Palacký University in Olomouc

Faculty of Science

Department of Cellular Biology and Genetics

and

Institute of Experimental Botany

Centre of Region Haná for Biotechnological and Agricultural Research

Petr Cápal

Analysis of plant genomes using flow-sorted chromosomes

Ph.D. thesis

Supervisor: Prof. Ing. Jaroslav Doležel, DrSc.

Olomouc 2016

Acknowledgment

I would like to thank my supervisor Prof. Ing. Jaroslav Doležel, DrSc. for his guidance and thoughtful insights during my whole Ph.D. study. Also the sole opportunity of working in his lab was a great life experience. My thanks go to all my colleagues at the Institute of Experimental Botany as well, for their help and friendship. It would like to thank my family for their inexhaustible love and support.

Declaration

I hereby declare that I have written this Ph.D. thesis independently, under the guidance of my supervisor Prof. Ing. Jaroslav Doležel, DrSc. and using only the information sources listed in the list of references.

……..…….

This work was supported by research grant awards from the National Program o Sustainability (award no. LO 1204) and the Czech Science Foundation (award no. P501-12-22 G090).

Bibliographic Identification

Author's name: Mgr. Petr Cápal

Title: Analysis of plant genomes using flow-sorted chromosomes

Type of thesis: Ph.D. thesis

Department: Cellular Biology and Genetics

Supervisor: Prof. Ing. Jaroslav Doležel, DrSc.

Year of presentation: 2016

Abstract:

Plant genomes exhibit an impressive range of sizes and very often the species that are in focus of plant scientists exhibit extraordinary large genomes. The next generation sequencing technology enabled a large number of plant genomes to be sequenced, generating vast amount of data. However the short-read sequences thus obtained are generally difficult to assemble into a faithful approximation of the original genome due to the nature of plant genetic information. The present thesis in the first chapter describes how the plants obtained such difficult-to-analyze genomes and describes the three main drivers behind this phenomenon – the series of whole genome duplication events that occurred during their evolution, abundance of repetitive elements either scattered or clustered in their genomes and finally the presence of similar but not identical subgenomes in case of polyploid plants. The analysis of complex polyploid and highly repetitive plant genomes can be significantly simplified by dissecting them into their natural subunits – chromosomes – by flow cytometry. The analysis of plants by flow cytometry and uses of flow-sorted chromosomes are the topics of the second chapter. The sorting of chromosomes in plants has its limitations due to their similar size and DNA content. To overcome this limitation a method for obtaining DNA from single copies of chromosome was developed. The third and final part of this thesis provides a review of single cell and single chromosome methods and their possible outcomes for analyzing plant

genomes and broadening the possibilities of application of sorted chromosomes in plant genomics.

Keywords:

Plant genomics, flow cytometry, chromosome sorting, single-cell / chromosome genomics

Number of pages/appendices: 73/V

Language: English

Bibliografická identifikace

Jméno: Mgr. Petr Cápal

Název práce: Analýza genomu rostlin pomocí tříděných chromozómů

Typ práce: Disertační

Katedra: Buněčné biologie a genetiky

Školitel: Prof. Ing. Jaroslav Doležel, DrSc.

Rok obhajoby: 2016

Abstrakt:

Pro genomy rostlin je mimo jiné charakteristické velké rozpětí jejich velikostí. Druhy rostlin, které jsou předmětem vědeckého zkoumání, mají často značně velký genom. Pokročilé sekvenační metody umožnily prozkoumání genomů řady zástupců rostlinné říše a vygenerovaly obrovská množství dat. Fragmentovaná sekvenační data je však obecně velmi obtížné sestavit do alespoň přibližného obrazu původního genomu a to díky vlastnostem rostlinných genomů. Předkládaná práce v první kapitole přibližuje, jak rostliny přišly k tak složitým genomům a popisuje tři klíčové fenomény – série celogenomových duplikací, které se odehrály v průběhu evoluce a speciace, nahromadění repetetivních elementů v jejich genomech a v neposlední řadě přítomnost navzájem podobných subgenomů u polyploidních rostlin. Analýza komplexních genomů může být zjednodušena jejich rozdělením na přirozené podjednotky – chromozomy – pomocí průtokové cytometrie. Tyto analýzy a využití tříděných chromozomů v genomice rostlin jsou tématem druhé kapitoly. Třídění chromozomů je u mnoha druhů rostlin komplikováno malými rozdíly v jejich velikosti. S cílem překonat toto omezení byla vyvinuta metoda pro získávání chromozomově specifické DNA z jednoho jediného tříděného chromozomu. Třetí a poslední kapitola této práce shrnuje metody analýzy jedné buňky či jednoho chromozomu a nastiňuje jejich využití v genomice rostlin.

Klíčová slova:

Genomika rostlin, průtoková cytometrie, třídění chromozómů, genomika na úrovni jedné buňky/chromozomu

Počet stran/příloh: 73/V

Jazyk: Anglický

Contents

1 Introduction **11**

2 Plant genomics **12**

2.1 Nuclear genome evolution in flowering plants by whole-genome duplications **12**

2.2 Contribution of transposable elements to genome complexity **16**

2.3 The case of homoeologs **18**

3 Flow cytometry **21**

3.1 Overview and principles of flow cytometry **21**

3.2 Flow cytometry in plants **25**

3.3 Flow cytometric analysis of plant nuclei **26**

3.4 Flow cytometric analysis of chromosomes **28**

3.5 Chromosome sorting and uses of flow sorted chromosomes **29**

3.6 Physical mapping **30**

3.7 Genetic mapping **32**

3.8 Sequencing **34**

3.9 Isolating non-discriminable chromosomes from flow karyotype **37**

3.10 Cytogenetic stocks **38**

3.11 Labeling of chromosomes in suspension **41**

3.12 Single chromosome sorting **43**

4 Single cell / chromosome genomics **45**

4.1 Obtaining and analyzing minute amounts of DNA **45**

4.2 Applications of single-cell genomics **48**

4.3 Emergence of single-cell/chromosome approaches in plants **52**

5 Aims **54**

6 General conclusion **55**

7 Publications **56**

7.1 Multiple displacement amplification of the DNA from single flow–sorted plant chromosome **56**

7.2 The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat **57**

7.3 Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley **58**

7.4 Advances in plant chromosome genomics **59**

7.5 Flow cytometry in plant research. In Applied plant cell biology **60**

8 References **61**

9 Abbreviations **73**

10 Supplements **74**

1 Introduction

One copy of a complete genetic information of an organism is called the genome. Many important crops and other plant species beneficial to humankind harbor very large genomes. Moreover, plant genomes are rich in repetitive DNA elements and very often are polyploid and contain several subgenomes in their nuclei. However, working with plants with large genomes in the era of genomics is a disadvantage, because sequencing a huge and complex genome is costly and subsequent assembly is hampered by its size and presence of repeats. In case of allopolyploids, the complexity is further augmented due to the occurrence of homoeologous subgenomes. Luckily enough, the nuclear genomes are divided into their natural subunits – chromosomes. When released into aqueous suspension, mitotic chromosomes can be analyzed by flow cytometry. If they can be discriminated one from another, they can be even sorted in high numbers. This perk can be readily utilized to achieve lossless reduction of genome complexity by dealing with one chromosome at a time. This thesis briefly elucidates how plants obtained large genomes during their evolutionary history and how their analysis can be simplified using flow cytometry. Principles of flow cytometry and chromosome genomics are explained, including the uses of flow-sorted chromosomes in various applications. Special attention is paid to single chromosome genomics as a tool for the investigation of chromosomes that cannot be conventionally resolved from other chromosomes in a karyotype. Single cell and single chromosome genomics is discussed in the last chapter, however, due to scarce uses in plants, mainly the cases from human biology and microbiology are listed.

2 Plant genomics

The goal of this chapter is to explain how some plant species ended up with large and complex genomes during their evolution, which events and phenomena shaped them and how modern science copes with the challenging task of unraveling what they encode. Several examples of plant sequencing projects will be mentioned as the topic is explained. Emphasis is put on flowering plants (Angiosperms) due to their beneficial role in human history and interest scientist have taken in them, supported by the fact that 82% of all sequenced plants belong to the Angiosperms clade, remaining 18% is largely predominated by Algae with only a minority of sequenced plant species left for Gymnosperms and Bryophytes (http://www.ncbi.nlm.nih.gov/genome/browse).

2.1 Nuclear genome evolution in flowering plants by whole-genome duplications

The origin of flowering plants was described by Charles Darwin as "an abominable mystery". Indeed, it is the example of one of the world's biggest radiations that resulted in more than 250,000 species. Shortly after their origin, flowering plants rose to ecological dominance, currently with more than 350,000 described extant species occupying almost all terrestrial ecosystems (De Bodt et al. 2005). With the advent of sequencing technologies, this mystery has just begun to unravel. The evolution of plants is very tightly linked with the evolution of their genomes. The first phenomenon contributing largely to plant genome evolution mentioned here will be the whole genome duplication.

As the name suggests, genome duplication events double the amount of genetic code thus creating a material on which the evolution can work. The functional divergence of duplicated genes is important for biological evolution and for increasing the organism complexity (Taylor and Raes 2004). Polyploidy can contribute to evolution in other ways as well. By having different number of alleles, polyploids differ in gene expression levels from their diploid progenitors.

In genes with allele-dosage effects, the increased variation could provide selective advantage for polyploid plants over the diploids (Osborn et al. 2003). With the effects on developmental processes, metabolism and breeding systems, polyploidy events often result in speciation (Ramsey and Schemske 1998).

The analysis of the first plant genomes supplied surprising evidence for whole genome duplications in species that have been considered true diploids due to their genetics. Nobody expected that the genome of *Arabidopsis thaliana*, the first sequenced land plant (The *Arabidopsis* Genome Initiative 2000), chosen among other reasons for its small genome as compared to other representatives of plant kingdom, was in fact an ancient polyploid. Several years after the release of the Arabidopsis genome sequence, enough evidence was gathered to proclaim that its genome (or rather the genome of *Arabidopsis* ancestors) was duplicated three times during the past 250 million years (Simillion et al. 2002).

