombination is unclear [Akhunov et al., 2003; Mézard, 2006]. In this work, recombination hotspots and cold spots were not frequently observed. Hotspots occurred in the central parts of chromosome arms and were usually located closer to telomeres. Hotspots with a frequency of recombination 4 times higher than the average occur at a physical distance of 25–30% in chromosome 4 and 85–90% in chromosome 6 (fig. 3). Cold spots, or absence of recombination, were generally localized to regions neighboring centromeres. Moreover, several other parts of chromosomes displayed reduced frequency of recombination. Unexpected cold spots occurred for example at a physical distance of 75–80% of chromosome 5 and between 30 and 35% of chromosome 6 (fig. 3). The cumulative length of regions with no recombination, excluding the most distal segments, ranged from 0% in chromosomes 1, 4 and 5 to 15% in chromosomes 6 and 7. The reduction of recombination frequency in the regions around the NOR in chromosomes 2 and 7 was already mentioned.

This study provides perhaps the most detailed map of homoeologous recombination in the *Festuca-Lolium* complex. In general, recombination frequency is the highest in the distal chromosome regions and decreases toward the centromere. In this sense, the pattern is similar to that observed in numerous other species. How valid this pattern is for homologous recombination is not clear. While the *Festuca-Lolium* complex offers a unique opportunity to visualize the parental genomes, chromosomes, and small chromosome segments in large numbers of progenies produced in a variety of ways, no such system exists for homologous recombination. The relationship between homologous and homoeologous recombination has never been established in any convincing form. In wheat, homoeologous recombination follows the same exact pattern of frequency distribution from the telomere toward the centromere, with the exception of


Chromosome 1
Cen F
Cen L
Frequency
0
5
10
15
20
25
0–5 5–10 10–15 15–20 20–25 25–30 30–35 35–40 40–45 45–50 50–55 55–60 60–65 65–70 70–75 75–80 80–85 85–90 90–95 95–100
Physical distance (from telomere of short arm to telomere of long arm)


Chromosome 2
Cen L
Cen F
NOR L
Frequency
0
5
10
15
20
25
0–5 5–10 10–15 15–20 20–25 25–30 30–35 35–40 40–45 45–50 50–55 55–60 60–65 65–70 70–75 75–80 80–85 85–90 90–95 95–100
Physical distance (from telomere of short arm to telomere of long arm)


Chromosome 3
NOR F
NOR L
Cen F
Cen L
Frequency
0
5
10
15
20
25
0–5 5–10 10–15 15–20 20–25 25–30 30–35 35–40 40–45 45–50 50–55 55–60 60–65 65–70 70–75 75–80 80–85 85–90 90–95 95–100
Physical distance (from telomere of short arm to telomere of long arm)


Chromosome 4
Cen F
Cen L
Frequency
0
5
10
15
20
25
0–5 5–10 10–15 15–20 20–25 25–30 30–35 35–40 40–45 45–50 50–55 55–60 60–65 65–70 70–75 75–80 80–85 85–90 90–95 95–100
Physical distance (from telomere of short arm to telomere of long arm)

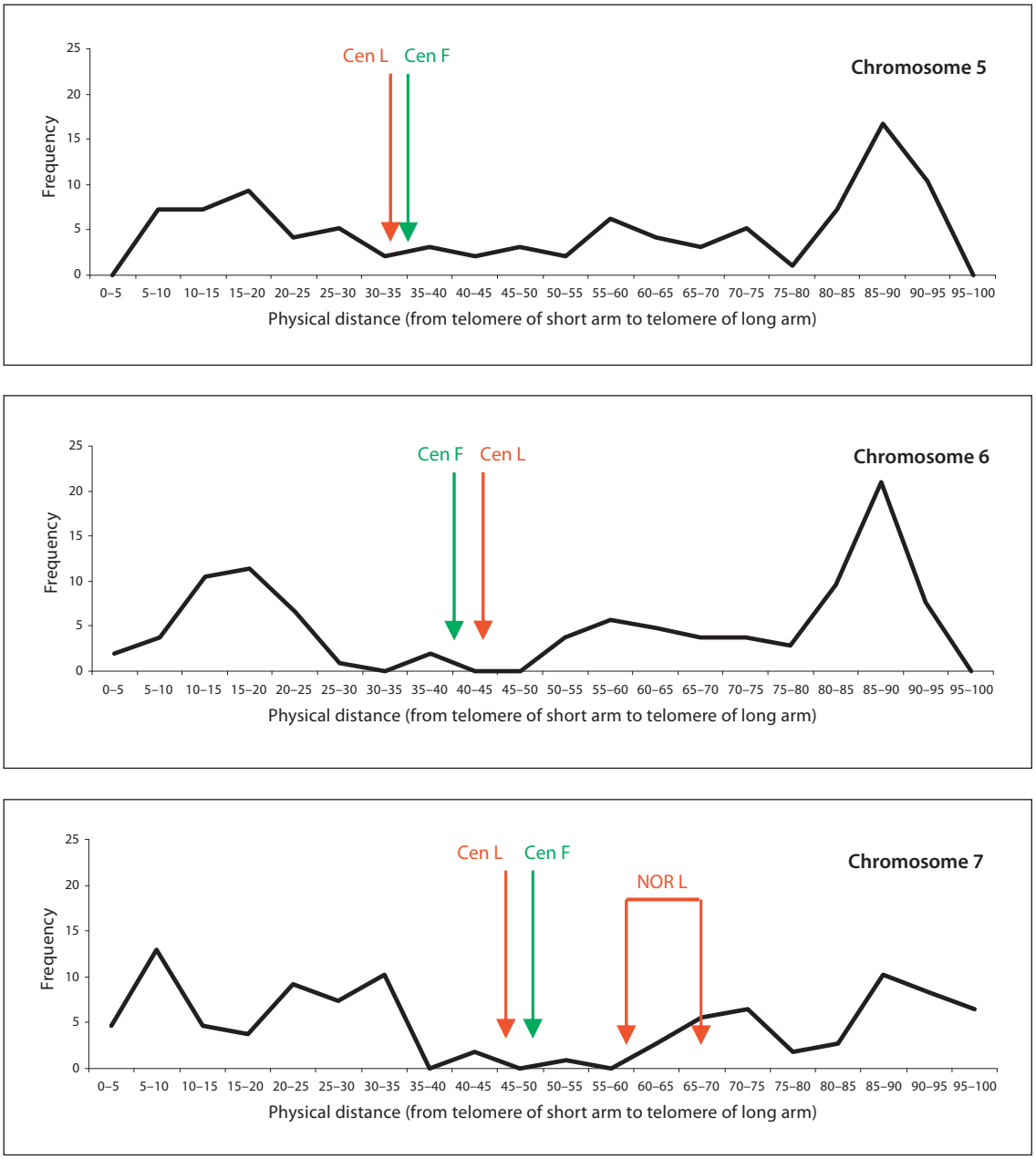


Fig. 3. Distribution of homoeologous recombination events along individual chromosomes of *L. multiflorum*/*F. pratensis* substitutions. Each chromosome is divided into 20 segments, each representing 5% of the physical length. The position of centromeres and NORs in parental species is marked by red *(L. multiflorum)* and green arrows *(F. pratensis)*.

the virtual absence of multiple crossovers per arm [Lukaszewski, 1995], but the pattern does not necessarily indicate much similarity at the DNA level. Moreover, homoeologous recombination in wheat is infrequent, even in the absence of the *Ph1* locus. In the grasses studied here, homoeologues pair and recombine so frequently that pairing preferences require considerable samples to be detected [Kopecký et al., 2008]. From this perspective, chances that similar chromosome regions are involved in homologues as in homoeologues seem higher. But then again, the human-chimpanzee example of recombination hotspots demonstrates that their relative importance

may change over time and so in the species studied here they may not be equivalent or even the same.

While this study detected some regions of higher and lower recombination, crossover events occurred along the entire lengths of chromosomes. Thus, in principle, introgression of any chromatin segment from the donor parental species *(Festuca)* into the recipient one *(Lolium)* should be possible, but not necessarily equally simple. However, introgression breeding can attempt to handle any trait of interest considering that the final success will depend on the ability to control the size of introgressed segments to avoid linkage drag.

Acknowledgement

This work was supported by the Ministry of Agriculture of the Czech Republic (grant award NAZV QH71267) and by the Czech Science Foundation (grant award 521/07/P479).

.

References

Akhunov ED, Akhunova AR, Linkiewicz AM, Dubcovsky J, Hummel D, et al: Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. Proc Natl Acad Sci USA 100:10836–10841 (2003a).

Akhunov ED, Goodyear AW, Geng S, Qi LL, Echalier B, et al: The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. Genome Res 13:753–763 (2003b).

Anderson LK, Lai A, Stack SM, Rizzon C, Gaut BS: Uneven distribution of expressed sequence tag loci on maize pachytene chromosomes. Genome Res 16:115–122 (2006).

Buard J, de Massy B: Playing hide and seek with mammalian meiotic crossover hotspots. Trends Genet 23:301–309 (2007).

Catalán P, Torrecilla P, Rodriguez JAL, Olmstead RG: 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 (2004).

Černoch V, Našinec I, Šramek P: Share of grassland on landscape forming in the Czech Republic. Czech J Genet Plant Breed 39:158–162 (2003).

Curtis CA, Lukaszewski AJ: Genetic linkage between C-bands and storage protein genes of tetraploid wheat. Theor Appl Genet 81:245–252 (1991).

Doležel J, Greilhuber J, Lucretti S, Meister A, Lysak M, et al: Plant genome size estimation by flow cytometry: inter-laboratory comparison. Ann Bot 82:17–26 (1998).

Dooner HK, Martinez-Ferez IM: Recombination occurs uniformly within the bronze gene, a meiotic recombination hotspot in the maize genome. Plant Cell 9:1933–1646 (1997).

Drouaud J, Camilleri C, Bourguignon PY, Canaguier A, Bérard A, et al: Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination 'hotspots'. Genome Res 16:106–114 (2006).

Dvořák J, Chen KY: Distribution of nonstructural variation between wheat cultivars along chromosome 6Bp: evidence from the linkage map and physical map of the arm. Genetics 106:325–333 (1984).

Erayman MD, Sandhu D, Sidhu H, Dilbirligi M, Baenziger PS, Gill KS: Demarcating the gene-rich regions of the wheat genome. Nucleic Acids Res 32:3546–3565 (2004).

Fojtík A: Methods of grass improvement used at the plant breeding station Hladké Životice. Genet Pol 35A:25–31 (1994).

Francis KE, Lam SY, Harrison BD, Bey AL, Berchowitz LE, Copenhaver GP: Pollen tetrad-based visual assay for meiotic recombination in *Arabidopsis*. Proc Natl Acad Sci USA 104: 3913–3918 (2007).

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E: Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 220:1049–1051 (1983).

Gill KS, Gill BS, Endo TR, Taylor T: Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. Genetics 144:1883–1891 (1996).

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

Hayward MD, Forster JW, Jones JG, Dolstra O, Evans C, et al: Genetic analysis of *Lolium* I. Identification of linkage groups and the establishment of a genetic map. Plant Breed 117:451–455 (1998).

Jauhar PP: Cytogenetics of the *Festuca-Lolium* complex. Relevance to Breeding. Monographs on Theor Appl Genet Vol. 18, p 255 (Springer-Verlag, Berlin 1993).

Jeffreys AJ, Neumann R: Factors influencing recombination frequency and distribution in a human meiotic crossover hotspot. Hum Mol Genet 37:601–606 (2005).

Jenczewski E, Eber F, Grimaud A, Huet S, Lucas MO, et al: *PrBn*, a major gene controlling homeologous pairing in oilseed rape *(Brassica napus)* haploids. Genetics 164:645–653 (2003).

Jones ES, Dupal MP, Dumsday JL, Hughes LJ, Forster JW: An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.). Theor Appl Genet 105:577–584 (2002).

Jones LE, Rybka K, Lukaszewski AJ: The effect of a deficiency and a deletion on recombination in chromosome 1BL in wheat. Theor Appl Genet 104:1204–1208 (2002).

Karp A, Jones RN: Cytogenetics of *Lolium perenne*. Part 2. Chiasma distribution in inbred lines. Theor Appl Genet 64:137–145 (1983).

King J, Armstead IP, Donnison IS, Thomas HM, Jones RM, et al: Physical and genetic mapping in the grasses *Lolium perenne* and *Festuca pratensis*. Genetics 161:315–324 (2002).

King J, Armstead IP, Donnison IS, Harper JA, Roberts LA, et al: Introgression mapping in the grasses. Chromosome Res 15:105–113 (2007).

Kopecký D, Lukaszewski AJ, Doležel J: Genomic constitution of Festulolium cultivars released in the Czech Republic. Plant Breed 124:454–458 (2005).

Kopecký D, Lukaszewski AJ, Doležel J: Meiotic behavior of individual chromosomes of *Festuca pratensis* in tetraploid *Lolium multiflorum*. Chromosome Res 16:987–998 (2008).

Kopecký D, Bartoš J, Zwierzykowski Z, Doležel J: Chromosome pairing of individual genomes in tall fescue (*Festuca arundinacea* Schreb.), its progenitors, and hybrids with Italian ryegrass (*Lolium multiflorum* Lam.). Cytogenet Genome Res 124:170–178 (2009a).

Kopecký D, Bartoš J, Lukaszewski AJ, Baird JH, Cernoch V, et al: Development and mapping of DArT markers within the *Festuca - Lolium* complex. BMC Genomics 10:473 (2009b).

Künzel G, Korzum L, Meister A: Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics 154:397–412 (2000).

Kurata N, Nagamura Y, Yamamoto K, Harushima Y, Sue N: A 300-kilobase interval genetic map of rice including 883 expressed sequences. Nat Genet 8:365–372 (1994).

Lawrence GJ, Appels R: Mapping the nucleolar organiser region, seed protein loci and isozyme loci on chromosome 1R in rye. Theor Appl Genet 71:742–749 (1986).
Leitch IJ, Bennett MD: Genome downsizing in polyploid plants. Biol J Linn Soc 82:651–663 (2004).
Leitch IJ, Heslop-Harrison JS: Physical mapping of 4 sites of 5S rRNA sequences and one site of the *α-amylase-2* gene in barley *(Hordeum vulgare)*. Genome 36:517–523 (1993).
Lenormand T, Dutheil J: Recombination difference between sexes: a role for haploid selection. PLoS Biol 3:e63 (2005).
Loureiro J, Kopecký D, Castro S, Santos C, Silveira P: Flow cytometric and cytogenetic analyses of Iberian Peninsula *Festuca* spp. Plant Syst Evol 269:89–105 (2007).
Lukaszewski AJ: A comparison of physical distribution of recombination in chromosome 1R in diploid rye and hexaploid triticale. Theor Appl Genet 83:1048–1053 (1992).
Lukaszewski AJ: Physical distribution of translocation breakpoints in homoeologous recombinants induced by the absence of the *Ph1* gene in wheat and triticale. Theor Appl Genet 90:714–719 (1995).
Lukaszewski AJ, Curtis CA: Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. Theor Appl Genet 86:121–127 (1993).
Lukaszewski AJ, Lapinski B, Rybka K: Limitations of in situ hybridization with total genomic DNA in routine screening for alien introgressions in wheat. Cytogenet Genome Res 109:373–377 (2005).
Maguire MP: The pattern of pairing that is effective for crossing over in complex B-A chromosome rearrangements in maize. III. Possible evidence for pairing centers. Chromosoma 94:71–85 (1986).
Masoudi-Nejad A, Nasuda S, McIntosh RA, Endo TR: Transfer of rye chromosome segments to wheat by a gametocidal system. Chromosome Res 10:349–357 (2002).
Mézard C: Meiotic recombination hotspots in plants. Biochem Soc Trans 34:531–534 (2006).
Okagaki RJ, Weil CF: Analysis of recombination sites within the maize waxy locus. Genetics 147:815–821 (1997).
Otto F: DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA, in Crissman HA, Darzynkiewicz Z (eds): Methods in Cell Biology, vol. 33, pp 105–110 (Academic Press, New York 1990).
Qi LL, Friebe B, Gill BS: A strategy for enhancing recombination in proximal regions of chromosomes. Chromosome Res 10:645–654 (2002).
Rees H, Dale PJ: Chiasmata variability in *Lolium* and *Festuca* populations. Chromosoma 47: 335–351 (1974).
Riley R, Chapman V: Genetic control of the cytologically diploid behaviour of hexaploid wheat. Nature 182:713–715 (1958).
Saintenac C, Falque M, Martin OC, Paux E, Feuillet C, Sourdille P: Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). Genetics 181: 393–403 (2009).
Sall T, Flink J, Bengtsson BO: Genetic-control of recombination in barley. 1. Variation in recombination frequency measured with inversed heterozygotes. Hereditas 112:157–170 (1990).
Seal AG: DNA variation in *Festuca*. Heredity 50: 225–236 (1983).
Sears ER, Okamoto M: Intergenomic chromosome relationships in hexaploid wheat. Proc 10th Int Congr Genet, Montreal 2:258–259 (1958).
Shah MM, Hassan A: Distribution of genes and recombination on wheat homoeologous group 6 chromosomes: a synthesis of available information. Mol Breed 15:45–53 (2005).
Sinha RP, Helgason SB: Action of actinomycin D and diepoxybutane on recombination of 2 closely linked loci in *Hordeum*. Canad J Genet Cytol 11:745 (1969).
Sybenga J: The zygomere as hypothetical unit of chromosome pairing initiation. Genetica 37: 186–198 (1969).
Sybenga J: Meiotic Configurations (Springer-Verlag, Berlin 1975).
Thomas HM, Morgan WG, Meredith MR, Humphreys MW, Thomas H, Leggett JM: Identification of parental and recombined chromosomes in hybrid derivatives of *Lolium multiflorum* × *Festuca pratensis* by genomic in situ hybridization. Theor Appl Genet 88: 909–913 (1994).
Winckler W, Myers SR, Richter DJ: Comparison of fine-scale recombination rates in humans and chimpanzees. Science 308:107–111 (2005).
Zwierzykowski Z, Lukaszewski AJ, Lesniewska A, Naganowska B: Genomic structure of androgenic progeny of pentaploid hybrids, *Festuca arundinacea* × *Lolium multiflorum*. Plant Breed 117:457–462 (1998).
Zwierzykowski Z, Lukaszewski AJ, Naganowska B, Lesniewska A: The pattern of homoeologous recombination in triploid hybrids of *Lolium multiflorum* with *Festuca pratensis*. Genome 42:720–726 (1999).

APPENDIX IV

A chromosomal genomics approach to assess and validate the *desi* and *kabuli* draft chickpea genome assemblies

Ruperao P., Chan Ch.-K.K., Azam S., Karafiátová M., Hayashi S., Čížková J., Saxena R.K., Song Ch., Vrána J., Chitikineni A., Visendi P., Gaur P.M., Milán T., Singh K.B., Taran B., Wang J., Batley J., Doležel J., Varshney R.K., Edwards D.

Plant Biotechnology Journal, DOI:10.1111/pbi.12182, 2014

IF: 5.181

Plant Biotechnology Journal

Plant Biotechnology Journal (2014), pp. 1–9 doi: 10.1111/pbi.12182

A chromosomal genomics approach to assess and validate the *desi* and *kabuli* draft chickpea genome assemblies

Pradeep Ruperao[1,2,3], Chon-Kit Kenneth Chan[1], Sarwar Azam[3], Miroslava Karafiátová[4], Satomi Hayashi[1], Jana Čížková[4], Rachit K. Saxena[3], Hana Šimková[4], Chi Song[5], Jan Vrána[4], Annapurna Chitikineni[3], Paul Visendi[1,2], Pooran M. Gaur[3], Teresa Millán[6], Karam B. Singh[7,8], Bunyamin Taran[9], Jun Wang[5], Jacqueline Batley[1], Jaroslav Doležel[4,*], Rajeev K. Varshney[3,*] and David Edwards[1,2,*]

[1]*University of Queensland, St. Lucia, Queensland, Australia*

[2]*Australian Centre for Plant Functional Genomics, University of Queensland, St. Lucia, Queensland, Australia*

[3]*International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, Andhra Pradesh, India*

[4]*Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc-Holice, Czech Republic*

[5]*Beijing Genomics Institute (BGI), Shenzhen, China*

[6]*Department of Genetics, University of Cordoba, Cordoba, Spain*

[7]*The University of Western Australia Institute of Agriculture, The University of Western Australia, Crawley, Australia*

[8]*CSIRO Plant Industry, Private Bag 5, Wembley, WA, Australia*

[9]*Crop Development Centre, Plant Sciences, University of Saskatchewan,Saskatoon, Saskatchewan, Canada*

Received 2 October 2013;
revised 21 January 2014;
accepted 9 February 2014.
**Correspondence* (Tel +420 585 238 703, +91 4030713305 and +61 7 3346 7084; fax +420 585 238 704, +91 4030713071 and +61 7 3365 1176; emails dolezel@ueb.cas.cz; r.k.varshney@cgiar.org; Dave.Edwards@uq.edu.au)

Keywords: chickpea, genome assembly, cytogenetics, cicer.

Summary

With the expansion of next-generation sequencing technology and advanced bioinformatics, there has been a rapid growth of genome sequencing projects. However, while this technology enables the rapid and cost-effective assembly of draft genomes, the quality of these assemblies usually falls short of gold standard genome assemblies produced using the more traditional BAC by BAC and Sanger sequencing approaches. Assembly validation is often performed by the physical anchoring of genetically mapped markers, but this is prone to errors and the resolution is usually low, especially towards centromeric regions where recombination is limited. New approaches are required to validate reference genome assemblies. The ability to isolate individual chromosomes combined with next-generation sequencing permits the validation of genome assemblies at the chromosome level. We demonstrate this approach by the assessment of the recently published chickpea *kabuli* and *desi* genomes. While previous genetic analysis suggests that these genomes should be very similar, a comparison of their chromosome sizes and published assemblies highlights significant differences. Our chromosomal genomics analysis highlights short defined regions that appear to have been misassembled in the *kabuli* genome and identifies large-scale misassembly in the draft *desi* genome. The integration of chromosomal genomics tools within genome sequencing projects has the potential to significantly improve the construction and validation of genome assemblies. The approach could be applied both for new genome assemblies as well as published assemblies, and complements currently applied genome assembly strategies.

Introduction

Efforts to sequence and characterize crop genomes have been boosted in recent years by unprecedented developments in next-generation DNA sequencing (NGS). These technologies have dramatically reduced the cost of generating genome sequence data and present exciting new opportunities for crop genetics and breeding (Edwards and Batley, 2010; Varshney *et al.*, 2009). NGS technologies, currently dominated by the Illumina sequencing platforms, have seen a steady increase in read length, data quality and data quantity since their introduction less than a decade ago. The bioinformatics analysis of this data has been a challenge (Batley and Edwards, 2009); however, an increasing number of tools are now available to interrogate and analyse these data (Lai *et al.*, 2012b; Lee *et al.*, 2012; Marshall *et al.*, 2010).

One consequence of the growth of genome sequencing projects is a general decrease in accepted genome quality. The aim of any genome sequencing project should be to produce a genome that is fit for purpose, and often rough drafts are all that are required to answer important biological questions. Basic assemblies that produce the sequence of all genes, promoters and low copy or unique regions are relatively inexpensive and provide valuable biological insights, while more robust pseudomolecule assemblies have greater utility in the identification of gene variation underlying traits, and for use in genomics-assisted

Please cite this article as: 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. and Edwards, D. (2014) A chromosomal genomics approach to assess and validate the *desi* and *kabuli* draft chickpea genome assemblies. *Plant Biotechnol. J.*, doi: 10.1111/pbi.12182

breeding (Duran *et al.*, 2010; Varshney *et al.*, 2005). However, the production of valid pseudomolecules representing individual chromosomes is the ultimate aim of many genome projects and remains a significant challenge, even in the age of NGS (Imelfort and Edwards, 2009).

Since the sequencing of the first plant genome, *Arabidopsis thaliana* (Arabidopsis_Genome_Initiative, 2000), and the first crop genome, rice (Yu *et al.*, 2002), genome sequencing methods have advanced significantly (Berkman *et al.*, 2012a; Edwards and Batley, 2010; Edwards and Wang, 2012; Edwards *et al.*, 2013). Maize was the first large crop genome to be published (Schnable *et al.*, 2011), and maize genome resequencing has demonstrated a huge diversity in the genome structure between different varieties. Other less complex crop genomes have been sequenced, including the 1.1 Gbp soybean genome (Schmutz *et al.*, 2010) and the 844 Mbp autotetraploid genome of potato (Xu *et al.*, 2011). The soybean genome was sequenced using a whole-genome shotgun approach, while the relatively small potato genome was resolved by sequencing a homozygous doubled-monoploid potato clone using data from the Illumina and Roche 454 platforms.

Generating draft genome sequence assemblies of the simpler crop genomes, such as pigeonpea, are feasible and almost routine using whole-genome shotgun sequencing and Illumina sequencing technology (Varshney *et al.*, 2012). For complex genomes such as bread wheat, the complexity and size of the 17 Gbp genome, comprising three homoeologous subgenomes, necessitates alternative approaches to whole-genome *de novo* sequencing. These include the isolation of individual chromosome arms using flow cytometry and a two-stage sequencing approach which aims to initially generate draft shotgun assemblies of individual isolated chromosome arms (Berkman *et al.*, 2011, 2012b, 2013; Hernandez *et al.*, 2012), followed by the sequencing of BAC tiling paths representing each of these arms (Lai *et al.*, 2012a). The highly complex canola genome, which combines polyploidy with recent triplication in the diploid progenitors, presents a significant challenge for assembly. A public assembly of one diploid progenitor genome was published in 2011 (Wang *et al.*, 2011), while the second is near completion (http://www.brassica.info/). An initial draft genome for canola was produced in 2009, although this remains proprietary and efforts are currently underway to produce a public canola genome sequence (http://www.brassica.info/).

Chickpea (*Cicer arietinum*) is the second most important grain legume crop in the world, grown on about 12 million hectares in Asia, Latin America and Australia. This crop is represented by two main market types: large seeded *kabuli* and small seeded *desi*. These two types share a common ancestry, with *kabuli* evolving from *desi* in the Mediterranean basin, with subsequent selection for traits such as flower colour and seed tannins (Jana and Singh, 1993; Maesen, 1972; Moreno and Cubero, 1978). Genome assemblies have recently become available for both *kabuli* (Varshney *et al.*, 2013) and *desi* (Jain *et al.*, 2013) types. Surprisingly, these genome assemblies appear to be significantly different. To resolve these differences, we have developed and applied a chromosomal genomics approach for genome assembly validation. Using flow cytometry, we isolated individual chromosomes of chickpea for the generation of Illumina NGS sequence data. Mapping the resulting sequence reads from isolated *kabuli* and *desi* chickpea chromosomes to the reference genome assemblies allowed us to assess the quality of assembly of the two published genome sequences.

Results and discussion

Estimation of nuclear genome size

Knowledge of genome size is critical to estimate the quality of a genome sequence assembly. To estimate the genome size of both *desi* and *kabuli* chickpea types, we used DNA flow cytometry, which is currently considered the most reliable method (Doležel and Bartoš, 2005). This analysis revealed that chickpea has a medium-sized genome of less than 900 Mbp and that both types of chickpea do not differ significantly in genome size (Table 1). Our estimates are similar to the 1.9 pg DNA/2C (929 Mbp/1C) reported by Bennett and Smith (Bennett and Smith, 1976), greater than the kmer-based estimate of CDC Frontier (Varshney *et al.*, 2013), but significantly lower than the average 2C value of 3.41 pg DNA as predicted by Ohri and Pal (Ohri and Pal, 1991). This difference may be attributed to different methods (Feulgen microdensitometry was used in the older study) and to different reference standards (Doležel and Bartoš, 2005). Nevertheless, it is worth noting that Ohri and Pal (Ohri and Pal, 1991) did not observe significant differences in genome size between *kabuli* and *desi*.

Comparison of the published *kabuli* and *desi* chickpea draft genomes

An initial comparison of assembly statistics for the two draft chickpea genomes suggests differences in assembly quality. Both draft genomes were assembled from NGS data. The *kabuli* assembly was constructed mostly from Illumina data (Varshney *et al.*, 2013) supported by BAC-end sequences generated using Sanger-based methods, while the *desi* assembly applied a hybrid approach, combining Roche/454 and Illumina data. The *kabuli* assembly captured 532 Mbp (60.3% of the estimated genome size) in scaffolds greater than 1000 bp compared to 519 Mbp for *desi* (59.8% of the estimated genome size) in scaffolds greater than 200 bp. Thus, both assemblies represented similar genome fractions. However, the *desi* genome assembly was far more fragmented, with a total of 32 935 scaffolds greater than 1000 bp and an N50 of 106 Kbp, compared to 7163 scaffolds and an N50 of 39 989 Kbp for *kabuli* (Table 2). The method applied to place the scaffolds into pseudomolecules was similar for both genomes, although genotyping by sequencing (GBS) markers were included to validate the *kabuli* assembly.

Pairwise comparison of each of the pseudomolecules from the two assemblies revealed numerous structural variations (Figure 1). These differences include both long and short regions where the orientations of the sequence differed, for example the region from 9.33 Mb to 24.96 Mb on *kabuli* pseudomolecule

Table 1 Estimation of 2C DNA amounts and genome size in chickpea

Cultivar / Genotype	Type	2C DNA amount (pg)		Mean genome size (Mbp/1C)
		Mean	±SD	
ICC 1882	*desi*	1.773	0.012	867
ICC 283	*desi*	1.741	0.009	851
ICC 8261	*desi*	1.793	0.009	877
ICC 4958*	*desi*	1.775	0.008	868
CDC Frontier*	*kabuli*	1.803	0.007	882

*Genotypes used for chromosome sorting and sequencing.

Table 2 A comparison of *desi* and *kabuli* reference genome assembly statistics

Features	*desi* draft genome	*kabuli* draft genome
Total assembly size (Mb)	456 (52.5%*)	532 (60.3%**)
Number of scaffolds	32 935	7163
Minimum reported scaffold length (bp)	1000	1000
Maximum scaffold length (Kbp)	23 376	59 460
Average scaffold length (bp)	13 857	74 311
N50 length (Kb)	106	39 989
GC content (%)	25.6	30.8%
Genome captured in pseudomolecule (Kbp)	124 386 (14.33%*)	347 247 (39.37%**)
Protein coding genes	27 571	28 269
Average gene length (bp)	3122	3055
Average coding sequence length (bp)	962	1166

*Considering 1C = 868 Mbp (Table 1).

**Considering 1C = 882 Mbp (Table 1).

Ca1 is inverted compared to the equivalent region on the *desi* assembly. There were differences in the position of regions within a pseudomolecule, for example the first half of *desi* pseudomolecule Ca5 is inverted and matches the centre of *kabuli* pseudomolecule Ca5. Of particular interest, we observed several large regions of similarity between unrelated pseudomolecules. These include *desi* pseudomolecule Ca8 matching a region at the start of *kabuli* pseudomolecule Ca7, while *kabuli* pseudomolecule Ca8 matches the last third of *desi* pseudomolecule Ca3. A large portion of *kabuli* pseudomolecule Ca6 matched the second half of *desi* pseudomolecule Ca2. These differences suggest misassembly of one or both draft genome assemblies.

Isolation and sequencing of chickpea chromosomes

To assess and validate the assembled pseudomolecules from the two genome assemblies, we isolated and sequenced individual chromosomes from both *kabuli* and *desi* varieties of chickpea and mapped the resulting sequence reads to the published reference assemblies. For shotgun sequencing, all chromosomes were flow sorted from the sequenced reference *kabuli* 'CDC Frontier', with chromosomes D and E sorted together as a group, while chromosomes A, B and H were flow sorted from the sequenced reference *desi* 'ICC 4958', (See Appendix 1 for details). DNA from these isolated chromosomes was amplified to produce samples suitable for sequencing using Illumina technology. All chromosome isolates could be sorted at high purity from both genotypes as determined by microscopic observation.

Estimation of molecular sizes of chickpea chromosomes

We estimated the molecular size of individual chromosomes based on relative chromosome lengths at mitotic metaphase. The results indicate differences in size between *desi* and *kabuli* chromosomes as large as 10 Mbp for chromosomes A and B and as small as several hundred Kbp for chromosome F (Table 3). Although the differences between the two types of chickpea may be ascribed in part to differences in chromatin condensation, they correspond well to differences between flow karyotypes of *desi* and *kabuli* and differences in chromosome peak positions (Figure 2). For example, chromosomes F and G of *desi* 'ICC 4958' differ by about 7 Mbp (7%), and their peaks cannot be discriminated based on flow karyotype. In *kabuli* 'CDC Frontier', the two chromosomes differ by about 10 Mbp (11%) and can be discriminated.

Comparison of pseudomolecule assemblies

A much greater portion of the *kabuli* assembly could be placed into pseudomolecules (347 247 Kbp) compared with *desi* (124 386 Kbp). The length of each of the pseudomolecules for *kabuli* was higher than for *desi,* and the pseudomolecules represented 39.37% and 14.33% of the estimated genome size in *kabuli* and *desi*, respectively (Table 2). Individual pseudomolecules differed in size and their representation of their predicted chromosome size (Table 4). Striking discrepancies were observed for *kabuli* chromosomes A, B and H, whose pseudomolecules represented on average only about 26% of their predicted size, compared to an average 50%. The smaller than expected pseudomolecule size of these three chromosomes could be explained by the presence of satellite CaRep2 on chromosomes A and B, satellite CaSat2 on chromosomes A and H, and the 45S rDNA locus on chromosome A (Zatloukalová *et al.*, 2011). These highly repetitive regions are likely to collapse into shorter representative regions during de Bruijn graph-based whole-genome assembly.

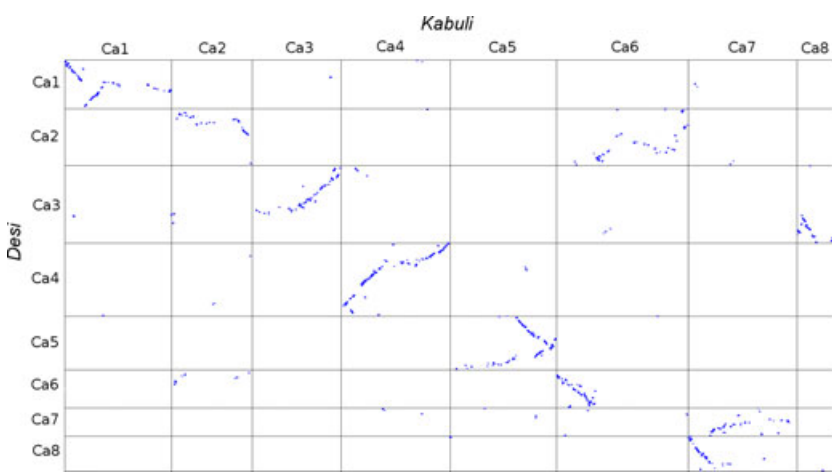


Figure 1 Dot plot matrix of a comparison of the *kabuli* and *desi* and *kabuli* draft chromosome assemblies.

Table 3 Chickpea *desi* and *kabuli* chromosome nomenclature, their assignment to peak on flow karyotypes, linkage groups, corresponding pseudomolecules and molecular chromosome sizes as determined cytologically

Cicer arietinum								
Desi 'ICC 4958'		*Kabuli* 'CDC Frontier'			Relative chromosome length [%]		Molecular chromosome size [Mbp]*	
Peak	Chromosome	Peak	Chromosome	Pseudomolecule	*Desi* '4958'	*Kabuli* 'Frontier'	*Desi* '4958'	*Kabuli* 'Frontier'
I	H	I	H	Ca8	7.2	7.8	62.5	68.80
II	G	II	G	Ca1	9.9	9.4	85.93	82.91
	F	III	F	Ca2	10.7	10.5	92.88	92.61
III	E	IV	E	Ca4	11.5	11.1	99.82	97.90
IV	D		D	Ca7	12.6	11.8	109.37	104.01
	C	V	C	Ca6	13.2	12.8	114.58	112.90
V	B	VI	B	Ca3	15.8	16.7	137.14	147.29
VI	A	VII	A	Ca5	19.0	19.8	164.92	174.64

*Calculated based on nuclear genome size and relative chromosome length.

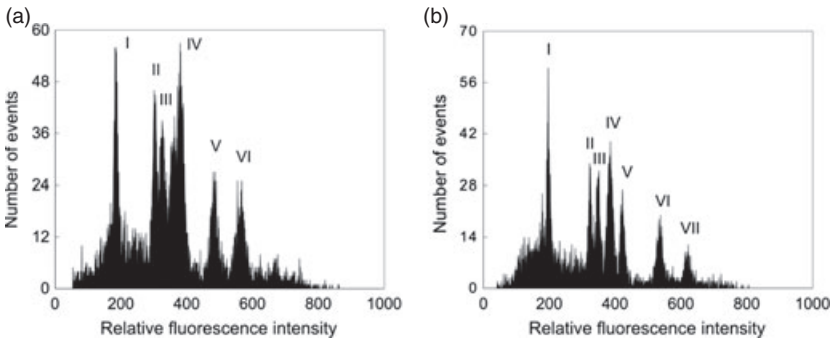


Figure 2 Histograms of relative fluorescence intensity obtained after flow cytometric analysis of DAPI-stained liquid suspensions of mitotic metaphase chromosomes prepared from chickpea *desi* 'ICC 4958' (a) and *kabuli* 'CDC Frontier' (b). For chromosome assignment of peaks on flow karyotypes, please see Table 3.

Table 4 Pseudomolecule size and percentage of predicted chromosome size

Pseudomolecule (chromosome)	*Desi* (%)	*Kabuli* (%)
Ca1 (G)	14 791 696 (15.9)	48 359 943 (52.2)
Ca2 (F)	17 304 114 (20.1)	36 634 854 (44.1)
Ca3 (B)	23 376 002 (17.0)	39 989 001 (27.1)
Ca4 (E)	22 093 647 (22.1)	49 191 682 (50.2)
Ca5 (A)	16 301 343 (9.8)	48 169 137 (27.5)
Ca6 (C)	11 482 212 (10)	59 463 898 (52.6)
Ca7 (D)	8 461 617 (7.7)	48 961 560 (47.0)
Ca8 (H)	10 574 966 (16.9)	16 477 302 (23.9)

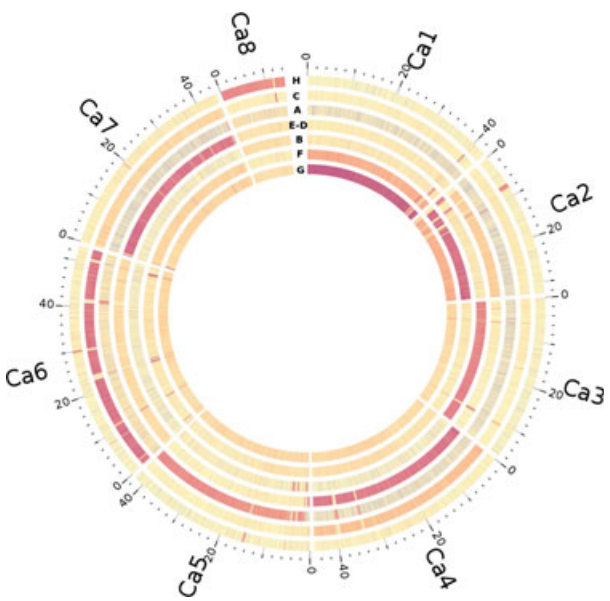


Figure 3 Circos heat map of the *kabuli* reference pseudomolecules demonstrating density of mapped Illumina paired sequence reads (red colour) from isolated *kabuli* chromosomes G, F, B, (E,D), A, C and H.

The *kabuli* reference contains short defined misassembled regions

Mapping each of the *kabuli* isolated chromosome sequence data sets to the *kabuli* reference genome assembly demonstrated that the majority of the reads matched to their respective pseudomolecule with the exception that chromosome F and G reads map to pseudomolecules Ca2 and Ca1, respectively, the inverse of the earlier assignments to genetic linkage experiments (Millan *et al.*, 2010; Thudi *et al.*, 2011; Zatloukalová *et al.*, 2011). Inspection of the read mapping density (Figure 3) suggested that chromosome F data included sequences specific for pseudomolecule G and *vice versa*. Chromosome C and the chromosome D/E group also shared contamination, while chromosomes A, B and H demonstrated a greater purity. The proportion of contamination of chromosome isolates with other chromosomes matched what was expected from the isolation method, with contamination between chromosomes from adjacent flow-sorted peaks.

In addition to the cross-mapping of reads due to chromosomal contamination, we observed regions in the reference pseudomolecules where few reads mapped from the respective chromosome sequence data (Figure 3). For example, a region from 40 141 642 to 40 436 753 bp on pseudomolecule Ca1 had very few reads mapping from the corresponding isolated chromosome G. Interestingly, this region had high mapped read coverage from isolated chromosome C (Ca6). A similar pattern was observed for other gaps across the pseudomolecules and suggests that there are numerous small regions across the *kabuli* pseudomolecule assembly which were misplaced. In total, we observed 46 regions ranging in size from 57 to 1371 Kbp and representing 16 164 Kbp (3.0%) of the pseudomolecule assemblies that were placed into the wrong pseudomolecule (Table 5). Pseudomolecule Ca8 appears to be the most accurate assembly with only a single region of 341 Kbp which should be located on pseudomolecule Ca6 (Figure 4). In contrast, pseudomolecule Ca6 contains 11 blocks of sequence which should be relocated onto other pseudomolecules.

Some misassembled regions appeared to be contigs misplaced during the scaffolding process, while others appeared within contigs suggesting chimeric contig assembly. Many of the misassembled regions were also flanked by highly repetitive retrotransposon sequences, although there was no clear correlation between the presence of these sequences and the type of misassembly.

An advantage of applying chromosomal genomics approaches to identify genome misassembles is the exceptional resolution provided by NGS read mapping. This resolution will greatly facilitate the relocation of these regions into their correct pseudomolecule. One of the limitations of this approach, however, is the inability to identify intrachromosomal misassembly or misassembles between chromosomes which cannot be separated physically by flow sorting. In this case, chromosomes D and E could only be isolated as a pool, and while we identified several regions on these chromosomes which should be placed on other chromosomes, we could not identify chromosome E (Ca4) regions which were misplaced onto pseudomolecule Ca7 (D) and *vice versa*.

Large-scale misassemblies in *desi* reference genome

To determine whether the differences between the two draft genome sequences reflect true structural genome variation or pseudomolecule misassembly, we isolated and sequenced chromosomes A, B and H from *desi* type chickpea and mapped these reads, together with the related *kabuli* chromosome-specific reads to the *desi* reference pseudomolecules (Figure 5) as well as the kabuli pseudomolecules (Figure S1). Sequence reads from both *desi* and *kabuli* isolated chromosomes demonstrated almost identical mapping patterns on the pseudomolecules suggesting that the physical genomes, at least for these three chromosomes, are highly similar between *desi* and *kabuli*. In contrast to the results from mapping *kabuli* chromosome reads to the *kabuli* pseudomolecules, we observed that the chromosome B (Ca3) reads from *kabuli* and *desi* only matched the first portion of *desi* pseudomolecule Ca3. Sequence reads from isolated chromosome H (Ca8) preferably mapped to the remaining portion of pseudomolecule Ca3 and not to pseudomolecule Ca8. This analysis suggested that the observed differences between the *desi* and *kabuli* reference genome assemblies are not due to structural genome differences but are due to misassembly of the *desi* reference genome.