Another piece for the plant evolutionary puzzle was brought by the grapevine sequencing project accomplished by the French-Italian public consortium (2005). Among other important revelations such as the expansion of gene families underlying synthesis of aromatic compounds and their high copy numbers, the analysis revealed ancestral whole genome duplication event, leading to present-day grapevine genome, which evolved from three ancestral genomes. The grapevine may therefore be described as a paleohexaploid organism; however, whether a genome triplication or two separated whole genome duplications occurred is still a question. This finding is based on the presence of paralogous regions in grapevine genome, with almost 95% of gene regions having two different paralogous clusters. Compared to the *Arabidopsis* genome, each component of triplicated regions present in grapevine genome corresponds to four segments in *Arabidopsis*. This shows that the ancestral hexaploidy was present in the common ancestor of grapevine and *Arabidopsis* as well as *Arabidopsis* contrary to the grapevine underwent another two rounds of whole genome duplication after their separation (Blanc et al. 2003). Comparison of grapevine and rice (representing monocots) genome enabled dating the

paleohexaploidization event after the divergence of monocotyledonous from dicotyledonous plants, with one genomic segment in rice having three orthologous matches in grapevine. The evolution of monocotyledonous plants and whole genome duplications specific for this clade are briefly described in the banana genome publication (D'Hondt et al. 2012), again based on paralogous gene clusters and their sharing with (in case of banana) three other gene clusters, suggesting two whole genome duplication events during the evolution of *Musa*. The most recent and arguably the most profound investigation of whole genome duplication's role in flowering plant evolution was brought by sequencing the *Amborella* genome (*Amborella* Genome Project 2013). *A. trichopoda* is believed to be the single living species of the sister lineage to all other extant flowering plants. This perk makes it a unique reference species for inferring the structure and genome content of the most recent common ancestor of all living angiosperms. An ancient genome duplication predating the diversification of angiosperms was found, while no evidence for subsequent genome duplication in *Amborella* lineage was found since it diverged from the rest of angiosperms. An overview of the whole genome multiplications across the plant kingdom is shown in figure 1. The goal of the schematics is to illustrate the complexity of plant genomes caused solely by whole genome duplications, while the next part of this chapter will increase the complexity by adding repeat elements and dynamics of their expansion into the equation.

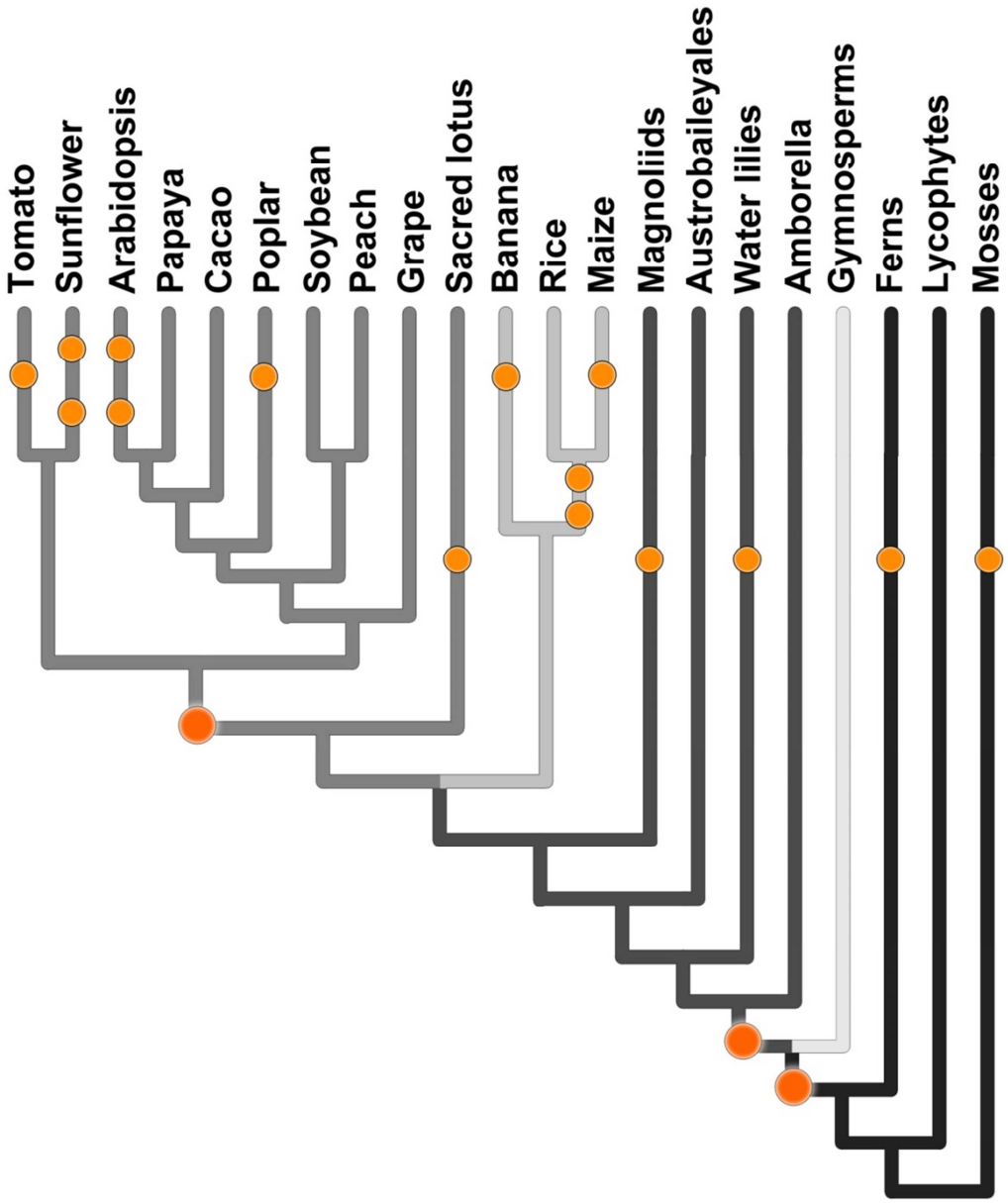

Figure 1: Overview of the land plant phylogeny. Special attention is paid to the whole genome duplication events, which are marked by the orange dots. Three major WGDs are marked by larger bright orange dots. Going forward in the evolutionary history, they predate the appearance of seed plants, divergence of Angiosperms and Gymnosperms, and divergence of the Rosid clade. The lineage specific WGDs are marked by the smaller dots. Modified from *Amborella* Genome project publication (2013).

2.2 Contribution of transposable elements to genome complexity

Transposable elements (TEs) may have major impact on genome structure and functions of genes. In every plant species, transposable elements are the prevalent type of non-genic DNA. In plants with large genomes, transposable elements represent a majority of the genome, in extreme cases they build up for more than 80% of its size (IWGSC 2014, Martis et al. 2013). Table 1 gives an overview of the proportion of TEs in genomes of sequenced plants to start this chapter with.

Table 1: Proportion of transposable elements in plant genomes

Species	Common name	Genome size [Mbp]	Proportion of transposable elements [%]
Arabidopsis thaliana	Thale cress	120	14
Fragaria vesca	Wild strawberry	240	23
Cucumis sativus	Cucumber	243	24
Carica papaya	Papaya	372	52
Medicago truncatula	Clover	375	10
Oryza sativa	Rice	420	35
Theobroma cacao	Cocoa tree	430	26
Populus trichocarpa	Poplar	485	42
Vitis vinifera	Grapevine	487	17
Brassica oleracea	Wild cabbage	600	20
Sorghum bicolor	Great millet	740	62
Gossypium raimondii	Cotton	880	54
Glycine max	Soya	1,100	59
Zea mays	Maize	2,045	76
Hordeum vulgare	Barley	5,439	80
Triticum aestivum	Bread wheat	16,979	80
Pinus taeda	Loblolly pine	21,516	80

Table modified from Kejnovský et al. 2012

All transposable elements have two properties in common. The first is the ability to move from one genomic locus to another and the second is the ability to amplify their copy number during the transposition. There are two main classes of transposable elements – the DNA transposons and the retroelements. First to

be identified in plants were the DNA transposons, namely the Ac/Ds and mutator transposons in maize (Flavell et al. 1994). Majority of these elements exhibit sizes up to 10 kbp. A somewhat smaller in size, however very abundant family of DNA mobile elements are miniature inverted-repeat transposable elements, abbreviated MITEs (Bureau and Wessler 1994). DNA transposons are characterized by the presence of terminal inverted repeats (TIRs), which also serve scientist to distinguish between the diverse families of mobile elements (Bennetzen and Wang 2014). In a successful transposition event, the enzyme called transposase recognizes the specific TIR, excises the transposable element and facilitates its insertion into another place in the genome. The gap after the excision is then repaired by ligation (leading to excision of the element) or by recombinational gene conversion, effectively leading to doubling the current number of the involved transposon (Bennetzen and Wang 2014).

The largest mass of transposable elements in plants is formed by the second group, retroelements. The retroelements transpose themselves *via* an RNA intermediate, avoiding the excision step and creating a copy that is inserted into the genome instead. Out of five known groups of retroelements, four were found in plants (Bennetzen 1993), while the most numerous class of retroelements in plants is long terminal repeats (LTRs) retrotransposons (Grandbastien 1992). Their size could reach over 10 kbp per unit and constitute a majority of large plant genomes, especially two predominant subclasses called *Gypsy* and *Copia* (SanMiguel and Bennetzen 1998). For example 75% of maize genome is constituted solely by LTR retrotransposons (Schnable et al. 2009).

The amount and rate of insertion and deletion of TE plays a major role in plant genome size evolution. Amplification of transposable elements may lead to extremely high copy number in the plant genomes. These events could happen in relatively short evolutionary time span. A short list of examples demonstrating how quickly can genomes grow by the means of TE amplification follows: Half of the present maize genome has been generated over the last six million years by bursts of repeated retrotransposon amplification (SanMiguel et al. 1998). Three-

fold increase in genome size of diploid cotton in the last 10 million years happened because of the amplification of LTR retrotransposons in its genome (Hawkins et al. 2006). Rice (*Oryza sativa*) genome size is half of its wild Australian sister *O. australiaensis* as a consequence of LTR retrotransposon amplification in the latter specie (Piegu et al. 2006). Similar situation is found in well-studied *Arabidopsis* species, where *A. lyrata* harbors almost three times as much TEs as *A. thaliana*, making its genome considerably larger (Hollister et al. 2011). The genome of *Brassica oleracea* has been hugely expanded by the amplification of DNA transposon families (Zhang and Wessler 2004), to list at least one example where the LTR retrotransposons are not to blame.