Table 5 Positions and sizes of misassembled genome blocks on the pseudomolecules of *kabuli* chickpea CDC Frontier, together with their correct chromosome (pseudomolecule) location

Pseudo-molecule	Start	End	Length (bp)	Chromosome (pseudomolecule)
Ca1	17 709 355	17 768 444	59 089	H (Ca8)
Ca1	39 419 875	39 639 265	219 390	F (Ca2)
Ca1	39 875 495	40 137 184	261 689	F (Ca2)
Ca1	40 141 642	40 436 753	295 111	C (Ca6)
Ca1	40 737 888	41 342 099	604 211	B (Ca3)
Ca2	1	1 370 632	1 370 631	B (Ca3)
Ca2	4 000 604	4 701 466	700 862	B (Ca3)
Ca2	5 875 970	5 981 305	105 335	G (Ca1)
Ca2	6 978 984	7 977 546	998 562	H (Ca8)
Ca2	8 329 465	8 839 803	510 338	D/E (Ca7/Ca4)
Ca2	9 713 165	10 056 176	343 011	A (Ca5)
Ca3	2 147 981	2 291 454	143 473	D/E (Ca7/Ca4)
Ca3	5 222 428	5 315 148	92 720	G (Ca1)
Ca3	5 817 530	5 985 956	168 426	D/E (Ca7/Ca4)
Ca3	16 652 391	16 736 790	84 399	C (Ca6)
Ca4	33 548 471	34 015 816	467 345	A (Ca5)
Ca4	39 258 943	40 092 009	833 066	A (Ca5)
Ca5	1	475 756	475 755	D/E (Ca7/Ca4)
Ca5	479 469	967 381	487 912	B (Ca3)
Ca5	1 056 592	1 302 990	246 398	D/E (Ca7/Ca4)
Ca5	1 443 073	1 575 871	132 798	B (Ca3)
Ca5	3 057 844	3 436 790	378 946	B (Ca3)
Ca5	4 300 510	4 897 203	596 693	B (Ca3)
Ca5	5 459 276	5 583 037	123 761	H (Ca8)
Ca5	14 570 063	14 984 575	414 512	H (Ca8)
Ca6	1	95 085	95 084	F (Ca2)
Ca6	1 273 532	1 471 783	198 251	F (Ca2)
Ca6	8 091 437	8 171 394	79 957	H (Ca8)
Ca6	10 834 481	11 048 871	214 390	F (Ca2)
Ca6	11 049 103	11 286 027	236 924	D/E (Ca7/Ca4)
Ca6	22 607 989	23 187 652	579 663	F (Ca2)
Ca6	23 191 730	23 605 994	414 264	A (Ca5)
Ca6	23 607 585	23 929 615	322 030	F (Ca2)
Ca6	29 686 973	30 164 801	477 828	H (Ca8)
Ca6	40 809 786	41 720 980	911 194	A (Ca5)
Ca6	50 526 024	51 189 107	663 083	F (Ca2)
Ca7	259 820	579 136	319 316	G (Ca1)
Ca7	10 581 321	10 638 294	56 973	G (Ca1)
Ca7	19 431 822	19 639 120	207 298	H (Ca8)
Ca7	24 424 554	24 528 692	104 138	F (Ca2)
Ca7	31 258 312	31 438 623	180 311	B (Ca3)
Ca7	34 418 700	34 570 247	151 547	A (Ca5)
Ca7	37 738 053	37 860 082	122 029	G (Ca1)
Ca7	37 864 185	37 939 907	75 722	A (Ca5)
Ca7	44 316 766	44 615 144	298 378	A (Ca5)
Ca8	11 734 308	12 075 396	341 088	C (Ca6)

Interestingly, there were regions of the *desi* reference pseudomolecules where no reads mapped. We investigated these regions further by mapping *desi* whole-genome sequence data to the *desi* pseudomolecules (Figure 5). Surprisingly, again no reads mapped to these regions. To assess whether these regions reflect highly rearranged misassembled chickpea sequence data, for example due to concatenation of reads from the Roche 454

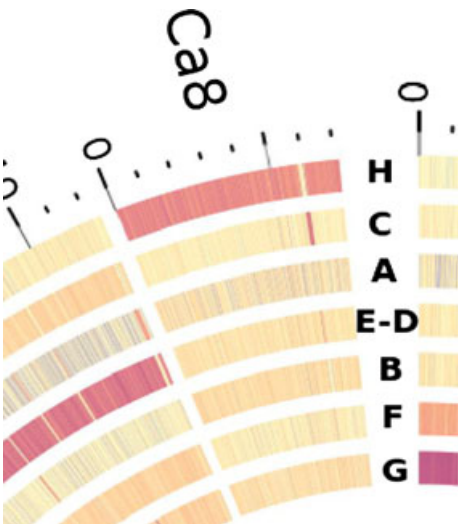


Figure 4 Circos heat map plot of the *kabuli* reference pseudomolecules demonstrating a high density of sequence reads (red colour) from *kabuli* chromosome H mapping to pseudomolecule Ca8. A small region on pseudomolecule Ca8 which lacks chromosome H reads is covered by chromosome C reads.

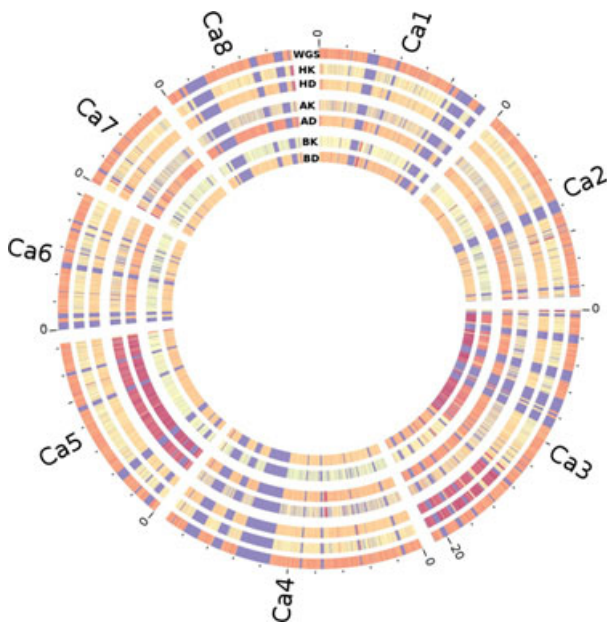


Figure 5 Circos heat map plot of the *desi* reference pseudomolecules demonstrating a high density of sequence reads (red colour) from *kabuli* and *desi* (D=*desi*, K=*kabuli*) chromosomes B, A, H and whole-genome sequence (WGS) reads of *desi*.

sequencing platform used in the assembly of the draft *desi* genome, we remapped the Illumina *desi* whole-genome data and isolated chromosome data to the *desi* pseudomolecules at a low stringency. This again failed to produce specific read mapping, and we therefore concluded that these regions of the *desi* reference pseudomolecules do not reflect the physical content of the *desi* genome. Extraction of the sequence for these regions and comparison with the swissprot gene database failed to identify a significant number of genes (data not shown), again suggesting that these regions are not true genome sequences.

Conclusions

The expansion of genome sequencing projects and variable quality of published genomes highlights the need for additional approaches to validate and finish high-quality genome assemblies. We have established and assessed a chromosomal genomics approach to validate and compare reference genome assemblies. Overall, the assembly quality of the *kabuli* genome is high, with relatively few regions in the reference pseudomolecules which appear to have been misassembled into scaffolds on the wrong pseudomolecule. The high-resolution identification of these misplaced regions will aid their relocation on their correct pseudomolecule and the production of an improved reference genome assembly. Observed differences between the *kabuli* and *desi* published reference sequences contrast with our previous understanding of the similarity between the genomes. Our chromosomal genomics analysis suggests that the physical genomes of *kabuli* and *desi* chickpea types are in fact very similar and the observed differences in the sequence assemblies are due to major errors in the *desi* genome assembly, including the misplacement of whole chromosomes, portions of chromosomes and the inclusion of a large portion of sequence assembly which does not appear to be from the genome of chickpea. In addition to validating and assessing the genomes of chickpea, chromosomal genomics can be applied to validate and assist in the accurate assembly of other genome references where chromosomes can be isolated using flow sorting and thereby provide more robust genome assemblies that can provide a higher level of value for the many end-users of a particular genome assembly.

Experimental procedures

Estimation of genome size

Nuclear genome size was estimated using flow cytometry according to Doležel *et al.* (2007) (Doležel *et al.*, 2007). Approximately 30 mg of young chickpea leaf and 10 mg of leaf of soybean (*Glycine max* L. cv. Polanka, 2C = 2.5 pg DNA), which served as internal standard (Doležel *et al.*, 1994), were used for sample preparation. Suspensions of cell nuclei were prepared by simultaneous chopping of leaf tissues of chickpea and soybean in a glass Petri dish containing 500 μL Otto I solution (0.1 M citric acid, 0.5% v/v Tween 20). Crude homogenate was filtered through a 50-μm nylon mesh. Nuclei were then pelleted (300 g, 5 min) and resuspended in 300 μL Otto I solution. After 30-min incubation at room temperature, 900 μL Otto II solution (0.4 M Na_2HPO_4) (Otto, 1990) supplemented with 50 μg/mL RNase and 50 μg/mL propidium iodide was added. Samples were analysed using a Partec PAS flow cytometer (Partec GmbH, Münster, Germany) equipped with a 488-nm argon laser. At least 5000 nuclei were analysed per sample. Three individuals were analysed for each chickpea accession, and each individual was measured three times on three different days. Nuclear DNA content was then calculated from individual measurements following the formula: 2C nuclear DNA content [pg] = 2.5 × G_1 peak mean of chickpea / G_1 peak mean of soybean. Mean nuclear DNA content was then calculated for each plant. Genome size (1C value) was then determined considering 1 pg DNA is equal to 0.978×10^9 bp (Doležel *et al.*, 2003).

Molecular sizes of chickpea chromosomes

We determined the relative chromosome lengths in chickpea *desi* 'ICC 4958' and *kabuli* 'CDC Frontier'. Mitotic metaphase plates were prepared using synchronized root tip meristems (Vláčilová *et al.*, 2002). Root tips were fixed in 3:1 fixative (absolute ethanol: glacial acetic acid) for a week at 37°C and stained in 2% acetocarmine solution. Chromosome preparations were made according to Masoudi-Nejad *et al.* (Masoudi-Nejad *et al.*, 2002). The preparations were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) in Vectashield (Vector Laboratories, Burlingame) and observed under a fluorescence microscope (Olympus AX70, Tokyo, Japan). Chromosome lengths were estimated using the MicroImage software (Olympus) in 15 complete metaphase plates in each genotype, and average values were determined for each chromosome. Molecular chromosome sizes were determined considering relative chromosome lengths and 1C nuclear genome sizes as shown in Table 3.

Flow cytometric chromosome sorting and sequencing

Actively growing roots were used for cell cycle synchronization and preparation of liquid chromosome suspensions according to Vláčilová *et al.* (Vláčilová *et al.*, 2002). Chromosomes in suspension were stained with 2 μg/mL DAPI and sorted using a FACSAria flow cytometer (BD Biosciences, San José). The identification of the sorted chromosomes A and B was performed using fluorescent *in situ* hybridization (FISH) following the protocol of Vláčilová *et al.* (Vláčilová *et al.*, 2002), using tandem repeat probe CaSat1. The purity of the chromosome H fraction was determined based on chromosome morphology without a specific probe. The chromosomal fractions were sorted with the following purities: A: 93.75% (88.8%), B: 93.50% (91%) and H: 96% (92%) for *desi* (and *kabuli*), respectively. For whole-genome amplification, aliquots of 100 000–180 000 chromosomes (corresponding to ~20 ng DNA) were sorted into PCR tubes containing 10 μL of deionized water. Chromosomal DNA was purified as described in Šimková *et al.* (Šimková *et al.*, 2008) using increased proteinase K concentration (300 ng/μL). The purified DNA was amplified using the Illustra GenomiPhi V2 DNA amplification kit (GE Healthcare, New York).

A total of 1 μg of amplified DNA was used to prepare an Illumina TruSeq DNA HT library for each isolated chromosome, according to the manufacturer's instructions, and sequenced on the Illumina Hiseq2000 platform using standard protocols (Table S1). Chromosomes D and E from *kabuli* were isolated and sequenced as a group.

Desi and *kabuli* genome comparison

A pairwise comparison of all *desi* pseudomolecules with all *kabuli* pseudomolecules (Figure 1) was produced using the synteny block and anchor filtering algorithms in SyMap v4.0 (Soderlund *et al.*, 2011). SOAP2.21 was applied to map Illumina sequence data to the draft reference genome assemblies. For high-confidence mapping, only paired reads mapping uniquely to the reference was considered. For low stringency mapping, single and nonunique mappings were permitted. Circos v0.56 (Krzywinski *et al.*, 2009) was used to produce circular heatmaps using modified reference genomes with all 'N' nucleotides removed. Custom perl scripts soap2nc.pl and nc2circos.pl were used to convert SOAP output to Circos format. The boundaries of misassembled regions were determined manually by visual examination of the BAM file of mapped reads.

Acknowledgements

The authors would like to acknowledge funding support from the Australian Research Council (Projects LP0882095, LP0883462, LP110100200 and DP0985953), the Australian India Strategic Research Fund (AISRF) Grand Challenge fund (GCF010013), the Australian Grains Research and Development Corporation (UWA00151), CGIAR Generation Challenge Programme (Theme Leader Discretionary grant), Czech Science Foundation (P501/12/G090) and by the grant award LO1204 from the National Program of Sustainability I, the Australian Genome Research Facility (AGRF), the Queensland Cyber Infrastructure Foundation (QCIF) and the Australian Partnership for Advanced Computing (APAC) and the Center of Excellence in Genomics (CEG) of ICRISAT. The part of this work has been undertaken as part of the CGIAR Research Program on Grain Legumes. ICRISAT is a member of CGIAR Consortium. We thank our colleagues M. Kubaláková, J. Číhalíková, R. Šperková and Z. Dubská from IEB for assistance in chromosome sorting.

References

Arabidopsis_Genome_Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796–815.

Batley, J. and Edwards, D. (2009) Genome sequence data: management, storage, and visualization. *Biotechniques*, **46**, 333–336.

Bennett, M.D. and Smith, J.B. (1976) Nuclear DNA amounts in angiosperms. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **274**, 228–274.

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. and 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.

Berkman, P.J., Lai, K., Lorenc, M.T. and Edwards, D. (2012a) Next generation sequencing applications for wheat crop improvement. *Am. J. Bot.* **99**, 365–371.

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. and Edwards, D. (2012b) Sequencing wheat chromosome arm 7BS delimits the 7BS/4AL translocation and reveals homoeologous gene conservation. *Theor. Appl. Genet.* **124**, 423–432.

Berkman, P.J., Visendi, P., Lee, H.C., Stiller, J., Manoli, S., Lorenc, M.T., Lai, K., Batley, J., Fleury, D., Šimková, H., Kubaláková, M., Weining, S., Doležel, J. and Edwards, D. (2013) Dispersion and domestication shaped the genome of bread wheat. *Plant Biotechnol. J.* **11**, 564–571.

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

Doležel, J., Doleželová, M. and Novák, F.J. (1994) Flow cytometric estimation of nuclear-DNA amount in diploid bananas (*Musa acuminata* and *M. balbisana*). *Biol. Plant.* **36**, 351–357.

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

Doležel, J., Kubaláková, M., Paux, E., Bartoš, J. and Feuillet, C. (2007) Chromosome-based genomics in the cereals. *Chromosome Res.* **15**, 51–66.

Duran, C., Eales, D., Marshall, D., Imelfort, M., Stiller, J., Berkman, P.J., Clark, T., McKenzie, M., Appleby, N., Batley, J., Basford, K. and Edwards, D. (2010) Future tools for association mapping in crop plants. *Genome*, **53**, 1017–1023.

Edwards, D. and Batley, J. (2010) Plant genome sequencing: applications for crop improvement. *Plant Biotechnol. J.* **7**, 1–8.

Edwards, D. and Wang, X. (2012) Genome Sequencing Initiatives. In: *Genetics, Genomics and Breeding of Oilseed Brassicas* (Edwards, D., Parkin, I.A.P. and Batley, J. eds), pp. 152–157. New Hampshire, USA: Science Publishers Inc.

Edwards, D., Batley, J. and Snowdon, R. (2013) Accessing complex crop genomes with next-generation sequencing. *Theor. Appl. Genet.* **126**, 1–11.

Hernandez, P., Martis, M., Dorado, G., Pfeifer, M., Gálvez, S., Schaaf, S., Jouve, N., Šimková, H., Valárik, M., Doležel, J. and 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.

Imelfort, M. and Edwards, D. (2009) De novo sequencing of plant genomes using second-generation technologies. *Brief. Bioinform.* **10**, 609–618.

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. and Chattopadhyay, D. (2013) A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *Plant J.* **74**, 715–729.

Jana, S. and Singh, K.B. (1993) Evidence of geographical divergence in kabuli chickpea from germplasm evaluation data. *Crop Sci.* **33**, 626–632.

Krzywinski, M., Schein, J., Birol, İ., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J. and Marra, M.A. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645.

Lai, K., Berkman, P.J., Lorenc, M.T., Duran, C., Smits, L., Manoli, S., Stiller, J. and Edwards, D. (2012a) WheatGenome.info: an integrated database and portal for wheat genome information. *Plant Cell Physiol.* **53**, 1–7.

Lai, K., Lorenc, M.T. and Edwards, D. (2012b) Genomic databases for crop improvement. *Agronomy*, **2**, 62–73.

Lee, H., Lai, K., Lorenc, M.T., Imelfort, M., Duran, C. and Edwards, D. (2012) Bioinformatics tools and databases for analysis of next generation sequence data. *Brief. Funct. Genomics*, **2**, 12–24.

Maesen, L.J.G.V.D. (1972) Cicer L., A monograph of the genus, with special reference to the chickpea (*Cicer arietinum* L.), its ecology and cultivation. *Maded. Landbouw. Wageningen*, **72**, 1–342.

Marshall, D., Hayward, A., Eales, D., Imelfort, M., Stiller, J., Berkman, P., Clark, T., McKenzie, M., Lai, K., Duran, C., Batley, J. and Edwards, D. (2010) Targeted identification of genomic regions using TAGdb. *Plant Methods*, **6**, 19.

Masoudi-Nejad, A., Nasuda, S., McIntosh, R.A. and Endo, T.R. (2002) Transfer of rye chromosome segments to wheat by a gametocidal system. *Chromosome Res.* **10**, 349–357.

Millan, T., Winter, P., Jungling, R., Gil, J., Rubio, J., Cho, S., Cobos, M.J., Iruela, M., Rajesh, P.N., Tekeoglu, M., Kahl, G. and Muehlbauer, F.J. (2010) A consensus genetic map of chickpea (*Cicer arietinum* L.) based on 10 mapping populations. *Euphytica*, **175**, 175–189.

Moreno, M.T. and Cubero, J.I. (1978) Variation in *Cicer arietinum* L. *Euphytica*, **27**, 465–485.

Ohri, D. and Pal, M. (1991) The origin of chickpea (*Cicer arietinum* L) - karotype and nuclear DNA amount. *Heredity*, **66**, 367–372.

Otto, F. (1990) DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. *Methods Cell Biol.* **33**, 105–110.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J.X., Mitros, T., Nelson, W., Hyten, D.L., Song, Q.J., Thelen, J.J., Cheng, J.L., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S.Q., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J.C., Tian, Z.X., Zhu, L.C., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C. and Jackson, S.A. (2010) Genome sequence of the palaeopolyploid soybean. *Nature*, **463**, 178–183.

Schnable, J.C., Springer, N.M. and Freeling, M. (2011) Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl Acad. Sci.* **108**, 4069–4074.

Šimková, H., Svensson, J.T., Condamine, P., Hřibová, E., Suchánková, P., Bhat, P.R., Bartoš, J., Šafář, J., Close, T.J. and Doležel, J. (2008) Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. *BMC Genomics*, **9**, 237.

Soderlund, C., Bomhoff, M. and Nelson, W.M. (2011) SyMAP v3.4: a turnkey synteny system with application to plant genomes. *Nucleic Acids Res.* **39**, e68.

Thudi, M., Bohra, A., Nayak, S.N., Varghese, N., Shah, T.M., Penmetsa, R.V., Thirunavukkarasu, N., Gudipati, S., Gaur, P.M., Kulwal, P.L., Upadhyaya, H.D., KaviKishor, P.B., Winter, P., Kahl, G., Town, C.D., Kilian, A., Cook, D.R. and Varshney, R.K. (2011) Novel SSR Markers from BAC-End Sequences, DArT Arrays and a Comprehensive Genetic Map with 1,291 Marker Loci for Chickpea (*Cicer arietinum* L.). *PLoS ONE*, **6**, e27275.

Varshney, R.K., Graner, A. and Sorrells, M.E. (2005) Genomics-assisted breeding for crop improvement. *Trends Plant Sci.* **10**, 621–630.

Varshney, R.K., Nayak, S.N., May, G.D. and Jackson, S.A. (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends Biotechnol.* **27**, 522–530.

Varshney, R.K., Chen, W.B., Li, Y.P., Bharti, A.K., Saxena, R.K., Schlueter, J.A., Donoghue, M.T.A., Azam, S., Fan, G.Y., Whaley, A.M., Farmer, A.D., Sheridan, J., Iwata, A., Tuteja, R., Penmetsa, R.V., Wu, W., Upadhyaya, H.D., Yang, S.P., Shah, T., Saxena, K.B., Michael, T., McCombie, W.R., Yang, B.C., Zhang, G.Y., Yang, H.M., Wang, J., Spillane, C., Cook, D.R., May, G.D., Xu, X. and Jackson, S.A. (2012) Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* **30**, 83–89.

Varshney, R.K., Song, C., Saxena, R.K., Azam, S., Yu, S., Sharpe, A.G., Cannon, S.B., Baek, J., Tar'an, B., Millan, T., Zhang, X., Rosen, B., 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., Upadhyaya, H.D., Luo, M., Singh, N.P., Lichtenzveig, J., Gali, K.K., Rubio, J., Nadarajan, N., Thudi, M., Doležel, 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. and Cook, D. (2013) Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat. Biotechnol.* **31**, 240–246.

Vláčilová, K., Ohri, D., Vrána, J., Číhalíková, J., Kubaláková, M., Kahl, G. and Doležel, J. (2002) Development of flow cytogenetics and physical genome mapping in chickpea (*Cicer arietinum* L.). *Chromosome Res.* **10**, 695–706.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.-H., Bancroft, I., Cheng, F., Huang, S., Li, X., Hua, W., Wang, J., Wang, X., Freeling, M., Pires, J.C., Paterson, A.H., Chalhoub, B., Wang, B., Hayward, A., Sharpe, A.G., Park, B.-S., Weisshaar, B., Liu, B., Li, B., Liu, B., Tong, C., Song, C., Duran, C., Peng, C., Geng, C., Koh, C., Lin, C., Edwards, D., Mu, D., Shen, D., Soumpourou, E., Li, F., Fraser, F., Conant, G., Lassalle, G., King, G.J., Bonnema, G., Tang, H., Wang, H., Belcram, H., Zhou, H., Hirakawa, H., Abe, H., Guo, H., Wang, H., Jin, H., Parkin, I.A.P., Batley, J., Kim, J.-S., Just, J., Li, J., Xu, J., Deng, J., Kim, J.A., Li, J., Yu, J., Meng, J., Wang, J., Min, J., Poulain, J., Hatakeyama, K., Wu, K., Wang, L., Fang, L., Trick, M., Links, M.G., Zhao, M., Jin, M., Ramchiary, N., Drou, N., Berkman, P.J., Cai, Q., Huang, Q., Li, R., Tabata, S., Cheng, S., Zhang, S., Zhang, S., Huang, S., Sato, S., Sun, S., Kwon, S.-J., Choi, S.-R., Lee, T.-H., Fan, W., Zhao, X., Tan, X., Xu, X., Wang, Y., Qiu, Y., Yin, Y., Li, Y., Du, Y., Liao, Y., Lim, Y., Narusaka, Y., Wang, Y., Wang, Z., Li, Z., Wang, Z., Xiong, Z. and Zhang, Z. (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **43**, 1035–1039.

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, J., Orjeda, G., Guzman, F., Torres, M., Lozano, R., Ponce, O., Martinez, D., De la Cruz, G., Chakrabarti, S.K., Patil, V.U., Skryabin, K.G., Kuznetsov, B.B., Ravin, N.V., Kolganova, T.V., Beletsky, A.V., Mardanov, A.V., Di Genova, A., Bolser, D.M., Martin, D.M.A., Li, G., Yang, Y., Kuang, H., Hu, Q., Xiong, X., Bishop, G.J., Sagredo, B., Mejia, N., Zagorski, W., Gromadka, R., Gawor, J., Szczesny, P., Huang, S., Zhang, Z., Liang, C., He, J., Li, Y., He, Y., Xu, J., Zhang, Y., Xie, B., Du, Y., Qu, D., Bonierbale, M., Ghislain, M., del Rosario Herrera, M., Giuliano, G., Pietrella, M., Perrotta, G., Facella, P., O'Brien, K., Feingold, S.E., Barreiro, L.E., Massa, G.A., Diambra, L., Whitty, B.R., Vaillancourt, B., Lin, H., Massa, A., Geoffroy, M., Lundback, S., DellaPenna, D., Buell, C.R., Sharma, S.K., Marshall, D.F., Waugh, R., Bryan, G.J., Destefanis, M., Nagy, I., Milbourne, D., Thomson, S.J., Fiers, M., Jacobs, J.M.E., Nielsen, K.L., Sonderkaer, M., Iovene, M., Torres, G.A., Jiang, J., Veilleux, R.E., Bachem, C.W.B., de Boer, J., Borm, T., Kloosterman, B., van Eck, H., Datema, E., Hekkert, B.T.L., Goverse, A., van Ham, R.C.H.J., Visser, R.G.F. and Potato Genome Sequencing, C. (2011) Genome sequence and analysis of the tuber crop potato. *Nature*, **475**, 189–195.