Even from the simple perspective of their numbers, the transposable elements should play important role in determining the structure of plant genomes just by the fact that they constitute large parts of them. Almost of the same significance is the fact that every aspect of their life strategy has a potential for changing the genome structure and alter gene expression. Mobile elements have the ability to remodel genomes and alter gene expression via several mechanisms, resulting in chromosomal rearrangements, including duplications and deletions, regulation of expression of neighboring genes, insertional mutation, and many others (Kejnovsky et al. 2012).

2.3 The case of homoeologs

To add yet another level of complexity to plant genomes, many plants harbor several subgenomes in their nuclei. This topic was touched at the beginning of this chapter concerning whole genome duplications, here however, will be dealt with specific type of WGD by interspecific or even intergeneric hybridization followed by genome doubling, generally referred to as allopolyploidization, because it is this process that results in the presence of homoeologous chromosomes in one nucleus. Homoeology is clearly and interesting phenomenon and is receiving more and more attention as the number of sequenced polyploid plant genomes is increasing. To name just a few

outstanding examples, both bread and pasta wheat, tobacco, cotton, rapeseed, coffee or oat, all of which are allopolyploids (Soltis and Soltis 2009).

Over the time, the definition of homoeology has been evolving. The first use of this term comes from cytogenetic study of allopolyploid wheat, where the homoeologous chromosomes are described as phylogenetically similar, but not strictly homologous (Huskins, 1931). Later they were defined as chromosomes homologous in limited parts of their length (Knight, 1949). Due to the biological alternations between genes that arise by mere duplication and those arising due to speciation (Tatusov et al. 1997), the term homoeology needs to be stated precisely. So far the most recent, evolutionary precise and unifying definition of homoeologs is that it represents pairs of genes or chromosomes in the same species that originated by speciation and were brought back together into the same genome by allopolyploidization (Glower et al. 2016).

Because of the numerous important crops being allopolyploid, there is a need for reliable methods of the homoeologs inference. As for the whole genome approach, several techniques already exist, both wet lab-based and with the advent of sequencing also the computational ones.

A possibility of separating homoeologous sequences from the mixture of molecules is by PCR amplification of the desired locus and then subcloning the resulting mixture of products into the bacteria. As each bacterium is transformed with only a single molecule, the bacterial colonies can be isolated and sequenced. By assigning these individual sequences to diploid progenitors, the homoeologs can be resolved (Small et al. 1999). Second possibility is to separate restriction enzyme-digested DNA on agarose or polyacrylamide gel, where the differences between homoeologs could alter the restriction sites and thus the size of the resulting fragments (Cronn and Wendell, 1998). This procedure can be followed by cutting the bands from gel and cloning the DNA into bacteria with a subsequent analysis similar to that described above. Larger-scale method for homoeologous sequence separation is based on hybridization of genomic DNA to an array, capturing hundreds to thousands of genes followed by NGS

sequencing to distinguish sequence polymorphism in homoeologs (Salmon et al. 2012).

In theory, high-throughput sequencing should provide a tool for inferring homoeology computationally at genome-wide level. However, with the level of complexity of polyploid plant genomes and the state of their assemblies (Schatz et al. 2012), this is not entirely possible. Moreover, one usually needs sequencing information (e.g. exome sequence, transcriptome sequence, etc.) of diploid progenitors to faithfully discriminate the homoeologous reads as exemplified in allotetraploid cotton (Udall et al. 2006). In alloxexaploid wheat, transcriptome reads were first assembled into clusters containing homoeologous sequences and in the second step, these clusters were reassembled at higher stringency to separate homoeologs (Schreiber et al. 2012). Again, the separated reads then had to be mapped onto the diploid progenitor genomes, *Triticum urartu* and *Aegilops tauschii* in this particular case, to reveal to which subgenome they belong (Li et al. 2014).

All of the above-mentioned approaches rely on the use of whole genomic DNA. In the next chapter, it will be described, how flow cytometry can be utilized to deal with the presence of homoeologous chromosomes in the genome and how it can greatly decrease the complexity of subsequent analyses, even for plants with "standard" diploid genetics by means of chromosome analysis and sorting. But first an introduction to flow cytometry might be useful.

3 Flow cytometry

Flow cytometry proved to be immensely useful technique with countless applications in biomedical research and practice, including immunology, hematology, pathology and oncology, and many other fields. It has also found its way into plant sciences where it has been serving scientists for almost forty years. Numerous areas of plant research made fast progress thanks to utilizing flow cytometry. In this chapter, a brief overview of the principles of flow cytometry is provided and utilization in plant sciences outlined. The main emphasis is put on plant chromosome sorting and their successful uses in plant genomics, where it helps to resolve genome complexity described in the previous chapter.

3.1 Overview and principles of flow cytometry

Flow cytometry walked a long way since its inception, starting as device that detected bacteria in aerosols in 1930s, exchanging flowing air for stream of fluid in late 1940s to utilization of fluorescence measurement in 1960s (exhausting overview of history is presented in Shapiro's Practical flow cytometry, 4th edition, 2003). Flow cytometry played major roles in a broad spectrum of life sciences. Without debate, the most profound impact was in biomedical research. Plant scientists weren't left behind, when they realized the potential uses in their research. The main reasons to incorporate flow cytometry in research pipelines are high throughput (thousands of particles per second can be analyzed), high resolution and accuracy, small sample volume/weight requirements and low operating costs (after the initial investment to purchase an instrument). Although flow cytometers in general do not provide visual information of the analyzed particles (imaging flow cytometers being the exception and exceeding the frame of this chapter, for review see Barteneva et al. 2012), it is more than balanced with the speed of analysis together with the ability to measure several parameters simultaneously. With ever expanding choices in fluorescent dyes, fluorescently tagged antibodies and other compounds used to highlight optical differences of

the analyzed particles, it is possible to measure more than 18 parameters simultaneously (Chattopadhyay et al. 2008).

Flow cytometry is both a preparative and analytical tool that provides high precision and high throughput. Its main advantage over other methods is the ability to analyze hundreds to thousands of events per seconds, each with multiple optical parameters at once, creating statistically relevant set of data. The basic principle of flow cytometer is shown in Figure 2. Individual particles in the sample are forced by sheath fluid to pass in the innards of the cytometer in a single file, one particle after another. They pass individually through a flow chamber, where they interact with a beam of light (usually several laser beams in a row with precisely defined delay times). The interactions between light and particles generate optical signals – light scatter and fluorescence pulses – that are collected by detectors. Each detector collects a defined part of optical spectrum and converts pulses of light into electrical pulses, which are then digitally processed. The optical properties of each analyzed particle converted into digital signal are then displayed in single-parametric histograms or multiparametric dot-plots.

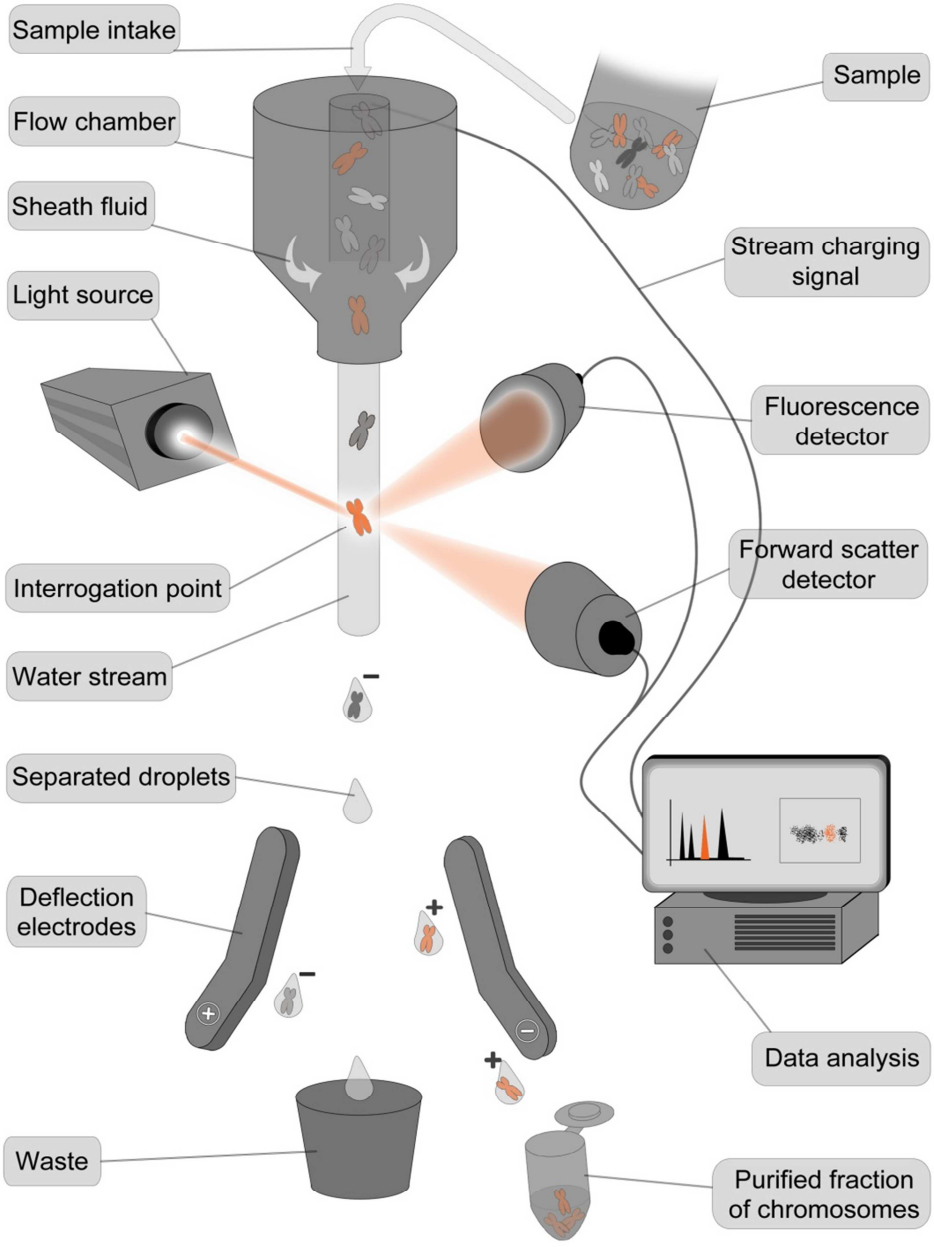

Figure 2: Schematics of a flow cytometer and sorter. Microscopic particles (stained chromosomes in this case) in liquid suspension are forced by the shape of the flow chamber and by a continuous stream of sheath fluid to pass the innards of the cytometer one-by-one in a single file. This ensures that individual chromosomes interact with laser beam at the interrogation point. Resulting pulses of scattered light and pulses of emitted fluorescence are collected by photodetectors and converted into digital signals. If the chromosome of interest

differs in optical properties from the rest, it can be identified and sorted, by breaking the liquid stream at precisely defined time point into individual droplets and charging only the droplets carrying the desired chromosome. The charged droplets are deflected as they pass through electrostatic field between two electrodes. Purified fractions of chromosomes are collected in appropriate vessel (exemplified here by a PCR tube).