Yu, J., Hu, S.N., Wang, J., Wong, G.K.S., Li, S.G., Liu, B., Deng, Y.J., Dai, L., Zhou, Y., Zhang, X.Q., Cao, M.L., Liu, J., Sun, J.D., Tang, J.B., Chen, Y.J., Huang, X.B., Lin, W., Ye, C., Tong, W., Cong, L.J., Geng, J.N., Han, Y.J., Li, L., Li, W., Hu, G.Q., Huang, X.G., Li, W.J., Li, J., Liu, Z.W., Li, L., Liu, J.P., Qi, Q.H., Liu, J.S., Li, L., Li, T., Wang, X.G., Lu, H., Wu, T.T., Zhu, M., Ni, P.X., Han, H., Dong, W., Ren, X.Y., Feng, X.L., Cui, P., Li, X.R., Wang, H., Xu, X.,

Zhai, W.X., Xu, Z., Zhang, J.S., He, S.J., Zhang, J.G., Xu, J.C., Zhang, K.L., Zheng, X.W., Dong, J.H., Zeng, W.Y., Tao, L., Ye, J., Tan, J., Ren, X.D., Chen, X.W., He, J., Liu, D.F., Tian, W., Tian, C.G., Xia, H.G., Bao, Q.Y., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W.M., Li, P., Chen, W., Wang, X.D., Zhang, Y., Hu, J.F., Wang, J., Liu, S., Yang, J., Zhang, G.Y., Xiong, Y.Q., Li, Z.J., Mao, L., Zhou, C.S., Zhu, Z., Chen, R.S., Hao, B.L., Zheng, W.M., Chen, S.Y., Guo, W., Li, G.J., Liu, S.Q., Tao, M., Wang, J., Zhu, L.H., Yuan, L.P. and Yang, H.M. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp *indica*). *Science*, **296**, 79–92.

Zatloukalová, P., Hřibová, E., Kubaláková, M., Suchánková, P., Šimková, H., Adoracion, C., Kahl, G., Millan, T. and Doležel, J. (2011) Integration of genetic and physical maps of the chickpea (*Cicer arietinum* L.) genome using flow-sorted chromosomes. *Chromosome Res.* **19**, 729–739.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Circos heat map plot of the *kabuli* reference pseudomolecules demonstrating a high density of sequence reads (red colour) from *kabuli* and *desi* (D=*desi*, K=*kabuli*) chromosomes B, A and H.

Table S1 Chromosome sequence data generated.

Appendix S1 Detailed analysis of chickpea chromosome sorting.

APPENDIX V

Flow sorting and sequencing meadow fescue chromosome 4F

Kopecký D., Martis M., Číhalíková K., Hřibová E., Vrána J., Bartoš J., Kopecká J., Cattonaro F., Stočes Š., Novák P., Neumann P., Macas J., Šinková H., Studer B., Asp T., Baird J.H., Navrátil P., Karafiátová M., Kubaláková M., Šafář J., Mayer K., Doležel J.

Plant Physiology Vol.163, 1323-1337, 2013

IF: 7.084

Flow Sorting and Sequencing Meadow Fescue Chromosome 4F[1][C][W]

David Kopecký*, Mihaela Martis, Jarmila Číhalíková, Eva Hřibová, Jan Vrána, Jan Bartoš, Jitka Kopecká, Federica Cattonaro, Štěpán Stočes, Petr Novák, Pavel Neumann, Jiří Macas, Hana Šimková, Bruno Studer, Torben Asp, James H. Baird, Petr Navrátil, Miroslava Karafiátová, Marie Kubaláková, Jan Šafář, Klaus Mayer, and Jaroslav Doležel

Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, 78371 Olomouc, Czech Republic (D.K., J.Č., E.H., J.V., J.B., J.K., Š.S., H.Š., P.Na., M.Ka., M.Ku., J.Š., J.D.); Munich Information Center for Protein Sequences/Institute of Bioinformatics and Systems Biology, Helmholtz Center Munich, 85764 Neuherberg, Germany (M.M., K.M.); Istituto di Genomica Applicata and Istituto di Genomica Applicata Technology Services, 33100 Udine, Italy (F.C.); Biology Centre Academy of Sciences of the Czech Republic, Institute of Plant Molecular Biology, 37005 Ceske Budejovice, Czech Republic (P.No., P.Ne., J.M.); Institute for Agricultural Sciences, 8092 Zurich, Switzerland (B.S.); Aarhus University, Department of Molecular Biology and Genetics, Research Centre Flakkebjerg, DK–4200 Slagelse, Denmark (T.A.); and Department of Botany and Plant Sciences, University of California, Riverside, California 92521 (J.H.B.)

ORCID ID: 0000-0002-2834-1734 (D.K.).

The analysis of large genomes is hampered by a high proportion of repetitive DNA, which makes the assembly of short sequence reads difficult. This is also the case in meadow fescue (*Festuca pratensis*), which is known for good abiotic stress resistance and has been used in intergeneric hybridization with ryegrasses (*Lolium* spp.) to produce Festulolium cultivars. In this work, we describe a new approach to analyze the large genome of meadow fescue, which involves the reduction of sample complexity without compromising information content. This is achieved by dissecting the genome to smaller parts: individual chromosomes and groups of chromosomes. As the first step, we flow sorted chromosome 4F and sequenced it by Illumina with approximately 50× coverage. This provided, to our knowledge, the first insight into the composition of the fescue genome, enabled the construction of the virtual gene order of the chromosome, and facilitated detailed comparative analysis with the sequenced genomes of rice (*Oryza sativa*), *Brachypodium distachyon*, sorghum (*Sorghum bicolor*), and barley (*Hordeum vulgare*). Using GenomeZipper, we were able to confirm the collinearity of chromosome 4F with barley chromosome 4H and the long arm of chromosome 5H. Several new tandem repeats were identified and physically mapped using fluorescence in situ hybridization. They were found as robust cytogenetic markers for karyotyping of meadow fescue and ryegrass species and their hybrids. The ability to purify chromosome 4F opens the way for more efficient analysis of genomic loci on this chromosome underlying important traits, including freezing tolerance. Our results confirm that next-generation sequencing of flow-sorted chromosomes enables an overview of chromosome structure and evolution at a resolution never achieved before.

Meadow fescue (*Festuca pratensis*) is a grass typically found in hygrophylic and mesophylic meadows and pastures. It provides high-quality forage in pure stands and in grass-clover mixtures. Besides its usage on temporary meadows and pastures, this species is preferentially used as a component of permanent grasslands, where it produces forage during the first 3 years. Its outstanding winter survival is why this species is the major component of grasslands in northern Europe and Canada, and for the same reason, it is used widely in mountain regions of central and western Europe (Germany, Austria, Switzerland, Czech Republic, and Slovakia). However, meadow fescue is less adapted and competitive in lowland coastal regions in the temperate zone. In such climates, it is being replaced by interspecific hybrids of meadow fescue (or tall fescue) with ryegrasses (*Lolium* spp.), especially with *Lolium multiflorum*.

Despite the agronomic importance of meadow fescue, either as a stand-alone crop or as a parent in intergeneric hybridization, the progress in genetics and genomics of meadow fescue is lagging far behind most other crops. Genomic studies in meadow fescue are complicated by the large genome size (3,175 Mb/1C; Kopecký et al., 2010) and a rather symmetrical karyotype. Relative to cereals, and even to closely related ryegrasses, genetic and genomic resources of meadow fescue are limited.

[1] This work was supported by the Czech Science Foundation (grant no. P501/11/0504), the Ministry of Education, Youth, and Sports of the Czech Republic (grant no. OC10037), and the European Regional Development Fund (Operational Programme Research and Development for Innovations grant no. ED0007/01/01).

* Address correspondence to kopecky@ueb.cas.cz.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: David Kopecký (kopecky@ueb.cas.cz).

[C] Some figures in this article are displayed in color online but in black and white in the print edition.

[W] The online version of this article contains Web-only data.

www.plantphysiol.org/cgi/doi/10.1104/pp.113.224105

The only existing genetic map was developed by Alm et al. (2003) and consists of 466 RFLP, amplified fragment length polymorphism (AFLP), and isozyme markers with a total length of 658.8 centimorgan (cM). This map was recently enriched by 149 Diversity Arrays Technology (DArT) markers (Bartoš et al., 2011). So far, two bacterial artificial chromosome (BAC) libraries have been developed for meadow fescue. Donnison et al. (2005) constructed a five-genome coverage library and used it to clone candidate orthologous sequences to the CONSTANS-like rice (*Oryza sativa*) Heading date1 (*Hd1*; Photosensitivity1 [*Se1*]) gene in *Lolium perenne* and meadow fescue, while Kopecký et al. (2010) used a partial genomic BAC library (about one genome coverage) as a source of cytogenetic markers. The only existing DNA array was developed by Kopecký et al. (2009) using DArT technology (Jaccoud et al., 2001), and it consists of 7,680 probes derived from methyl-filtered genomic representations of five species from the *Festuca-Lolium* species complex (*L. perenne*, *L. multiflorum*, *F. pratensis*, *Festuca arundinacea*, and *Festuca glaucescens*). It has been used for diversity studies (Kopecký et al., 2009, 2011; Baird et al., 2012), genetic mapping (Bartoš et al., 2011; Tomaszewski et al., 2012), analysis of genomic composition of Festulolium cultivars (Kopecký et al., 2011), and association analysis of agronomic traits (Bartoš et al., 2011). Sequencing 620 genetically mapped DArT markers provided the first large source of DNA sequences for meadow fescue. However, DArT markers are believed to originate primarily from genic regions due to a methyl-filtration step in the preparation of genomic representations. This was proved by BLAST analysis against EST and hypothetical protein databases (Bartoš et al., 2011). Thus, the DArT markers did not provide a complete insight into the genome composition. Recently, 454 transcriptome sequencing of two mapping parents of meadow fescue has been reported by Vigeland et al. (2013).

Recent advances in next-generation sequencing (NGS) technologies, which enable the sequencing of entire genomes, make it possible to analyze genome structure at high resolution and perform detailed comparative studies across species. NGS data have been widely used for molecular marker development, phylogenetic and ecological studies, and analysis of transcriptomes using RNA sequencing (Egan et al., 2012). Unfortunately, the NGS of whole genomes has limited use in comparative genomic studies in species with no or poor genetic maps and in species for which a reference genome sequence is not available. In such cases, sequencing flow-sorted chromosomes was found to be an extremely valuable approach. For example, NGS of sorted chromosomes has been used to study structural genome changes in cereals during their evolution and speciation. Mayer et al. (2009, 2011) used a novel approach called GenomeZipper to deduce the putative gene order in barley (*Hordeum vulgare*). This approach is based on the comparisons of chromosomal shotgun sequences against model genomes (rice, sorghum [*Sorghum bicolor*], and *Brachypodium distachyon*) to detect syntenic regions. Similarly, Vitulo et al. (2011), Wicker et al. (2011), and Hernandez et al. (2012) used GenomeZipper for comparative studies of wheat (*Triticum aestivum*) homologous group 1 and wheat chromosomes 5A and 4A with rice, sorghum, *B. distachyon*, and a virtual barley genome. This analysis revealed considerable restructuring of wheat chromosome 4A and confirmed the utility of this approach.

Besides the use of flow-sorted chromosomes for NGS, they can be used for several other applications. A number of chromosome-specific BAC libraries were constructed for hexaploid wheat and rye (*Secale cereale*; Šafář et al., 2010), and many of these libraries are used for physical mapping (Paux et al., 2008). Moreover, sorted chromosomes are valuable templates for cytogenetic mapping using fluorescence in situ hybridization (FISH; Kubaláková et al., 2003; Suchánková et al., 2006). Sorted chromosomes were also used as an efficient source of genetic markers for the saturation of genetic maps (Wenzl et al., 2010).

In this study, we set out to explore the possibilities of applying chromosome genomics in meadow fescue. We have developed flow cytogenetics in this species and used flow cytometric sorting to dissect the meadow fescue genome into individual chromosomes. NGS of isolated chromosomes provided, to our knowledge, the first insights into the genome structure of this species and permitted the analysis of collinearity at a high resolution level.

RESULTS

Cell Cycle Synchronization and Preparation of Chromosome Suspension

The experiments performed to optimize the induction of cell cycle synchrony and the accumulation of metaphases in meristem root tips were based on our previous experience with cereals and legumes. The optimal concentration of hydroxyurea (HU) was 1.5 mM; lower concentration (1 mM) did not completely block the cycle. On the other hand, the cell cycle did not recover after blockage by 2 mM HU. The highest metaphase index (55%) was reached 5 h after removal from the DNA synthesis inhibitor (recovery time).

Chromosome Analysis and Sorting

Flow cytometry of mitotic metaphase chromosomes permitted the separation of meadow fescue chromosomes based on their size (relative 4,6-diamino-2-phenylindole [DAPI] fluorescence intensity). The histogram of the chromosome DAPI fluorescence (flow karyotype) of meadow fescue ($n = 7$) consisted of three distinct peaks (Fig. 1). All seven chromosomes of meadow fescue can be individually identified based on their specific patterns of hybridization with various DNA probes (Kopecký et al., 2010). This permitted the identification of particles sorted from individual peaks on flow

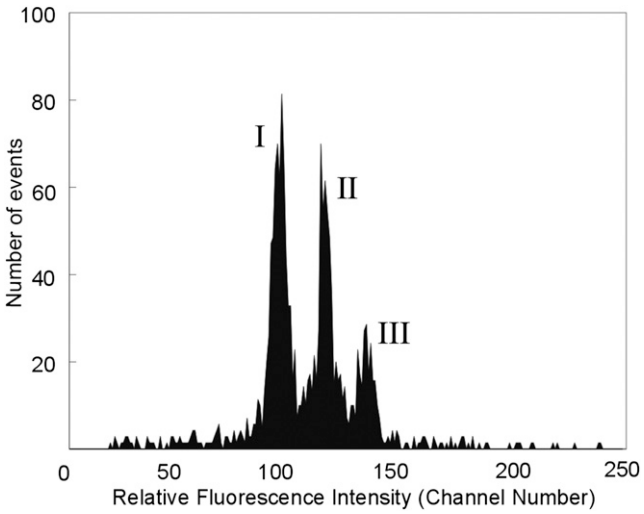


Figure 1. Flow karyotype. The distribution of relative chromosome fluorescence intensity (flow karyotype) was obtained after analysis of a DAPI-stained suspension of mitotic chromosomes of meadow fescue ($2n = 2x = 14$). The flow karyotype consists of two composite peaks, I and II, representing chromosomes 1F, 5F, and 6F and chromosomes 2F, 3F, and 7F, respectively, and peak III, representing chromosome 4F. This chromosome can be discriminated and sorted individually.

karyotype. Thus, chromosomes 1, 5, and 6 were sorted as a group from peak I, chromosomes 2, 3, and 7 formed peak II, and chromosome 4 formed peak III. On average, we were able to collect about 8×10^3 copies of chromosome 4F from a sample prepared from approximately 150 root tips. For DNA amplification and subsequent sequencing, we collected about 40,000 4F chromosomes. The purity of isolated chromosomes was estimated by FISH. Based on screening 300 sorted chromosomes from two randomly chosen samples, the average purity was 92.4%, with contamination by almost all remaining chromosomes in low frequencies (chromosome 1, 0.36%; chromosome 2, 2.18%; chromosome 3, 2.54%; chromosome 5, 0.36%; and chromosome 7, 2.18%).

Amplification and Sequencing of Chromosome 4F

An equivalent of 50 ng of DNA was obtained by flow cytometric sorting of 40,000 copies of chromosome 4F. This yielded 25 ng of purified DNA, which was used in three independent multiple displacement amplification (MDA) reactions, providing a total of 12.1 μg (4.1 μg + 4.2 μg + 3.8 μg) of DNA. For sequencing, individual samples of amplified DNA were combined to reduce a possible bias introduced by the MDA. A total of 4 μg of MDA DNA from chromosome 4F was used for Illumina HiSEquation 2000 sequencing. A total of 85,351,865 paired-end sequence reads with lengths of 101 bp were generated, yielding 28,553 Mb of sequence. Considering the size of chromosome 4F at 543 Mb (Kopecký et al., 2010), this represented approximately 50× coverage.

The assembly of the Illumina paired-end reads was performed using SOAPdenovo (Luo et al., 2012) with different k-mer sizes (i.e. 21–83 k-mer). The assembly with the best L50 (a weighted median statistic such that 50% of the entire assembly is contained in contigs equal to or larger than this value; 70-mer, 132 bp) and a maximum contig length of 15,919 bp was chosen for further analyses. The low L50 value, the large number of short contigs, and the high amount of repetitive elements indicate a highly fragmented assembly. Removing the repetitive sequences from the assembly, the L50 value increases considerably (1,623 bp). This suggests that the nonrepetitive genome space is well assembled and provides a useful resource for chromosome structure analysis, gene detection, and construction of a virtual ordered gene map.

Gene Content of Chromosome 4F

To estimate the number of genes on 4F, sequence comparisons against the genomes of *B. distachyon*, rice, and sorghum were made. We used stringent BLASTX searches that required 75% or greater (*B. distachyon*)/70% (rice, sorghum) sequence similarity over at least 30 amino acids (Table I). Between 2,629 and 3,056 significant matches were found. A total count of cumulative, nonredundant matches on 4F was 4,626. Considering the estimated molecular length of 4F at 543 Mb and the gene density on 4F representative for the entire meadow fescue genome (Kopecký et al., 2010), this would scale up to 27,048 genes. It should be mentioned that this number does not consider the putatively nonhomologous genes.

Table I. *GenomeZipper results*

General overview of the GenomeZipper results of the 4F chromosome. All numbers given are nonredundant.

Parameter	Chromosome 4F
No. of anchored gene loci	4,672
Lolium spp. SNP markers with match to syntenic genes	109 (out of 154)
No. of anchored barley full-length complementary DNAs	1,697
No. of anchored 4F contigs via first best hit	20,613
No. of anchored 4F contigs via bidirectional BLAST hit	3,700
No. of anchored *Festuca* spp. ESTs	1,444
No. of anchored *B. distachyon* genes	3,056
No. of anchored rice genes	2,629
No. of anchored sorghum genes	2,826

A Virtual Gene Map of Chromosome 4F

Because of the lack of suitable genetic markers for meadow fescue, 154 genetic markers from chromosome 4 of *L. perenne* were used as a backbone to construct a virtual gene map of chromosome 4F. *L. perenne* is a closely related species that evolved from the genus *Festuca* about 2.2 million years ago (Polok, 2007). Out of the 154 *L. perenne* markers, 109 (70.8%) could be associated to genes with conserved synteny under these conditions: (1) the best first hit, with (2) a minimal alignment length of 30 amino acids, and (3) an identity of 75% (*B. distachyon*)/70% (rice, sorghum). A virtual ordered gene map was developed using the GenomeZipper protocol (Mayer et al., 2011). With this approach, we identified 4,672 loci, of which 4,626 correspond to regions with conserved synteny in *B. distachyon*, rice, and sorghum (Table I). Among these, 1,515 (32.7%) loci were supported by genes in all three reference organisms, 855 (18.5%) in two of them, and 2,256 (48.8%) genetic loci were supported by only one gene with conserved synteny (Fig. 2). The higher number of syntenic *B. distachyon* genes reflects a closer phylogenetic relationship of meadow fescue to *B. distachyon* than to rice and sorghum. Furthermore, 1,697 barley full-length complementary DNAs were uniquely associated to either a marker sequence or at least one syntenic gene (Table I).

Comparative Analysis of Chromosome 4F with Relatives

Using GenomeZipper, we performed a detailed comparative analysis of chromosome 4F with sequenced genomes of *B. distachyon*, rice, sorghum, and barley. An ancient translocation differentiates chromosome 4F from 4H of barley (Fig. 3). Chromosome 4F is collinear

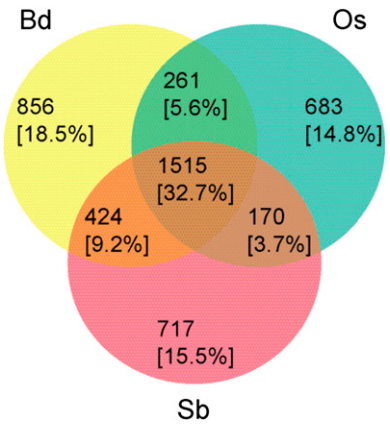


Figure 2. Model species genes anchored to 4F. The Venn diagram shows the number of *B. distachyon* (Bd), rice (Os), and sorghum (Sb) genes that are anchored in the virtual gene map of meadow fescue chromosome 4F. Overlaps of the circles show the number of genes anchored at the same locus. [See online article for color version of this figure.]

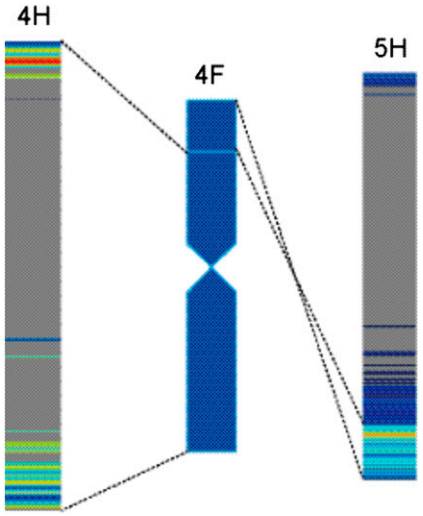


Figure 3. Collinearity of 4F with the barley genome. Meadow fescue 4F Illumina contigs were mapped on the physical map of barley chromosomes 4H and 5H. Syntenic regions are colored yellow-red, whereas nonsyntenic regions are colored dark blue. This approach clearly identifies the collinearity of chromosome 4F with barley chromosome 4H and the terminal part of 5HL. The connectors used indicate the orientations of the particular parts.

with the entire chromosome 4H but also with the terminal region of the long arm of chromosome 5H (5HL). Based on the position of genetic markers, the 5H syntenic segment is in an inverted orientation. For sequences generated from 4F, we identified in silico syntenic regions on *B. distachyon* chromosomes 1 and 4 (Fig. 4), rice chromosomes 3 and 11 (Supplemental Fig. S1), and sorghum chromosomes 1 and 5 (Supplemental Fig. S2).

Comparative mapping results between physical maps of barley chromosomes 4H and 5H, *Festuca* chromosome 4F, *Lolium* chromosome 4L, and wheat chromosome 4A show that the terminal segment originally from the ancestral chromosome 4 and now located at 5HL is not identical to the segment involved in the 4AL/5AL translocation (Naranjo et al., 1987; Supplemental Fig. S3). Our observation is similar to that of Alm et al. (2003). The location of the corresponding segment from 5HL is found on short arms of *Festuca* spp. 4F and *Lolium* spp. 4L chromosomes (both carrying ancestral type chromosome 4), while in wheat, the long arm of chromosome 4A is involved (Supplemental Fig. S3). The 5H regions syntenic with 4F and 4A are from 127.96 and 159.79 cM to the telomere (196.85 cM), respectively. While *Lolium* spp. 4L and *Festuca* spp. 4F appear structurally alike, a comparison of *Festuca* spp. 4F against wheat 4A indicates numerous inversions and translocations (Fig. 5). The extensive collinearity between *Lolium* and *Festuca* spp. along with the known genome structure of *Lolium* spp. (Pfeifer et al., 2013) suggest that the collinear 5HL segment on *Festuca* spp. chromosome 4 represents a

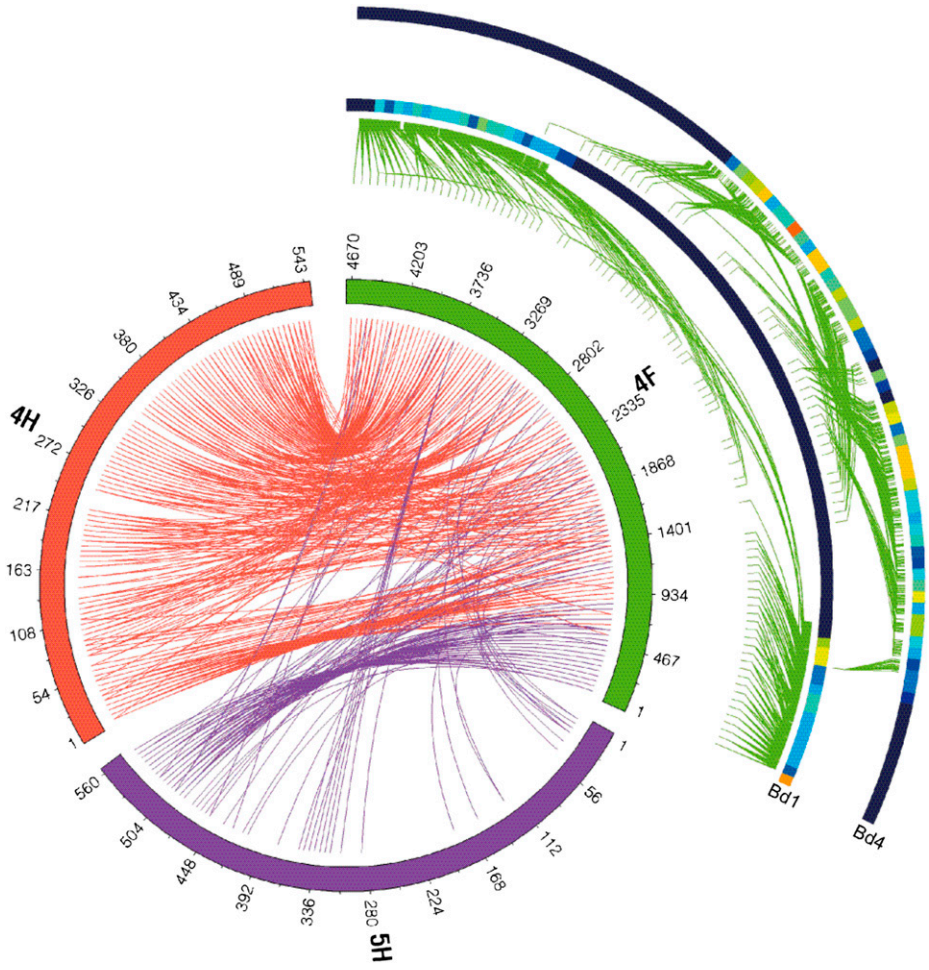


Figure 4. Collinearity of 4F with *B. distachyon*. High-density comparative analyses show the linear gene order of the meadow fescue GenomeZipper versus the sequenced genomes of barley and *B. distachyon*. The concentric circles are as follows. The inner circle represents chromosome 4F (green) and collinear chromosomes 4H (red) and 5H (purple) of barley. The outer circles represent homologous chromosomes of *B. distachyon* (Bd1 and Bd4). The heat maps illustrate the density of genes hit by the contigs from the 4F chromosome. Syntenic regions are colored red, whereas nonsyntenic regions are colored blue. Putative orthologs between 4F and barley chromosomes 4H and 5H are connected with lines. The scale is given in gene loci for chromosome 4F and in megabase pairs for barley chromosomes 4H and 5H.

unidirectional translocation in barley rather than a reciprocal translocation. In the case of reciprocal translocation, we would expect to observe some sequences being present in the terminal region of the short arm of barley chromosome 4H (4HS) and absent in 4F (Figs. 3 and 5A). However, this was not the case, and all sequences detected in the terminal region of 4HS were also present on chromosome 4F. This interpretation is consistent with observations made in *Lolium* spp. (Pfeifer et al., 2013). Since *Festuca* (and *Lolium*) spp. chromosome 4 is collinear with rice chromosome 3, which indicates more ancestral chromosomal structure, we hypothesize that this unidirectional translocation occurred in the Triticeae branch after the *Festuca/Lolium* genera split.

The linear ordered gene maps of *Festuca* spp. chromosome 4F, wheat chromosome 4A, and barley chromosome 4H allow one to analyze the number of syntenically conserved genes against the reference genomes of *B. distachyon*, rice, and sorghum (Fig. 6). The proportion of shared syntenic genes on all three chromosomes ranges between 23% and 28%, while between 19% and 24% of the genes are found for two of the chromosomes and between 49% and 57% are found for only one of the chromosomes analyzed. Strikingly higher percentages of genes conserved in *Festuca* spp. 4F but not for barley 4H and wheat 4A (30%–47%) were found illustrating the higher degree of syntenic conservation and, at least in part, a closer evolutionary relationship of *Festuca* spp. against the reference genomes used.

Repetitive DNA of Chromosome 4F

With the aim to determine the major repetitive DNA constituents of chromosome 4F, repeat reconstruction was done on Illumina data representing 1× coverage of chromosome 4F (550 Mb). Graph-based clustering resulted in 252,144 clusters containing 4,224,228 sequence reads (76%). A total of 855 clusters contained at least 100 sequences, from which 450 clusters contained more than 555 sequences (0.01% of analyzed reads; Supplemental Fig. S4). The largest clusters were manually annotated using a combination of multiple approaches based on similarity searches against sequence databases and graph layout analysis. The aim was to identify tandem organized repeats useful as

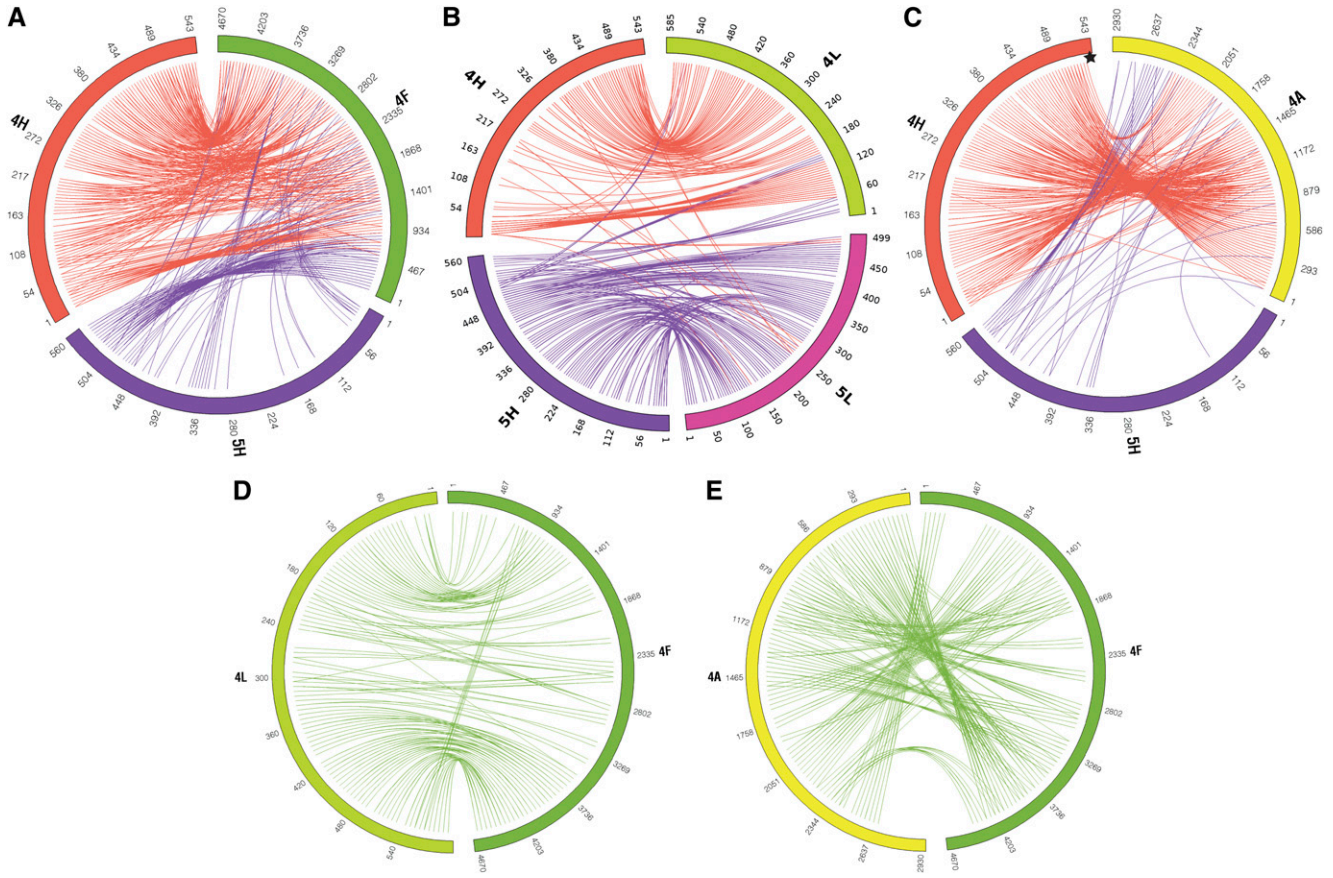


Figure 5. Chromosome 4/5 translocation. The top three circles (A–C) show the chromosome 4/5 translocation in three different grass chromosomes (*Festuca* spp. 4F, wheat 4A, and *Lolium* spp. 4L) in comparison with the corresponding homologous chromosomes 4 and 5 in barley. The location of the segment from 5HL for both *Festuca* (A) and *Lolium* (B) spp. is found on the short chromosome arm, while in wheat (C), the long arm of chromosome 4A is involved. The observed chromosome 4/5 translocation is a unidirectional translocation in *Festuca* and *Lolium* spp. rather than a reciprocal translocation, as found in wheat (Alm et al., 2003). The bottom two circles depict the collinearity between meadow fescue chromosome 4F and *L. perenne* chromosome 4L (D) and wheat chromosome 4A (E). While *Lolium* and *Festuca* spp. are highly collinear, the synteny between wheat and *Festuca* spp. is interrupted by several inversions and translocations.

species-specific and/or chromosome-specific cytogenetic markers.

Annotation of the largest clusters led to the classification of approximately 53% of the 4F chromosome sequence. The largest clusters contained mainly Ty3/Gypsy-like elements, which were found to be the most abundant repeats (28.64% of analyzed sequence reads) of chromosome 4F. Ty1/Copia-like elements represented less than 7% and DNA transposons represented about 3% of annotated sequences (Fig. 7). Interestingly, we were not able to characterize sequences within several large clusters that account for more than 8% of the analyzed sequence reads of chromosome 4F. These sequences can represent unknown DNA repeats and/or whole-genome amplification artifacts (Fig. 7).

Based on the graph layout (Fig. 8), contigs derived from clusters containing putative tandem repeats were identified and further investigated using similarity dot-plot analysis (Supplemental Fig. S5). The analysis revealed the presence of tandem or semitandem organized repeats in 43 clusters, which together represented about 2.71% of the chromosome 4F sequence.

Experimental Investigation of Tandem Organized Repeats

Out of 43 in silico-identified tandem repeats, sequences of 15 putative tandem repeats containing repetitive units of different length were investigated in more detail to study their organization in the *Festuca* and *Lolium* spp. genomes (Supplemental Table S1). A set of primers specific for tandem units was designed and used for the preparation of probes for FISH and for Southern hybridization. PCR products obtained after amplification on genomic DNA of meadow fescue as the template DNA were sequenced using dideoxy-chain termination reaction to check their authenticity. Moreover, PCR was done on flow-sorted chromosomes of

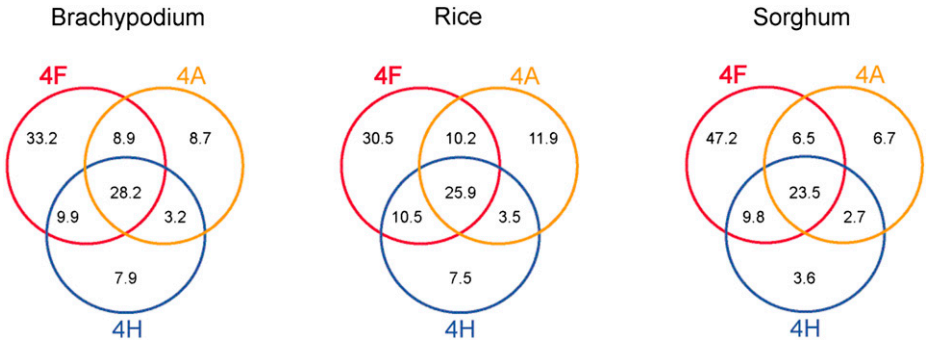


Figure 6. Conserved syntenic regions in the virtual linear ordered maps of *Festuca* spp., barley, wheat, and *Lolium* spp. chromosome 4 as defined by comparison with model grass genomes. The intersecting circles of the Venn diagrams depict the percentage of syntenic conserved reference genes (*B. distachyon*, rice, sorghum) shared by meadow fescue chromosome 4F (red), wheat chromosome 4A (orange), and barley chromosome 4H (blue). [See online article for color version of this figure.]

meadow fescue and on *L. perenne* genomic DNA to verify species and/or chromosome specificity. The results showed that all tandem repeats were present in all chromosome peaks of meadow fescue flow karyotype and that positive PCR products were obtained for all 15 selected tandem repeats also on *Lolium* spp. genomic DNA (Supplemental Table S1).

To confirm the tandem character of the repeats within the clusters, Southern hybridization was used to study the organization of repeats from the clusters that were successfully mapped on *Festuca* spp. mitotic chromosomes using FISH (see below). A ladder-like pattern that is typical for tandem organized repetitive units was obtained for all eight putative tandem repeats that gave visible signals after FISH (see below; Fig. 9). As shown in Figure 10, the tandem repeats identified in sequences of chromosome 4F of meadow fescue are also present in the nuclear genome of *L. perenne*, although at lower copy number.

Physical Localization of Repeats Using FISH

Out of 15 tandem repeats tested (fpTR1–fpTR15), we were able to physically map five fpTRs on chromosome 4F using FISH (Fig. 9; Supplemental Table S1). All of them produced hybridization patterns also on other chromosomes. Additionally, the other three fpTRs mapped on chromosomes other than 4F (fpTR6, fpTR11, and fpTR12; Supplemental Table S1). This might be caused by the sensitivity of FISH. None of the fpTRs produced signal

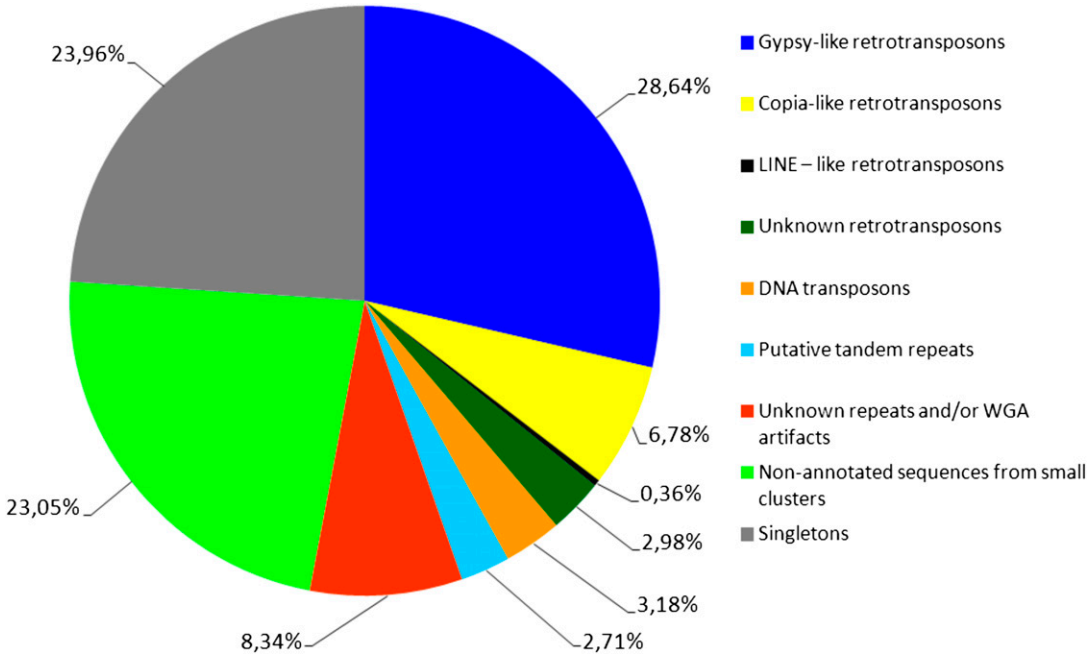


Figure 7. Repeats on 4F. The proportion of different repetitive elements in sequence reads of 4F chromosome is shown. [See online article for color version of this figure.]

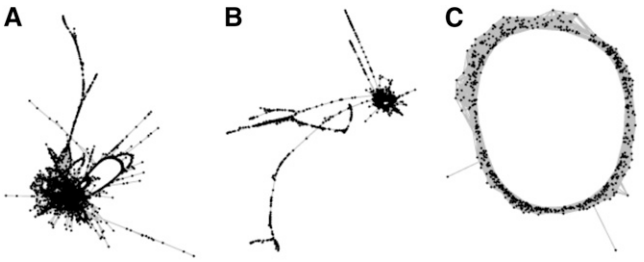


Figure 8. Graph layouts visualized by SeqGrapheR. Examples of graph layouts were calculated using the Fruhterman-Reingold algorithm for different types of tandem organized repeats: fpTR1 (A), fpTR5 (B), and fpTR15 (C). Dots and lines represent sequence reads and similarity hits between them, respectively.

exclusively on 4F; thus, they cannot be considered as chromosome specific. On the other hand, our observation opens the way for precise and robust molecular karyotyping of this species. As most of the fpTRs produced visible signals also on mitotic chromosomes of *L. perenne* and F1 hybrid *L. perenne* × meadow fescue (Fig. 9), the repeats may find even broader use as probes for FISH in cytogenetics of the *Festuca-Lolium* spp. complex.

Using FISH, we were able to localize a telomeric repeat on *Festuca* spp. chromosomes. Surprisingly, it produced signals not only in the telomeric regions of both chromosome arms of 4F but also interstitially (Fig. 9). Such interstitial signals were also detected on chromosome arms 2FL, 5FL, and 6FS. In general, FISH done on sorted chromosomes displayed higher sensitivity and produced more bands with higher resolution on individual chromosomes than on standard squashed preparations (Fig. 9).

DISCUSSION

Despite their considerable economic importance, forage and turf grasses lag far behind many other crops in terms of genetics, genomics, and bioinformatics. The delay is due to many factors, but outcrossing character, population-based breeding, and frequent aneuploidy do not help. Moreover, genomes of species within *Festuca-Lolium* are large and complex. The genome of meadow fescue was estimated at 1C = 3,175 Mb by DNA flow cytometry and is similar in size to that of human (Doležel et al., 2003; Kopecký et al., 2010). The ryegrass genome is just a little smaller (1C = 2,623 Mb for *L. perenne* and 1C = 2,567 Mb for *L. multiflorum*). Because of the genome complexity, any approach reducing sample complexity is welcome in genomic studies.

Partitioning the Grass Genomes

Analysis of a complex genome where a majority of DNA is represented by repeats can be simplified by several approaches. There are methods available that avoid sequencing of the repetitive parts of genomes. Of these, sequencing of complementary DNA to generate ESTs is among the most successful. However, ESTs fail to sample rare or conditional transcripts (Martienssen et al., 2004), and other methods were proposed to target the gene space in large and complex genomes, such as Cot fractionation and methyl filtration (Rabinowicz et al., 1999; Peterson et al., 2002). Unfortunately, the latter two methods did not meet the expectations.

Another alternative to tackle the complex genomes is to dissect them into smaller elements and sequence these elements individually. Working with naturally uniform and independent units, such as chromosomes, is perhaps the most powerful approach. In hexaploid wheat, individual chromosomes represent only 3.6% to 5.9% of the entire genome (Doležel et al., 2009), and even in a diploid species such as meadow fescue, isolation of individual chromosomes dissects the 3.2-Gb genome into 373- to 543-Mb units, each representing 11.7% to 17.1% of the entire genome (Kopecký et al., 2010).

Individual chromosomes can be isolated in two ways: by microdissection and flow sorting. Microdissection gives access to any chromosome or a chromosome segment. However, this is a very tedious approach, and for all practical purposes, the total yield is limited to only a few copies of a particular chromosome (Zhou and Hu, 2007). On the other hand, flow cytometry can isolate high copy numbers of the same chromosome with purity usually exceeding 90%. Flow sorting relies on differences in chromosome size or, rather, on the difference in relative fluorescence intensity. The output of flow cytometric analysis is a histogram of relative chromosome fluorescence intensity (reflecting chromosome size), which is called a flow karyotype (Fig. 1). Ideally, each chromosome is represented by a single peak on the flow karyotype. However, similarities in chromosome size within a genome usually result in the appearance of composite peaks representing two or more chromosomes. According to Doležel et al. (2009), there has to be at least a 10% difference in chromosome size to generate a separate peak on a flow karyotype. Unfortunately, in most plant species, chromosomes are not that different in size. In hexaploid wheat (21 chromosome pairs), the flow karyotype consists of only four peaks, and only one of those contains a single chromosome, 3B (Vrána et al., 2000). Similarly, the genome of meadow fescue was dissected here into three peaks, where only one peak represented a single chromosome type (chromosome 4F). However, the plasticity of plant genomes (especially in polyploids) makes it possible to develop special cytogenetic stocks with reconstructed karyotypes. Single chromosome substitution lines of *L. multiflorum*-meadow fescue, developed in our previous work, may provide an option for sorting other chromosomes of meadow fescue due to a sufficient difference in chromosome length (Kopecký et al., 2010).