Flow sorters are a more advanced category of flow cytometers. These instruments work as flow cytometers in terms of analysis and, as an important addendum, they are capable of physically separating analyzed particles. Figure 2 shows a principle of a droplet sorter, the most commonly used flow sorting principle. The fluid stream carrying the analyzed particles is broken into individual drops by a vibration of piezoelectric crystal when it leaves the flow chamber. When the particle of interest is detected, the instrument waits until the particle reaches the break-off point, the liquid stream is then charged and thus the separating droplet carrying desired particle is charged as well. When it passes downwards through the electric field between two deflection electrodes, it is pulled towards one of them and lands in a selected vessel (e.g., a collection tube, microscopic slide, beaker, PCR plate, and other).

Although the flow cytometry was initially developed to analyze intact cells, scientists quickly realized its potential for the analysis of subcellular organelles, such as nuclei, mitochondria and plastids, with the chromosome analysis soon to follow (Doležel et al. 1989, Böck et al. 1997, de Laat and Blaas 1983). In fact, particles as small as bacteria can be efficiently analyzed and sorted (Amann et al. 1990). It is stunning how many properties and attributes of analyzed particles can be inferred based just on a few optical parameters like scattered light, fluorescence and respective pulse profiles. To name just a few: particle size, internal structure (granularity), to some extent particle shape, DNA, RNA and protein content, pH and viability (Shapiro 2003). Coupled with fluorescent labeling, the number of possibilities rises dramatically and new applications continue to be explored. A broad range of flow cytometers are on the market, from simple table-top analyzers with one laser capable of measuring just a few

parameters to high-end instruments equipped with three lasers designed for multiparametric analyses and high speed sorting (Figure 3).

Figure 3: Commercially available flow cytometers. From left to right are displayed small table top analyzer, high-throughput analyzer with automated sample loader used in clinics and advanced high-speed flow sorter. Pictures taken from official websites of the manufacturers.

3.2 Flow cytometry in plants

Samples for flow cytometry have to be in a form of aqueous suspension of single particles. This puts a little bit of an obstacle in the way for a plant scientist. Plant cells are bound together in a complex matrix forming various tissues and organs, which hampers isolating single cells suitable for flow cytometric analysis. Moreover, plant cells have a rigid cell wall, which gives them irregular shape that disturbs the laminar flow. Together with autofluorescence, these properties effectively compromise flow analysis. If that's not enough, some plant cell types are larger than orifices of standard flow chambers, which are typically 50-130 μm. Two ways around this constrain were designed. The first was the analysis of plant protoplasts. A protoplast is a plant cell devoid of cell wall, which can be removed using hydrolytic enzymes. This forces the plant cell surrounded only by plasma membrane to acquire a regular shape suitable for flow cytometry. However, protoplast preparation is time consuming and not every plant species and tissue is suitable for their preparation. Although several investigations have been conducted using protoplasts in flow cytometry (reviewed in Galbraith et al. 2004) and the resolution of measurement of genome size using nuclei released

from protoplasts remains unparalleled (Galbraith et al. 1994), the fragility of protoplasts and the fact that removal of cell wall represents a considerable stress factor (Papadakis and Roubelakis-Angelakis 2002) precludes them for being a substitute for intact cells in most cases. The second way which arose more like a necessity from a fact that intact cells and protoplasts cannot be widely used, is the analysis of subcellular organelles – plastids, mitochondria, nuclei and ultimately chromosomes. Although the title of the present thesis aims directly at chromosome analysis, presentation of the topics would not be complete without mentioning the analysis of cell nuclei.

3.3 Flow cytometric analysis of plant nuclei

Nucleus is one of the most prominent organelles inside plant cells. Almost the whole plant genome is stored there and important processes occur inside, including DNA replication, transcription, repair, as well as response to various stimuli through binding of transcription factors. The main applications of flow cytometry in plant research are the measurement of absolute DNA content, e.g. genome size, and estimation of relative DNA content, inferring ploidy level of investigated plant (Vrána et al. 2014). Other applications include cell-specific gene expression and cell cycle analysis. The most used optical parameter in nuclei analysis is fluorescence. Fluorochromes used to stain nuclei can be chosen from a rich selection and, depending on the aim of the analysis, they include nucleic acid dyes (DAPI, propidium iodide or acridin orange), fluorescent tags (FITC or PE) and fluorescent proteins used to evaluate gene expression (GFP and its derivatives), with nucleic acid dyes being the most used ones (Petit et al. 1993).

Nuclear genome size is a unique attribute of every organism that can be utilized in a vast array of research areas, including taxonomy, evolutionary studies, population biology and ecology. Moreover, the data on genome size are essential for large sequencing programs to estimate their cost and decide proper management (Cardoso et al. 2012), especially in plants this point is double valid.

Despite this fact, recent counting of angiosperm plant species with estimated DNA content reached the number 6287 (Bennet and Leitch 2011), which corresponds roughly to only 2% of known angiosperms. Since 1950, the nuclear genome size is stated in C-values (Swift 1950), where 1C is equal to the amount of DNA contained within a non-replicated haploid chromosome se (n). The first published compilation on C-value data in plants was released in 1976 (Bennet and Smith 1976), it is being constantly updated by the same group, and since 1997 it is publicly available online (http://data.kew.org/cvalues/). The values are stated in picograms of DNA and to name just the extremes, the smallest C-value in the database is currently held by *Genlisea margaretae* with 1C=0.065 (Greilhuber et al. 2006) while the opposite end is occupied by *Paris japonica* with 1C value of 152.2 pg (Pellicer et al. 2010).

The most common use of flow cytometric analysis in plants to date has not been the estimation of absolute DNA content, but the estimation of relative DNA content for the purpose of estimating ploidy levels. The knowledge on the total number of chromosome sets present in a genome is important for taxonomy as well as for breeding programs (reviewed in Doležel et al. 1997). Flow cytometry is capable of distinguishing between individuals with different number of complete chromosome sets, but also minor changes can be detected, like those caused by gain or loss of chromosome in cases of aneuploidy. For example, a single telocentric chromosome addition was detected by flow cytometric analysis of wheat-rye addition line nuclei (Pfosser et al. 1995), aneuploidy in banana (Roux et al. 2003) or cases of trisomy and monosomy in hops (Šesek et al. 2000). Special application of the estimation of relative DNA amounts is the detection of endopolyploidy, where cells with nuclear DNA content 4C, 8C, 16C, etc., are present within one organism, which is quite common in angiosperms (Barrow 2006).

3.4 Flow cytometric analysis of chromosomes

Most of plant cells are in the interphase during the life span of a particular plant, chromosomes in their nuclei are decondensed and cannot be physically separated one from each other. This becomes possible as the cells enter the metaphase stage of cell division and the chromosomes condense. Cell cycle in tissues from which the samples will be prepared thus must be synchronized and mitosis subsequently blocked in order to achieve high number of cells in metaphase. In plants, reaching a synchrony of 50% of diving cells is considered a success (Vrána et al. 2000). Currently the predominant method is the preparation of chromosome samples from root tip meristems of young seedlings (Doležel et al. 1992); as an alternative the hairy root cultures can be used (Neumann et al. 1998). In order to be analyzed by flow cytometry, the chromosomes have to be released, which is achieved by mechanical homogenization of formaldehyde-fixed root tissues. Upon the release of chromosomes into a suitable isolation buffer (Doležel et al. 1992), they're stained by a DNA-specific fluorochrome (DAPI has been the most commonly used fluorochrome) that allows chromosomes to be classified according to their relative DNA content. This analysis results in representative distribution of relative DNA content, termed flow karyotype (as shown in Figure 4a).