Virtual Gene Order on Chromosome 4F

The GenomeZipper approach applied to individual chromosome sequences is a powerful tool for comparative studies. It significantly increases the resolution level relative to genetic markers. In the previous report

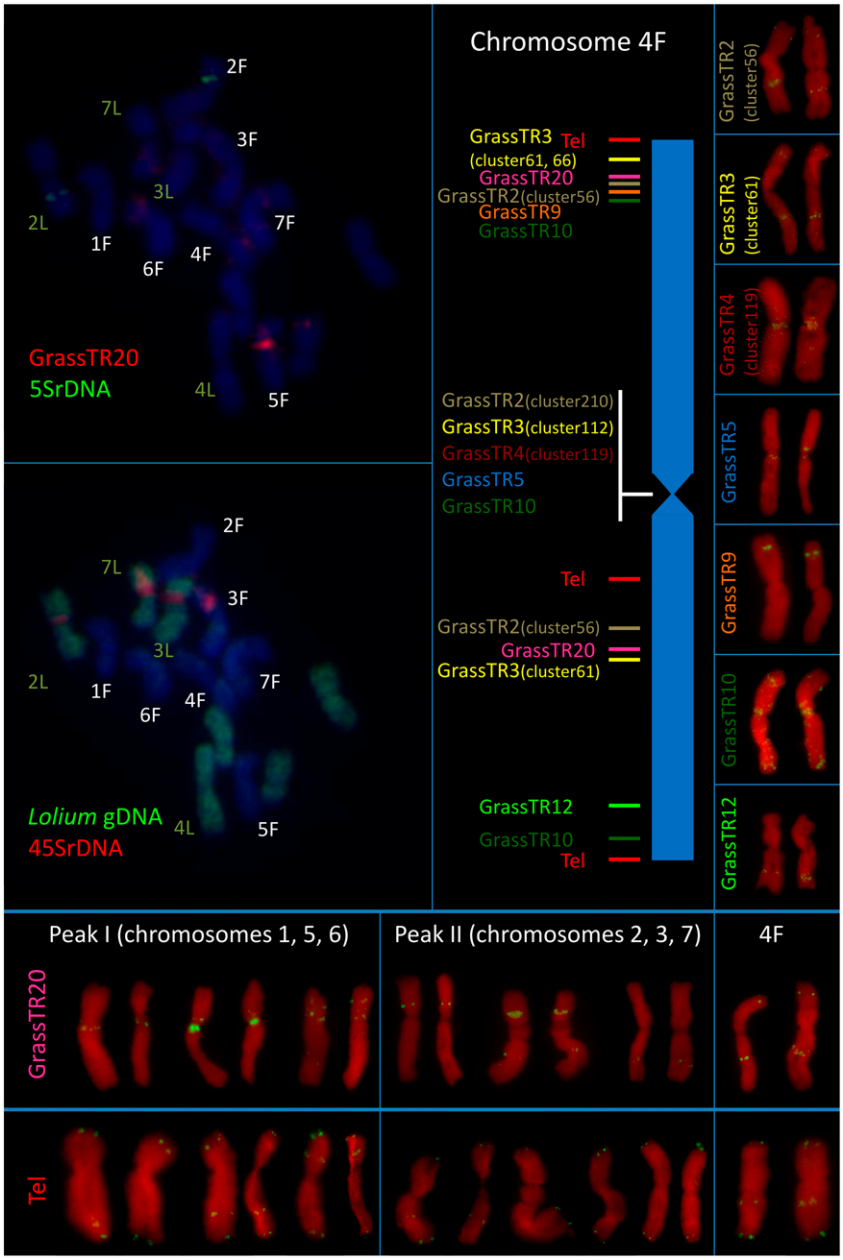


Figure 9. Cytogenetic mapping of repeats. Cytogenetic mapping shows newly identified tandem repeats and telomeric repeats on sorted chromosomes of meadow fescue and the F1 hybrid of *L. perenne* × meadow fescue.

on the collinearity of the meadow fescue genome with rice, sorghum, *Lolium* spp., oat (*Avena sativa*), maize (*Zea mays*), and Triticeae, a limited number of markers were used (Alm et al., 2003). Prior to this study, only 36 markers were available from chromosome 4F, with only a subset (seven to 30) suitable for any pairwise comparison. This contrasts with thousands of sequences from individual chromosomes (4F here) of model plant species (rice, sorghum, *B. distachyon*, and barley).

Estimation of Gene Content

The GenomeZipper approach permitted the estimation of gene content in meadow fescue and, by implication, all related *Festuca* spp. Assuming the estimated size of chromosome 4F at 543 Mb and the 4F gene density representative for the entire genus *Festuca* genome (Kopecký et al., 2010), the total genus *Festuca* genome contains about 35,000 genes. This is a slightly higher number that the estimation for barley (32,000 or more genes; Mayer et al., 2011) and less than that for the B genome of wheat (38,000 genes; Choulet et al., 2010). The recent estimate for the A genome of wheat is below that for *Festuca* spp. and barley (28,000 genes; Hernandez et al., 2012). In *L. perenne*, the most closely related species for which DNA sequence data are available, almost 25,700 matches with *B. distachyon* genes were obtained (Byrne et al., 2011).

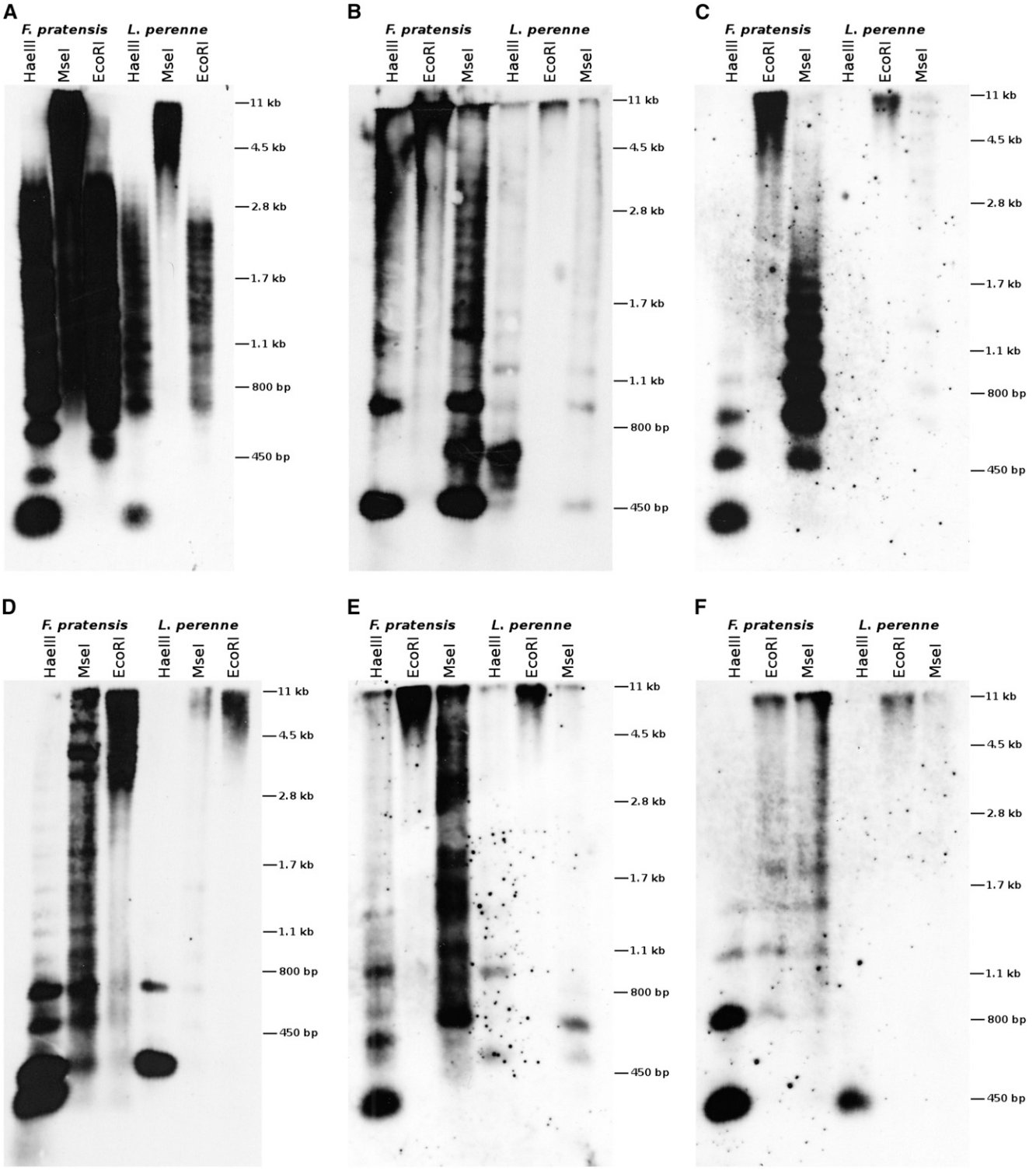


Figure 10. Southern hybridization with the probes for tandem repeats. Examples show Southern hybridization of meadow fescue and *L. perenne* genomic DNA digested using three different restriction endonucleases with the probes for tandem organized repeats fpTR1 (A), fpTR4 (B), fpTR5 (C), fpTR7 (D), fpTR12 (E), and fpTR15 (F). The ladder-like pattern indicates the presence of tandem organized repetitive units and supports the results of dotter analysis as well as the FISH analysis.

Apart from estimating gene content, the annotated sequences represent an invaluable genome resource for molecular biologists and breeders of fescue. For example, the sequences can be screened for the presence of resistance gene analogs. Because of a general lack of molecular markers for fescue, this new resource opens avenues for in silico identification of sequences suitable as DNA markers, including simple sequence repeat, presence/absence variation, and single-nucleotide polymorphism (SNP), and help in developing SNP platforms for marker-assisted breeding of this agriculturally valuable species.

4F Is an Ancestral Chromosome toward Modern Cereals

Based on the comparative analysis, chromosome 4F appears collinear with the entire barley chromosome 4H as well as with the terminal region of 5HL (Fig. 3). The observed homology of 4F to chromosome arms 4S, 4L, and 5L from Triticeae is consistent with previous observations by Alm et al. (2003) based on mapping of molecular markers. Chromosome 4F is collinear with the entire chromosome 3Os of rice (including the centromeric introgression of chromosome 11Os). On the other hand, 3Os is collinear with chromosome 4H and also with the terminal region of 5HL (Gale and Devos, 1998). Thus, meadow fescue chromosome 4F is likely more ancient than barley chromosome 4H. In cereal genomes (the A, B, and D genomes of wheat and the H genome of barley), the direct ancestral precursor that is syntenic with rice chromosome 3Os was broken and the terminal part of the short arm was translocated to the distal part of the long arm of chromosome 5, resulting in the present-day chromosome 5H. This happened after the divergence of the *Festuca* and *Lolium* genera from the grass lineage between 25 million years ago (divergence of Triticeae and Poeae) and 11.6 million years ago (divergence of barley and wheat genomes; Chalupska et al., 2008). In *Festuca* and *Lolium* spp., the direct ancestral precursor syntenic with 3Os of rice remained intact and resulted in the present-day 4F and 4L chromosomes of *Festuca* and *Lolium* spp., respectively, as originally proposed by Alm et al. (2003). Therefore, chromosome 4F, in its current form, can be assumed as a potential transient chromosome 4H′ proposed by Thiel et al. (2009). Interestingly, a 4-fold increase in recombination frequency was observed in the region surrounding the translocation break point (Kopecký et al., 2010). This may indicate an unstable region of the chromosome more prone to chromosome breakage.

Newly Identified Repeats Are a Valuable Source of Cytogenetic Markers

Cytogenetic mapping involves the physical localization of a sequence of interest directly on chromosomes. Cytogenetic mapping has many applications, ranging from the study of structural chromosome changes (Mandáková and Lysák, 2008) to the determination of the positions and orientation of unassembled BAC contigs to support the development of physical maps (Pedrosa-Harand et al., 2009). In our specific case, cytogenetic markers capable of precise and unambiguous identification of chromosomes are helpful for purity checks of sorted chromosome fractions.

Various types of tandem repeats are valuable sources of useful cytogenetic markers. Satellites are tandemly repeated sequences with a repeat unit greater than 25 bp. Satellites of the *Afa* family have been used for identification of the D-genome chromosomes in wheat (Rayburn and Gill, 1986) and of barley chromosomes (Tsujimoto et al., 1997). Other satellites were used for karyotyping in barley (Brandes et al., 1995), *Avena* spp. (Katsiotis et al., 1997), banana (*Musa* spp.; Hřibová et al., 2007), and other plant species (Sharma and Raina, 2005) because they tend to form clusters, which facilitates their detection and positive chromosome identification. In this study, 11 newly characterized satellites provided localized signals on one or more chromosomes of meadow fescue and could serve as new cytogenetic markers in this species and probably also in ryegrasses.

Chromosomes of *Festuca* and *Lolium* spp. can be readily discriminated in hybrids using genomic in situ hybridization (Thomas et al., 1994; Kopecký et al., 2008a, 2008b). This method enables the visualization of parental chromatin in any stage of the mitotic cycle in natural allopolyploids or synthetic wide hybrids (Schwarzacher et al., 1989). It is assumed that this discrimination is based on the presence of abundant species-specific repetitive elements. From this perspective alone, it is of considerable interest to analyze the frequency and distribution of major repeats in parental species of a wide hybrid, to determine which of them, and in what proportions, contribute to the discrimination of parental genomes in hybrids. From observations using PCR made in this study, all newly identified tandem repeats were present in both genera, *Festuca* and *Lolium*. Southern hybridization and FISH revealed slight differences in the presence or signal intensity of individual fpTRs. Hence, none or just a few of the 15 tested appears responsible for distinguishing parental genomes. This indicates that tandem repeats are not responsible for the discrimination capacity of genomic in situ hybridization. However, newly identified tandem repeats provided new sources of cytogenetic markers and significantly increased levels of chromosome identification.

CONCLUSION

To our knowledge, this is the first report on the dissection of a complex and large forage grass genome using chromosome sorting. We sorted chromosome 4F of meadow fescue and sequenced it by Illumina. This provided a unique and rich resource to study the

genome organization of the species. The sequence data were used to estimate gene content, construct virtual gene order, and characterize repetitive elements. This approach provided an opportunity to describe synteny between chromosome 4F and genomes of model species (*B. distachyon*, rice, sorghum, and barley). Finally, we demonstrated the potential of sorted chromosomes for cytogenetic mapping of various repeats. Our results demonstrate that coupling the chromosome sorting and NGS technologies is a powerful approach that provides insights on chromosome structure and evolution at superior resolution.

MATERIALS AND METHODS

Plant Material

For chromosome sorting, seeds of meadow fescue (*Festuca pratensis* 'Fure'; $2n = 2x = 14$) were obtained from Dr. Arild Larson (Graminor). Seeds were germinated in the dark at 25°C distributed in petri dishes filled with distilled water for 5 to 7 d to achieve optimal root length (approximately 2–3 cm). For additional in situ hybridization experiments, *Lolium perenne* 'SR4220', meadow fescue cv Fure, and a diploid F1 hybrid of *L. perenne* cv SR4220 × meadow fescue cv Skawa were used. The plants was transferred to a hydroponic culture of Hydroponex at 0.9 g L^{-1} (Hu-Ben); after 5 to 7 d, actively growing root tips were collected to ice water for approximately 28 h, fixed in a 3:1 mixture of absolute ethanol and glacial acetic acid at 37°C for 7 d, stained in 1% acetocarmine for 2 h, and squashed in a drop of 45% acetic acid on clean microscope slides (Masoudi-Nejad et al., 2002).

Cell Cycle Synchronization and Accumulation of Metaphases

For chromosome flow sorting, synchronization of the cell cycle and accumulation of metaphases in root tips were performed according to Vrána et al. (2000). Distilled water in petri dishes was replaced by Hoagland nutrient solution (Gamborg and Wetter, 1975) with 1, 1.5, or 2 mM HU. After 18 h of incubation, the solution was replaced by a HU-free Hoagland solution. Samples of root tips were taken at 1-h intervals for up to 6 h, and the cell cycle synchrony and mitotic activity were examined with a microscope. To accumulate cells at metaphase, the seedlings were treated for 2 h with Hoagland solution containing 5 μM oryzalin after recovery from HU. Mitotic activity and metaphase frequency were analyzed on Feulgen-stained squash preparations.

Preparation of Chromosome Suspension

Chromosome suspension was prepared according to Doležel et al. (1992) with minor modifications. Seedlings were rinsed in deionized water and fixed in 2% (v/v) formaldehyde fixative made in Tris buffer supplemented with 0.1% (w/v) Triton X-100 at 5°C for 20 min. After washing in Tris buffer, about 150 root tips (approximately 2 mm long) were cut and transferred to a vial tube containing 750 μL of LB01 buffer (Doležel et al., 1989). The chromosomes were isolated after homogenization with a Polytron PT1300 homogenizer (Kinematica) at 20,000 rpm for 15 s. To remove large cellular fragments, the suspension was passed through a 50-μm pore size nylon mesh. Prior to flow cytometry, isolated chromosomes were stained with DAPI adjusted to a final concentration of 2 μg mL^{-1}.

Chromosome Analysis and Sorting

Chromosome analysis and sorting were done on a FACSAria II SORP flow cytometer (BD Biosciences) equipped with a 100-mW, 488-nm laser for scattered light detection and a 100-mW, 355-nm laser for DAPI excitation. Suspension of chromosomes was analyzed at rates of 800 to 1,000 events per second. Approximately 20,000 to 50,000 chromosomes were analyzed in each sample. DAPI fluorescence was collected through a 450/50-nm band-pass filter. To discriminate doublets, dot plots of DAPI fluorescence area versus DAPI fluorescence width were used. To ensure high purities of the sorted chromosomes, a tight gate was created around the chromosome 4F population, and chromosomes were sorted at rates of four to six chromosomes per second. To evaluate the purity of chromosome sorting, approximately 1,000 chromosomes were sorted onto microscope slides and tested by FISH. For DNA amplification, 40,000 chromosomes were collected into individual tubes and stored at −20°C until used.

Test of the Purity of Sorted Chromosomes Using FISH

To identify flow-sorted chromosomes and determine the extent of contamination of the sorted chromosome fractions, we performed FISH with probes for the 5S and 45S ribosomal DNA and two BAC clones (1G18 and 2N14). The two BAC clones were selected from a partial BAC library of meadow fescue (Kopecký et al., 2010). The DNA was labeled by the DIG-Nick Translation Kit or the biotin-Nick Translation Kit (Roche Applied Science) or by PCR labeling with biotin- or digoxigenin-labeled nucleotides (Roche). In situ hybridization was performed according to Kubaláková et al. (2003). Detection of hybridization sites was by anti-digoxigenin-fluorescein isothiocyanate (Roche) and streptavidin-Cy3 (Amersham), and counterstaining was with DAPI in the Vectashield antifade solution (Vector Laboratories). Preparations were screened with an Olympus AX70 microscope with epifluorescence and a SensiCam B/W camera.

Amplification and Sequencing of Chromosome 4F

Flow-sorted chromosomes were treated with proteinase, and their DNA was subsequently purified using Microcon YM-100 columns (Millipore) as described by Šimková et al. (2008) with minor modifications. Chromosomal DNA was amplified by MDA using the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Bio-Sciences). Sequencing of amplified chromosomal DNA was performed with HiSEquation 2000 (Illumina). Four micrograms of MDA-amplified DNA was used to create the corresponding shotgun DNA-seq library. The library for the 4F chromosome was run in a single lane at Istituto di Genomica Applicata Technology Services.

Assembly of Chromosome 4F Sequences and Genome Zipper Analysis

The Illumina paired-end reads were assembled by SOAPdenovo (Luo et al., 2012) using different k-mer sizes. The result of the 70-mer run provided the assembly with the best sequence coverage and L50. Repetitive DNA content was identified using Vmatch (http://www.vmatch.de) against the MIPS-REdat Poaceae version 8.6.2 repeat library by applying the following parameters: 70% identity cutoff, 100-bp minimal length, seed length 14, exdrop 5, and e-value 0.001. The Munich Information Center for Protein Sequences repeat library contains known grass transposons from the Triticeae Repeat Database (http://wheat.pw.usda.gov/ITMI/Repeats) as well as de novo-detected long terminal repeat retrotransposon sequences from various grass species. After repeat masking and filtering out the repeats, 26.6% (1,081,325 sequences) of the sequences remained and were considered for the subsequent steps.

To assess the number of syntenic conserved genes present on meadow fescue chromosome 4, the repeat-filtered contigs were aligned against the protein sequences of *Brachypodium distachyon* (version 1.2), rice (*Oryza sativa*; rice RAP-DB genome build 4), and sorghum (*Sorghum bicolor*; version 1.4) by BLASTX. The following stringent filtering criteria were applied: (1) only the first best hit with (2) a minimal alignment length of 30 amino acids and (3) a minimal sequence identity of 75% (*B. distachyon*)/70% (rice, sorghum). A sliding-window approach (window size of 0.5 Mb, shift size of 0.1 Mb) was used to identify segments with conserved gene order, based on the density of homology matches between *Festuca* spp. and the reference genomes.

The extracted conserved genes from the three model grass genomes and the corresponding *Festuca* spp. contigs were structured and ordered to a virtual linear gene model by using the GenomeZipper approach (Mayer et al., 2011). The approach uses a genetic marker scaffold to compare and subsequently integrate and order syntenic conserved, homologous genes (orthologs) along the scaffold. Thereby, intervals defined by the genetic markers serve as anchor points to define the corresponding intervals from the reference genomes. Genes from the least distant reference genome (in this case, *B. distachyon*) get highest

priority for ordering, and subsequently, the additional reference genomes are overlaid onto this scaffold. Due to the absence of a high-quality genetic marker map of meadow fescue, genetic markers of *L. perenne* chromosome 4 were used as a scaffold to anchor the collinear segments from the model genomes (Studer et al., 2012).

To analyze chromosomal rearrangements between *Festuca* spp. and barley (*Hordeum vulgare*), chromosome 4F contigs were compared against the physical map of the barley genome (Mayer et al., 2012) using BLASTX. Filtering criteria were adjusted to the evolutionary distance between meadow fescue and barley (at least 85% identity and a minimum alignment length of 100 bp).

To test whether the observed *Festuca* spp. chromosome 4/5 translocation is structurally identical to the 4/5 translocation found in Triticeae (Naranjo et al., 1987), the *Festuca* spp. 4F contigs, ordered gene scaffolds for wheat (*Triticum aestivum*) 4A and *Lolium* spp. 4L, and barley chromosomes 4H and 5H were compared by bidirectional BLAST hits and visualized using Circos (Krzywinski et al., 2009).