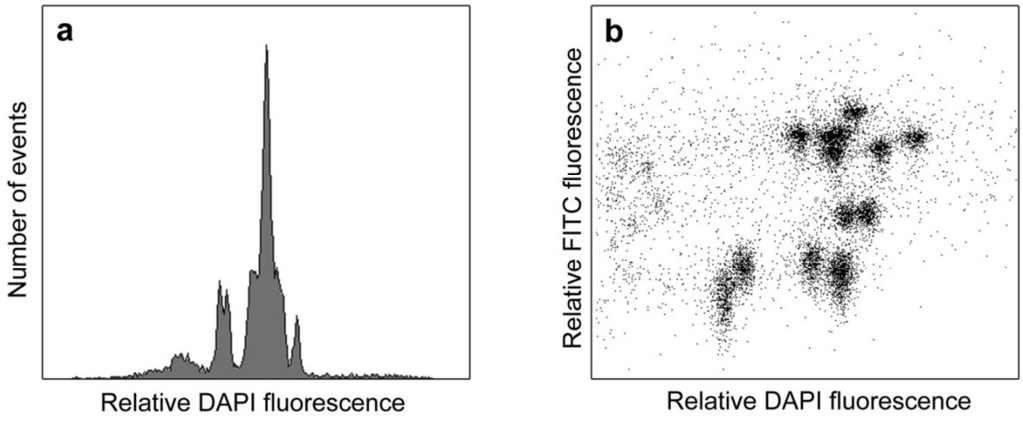

Figure 4: Output of the flow cytometric analysis of GAA-labeled plant chromosomes of durum wheat (*Triticum durum*). **a** Monoparametric analysis resulting in histogram of relative DAPI fluorescence (flow karyotype). **b** Biparametric analysis, where the DNA content represented by DAPI fluorescence

is plotted against the abundance of GAA-repeat in chromosomes, represented by FITC fluorescence (labeling of chromosomes is described later in section 3.11)

Flow karyotyping is a high-throughput quantitative approach for characterization of a karyotype and for detection of numerical and structural chromosome changes. With tens of thousands analyzed chromosomes per screening this approach provides statistical accuracy. In plants it has been used to estimate frequency of alien chromosomes in wheat-rye addition lines (Kubaláková et al. 2003) and for detection of trisomy of chromosome 6 in barley (Lee et al. 2000). Alien chromosomes were identified in oat-maize chromosome addition lines (Li et al. 2001) and alien chromosome arms were detected in wheat-rye telosome addition line (Suchánková et al. 2006). Translocation chromosomes were detected in garden pea and field bean (Neumann et al. 1998), however reciprocal translocations and interchromosomal rearrangements where identical amounts of DNA are exchanged cannot be identified by flow cytometry. Nonetheless, the method proved to be sensitive enough to detect chromosome deletions (Gill et al. 1999). Once flow karyotyped, the stage is set for chromosome sorting.

3.5 Chromosome sorting and uses of flow sorted chromosomes

A scheme of the most commonly used flow sorter is shown above (Fig. 2). The usefulness of flow-sorted chromosome fractions is determined by two factors – the purity in the sorted fraction and the quality of DNA. The purity depends on the degree to which the investigated chromosomes can be separated from other chromosomes in a karyotype and the debris (comprising chromatids, chromosomal fragments and chromosome clumps). The extent of contamination can be estimated microscopically after sorting chromosomes onto microscopic slide. Labelling sorted chromosomes by FISH using several probes provides the opportunity to identify sorted chromosomes (Kubaláková et al. 2003). The ability

to characterize chromosomes in sorted fractions is also important to assign individual chromosomes to peaks on a flow karyotype. Under favorable conditions, purities exceeding 95% can be achieved (Doležel et al. 2014). The morphology and DNA a proteins of sorted chromosomes remain intact after flow sorting (Doležel et al. 1992), rendering them suitable for a wide range of applications, including isolation of high molecular weight DNA for preparation of BAC libraries, long-read sequencing technologies and optical mapping, physical and genetic mapping, whole genome sequencing, as well as immunocytochemistry and proteomic analyses. Possible uses of flow sorted chromosomes are listed below, with respect to the outcomes of this thesis the emphasis is put on marker development and sequencing.

3.6 Physical mapping

Only one hundred to one thousand flow-sorted chromosomes are enough as a template for PCR reaction. These quantities can be sorted in the order of minutes, thus providing plausible way for assigning specific sequences to corresponding chromosomes. This approach has been adopted for mapping vicillin genes in field bean (Macas et al. 1993) and for mapping genes to sex chromosomes in *Silene latifolia* (Kejnovský et al. 2001). Genetic and physical maps were successfully integrated *via* PCR on sorted chromosomes in *Pisum sativum* and *Cicer arietinum* (Neumann et al. 2002, Vláčilová et al. 2002). Mapping to sub-chromosomal region is also possible, given that appropriate cytogenetic stock is used, for example mapping sequences to translocated regions between field pea and garden pea (Macas et al. 1993). PCR on flow-sorted chromosomes can serve also as a support for positional cloning. It is tempting especially in allopolyploid species, where flow-sorted chromosomes help to resolve position of investigated sequences between homoeologous chromosomes (Šimková et al. 2011).

If higher quantities of chromosomal DNA are needed, long sorting times can be significantly shortened by polymerase Phi29-mediated multiple displacement amplification (Dean et al. 2002). Microgram quantities of chromosome-specific

DNA with fragments exceeding 10kbp can be obtained from the equivalent of 10 ng DNA, which corresponds to circa 20,000 sorted chromosomes prior their purification (exact quantity depends on chromosome molecular size). DNA generated this way proved to be representative and 98.1% of SNPs present on high-density SNP map of barley chromosome 1H could be successfully detected (Šimková et al. 2008). To go down the road with physical mapping in barley, all the arms of the remaining barley chromosomes (2H-7H) were flow sorted from wheat-barley addition lines and amplified in the same manner. The resulting amplicons were used to assign 370 SNP loci from barley genetic map to individual arms (Muñoz-Amatriaín et al. 2011). SSR markers used to describe wheat-barley translocation event were prior to this evaluation assigned to both arms of chromosome 7H (Cseh et al. 2011). In wheat, the chromosome-based mapping was used for positional cloning of powdery mildew resistance gene (Jakobson et al. 2012)

Another category of physical mapping where flow-sorted chromosomes can be used is cytogenetic mapping. It is particularly useful for ordering and orienting of BAC clones and assembling sequence contigs during physical map construction and sequence whole genome assembly. FISH with cDNA probes applied on mitotic chromosomes is able to achieve just that (Karafiátová et al. 2013). The spatial resolution of this method can be further improved, if the flow-sorted chromosomes are stretched up to 100-fold (Valárik et al. 2004).

The etalon of physical mapping is the construction of physical map of an entire chromosome (or chromosomal arm). A physical map is a prerequisite for a successful clone-by-clone sequencing project. The construction of chromosome arm-specific large insert size libraries requires corresponding quantity and quality of DNA. For this purpose, as special protocol for isolating high molecular weight DNA from flow sorted chromosomes was developed (Šimková et al. 2003). Shortly after, the first chromosome-specific BAC library was constructed (Šafář et al. 2004). This BAC library consisted of 68,000 clones with mean insert size of 103 kb, standing for 6x coverage of the 3B chromosome. Protocols were

further improved and chromosome coverages exceeding 15x and insert sizes longer than 120 kbp were achieved (Šafář et al. 2010). As the International Wheat Genome Sequencing Consortium adopted the chromosome-based strategy for constructing a reference sequence of bread wheat (Feuillet and Eversole 2007), BAC libraries were constructed for all twenty remaining chromosome arms of bread wheat cultivar Chinese Spring. Chromosome 3B was also the first one for which the BAC library was fingerprinted and physical contig map was constructed (Paux et al. 2008). With 82% estimated coverage of the chromosome, the map was anchored with almost 1500 molecular markers. This result suggested that fingerprinting of the remaining BAC libraries were soon to follow (Lucas et al. 2012, Phillipe et al. 2013). Chromosome-specific BAC libraries were also constructed for cultivars of wheat other than Chinese Spring, in order to perform positional cloning of agronomically important genes that were not present in this cultivar (Janda et al. 2006). Chromosome-specific BAC libraries found its relevance particularly in polyploid species where dissecting genome to individual chromosomes overcomes the problems connected with homoeologous chromosomes. As a welcomed addendum, the exhaustive analyses can be divided into several labs at the same time, when the genome is processed chromosome by chromosome in form of BAC libraries.

3.7 Genetic mapping

Flow-sorted chromosomes proved to be a boon also in development of markers for genetic mapping. Abundance of markers is essential for the construction of saturated linkage maps. A high quality genetic map in turn accelerates the assembly of shotgun sequences and positional cloning. Out of the plethora of markers used for genetic mapping in plants, the most popular have been SSR, ISBP, SNP and DArT markers (reviewed in Poczai et al. 2013). Main advantage of chromosome-based strategy in genetic marker development is the marker specifics, which saves time and money. Chromosome-specific genetic markers can be developed both from BAC libraries and shotgun sequences.

SSR markers were developed from libraries of chromosome 1 of *Vicia faba* (Požárková et al. 2002), that were in turn used for construction of a genetic map (Román et al. 2004). Another successful example is the generation of a set of SSR markers from rye chromosome arm 1RS, both from shotgun sequences of amplified DNA of 1RS and from 1RS BAC library (Kofler et al. 2008). BAC-end sequences used in this work contained many junctions of two retroelements, which facilitated development of 1RS-specific ISBP markers (Bartoš et al. 2008). The same approach yielding both SSR and ISBP markers was used in wheat chromosome arm 1AL (Lucas et al. 2012).

DArT markers proved to be especially useful in saturating genetic maps (Wenzl et al. 2006). In wheat, the DArT array was developed for chromosome 3B and chromosome arm 1BS (Wenzl et al. 2010), with the majority of derived markers specific to the investigated chromosome/arm. The 510 3B-specific polymorphic markers developed in this study were used to generate genetic map of the 3B chromosome, doubling the number of loci in the map available. The same study demonstrated the efficiency of using DNA from flow sorted chromosomes – 510 polymorphic markers were produced after screening only 2,688 3B-derived clones (19%) while only 269 polymorphic markers were produced after screening 70,000 clones derived from whole genomic sequence (0.4 %).

With the ever-dropping prices of next generation sequencing, *in silico* identification of genetic markers became a plausible way of their development. Coupling chromosome flow sorting and whole genome amplification is the basis of this approach. Shatalina et al. (2013) used 3B-specific shotgun sequences to develop SNP map of the 3B chromosome. These SNPs were then used to genotype recombinant inbred line population between wheat cultivars Arina and Forno. More than 900,000 SNP loci arose from a comparative study of sequences of wheat homoeologous chromosomes 7A, 7B and 7D (Berkman et al. 2013). Thousands of potential SSR loci were identified in shotgun sequences of wheat chromosome arm 7DL (Nie et al. 2012) and rye chromosome arm 1RS (Fluch et al. 2012), both of which being used to genotype mapping populations. 19,000

potential ISBP markers were harvested from 1BL-specific shotgun sequences (Phillipe et al. 2013).

3.8 Sequencing

Modern next-generation sequencers are capable of grasping the whole genomes of plants and generate vast amounts of data to reach genome coverages desired for various analyses. The catch comes later when the short reads generated by these instruments are used to assemble a whole genome sequence. Even clone-by clone sequencing strategy could be painful in complex genomes just because of the sheer size of genomic BAC library. A feasible solution again lies in the utilization of flow sorted chromosomes, be it in form of chromosomal BAC libraries or DNA amplified by MDA (multiple displacement amplification). This approach is most useful in polyploid species, where it effectively eliminates issues associated with the presence of homoeologous chromosomes.

The 17 Gbp-sized hexaploid genome of bread wheat (2n=6x=42) can be readily dissected into 40 chromosomal arms and one chromosome by flow sorting using various cytogenetic stocks, including telosome lines of wheat (Šafář et al. 2010), reducing it without losing any information to fractions representing only 1.3% to 3.3% of the genome. Each of those were sequenced by illumina technology and assembled to produce survey sequences for each bread wheat chromosome (IWGSC 2014). The 124,201 gene loci spread almost evenly across all the chromosome arms were annotated. Comparison with its wild ancestors proved sequence conservation of all subgenomes with only a limited gene loss. The positional assignment of more than 75,000 genes is the basis for their future isolating and cloning that in turn has possibility to hasten and improve wheat breeding efficiency.

The survey sequences were produced based on chromosomal DNA amplified by MDA according to Šimková et al. (2008). The next step in sequencing wheat genome is the construction of reference sequences for all of its chromosomes,

which utilizes the clone-by-clone sequencing of MTP (minimum tilling path, smallest number of clones needed to cover the entire chromosome) clones from chromosome-specific BAC libraries. This is a more laborious approach, but the quality of the assembly is much higher as demonstrated by the first sequenced plant genomes of *Arabidopsis*, maize and rice compared to the rest (quality of the assemblies of clone-by-clone vs. whole genome shotgun approach is reviewed in Shangguan et al. 2013). The first reference sequence, that of the largest wheat chromosome - 3B, has already been published (Choulet et al. 2014). It was produced by sequencing 8542 bacterial clones. The resulting pseudomolecule spanned 774 Mbp (over 90% of the estimated size) with 5326 protein-coding genes and 1938 pseudogenes annotated, with 85% of the chromosome consisting of transposable elements. The availability of the sequence anchored to genetic map will help in characterization of genes responsible for agronomically important traits. In addition, both the draft sequences of all chromosomes and the one so far published reference sequence help to elucidate the interplay of three subgenomes present in one organism and its evolutionary history.

BAC-end sequencing (BES) itself proved to be source of interesting scientific outcomes and valuable information can be dug out by extrapolating these findings. A set of 1,536 bacterial clones from BAC library of chromosome 1RS was chosen for BES. With the 0.9% of sequences marked as coding ones, the estimate could be done that 1RS contain about 2000 genes, while the repeat content of the arm was estimated to be 84% (Bartoš et al. 2008). Likewise in wheat arm 1AL, BES was used to estimate the gene number to be roughly 4700, moreover the occurrence of two syntenic blocks was confirmed, while three new blocks of synteny were found (Lucas et al. 2012).

The most affordable and straightforward way of obtaining chromosome-specific sequences is to directly sequence DNA amplified from flow-sorted chromosomes. Mayer et al. (2009) sequenced the chromosome 1H of barley using 454 technology. Not only it allowed identification of some 5400 gene loci, also by comparing the data with published genome sequences of rice and sorghum, the

virtual gene order could be drawn along the chromosome based on collinearity between these related species, thus producing the so-called genome zipper (a linear catalog of genes) for the first barley chromosome. The same was done for the six remaining barley chromosomes just a few years later (Mayer et al. 2011), resulting in sequence-based gene map containing 86% of the estimated barley gene content. Owing to the success in barley and partly due to its relative simplicity and low cost, the genome zipper approach was used subsequently in wheat as well. A virtual gene order was constructed along the chromosome arm 1AL (Lucas et al. 2013), which identified putative translocations and confirmed the presence of genes non-syntenic for group 1 chromosomes of model grasses. Similarly was the genome zipper developed for chromosome 4A (Hernandez et al. 2012), ordering 85% of the genes present on this chromosome. Rye (*Secale cereale*) could not be left behind and its genome zipper for all chromosomes based on 454 data was generated as well (Martis et al. 2013), ordering 22,000 of its genes and elucidating the genome evolution of rye by revealing the translocations and whole genome duplications that played a part in this process.

With the recent global ascendance of illumina sequencing technology, several analyses have been done on sorted chromosomes. Berkman and colleagues analyzed in a stepwise manner the chromosomes of homoeologous group 7 in bread wheat. The first attempt was to assemble chromosome arm 7DS from short illumina reads, reaching approx. 40% coverage of the arm identifying some 1,500 genes (Berkman et al. 2011). Then the 7BS came to focus, comparing it directly to 7DS revealed 84% collinearity, while in comparison with model grass *Brachypodium distachon*, only 60% of the chromosome was collinear (Berkman et al. 2012). Finally, the research was broadened to all three chromosomes 7A, 7B and 7D (Berkman et al. 2013). Assembling genic regions of the homoeologous group revealed differential gene loss in the three chromosomes and brought evidence of shaping the wheat genome as it became dispersed around the world. Finally, the international effort resulted in illumina sequencing of all 40 chromosome arms and one chromosome of bread wheat, which so far lead to

release of reference sequence of chromosome 3B (Choulet et al. 2014) and draft sequences of all of them (IWGSC 2014).

3.9 Isolating non-discriminable chromosomes from flow karyotype

Successful application of flow-sorted chromosomes in genomic analyses described above relies on the ability to discriminate individual chromosome on a flow karyotype. In plants this effort is hampered by the occurrence of similarly sized chromosomes and similar GC/AT content (Doležel et al. 2012), effectively preventing mono- and bi-variate (GC/AT) chromosome sorting. To demonstrate that, Table 2 shows several agronomically important plant species and corresponding number of chromosomes that can be readily sorted from plants with standard karyotype. To overcome this limitation, three different approaches have been developed and will be separately discussed with their pros and cons.

Table 2: A list of agronomically important species with the number of chromosomes that can be sorted from their standard karyotype

Common name	Latin name	Number of chromosomes (haploid set)	Number of discriminated chromosomes*
Barley	*Hordeum vulgare*	7	1
Bread wheat	*Triticum aestivum*	21	1
Chickpea	*Cicer arietinum*	8	5
Durum wheat	*Triticum durum*	14	1
Field bean	*Vicia faba*	6	1
Maize	*Zea mays*	10	2
Meadow fescue	*Festuca pratensis*	7	1
Oat	*Avena sativa*	21	0
Pea	*Pisum sativum*	7	2
Rice	*Oryza sativa*	12	0
Rye	*Secale cereal*	7	1
Tobacco	*Nicotiana plumbagini*	10	0
Tomato	*Lycopersicon esculentum*	12	0

*Corresponding references are reviewed in Doležel et al. (2012)

3.10 Cytogenetic stocks

The utilization of cytogenetic stocks has been the most productive approach so far. For hexaploid wheat, where only one chromosome out of 21 can be discriminated in a standard karyotype (e.g., cv. Chinese Spring), a complete set of ditelosomic lines (lines possessing 20 standard chromosomes and the remaining chromosome in a form of two stably transmitted telocentric chromosomes) has been produced in 1950s by Sears (1954). Five decades later, these lines were found to represent an ideal material for chromosome flow sorting in wheat, therefore they have been utilized by IWGSC to isolate individual arms and to produce chromosome arm-specific libraries for 40 arms and chromosome 3B specific library (IWGSC 2014). Due to its polyploid nature, bread wheat genome

tolerates not only deletions, but also other forms of aneuploidy that have been exploited in flow sorting. For example only one chromosome in rye (1R) can be sorted from a standard karyotype, while the remaining six chromosomes can be sorted from the respective wheat-rye chromosome addition lines (Kubaláková et al. 2003). To cope with similar situation in barley, where only chromosome 1H can be readily sorted, the series of wheat-barley chromosome arm addition lines were prepared and utilized for sorting all chromosome arms of barley (Suchánková et al. 2006). Maize-oat chromosome addition lines were prepared for all ten maize chromosomes (Riera-Lizarazu et al. 1996), which facilitated their sorting and subsequent analyses (Li et al. 2001).

Translocations can be tolerated even by diploid plants and this ability was exploited for example in preparation of pure fractions of chromosomes in pea, where stable reciprocal translocation between two chromosomes enabled discrimination of four chromosomes as compared to two in a standard karyotype (Neumann et al. 1998) or field bean, where three reciprocal translocations facilitated sorting of all six chromosome types (Lucretti et al. 1993). Translocation events also enabled sorting of additional chromosomes in wheat and rye as well (Kubaláková et al. 2003). The convenience of cytogenetic stocks for chromosome sorting is demonstrated in figure 5.