To estimate the number of conserved syntenic genes shared by wheat 4A, barley 4H, and *Festuca* spp. 4F, the *B. distachyon*, rice, and sorghum homologs contained in the individual linear ordered gene maps were compared with each other. A Venn diagram to illustrate shared genes in all three chromosomes was generated for each reference genome.

Repeat Reconstruction and Annotation

A random data set representing 1× coverage of chromosome 4F (550 Mb) was extracted from the Illumina data and used for the reconstruction of repetitive elements using a graph-based method according to Novák et al. (2010) using the Louvain clustering algorithm (Blondel et al., 2008). Resulting clusters of sequence reads were manually annotated using several sources, including similarity searches with RepeatMasker, and BLASTX and BLASTN programs were used for similarity searches against public databases and also against the database of domains derived from plant mobile elements (http://repeatexplorer.umbr.cas.cz). Clusters represented as graphs were also analyzed using the SeqGraphR program (http://w3lamc.umbr.cas.cz/lamc/resources.php). Dotter (Sonnhammer and Durbin, 1995) was used to confirm the presence of tandem organized repeats.

Physical Localization of Repeats Using FISH

FISH on flow-sorted chromosomes was done according to Kubaláková et al. (2003), while the protocol of Masoudi-Nejad et al. (2002) was used for metaphase plates prepared from root tips of the F1 hybrid. Probes were prepared for 15 putative tandem repeats (fpTR1–fpTR15) identified in sequence data using PCR labeling with biotin- or digoxigenin-labeled nucleotides (Roche) and pairs of specific primers. For chromosome identification in metaphase of F1 hybrids, *L. perenne*, and meadow fescue, we used a 5S ribosomal DNA (rDNA) probe accompanied by reprobing of slides with a 45S rDNA probe and a probe made from genomic DNA of *L. perenne*. The latter two probes were made using the biotin-Nick Translation Kit and the DIG-Nick Translation Kit (Roche), respectively. For the 45S rDNA probe, DNA clone pTa71 (Gerlach and Bedbrook, 1979) containing a 9-kb *Eco*RI fragment of wheat rDNA, which carries the 18S-5.8S-26S cluster of ribosomal RNA genes, was used. The probe for 5S rDNA was prepared using PCR with a pair of specific primers (RICRGAC1 and RICRGAC2) that amplify 303 bp in rice (Fukui et al., 1994), using rice genomic DNA as a template. Additionally, a probe for telomeric repeats was prepared using PCR with $(AGGGTTT)_3$ and $(CCCTAAA)_5$ primers without a template. Reprobing was done as described by Schwarzacher and Heslop-Harrison (2000). Probe hybridization signals were detected by anti-digoxigenin-fluorescein isothiocyanate and avidin-Cy3 conjugates, and counterstaining was by 0.2 mg mL^{-1} DAPI in Vectashield antifade solution (Vector Laboratories). Slides were evaluated with an Olympus AX70 microscope equipped with epifluorescence and a SensiCam B/W camera. ScionImage and Adobe Photoshop software were used for processing of color images.

Southern Hybridization

Aliquots of genomic DNA samples corresponding to 3×10^8 of the nuclear genomes of meadow fescue cv Fure and *L. perenne* cv SR4220 were digested using restriction enzymes *Hea*III, *Mse*I, and *Eco*RI, size fractionated by 1.2% agarose gel electrophoresis, and transferred to Hybond N^+ nylon membranes (Amersham). Probes specific for selected putative tandem repeats were amplified using specific primers (Table I) and labeled with biotin. The Southern hybridization was done at 68°C overnight, and signals were detected using the BrightStar BioDetect kit according to the manufacturer's instructions to 90% stringency (Ambion), incubated with chemiluminescent substrate (CDP-Star; Amersham Biosciences), and exposed on x-ray film.

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers JX624129 to JX624136.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Collinearity of 4F with rice.

Supplemental Figure S2. Collinearity of 4F with sorghum.

Supplemental Figure S3. Collinearity of 4F, 4A of wheat, and 5H of barley.

Supplemental Figure S4. Distribution of chromosome 4F reads in sequence clusters.

Supplemental Figure S5. Dot plot of five tandem repeats.

Supplemental Table S1. Summary of the detection of tandem repeats.

ACKNOWLEDGMENTS

We are grateful to Dr. Arild Larsen and Dr. Vladimír Černoch for providing plant material. We thank Dr. Jarmila Číhalíková, Romana Nováková, and Zdeňka Dubská for assistance with chromosome sorting and DNA amplification. Special thanks go to Adam J. Lukaszewski for critical reading and valuable comments.

Received June 26, 2013; accepted October 4, 2013; published October 4, 2013.

LITERATURE CITED

Alm V, Fang C, Busso CS, Devos KM, Vollan K, Grieg Z, Rognli OA (2003) A linkage map of meadow fescue (*Festuca pratensis* Huds.) and comparative mapping with other Poaceae species. Theor Appl Genet **108:** 25–40

Baird JH, Kopecký D, Lukaszewski AJ, Green RL, Bartoš J, Doležel J (2012) Genetic diversity of turf-type tall fescue using diversity arrays technology. Crop Sci **52:** 408–412

Bartoš J, Sandve SR, Kölliker R, Kopecký D, Christelová P, Stočes S, Østrem L, Larsen A, Kilian A, Rognli OA, et al (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

Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech **2008:** P10008

Brandes A, Röder MS, Ganal MW (1995) Barley telomeres are associated with two different types of satellite DNA sequences. Chromosome Res **3:** 315–320

Byrne S, Panitz F, Hedegaard J, Bendixen C, Studer B, Farrell JD, Swain SC, Armstead I, Caccamo M, Asp T (2011) *De novo* genome sequencing of perennial ryegrass (*Lolium perenne*). *In* S Barth, D Milbourne, eds, Abstracts of the EUCARPIA 29th Fodder Crops and Amenity Grasses Section Meeting. Teagasc, Dublin, p 23

Chalupska D, Lee HY, Faris JD, Evrard A, Chalhoub B, Haselkorn R, Gornicki P (2008) Acc homoeoloci and the evolution of wheat genomes. Proc Natl Acad Sci USA **105:** 9691–9696

Choulet F, Wicker T, Rustenholz C, Paux E, Salse J, Leroy P, Schlub S, Le Paslier M-C, Magdelenat G, Gonthier C, et al (2010) Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. Plant Cell **22:** 1686–1701

Doležel J, Bartoš J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. Cytometry A **51:** 127–128, author reply 129

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

Doležel J, 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

Doležel J, Šimková H, Kubaláková M, Šafář J, Suchánková P, Číhalíková J, Bartoš J, Valárik M (2009) Chromosome genomics in the Triticeae. *In* C Feuillet, GJ Muehlbauer, eds, Genetics and Genomics of the Triticeae. Plant Genetics and Genomics: Crops and Models 7. Springer Science +Business Media, New York, pp 285–316

Donnison IS, O'Sullivan DM, Thomas A, Canter P, Moore B, Armstead I, Thomas H, Edwards KJ, King IP (2005) Construction of a *Festuca pratensis* BAC library for map-based cloning in *Festulolium* substitution lines. Theor Appl Genet **110:** 846–851

Egan AN, Schlueter J, Spooner DM (2012) Applications of next-generation sequencing in plant biology. Am J Bot **99:** 175–185

Fukui K, Kamisugi Y, Sakai F (1994) Physical mapping of 5S rDNA loci by direct-cloned biotinylated probes in barley chromosomes. Genome **37:** 105–111

Gale MD, Devos KM (1998) Plant comparative genetics after 10 years. Science **282:** 656–659

Gamborg OL, Wetter LR (1975) Plant Tissue Culture Methods. National Research Council of Canada, Saskatoon, Canada

Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Res **7:** 1869–1885

Hernandez P, Martis M, Dorado G, Pfeifer M, Gálvez S, Schaaf S, Jouve N, Šimková H, Valárik M, Doležel J, et al (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

Hřibová E, Dolezelová M, Town CD, Macas J, Doležel J (2007) Isolation and characterization of the highly repeated fraction of the banana genome. Cytogenet Genome Res **119:** 268–274

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

Katsiotis A, Hagidimitriou M, Heslop-Harrison JS (1997) The close relationship between the A and B genomes in *Avena* L. (Poaceae) determined by molecular cytogenetic analysis of total genomic, tandemly and dispersed repetitive DNA sequences. Ann Bot (Lond) **79:** 103–109

Kopecký D, Bartoš J, Christelová P, Černoch V, Kilian A, Doležel J (2011) Genomic constitution of *Festuca* × *Lolium* hybrids revealed by the DArTFest array. Theor Appl Genet **122:** 355–363

Kopecký D, Bartoš J, Lukaszewski AJ, Baird JH, Černoch V, Kölliker R, Rognli OA, Blois H, Caig V, Lübberstedt T, et al (2009) Development and mapping of DArT markers within the *Festuca-Lolium* complex. BMC Genomics **10:** 473

Kopecký D, Havránková M, Loureiro J, Castro S, Lukaszewski AJ, 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

Kopecký D, Lukaszewski AJ, Doležel J (2008a) Cytogenetics of Festulolium (Festuca × Lolium hybrids). Cytogenet Genome Res **120:** 370–383

Kopecký D, Lukaszewski AJ, Doležel J (2008b) Meiotic behaviour of individual chromosomes of *Festuca pratensis* in tetraploid *Lolium multiflorum*. Chromosome Res **16:** 987–998

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res **19:** 1639–1645

Kubaláková M, Valárik M, Barto 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

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, et al (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience **1:** 18

Mandáková T, Lysák MA (2008) Chromosomal phylogeny and karyotype evolution in x=7 crucifer species (Brassicaceae). Plant Cell **20:** 2559–2570

Martienssen RA, Rabinowicz PD, O'Shaughnessy A, McCombie WR (2004) Sequencing the maize genome. Curr Opin Plant Biol **7:** 102–107

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

Mayer KFX, Martis M, Hedley PE, Simková H, Liu H, Morris JA, Steuernagel B, Taudien S, Roessner S, Gundlach H, et al (2011) Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell **23:** 1249–1263

Mayer KF, Taudien S, Martis M, Simkova H, Suchankova P, Gundlach H, Wicker T, Petzold A, Felder M, Steuernagel B, et al (2009) Gene content and virtual gene order of barley chromosome 1H. Plant Physiol **151:** 496–505

Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K, Close TJ, et al (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature **491:** 711–716

Naranjo T, Roca A, Goicoecha PG, Giraldz R (1987) Arm homoeology of wheat and rye chromosomes. Genome **29:** 873–882

Novák P, Neumann P, Macas J (2010) Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. BMC Bioinformatics **11:** 378

Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, Korol A, Michalak M, Kianian S, Spielmeyer W, et al (2008) A physical map of the 1-gigabase bread wheat chromosome 3B. Science **322:** 101–104

Pedrosa-Harand A, Kami J, Gepts P, Geffroy V, Schweizer D (2009) Cytogenetic mapping of common bean chromosomes reveals a less compartmentalized small-genome plant species. Chromosome Res **17:** 405–417

Peterson DG, Wessler SR, Paterson AH (2002) Efficient capture of unique sequences from eukaryotic genomes. Trends Genet **18:** 547–550

Pfeifer M, Martis M, Asp T, Mayer KFX, Lübberstedt T, Byrne S, Frei U, Studer B (2013) The perennial ryegrass GenomeZipper: targeted use of genome resources for comparative grass genomics. Plant Physiol **161:** 571–582

Polok K (2007) Molecular Evolution of the Genus *Lolium* L. Studio Poligrafii Komputerowej, Olsztyn, Poland

Rabinowicz PD, Schutz K, Dedhia N, Yordan C, Parnell LD, Stein L, McCombie WR, Martienssen RA (1999) Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. Nat Genet **23:** 305–308

Rayburn AL, Gill BS (1986) Molecular identification of the D-genome chromosomes of beat. J Hered **77:** 253–255

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

Schwarzacher T, Heslop-Harrison P (2000) Practical in Situ Hybridization. BIOS Scientific Publishers, Oxford

Schwarzacher T, Leitch AR, Bennett MD, Heslop-Harrison JS (1989) *In situ* localization of parental genomes in a wide hybrid. Ann Bot (Lond) **64:** 315–324

Sharma S, Raina SN (2005) Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. Cytogenet Genome Res **109:** 15–26

Šimková H, Svensson JT, Condamine P, Hřibová E, Suchánková P, Bhat PR, Bartoš J, Šafář J, Close TJ, Doležel J (2008) Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genomics **9:** 294

Sonnhammer ELL, Durbin R (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene **167:** GC1–GC10

Studer B, Byrne S, Nielsen RO, Panitz F, Bendixen C, Islam MS, Pfeifer M, Lübberstedt T, Asp T (2012) A transcriptome map of perennial ryegrass (*Lolium perenne* L.). BMC Genomics **13:** 140

Suchánková P, Kubaláková M, Kovárová P, Bartoš J, Číhalíková J, Molnár-Láng M, Endo TR, Doležel J (2006) Dissection of the nuclear genome of barley by chromosome flow sorting. Theor Appl Genet **113:** 651–659

Thiel T, Graner A, Waugh R, Grosse I, Close TJ, Stein N (2009) Evidence and evolutionary analysis of ancient whole-genome duplication in barley predating the divergence from rice. BMC Evol Biol **9:** 209

Thomas HM, Morgan WG, Meredith MR, Humphreys MW, Thomas H, Leggett JM (1994) Identification of parental and recombined chromosomes in hybrid derivatives of *Lolium multiflorum* × *Festuca pratensis* by genomic *in situ* hybridization. Theor Appl Genet **88:** 909–913

Tomaszewski C, Byrne S, Foito A, Kildea S, Kopecký D, Doležel J, Heslop-Harrison JS, Stewart D, Barth S (2012) A genetic linkage map of an F2 perennial ryegrass population highly enriched with DArT markers suitable for mapping of crown rust resistence QTL. Plant Breed **131:** 345–349

Tsujimoto H, Mukai Y, Akagawa K, Nagaki K, Fujigaki J, Yamamoto M, Sasakuma T (1997) Identification of individual barley chromosomes based on repetitive sequences: conservative distribution of *Afa*-family repetitive sequences on the chromosomes of barley and wheat. Genes Genet Syst **72:** 303–309

Vigeland MD, Spannagl M, Asp T, Paina C, Rudi H, Rognli OA, Fjellheim S, Sandve SR (2013) Evidence for adaptive evolution of low-temperature stress response genes in a Pooideae grass ancestor. New Phytol **199:** 1060–1068

Vitulo N, Albiero A, Forcato C, Campagna D, Dal Pero F, Bagnaresi P, Colaiacovo M, Faccioli P, Lamontanara A, Šimková H, et al (2011) First survey of the wheat chromosome 5A composition through a next generation sequencing approach. PLoS ONE **6:** e26421

Vrána J, Kubaláková M, Simková H, Cíhalíková J, Lysák MA, Doležel J (2000) Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). Genetics **156:** 2033–2041

Wenzl P, Suchánková P, Carling J, Simková H, Huttner E, Kubaláková M, Sourdille P, Paul E, Feuillet C, Kilian A, et al (2010) Isolated chromosomes as a new and efficient source of DArT markers for the saturation of genetic maps. Theor Appl Genet **121:** 465–474

Wicker T, Mayer KFX, Gundlach H, Martis M, Steuernagel B, Scholz U, Šimková H, Kubaláková M, Choulet F, Taudien S, et al (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

Zhou RN, Hu ZM (2007) The development of chromosome microdissection and microcloning technique and its applications in genomic research. Curr Genomics **8:** 67–72

APPENDIX VI

Low-copy FISH on barley chromosomes

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Book of Abstracts “Olomouc Biotech 2011 - Plant Biotechnology: Green for Good”. P.49. Olomouc, 2011.

[Poster presentation]

LOW-COPY FISH ON BARLEY CHROMOSOMES

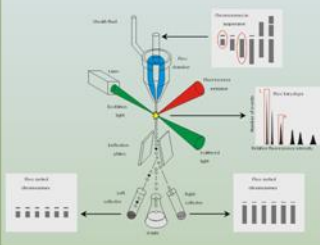
Havránková Miroslava[1], Kopecký David[1], Bartoš Jan[1],Stein Nils[2], Sato Kazuhiro[3], Doležel Jaroslav[1]

[1)] Centre of the Region Haná for Biotechnological and Agricultural Research,
Institute of Experimental Botany, Sokolovská 6, 772 00 Olomouc, Czech Republic
[2)] Leibniz Institute of Plant Gnetics and Crop Plant Research (IPK),
06466 Gatersleben, Germany
[3)] Barley Germplasm Center, Research Institute for Bioresources, Okayama University,
Kurashiki, 710-0046, Japan

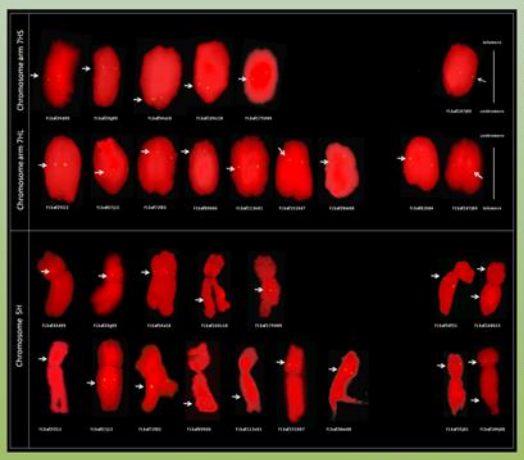
BACKGROUND

Barley *(Hordeum vulgare* L.; 2n = 2x = 14) is one of the major crops cultivated in the temperate zone. Large genome size (5.1Gbp/1C) with 80% of repetitive DNA makes studies of barley genome complicated. It is difficult to map DNA markers in regions with low frequency of recombination, especially in centromeric and pericentromeric regions. One of the methods, which allows to overcome this limitation, is the fluorescence *in situ* hybridization (FISH). Metaphase chromosome spreads are common targets for cytogenetic mapping. However, chromosomes purified by flow cytometric sorting can be used to increase the sensitivity and spatial resolution of FISH.

Chromosome sorting using flow cytometry



DAPI-stained chromosomes pass individually through a beam of intense light. Emitted fluorescence of stained chromosomes is quantified. Chromosomes can be purified by breaking the liquid stream into droplets and usage of an electrostatic field to deflect electrically charged droplets containing chromosomes of interest into a collection tube.



FISH on flow-sorted chromosome arms 7HS and 7HL and chromosome 5H of barley with cDNA probes. Signal localization is indicated by white arrows (red pseudocolor: DAPI, yellow pseudocolor: biotin).

MATERIAL AND METHODS

Mitotic chromosomes from synchronized root tips meristems were used for cytogenetic mapping. Set of 25 full-lenght cDNA were localized on metaphase chromosome spreads of barley (cv. Morex) and on flow-sorted chromosome 5H and chromosome arms 7HS and 7HL of barley. FISH was done according to protocol of Heslop-Harrison and Schwarzacher (2000) with minor modifications. The probes were prepared by Nick-translation and labelled with biotin. Signals were detected with Texas Red avidin after two rounds of amplification with biotinylated anti-avidin and Texas Red avidin. Chromosomes were counterstained with DAPI.

RESULTS

In this project, we localized 23 out of 25 full-lenght cDNA on barley metaphase spreads and flow-sorted chromosomes using FISH. These fl-cDNAs were genetically mapped in centromeric region of chromosome 7H. The lenght of cDNA clones ranged from 2 – 3.5 kb. Overall, 19 of these clones provide fluorescent signal on chromosome 7H. Out of these, 12 clones were detected in centromeric or pericentromeric regions. These results are in agreement with previous genetic studies. Unexpectedly, signals of 19 out of 25 cDNAs were also detected on chromosome 5H. This can be explained by the existence of ancient segmental duplication between chromosomes 5H and 7H.

FISH on metaphase chromosomes spreads of barley with cDNA clone FLbaf25I12 labelled with biotin (red). White arrows indicate localization of fluorescent signal on a pair of chromosomes 7H.

REFERENCES

Schwarzacher and Hespol-Harrison (2000): Practical *in situ* hybridization, BIOS Scientific Publishers Limited, Oxford, UK

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. CZ.1.05/2.1.00/01.0007).

APPENDIX VII

5H/7H segmental duplication in barley revealed by cDNA-FISH

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Abstracts of the 18th International Chromosome Conference. P. 85. Manchester, 2011.

[Poster presentation]

5H/7H segmental duplication in barley revealed by cDNA-FISH

Miroslava Havránková[1], David Kopecký[1], Jan Bartoš[1], Nils Stein[2], Kazuhiro Sato[3], Jaroslav Doležel[1]

[1)] Centre of the Region Haná for Biotechnological and Agricultural Research,
Institute of Experimental Botany, Sokolovská 6, 772 00 Olomouc, Czech Republic, email: havrankova@ueb.cas.cz
[2)] Leibniz Institute of Plant Genetics and Crop Plant Research,
06466 Gatersleben, Germany
[3)] Barley Germplasm Center, Research Institute for Bioresources, Okayama University,
Kurashiki, 710-0046, Japan

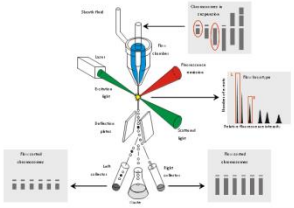
BACKGROUND

The evolution of plant genomes is closely associated with chromosome rearrangements and these structural changes play a crucial role in the establishment of reproductive isolation and of speciation. On the basis of DNA sequence analysis, multiple ancient polyploidization events were detected in several plant genomes including *Arabidopsis thaliana*, poplar and rice. In diploid species, polyploidization is probably followed by gene lost and gene silencing and only some duplications are conserved as traces of polyploidization. The ancient segmental duplications were found also in barley genome (*Hordeum vulgare* L., 2n = 2x = 14). One of them was proposed to involve chromosomes 5H and 7H. In this study, we used fluorescence *in situ* hybridization (FISH) with 25 full-length cDNA clones as probes to prove this duplication. To increase the efficiency of FISH, flow-sorted chromosome arms 5HS, 5HL, 7HS and 7HL were used as targets.

METHODS

Mitotic chromosomes were used as a target DNA for cytogenetic mapping. 25 full-length cDNA clones were localized on metaphase spreads and flow-sorted chromosome arms using FISH. Liquid suspensions of chromosomes were prepared from synchronized root tip meristems and the chromosome arms were sorted onto microscopic slides. The probes were prepared by Nick-translation and labeled with biotin.

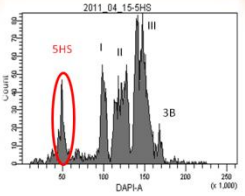
Chromosome sorting using flow cytometry



DAPI-stained chromosomes pass individually through a beam of intense light. Emitted fluorescence of stained chromosomes is quantified. Chromosomes can be purified by breaking the liquid stream into droplets and usage of an electrostatic field to deflect electrically charged droplets containing chromosomes of interest onto a microscopic slide.