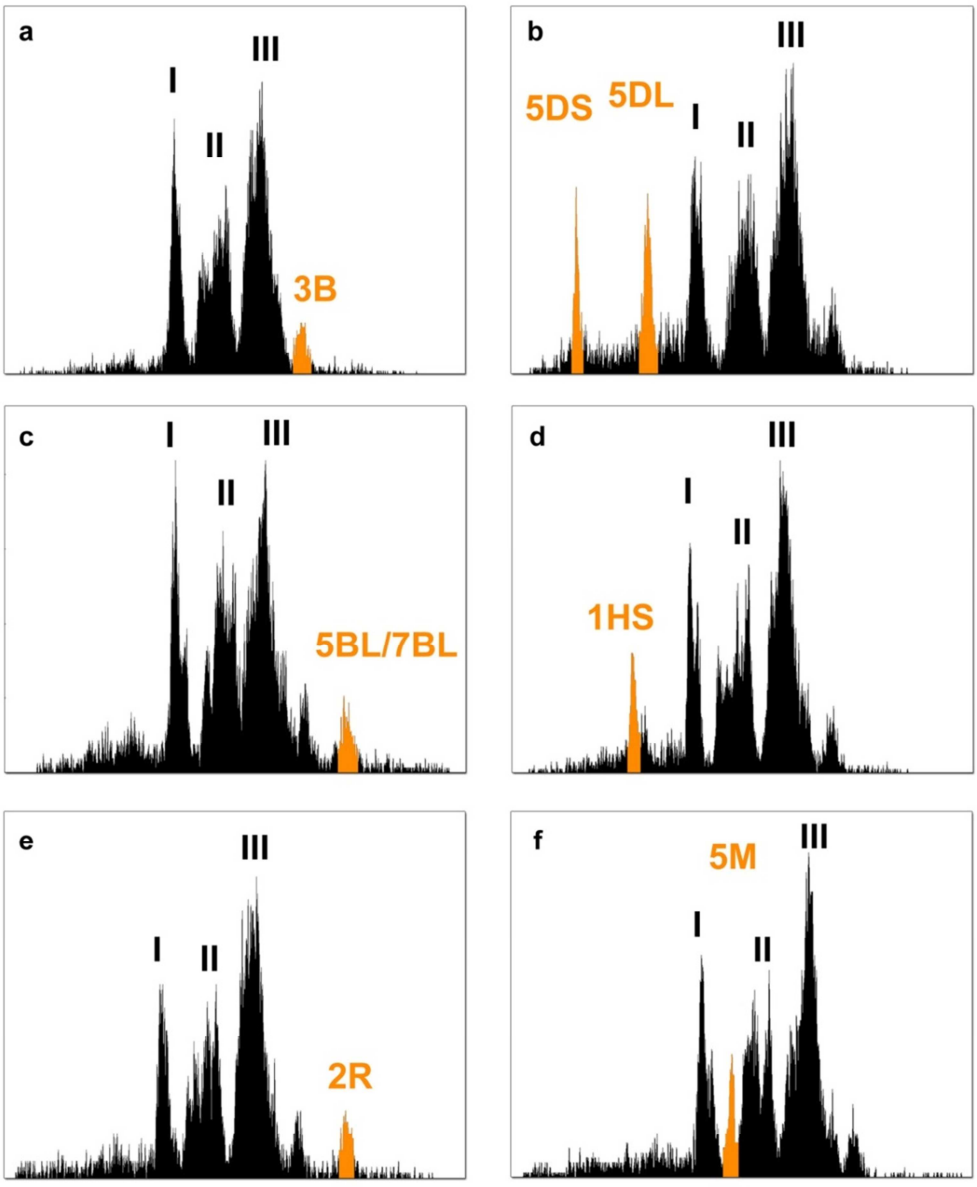

Figure 5: Flow karyotypes of bread wheat lines. Chromosomes readily sortable are highlighted by orange color. Three composite peaks representing groups of chromosomes are designated in each histogram. **a** Flow karyotype of euploid cultivar Chinese Spring. Only the largest chromosome 3B is distinguishable from the remaining twenty chromosomes. **b** The use of 5D double ditelosomic line allows sorting of chromosome 5D in form of its two arms, which are small enough to be discriminated from the rest. **c** Well-described translocation in cultivar Arina between chromosomes 5B and 7B enables sorting of the large chromosome made by fusion of two long arms. **d** Alien chromosome arm addition, in this case short arm of chromosome 1 of barley into the bread wheat background allows sorting of this arm, which is not possible in standard barley flow karyotype. **e** Large rye chromosomes form a composite peak in rye,

however rye chromosome addition into bread wheat allows their sorting, 2R-wheat addition line shown as example here. **f** Substitution of chromosome 5M of *Aegilops geniculata* for 5D chromosome of bread wheat allows isolation of the wild grass's chromosome.

While it is true that the use of various cytogenetic stocks brought a majority of results in the field of chromosome flow sorting and their application, their establishment is practically impossible in most plant species. In those where the possibility remains, it is still laborious and both time- and space-consuming effort. Therefore the scientists turned their attention to fluorescently labeling of chromosomes in order to discriminate between them on flow karyotype.

3.11 Labeling of chromosomes in suspension

Tagging particular (repetitive) DNA sequences should in theory provide a basis to identify any plant chromosome. The sample for flow cytometric analysis has to be in a form of particles in aqueous suspension, thusly labeling of these particles has to be performed in suspension as well. Several attempts discussed here prove that it is not a trivial task. Standard cytological approach to label DNA sequence is FISH (fluorescence *in situ* hybridization), so it does not come as a surprise that the first attempt was hybridization in suspension (Trask et al. 1985) performed with mouse cell nuclei. Analysis by flow cytometry enabled quantifying the amount of bound probe together with measuring the DNA content. The same was done on human chromosomes from hybrid cell line (Dudin et al. 1987), but unfortunately without analyzing the chromosomes by flow cytometry.

In plants the hindrances with FISH in suspension, which comprises several washing and pelleting steps and changes of solutions effectively prevented developing a robust protocol for sample preparation suitable for flow sorting. Answer for that was PRINS (primed *in situ* labeling) in suspension, which reduced the number of centrifugation and washing steps, and was developed for pea and field bean (Macas et al. 1995). This procedure was immediately used in field bean for discriminating and sorting of its chromosomes based on different

copy number of *FokI* repeat sequences (Pich et al. 1995). However this protocol has not been used since for any other repeat sequence and/or different plant species because of poor reproducibility. Just recently, after twenty years of attempts of labeling chromosomes in suspension a protocol named FISHIS (fluorescence *in situ* hybridization in suspension) was developed (Giorgi et al. 2013). By labeling GAA microsatellite repeats with fluorescent probe it enables discrimination of all chromosomes in tetraploid durum wheat (*Triticum durum*) and majority of hexaploid wheat chromosomes in bivariate analysis (figure 3b). So far this method is however restricted to a limited number of repeat sequences and hasn't found broader use due to some still unresolved issues (unpublished observations).

An alternative to FISH in suspension could be chromosome labeling by synthetic polyamide probes. These synthetic compounds are based on naturally occurring antibiotics distamycin A and netropsin and their ability to bind into minor groove of DNA in sequence-specific manner (Vaijayanthi et al. 2012). They can be conjugated to fluorophores and utilized as probes in microscopy and flow cytometry. Their specific binding has been proved by targeting telomeres of *Drosophila melanogaster* (TTAGG motif) and human HeLa cells (TTAGGG motif), allowing their visualization and estimation of relative telomere length in these species (Maeshima et al. 2001), the samples were however checked by the epifluorescence microscope only. Subsequently, a feasibility of fluorescently labeling chromosomes in suspension for flow cytometric analysis was demonstrated (Gygi et al. 2002). By targeting TTCCA motif the authors were able to discriminate chromosome 9 from chromosome 10, 11 and 12 in human, which form otherwise inseparable population. This method surprisingly has not been used by anybody else since 2002. Doležel's lab recently made an effort to label chromosomes of pea (targeting PisTR-B repeat) and wheat (targeting *afa* repeat) using synthetic polyamides with 18bp-long recognition site, but the expected labelling was not observed (unpublished results).

A lack of reproducible and generally usable method for labeling chromosomes in suspension for flow cytometric analysis led to exploring other approaches for discrimination of individual chromosomes and their sorting. Surprisingly the approach described in the next chapter avoids any discrimination at all.

3.12 Single chromosome sorting

To overcome the difficulties with discrimination and flow sorting individual plant chromosomes, a method for single chromosome sorting and amplification of DNA from single copies of chromosomes was developed. Although one of the biggest advantages of flow sorters is their high throughput with hundreds to thousands particles analyzed per second, they can be utilized for sorting individual single particles. This effectively alleviates the need to resolve chromosomes on a flow karyotype – in plain words it is a blind sorting, one never knows which chromosome landed in the tube, but for sure it's there alone. The identity of the chromosome is ascertained once the DNA of the chromosome is amplified by PCR with chromosome specific primers, or after sequencing.

The protocol is capable of producing microgram quantities of chromosome-specific DNA from (sub)picogram amounts of DNA. A slight advantage is taken in this protocol with utilizing mitotic chromosomes, thus the amount of DNA is doubled and every locus is present in two copies. Nevertheless, just to illustrate the sizes of chromosomes, chickpea chromosome H is roughly 120 Mbp, which corresponds to 0.24 pg DNA while 3B of bread wheat is 860 Mbp corresponding to 1.7 pg (1 pg DNA equals 0.978 Gbp according to Doležel et al. 2003) in both cases, this is the amount that has to be amplified.

To test the performance and representativeness of DNA amplification, chromosome 3B of bread wheat was chosen as a model. The results showed that it is possible to obtain DNA in sufficient quantity and quality for various downstream analyses including next-generation sequencing (Cápal et al. 2015). This approach permits to dissect nuclear genome of any plant where the

suspension of intact mitotic chromosomes can be prepared into fractions that are much cheaper to sequence and easier to assemble and analyze, albeit the resulting coverage obtained by this method is limited. The same approach can be used for flow-sorted nuclei as well, which was demonstrated by analyzing meiotic recombination rates in pollen nuclei of barley (Dreissig et al. 2015). Most recently, the single chromosome amplification was utilized in three wheat transgenic lines to identify chromosomes into which a transgene was integrated (Cápal et al. 2016). As the single chromosome genomics is the leitmotiv of this thesis and the main result of its author, a separate chapter will be dedicated to this topic.

4 Single cell / chromosome genomics

Single cell methods and approaches have the potential to change many areas of biology and medical research. They help to tackle genomic and expression diversity of cells within tissues of multicellular organisms, discover heterogeneity in microbial samples that was formerly obscured by metagenomics approaches, or infer metabolic properties of uncultured microbes, they could phase the alleles along the chromosomes in order to infer specific haplotypes as well as reveal clonal diversity in cancer tissues. It was no surprise that single cell sequencing was chosen a method of the year 2013 by the scientific journal Nature Methods in its annual announcement. As the single cell and single chromosome genomics in plant kingdom is just being established, this chapter summarizes the achievements reached mainly in the fields of microbiology and human biology. Nevertheless, the possibilities of applying single cell and mainly single chromosome approach for analyzing complex plant genomes and recent achievements in this field will be outlined as well.