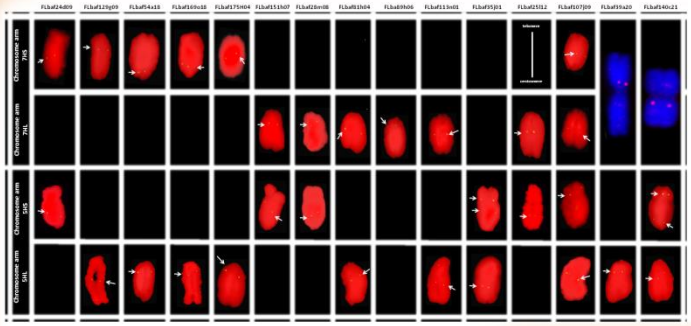
MATERIAL

- Barley cv. „Morex" (2n=14)
- Barley addition lines of common wheat cv. Chinese Spring
  - Barley 5HS addition line (21"+t", 2n=44)
  - Barley 5HL addition line (21"+t", 2n=44)
  - Barley 7HS addition line (21"+t", 2n=44)
  - Barley 7HL addition line (21"+t", 2n=44)

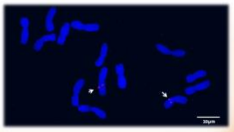


Example of sorting barley chromosome arms. Flow karyotype obtained after analyzing wheat-barley 5HS addition line contains well-resolved peak of chromosome arm 5HS, and four peaks representing chromosomes of wheat (three composite peaks (I, II, III) and a peak representing chromosome 3B.

RESULTS



We localized 19 of 25 fl-cDNA clones on barley metaphase spreads and/or on flow-sorted chromosome arms using FISH. 14 (56 %) cDNA clones of these provided a visible signal on both chromosomes 5H and 7H. Signal of 2 clones occurred only on chromosome 5H and other 3 clones just on chromosome 7H. Two cDNAs provided signal on sorted arms of chromosome 5H and on 7H of metaphase spread, but not on flow-sorted chromosome arms 7HS/7HL (last two columns in Figure). Despite of the less condensed chromatin of flow-sorted chromosomes, it was not possible to order linearly these clones on particular chromosome arms .



FISH on mitotic metaphase chromosome spreads of barley with cDNA clone FLbaf25I12 labeled with biotin (red). Arrows indicate localization of fluorescent signals on a pair of chromosomes 7H.

CONCLUSION

Our results confirmed the ancient segmental duplication between chromosomes 5H/7H and indicated suitability of this approach to verify similar chromosome rearrangements.

IN PROGRESS:

- Mapping cDNA clones by FISH using super-stretched flow-sorted barley chromosomes
- Localization of PCR based probes of the identical cDNA collection

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. CZ.1.05/2.1.00/01.0007).

APPENDIX VIII

Mapping of non-recombining region of barley chromosome 7H using multicolor FISH

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Abstracts of the 10th International Ph.D. Student Conference on Experimental Plant Biology. – Bulletin of the Czech Society of Experimental Plant Biology and the Physiological Section of the Slovak Botanical Society. P. 109. Brno, 2012.

[Poster presentation]

Mapping non-recombining region of barley chromosome 7H using multicolor FISH

Miroslava Havránková[1], David Kopecký[1], Lu Ma[2], Jan Bartoš[1], Kazuhiro Sato[3], Houben Andreas [2], Nils Stein[2], Jaroslav Doležel[1]

[1)] Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Sokolovská 6, 772 00 Olomouc, Czech Republic, email: havrankova@ueb.cas.cz
[2)] Leibniz Institute of Plant Genetics and Crop Plant Research, 06466 Gatersleben, Germany
[3)] Barley Germplasm Center, Research Institute for Bioresources, Okayama University, Kurashiki, 710-0046, Japan

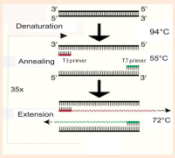
BACKGROUND

Mapping the markers or sequences is common approach to study the genome stucture and organization of different plant species. Genetic mapping, which is based on cross-over events, has a limited resolution in non-recombining regions. This limitation could be overcome using some technique of physical mapping as radiation hybrid mapping, optical mapping or fluorescence *in situ* hybridization (FISH). FISH is useful and quick technique to map DNA sequences directly on chromosomes. The probes longer than 10 kb are localized routinely, but localization of the short probes is still quite challenging. In this study, we prove that FISH with probes shorter than 3kb is possible and provide the stable fluorescent signal in plant species. It opens the possibility to use coding sequences without repetitive patterns for mapping which results in lower hybridization background.

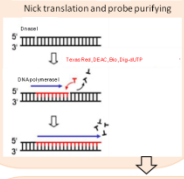
METHODS

1.Probe preparing

Plasmid isolation and cDNA amplification

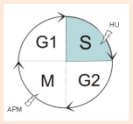


Nick translation and probe purifying

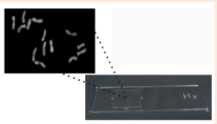


2. Chromosome preparing

Cell cycle synchronization and root tips fixation (plant material: barley cv. „Morex")

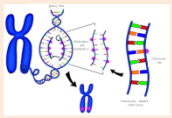


Metaphase spreads and postfixation



3. Fluorescence *in situ* hybridization (FISH)

Fluorescence *in situ* hybridization



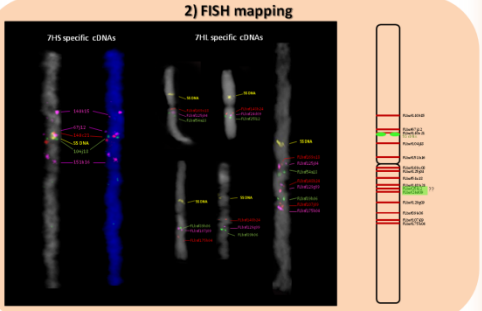
Microscope analyzing

Final image adjustment (Adobe Photoshop 6.0 software) and chromosome straightening (Image J software)

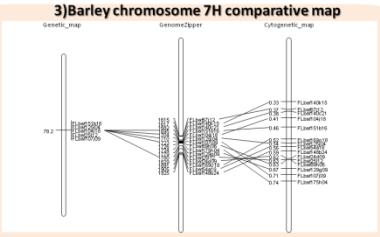
RESULTS

1) Blast results

2) FISH mapping



3)Barley chromosome 7H comparative map



CONCLUSION

- Using FISH, we solved the order of the 13 cDNA clones, which were previously mapped in the identical position on genetic map. Clones FLbaf24d09 and FLbaf25l12 are too close to distinguish their mutual position on prometaphase chromosomes.
- According to the cytogenetic map, non-recombining region spanning the pericentromeric area of chromosome 7H represent more than 30 % of the chromosome length
- In this study, we confirm the crucial importance of FISH for physical mapping and proved that low-copy FISH provides the alternative approach to overcome the troubles of mapping in the region with the low recombination frequency.

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. CZ.1.05/2.1.00/01.0007).

APPENDIX IX

Physical mapping of non-recombining genomic regions in barley using FISH

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Abstracts of the International Conference “Plant and Animal Genome XXI”. P. 218. Sherago International, Inc., San Diego, 2013

[Poster presentation]

Physical mapping of non-recombining genomic regions in barley using FISH

Havránková M. [1], Bartoš J.[1], Kopecký D.[1], Lu M.[2], Sato K.[3], Stein N.[2], Doležel J.[1]*

1) Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Sokolovská 6, 772 00 Olomouc, Czech Republic
2) Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Seeland OT Gatersleben, Germany
3) Barley Germplasm Center, Research Institute for Bioresources, Okayama University, Kurashiki, 710-0046, Japan

Background

Physical map assembly relies on ordering contigs of overlapping BAC clones. The contigs are oriented and ordered along chromosomes after integration with genetic maps. Unfortunately, this approach fails in genomic regions with highly reduced recombination frequency, where genetic maps suffer from poor resolution. This limitation could be overcome using other methods such as radiation hybrid mapping, optical mapping and fluorescence *in situ* hybridization (FISH). FISH is a useful and rapid technique to map DNA sequences directly on chromosomes. However, while he probes longer than 10 kb are localized routinely, localization of short probes remains a challenging task. Here we demonstrate that FISH with single-copy sequences shorter than 3kb such as cDNAs is feasible in barley. This advance opens a way for using coding sequences free of repetitive DNA elements for FISH, which results in lower hybridization background. Our observations confirm the constraints of using genetic markers in anchoring physical maps through significant parts of large plant genomes, underline a need for alternative approaches, and demonstrate that FISH is a feasible approach to anchor BAC contigs.

Plant material

- *Hordeum vulgare* L. (2n=14), cv. Morex

Target DNA:
- Mitotic metaphase and prometaphase chromosome plates prepared from synchronized root tip meristem cells

Methods

Probe preparation
Plasmid isolation
↓
cDNA amplification
↓
Nick translation
↓
Probe purification

Chromosome preparation
Cell cycle synchronization
↓
Root tip fixation
↓
Squashing
↓
Postfixation

Fluorescence *in situ* hybridization (FISH)

Image processing
Final adjustment (Adobe Photoshop 6.0.)
↓
Chromosome straightening (image processing program ImageJ v1.43)

Conclusions

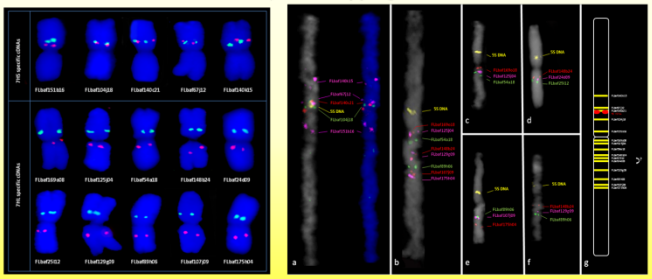
- Using FISH, we determined the order of 13 out of 15 cDNAs clones, which were previously mapped in identical position on a genetic map of barley chromosome 7H. cDNAs FLbaf24d09 and FLbaf25l12 are too close to distinguish their mutual position on prometaphase chromosomes.
- According to the cytogenetic map, non-recombining region spanning the pericentromeric area of chromosome 7H represents more than 30% of the chromosome length.
- We confirm the potential of FISH for physical mapping and prove that low-copy FISH provides the alternative approach to overcome the difficulties in physical mapping genomic regions with low recombination frequency.

Results

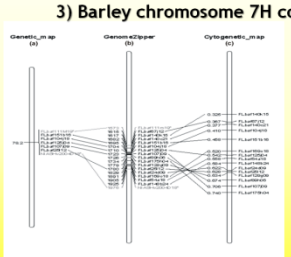
1) Blast results

Full-length cDNA	Accesion number	Length (bp)	Blast results
FLbaf67j12	AK250219	2300	7HS
FLbaf140c21	AK252013	3400	7HS
FLbaf104j18	AK251038	2475	7HS
FLbaf140k15	AK248620	2127	7HS
FLbaf151b16	AK252317	2328	7HS
FLbaf169o18	AK248228	2488	7HS
FLbaf24d09	AK249246	3499	7HL
FLbaf25l12	AK249387	2447	7HL
FLbaf54a18	AK249749	2661	7HL
FLbaf89h06	AK250597	3083	7HL
FLbaf107j09	AK248217	2038	7HL
FLbaf125j04	AK251498	3382	7HL
FLbaf129g09	AK251673	2295	7HL
FLbaf148b24	AK252034	2568	7HL
FLbaf175h04	AK252946	3101	7HL

2) FISH mapping



3) Barley chromosome 7H comparative map



a) Out of 15 cDNA clones originating from centromeric region of 7H, seven cDNAs were mapped genetically to the same locus (Mayer *et al.*, 2011)
b) The application of GenomeZipper enabled tentative ordering all 15 cDNAs (Mayer *et al.*, 2011)
c) Our results demonstrate that cytogenetic mapping using FISH can improve the ordering of markers in the area of interest

Acknowledgements

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programe Research and Development for Innovations No. ED0007/01/01 and Internal Grant Agency of Palacky University (Grant award No. IGA PrF/2012/001).

Reference: Mayer K.F.X., Martis M., Hedley P.E., Šimková H. *et al.* (2011): Unlocking the Barley Genome by Chromosomal and Comparative Genomics. The Plant Cell, 23, 1249-1263.

APPENDIX X

The potencial of low-copy FISH in physical mapping

Havránková M., Kopecký D., Bartoš J.,Stein N., Sato K., Doležel J.

In: Book of Abstracts “Olomouc Biotech 2013 - Plant Biotechnology: Green for Good”. P.70. Olomouc, 2013.

[Poster presentation]

THE POTENCIAL OF LOW-COPY FISH IN PHYSICAL MAPPING

Havránková M. [1], Bartoš J.[1], Kopecký D.[1], Lu M.[2], Sato K.[3], Stein N.[2], Doležel J.[1]

1) Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Sokolovská 6, 772 00 Olomouc, Czech Republic
2) Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Seeland OT Gatersleben, Germany
3) Barley Germplasm Center, Research Institute for Bioresources, Okayama University, Kurashiki, 710-0046, Japan

Background

The process of creating high resolution physical maps relies on ordering contigs of overlapping BAC clones and integration with genetic maps. Unfortunately, the ordering contigs is very difficult in the region with highly reduced frequency of recombination, where genetic maps suffer from the poor resolution. The existence of non-recombining regions underline the necessity of additional approach, which could help to overcome this imperfection. FISH (fluorescence *in situ* hybridization) is one of the possible techniques, which is not limited by the presence of crossing overs and could increase the number of the markers in these problematic areas. In this work, we concentrated on mapping of short sequences and our results indicate, that single-copy FISH in barley is feasible. Furthermore, this progress opens the possibilities also for other application. Recently, genomes or their portions are massively sequenced and FISH with short, repeat-free probes, is an attractive alternative for anchoring contigs and could facilitates the genome sequence assembling.

Plant material

- *Hordeum vulgare* L. (2n=14), cv. Morex

Target DNA:
- Mitotic metaphase and prometaphase chromosome plates prepared from synchronized root tip meristem cells

Methods

Probe preparation

Plasmid isolation
↓
cDNA amplification
↓
Nick translation
↓
Probe purification

Chromosome preparation

Cell cycle synchronization
↓
Root tip fixation
↓
Squashing
↓
Postfixation

Fluorescence *in situ* hybridization (FISH)

Image processing

Final adjustment (Adobe Photoshop 6.0.)
↓
Chromosome straightening (image processing program ImageJ v1.43)

Conclusions

- We focused on ordering single copy sequences in non-recombining region of barley chromosome 7H

- Using FISH, we ordered 13 out of 15 cDNA clones. Signals of the clones FLbaf129g09 and FLbaf25l12 co-localized to the same positions and clones FLbaf24d09 and FLbaf25l12 provided ambiguous results.

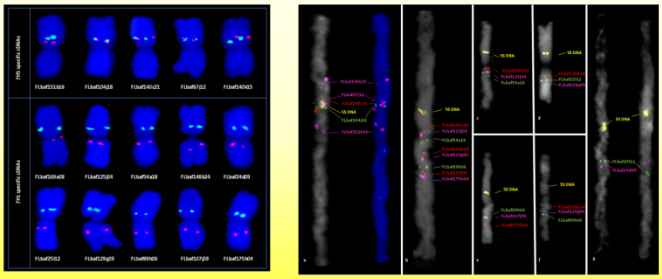
- In this study, we underline the importance of cytogenetics in physical mapping. We proved that cytogenetic tools could significantly increase the resolution of the physical map in non-recombining regions.

Results

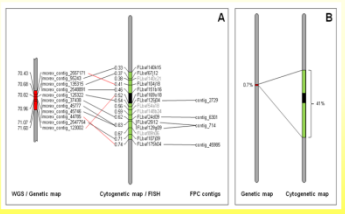
1) Blast results

Full-length cDNA	Accession number	Length (bp)	Signal location	Relative position*	Hit name	Chromosome ID	Genetic position (cM)	FPC_contig
FLbaf140k15	AK248620	2,127	7HS	0.33	morex_contig_55243	7H	70.68	---
FLbaf67j12	AK250219	2,3	7HS	0.37	morex_contig_135315	7H	70.68	---
FLbaf140c21	AK252013	3,4	7HS	0.38	morex_contig_160232	---	---	---
FLbaf104j18	AK251038	2,475	7HS	0.41	morex_contig_2557171	7H	70.43	---
FLbaf151b16	AK252317	2,328	7HS	0.46	morex_contig_2548891	7H	70.68	---
FLbaf169o18	AK248228	2,488	7HL	0.52	morex_contig_123002	7H	71.60	---
FLbaf125j04	AK251498	3,382	7HL	0.54	morex_contig_125322	7H	70.82	contig_2729
FLbaf54a18	AK249749	2,661	7HL	0.56	morex_contig_359593	---	---	---
FLbaf148b24	AK252034	2,568	7HL	0.59	morex_contig_1564111	---	---	---
FLbaf24d09	AK249246	3,499	7HL	0.62	morex_contig_37438	7H	70.96	contig_6301
FLbaf25l12	AK249387	2,447	7HL	0.63	morex_contig_45777	7H	70.96	contig_714
FLbaf129g09	AK251673	2,295	7HL	0.63	morex_contig_45746	7H	70.96	contig_714
FLbaf89h06	AK250597	3,083	7HL	0.67	morex_contig_6028	---	---	---
FLbaf107j09	AK248217	2,038	7HL	0.71	morex_contig_2547754	7H	71.07	---
FLbaf175h04	AK252946	3,101	7HL	0.74	morex_contig_44785	7H	70.96	contig_45985

2) FISH mapping



3) Barley chromosome 7H comparative map



A: Comparison of the cytogenetic map and genetic order of fl-cDNA (TIBGSC, 2012). We corrected the position of three cDNAs using FISH.

B: According to our results, the locus which was considered to represent 0.7 % of the genetic map, actually make up more than 40 % of the chromosome length.

Acknowledgements

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programe Research and Development for Innovations No. ED0007/01/01) and Internal Grant Agency of Palacky University (Grant award No. IGA PrF/2012/001).

References: TIBGSC (2012) A physical, genetical and functional sequence assembly of the barley genome.. Nature 491, 711-716

APPPENDIX XI

Chromosome painting in barley – a new milestone in cytogenetics of cereals

Havránková M., Knauft M., Bartoš J., Vrána J., Kubaláková M., Stein N., Doležel J.

In: Abstracts of the 19th International Chromosome Conference. P. 78. Bologna, 2013.

[Poster presentation]

CHROMOSOME PAINTING IN BARLEY - A NEW MILESTONE IN CYTOGENETICS OF CEREALS

Havránková M.[1], Knauft M.[2], Bartoš J.[1], Vrána J.[1], Kubaláková M.[1], Stein N.[2], Doležel J.[1]

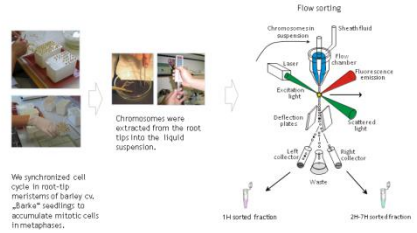
1) Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Sokolovská 6, 772 00 Olomouc, Czech Republic
2) Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Seeland OT Gatersleben, Germany

Background

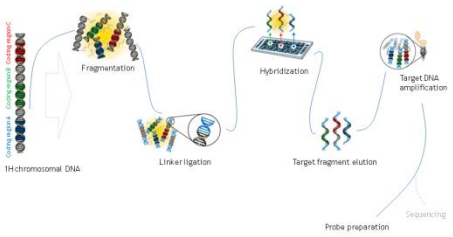
Chromosome painting is an appealing technique, which significantly increased the attractiveness of cytogenetics. It is not only a technique of basic research, but an important tool of clinical cytogenetics, which helps to uncover serious diseases and save human lives. The history of this technique covers several decades, has gone through changes, diversified a lot and was adapted to current demands. The principle of the method is generating fluorescently labeled probes from whole chromosomes, which are used for fluorescence *in situ* hybridization on metaphase chromosomes and interphase nuclei. The technique was originally developed for analysis of human cells. The attempts to utilize chromosome painting with composite chromosome probes in plants failed, mainly due to the presence of dispersed repeats. Here we describe a novel approach suitable for chromosome painting in plants with large genomes. The method relies on the ability to prepare chromosome painting probes composed mainly from low-copy coding sequences, which are obtained after gene capture from chromosomes isolated by flow cytometric sorting.

Material and Methods

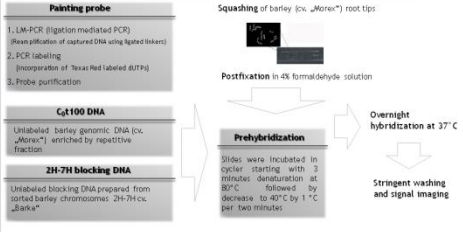
Flow cytometric chromosome sorting



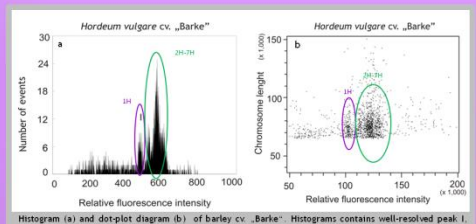
Gene capture



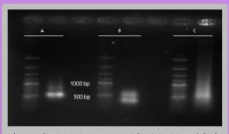
Probe preparation and FISH



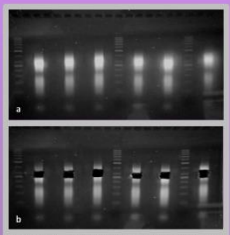
Results



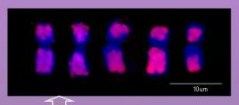
Histogram (a) and dot-plot diagram (b) of barley cv. „Barke". Histograms contains well-resolved peak 1, which comprise barley chromosome 1H. Due to the difference in DNA content, 1H chromosome can be discriminated from other chromosomes. In total, 140 000 chromosomes were sorted for one gene capture experiment.



Electrophoretic separation of: (A) reamplified copy 1 of captured DNA (B) original captured DNA of barley chromosome 1H (C) probe labeled with Texas Red.

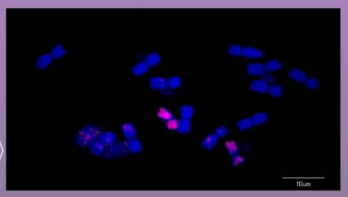


LM-PCR of captured DNA eluted from the microarray (a) and fragment size selection for library construction and chromosome painting (b).



The highly specific probe consists of coding regions and regularly provides weak or no signal in centromeric and telomeric region, which are supposed to be highly repetitive.

Chromosome painting with 1H whole chromosome probe on complete metaphase spreads of barley.



Conclusion

Here we present for the first time a protocol for chromosome painting in plants using composite chromosome probes. This advance opens avenues for the study of behavior of particular chromosomes during mitotic cell cycle, meiosis, and their organization in interphase nuclei. The ability to paint particular chromosomes provides an opportunity to study structural chromosome changes that accompanied the evolution and speciation. As chromosome sorting using flow cytometry has been described in more than 30 plant species, our new approach is not limited to barley, which was used as a model in the present study.

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01 and Internal Grant Agency of Palacky University (Grant award No. IGA PrF/2012/001).