4.1 Obtaining and analyzing minute amounts of DNA

Acquisition of high quality single-cell or single-chromosome data has several technical challenges. The first issue encountered is how to reliably obtain the genetic material to work with. Currently there are three predominant techniques for single cell or chromosome isolation – flow sorting, microdissection and custom microfluidics or micromanipulation devices. Although among the greatest advantages of flow sorters is their high-throughput with thousands of particles analyzed per second, they are also capable of reliably sorting only one particle. In the context of single cell collection methods, the high-throughput is again the advantage of FACS (fluorescence activated cell sorter) with a capacity for sorting tens of cells/nuclei/chromosomes into individual wells (usually in 96 or 384 plate format) in a short time (Navin et al. 2011, Evrony et al. 2012, Dreissig et al. 2015). Microdissection is the slowest approach, but allowing for visual

inspection of isolated particles, in most published cases chromosomes (Gribble et al. 2005, Kosyakova et al. 2013), which could be invaluable for obtaining chromosome-specific DNA. However, due to strong fixation of tissues prior to microdissection, the material collected this way is not ideal for multiple displacement amplification using Phi29 enzyme (Little et al. 2005). Although major producers of the amplification kits offer solutions for formalin-fixed (and paraffin-embedded for that matter) samples, their prerequisite is the high amount of input DNA, definitely not a single particle, be it a single cell, nucleus, or chromosome. Custom microfluidic devices are probably the best option for single particle capture, offering both parallelization and visual control of the isolation process, as well as limited space for contamination (Fan et al. 2011, Wu et al. 2013, Macosko et al. 2015). Moreover, on-chip DNA amplification in nanoliter volumes improves the amplification process, but this is the topic of next paragraph.

The amount of DNA in a nucleus of a typical diploid human cell is ~6 pg, mitotic plant chromosomes usually contain less than one picogram of DNA, while bacterial genomes are in range of femtograms (Doležel et al. 2003, Trevors 1996). These minute amounts of DNA have to be amplified in order to successfully perform any subsequent analyses, which usually require nanogram or even microgram amounts of input DNA. Amplifying faithfully one copy of the genome (or in the best case two copies of the genome) without introducing any artifacts such as amplification bias, allele dropout or chimaera formation is the ultimate goal and there has been substantial progress in the last few years in order to reach it.

Current approaches may be grouped into three camps. The first and the earliest group of the methods is based on PCR with either degenerate oligonucleotide primers (DOP-PCR), or a common sequence ligated to fragmented genomes, the linker-adaptor-PCR (LA-PCR). As a result of mechanism of these approaches, the genome representation is low, typically <10% (Gawad et al. 2016). Furthermore, they rely on thermostable polymerases which are error prone and introduce *de*

novo mutations during the amplification process. On the other hand, they allow amplification of material from fixed tissues and more importantly, they are more suited for copy number variation (CNV) studies (Navin et al. 2011). The second method, currently the predominant one, is based on isothermal amplification by Phi29 polymerase primed by random hexamers, called multiple displacement amplification (MDA) (Dean et al. 2002). The Phi29 enzyme is a highly processive enzyme with proof-reading activity, resulting in greater genome coverage with low error rate, as compared to PCR-based methods; also the yield is ~ ten times higher (deBourcy et al. 2014). The majority of fragments generated by MDA exceed 10 kbp, which broadens the spectrum of possible applications. The exponential amplification leads to overrepresentation of loci that were amplified first, leading to a great bias, which is the biggest disadvantage of MDA (Gawad et al. 2016). The third group of methods tries to address the issues of low coverage in PCR-based methods and the lack of uniform amplification introduced by MDA. Two hybrid methods were developed in order to solve that – displacement DOP-PCR (Langmore 2002) and multiple annealing and looping-based amplification cycles, abbreviated MALBAC (Zong et al. 2012), which immediately replaced the former. MALBAC comprises a pre-amplification step with a pool of random primers that share a common part of the sequence and allow amplicons to form loops and prevent their further nonlinear amplification; subsequently they are subjected to PCR amplification. These steps combined reduce significantly the amplification bias by three to four orders of magnitude (Lasken 2013) and in words of its inventor "MALBAC makes copies of the original instead of copies of another copies as in case of MDA". On the other hand, the use of two more error-prone polymerases, *Bst* and *Taq*, renders MALBAC more prone to calling false positives, which was reported to be 40x higher than MDA (Lasken 2013).

The amount of input DNA is a critical parameter of the amplification reactions. Genotyping performance raises dramatically, as the amount of starting material is increased beyond 1 ng in MDA reactions (Bergen et al. 2005), while if the amount exceeds 10 ng, the amplified DNA is equivalent to genomic DNA in

terms of detectable loci and genotyping performance (Šimková et al. 2008). It has also been proven in several studies that performing amplification reactions in nanoliter scale on microfluidic chips improves both accuracy of the amplification, reduces bias and reduces the risk of contamination (Marcy et al. 2007a, Fan et al. 2011, Wang et al. 2012, Han et al. 2013, Macosko et al. 2015).

4.2 Applications of single-cell genomics

It is almost unthinkable that genomes of any two cells in one individual are completely identical due to the mutations that had accumulated during DNA replication or as a result of various mutagenic sources (Frumkin et al. 2015). These genomic alternations are however masked when analysing bulk of cells from a particular tissue. In order to reveal them, one has to go to single cell level. Evrony et al. (2012) analysed genomic heterogeneity in human neuron nuclei by profiling LINE-1 retrotransposon insertions in 300 individual neuronal cells from cerebral cortex and caudate nucleus. Their results confirmed a diverse rate of insertions between individual cells, however showed that L1 retrotransposition is not the main driver behind the genomic diversity of neurons in human brain. A year later another study utilizing single cell sequencing was conducted on human neurons, proving that CNVs are abundant in human brain, with 13 – 41% of neurons having at least one *de novo* CNV spanning more than 1 Mbp and that deletions are twice as abundant as duplications (McConnel et al. 2013). Let's stay a while with heterogeneity of cell populations, but switch from healthy cells to cancer ones. Genomic instability is one of the main hallmarks of cancer cells (Park et al. 2010). The presence of several clonal subpopulations and rare cells in tumor tissue is difficult to deconvolute from bulk tissue sequencing. This has been recently addressed by single cell sequencing, by identifying three distinct clonal subpopulations in breast cancer tissue, representing sequential clonal expansions of the tumor (Navin et al. 2011). Li et al. (2012) carried out exome sequencing of 66 single tumour cells from invasive bladder carcinoma which lead

to revelation of two distinct tumour cell subpopulations, both being ancestry of single tumour cell.

Preimplantation genetic diagnostics (PGD) has been introduced to lower the risk for couples with known genetically inherited diseases to transmit it to their progeny *via in vitro* fertilization. PGD has experienced a huge leap forward in methodology since the technologies for single cell genetic analysis reached accuracy and precision of the bulk methods and enabled multiple analyses to be performed starting from single cell (Fiorentino 2012). Isothermal whole genome amplification was reported from single blastomeres isolated from cleavage stage embryo, followed by PCR analysis of 20 loci involved in severe human diseases, including deletion causing cystic fibrosis (Handyside et al. 2004). Le Caignec et al. (2006) were able to detect chromosomal imbalances from single blastomeres in preimplantation embryos at eight-cell stage. Recently, the MALBAC-based preimplantation screening was performed, where sequencing of the first and second polar bodies from fertilized eggs allowed for relatively low-cost procedure to select eggs for transfer, free of maternal aneuploidy and alleles associated with diseases (Hou et al. 2013). A feasibility of whole genome amplification has also been demonstrated for single sperm cells with 93% of the sequence amplified more than 250-fold in a microliter reaction volumes (Jiang et al. 2005). The resulting MDA products proved to be suitable for SNP and STR (short tandem repeat) genotyping as well as for sequencing. Wang et al. (2012) went further and implemented the WGA of single sperms onto microfluidic chip to minimize non-specific amplification and improve accuracy and by calling SNP variants at 1.2 million loci creating a personal recombination map of an individual. In the light of this performance, Kirkness et al. (2013) amplified and low-pass sequenced 96 sperm cells separated by micromanipulation in order to phase the heterozygous loci, fully reconstructing the chromosome-length haplotypes in human.

High-throughput sequencing of whole transcriptome, termed RNA-seq, is well-established method for gene expression profiling of bulk tissues (Wang et al.

2009). There is a growing interest in methods capable of transcriptional profiling single cells that could differentiate investigated subpopulations of cell from the bulk tissues (Kalisky et al. 2011). The averaged expression level of tissue can be skewed by a small number of cells with high expression; the level of specific transcripts can differ as much as 1000-fold between presumably equivalent cell (Bengtsson et al. 2005). Wu et al. (2014) proved that single-cell RNA-seq is capable of accurate quantitative measurement of transcriptome from single cells from relatively small number of sequencing reads; moreover that sequencing of larger number of single cell and subsequent pooling of the reads faithfully represents the bulk transcriptome complexity. Later, a proof of principle for similar transcriptomic analysis was published, except starting with single nuclei instead of single cell (Grindberg et al. 2013), capable on average to capture 16,000 of 24,057 mouse protein coding genes from single nuclei. Although performed on mouse CNS cells, one of the largest promises of this method is its feasibility for applying to different kingdoms of eukaryots - in the context of this thesis – plant nuclei. A year later was the single-cell transcriptomic analysis coupled with the analysis of genomic DNA of the very same cell (Han et al. 2014). A microfluidic device was constructed for that effort, capable of separating cytoplasmic and nuclear contents and on-chip amplification of genomic DNA and cytoplasmic mRNA, subsequently allowing off-chip co-detection and sequencing the panel of forty genes expressed by that cell, opening avenue for studying non-genetic cell-to-cell variability.

Arguably the field which was revolutionized the most by single-cell approaches is microbiology. It is estimated that currently more than 99% of bacterial taxons cannot be cultivated, owing to unknown growth conditions (Lasken and McLean 2014), thus sequencing these bacterial species relies on metagenomics approaches. Given the complexity of environmental samples, the assembly of genes and individual bacterial genomes remains a challenge. Although metagenomic investigations are improving, driven by new bioinformatics methods (reviewed in Blainey 2013), the important link between gene and organism is often lost in these studies, which is exactly where investigations on a

cell by cell basis help to shed light and unambiguously assign genes to a particular microbe. The first bacterial species that were sequenced *via* single-cell approach without prior axenic culture were the members of TM7 phylum, isolated from human oral cavity (Marcy et al. 2007b). Studies on other phyla were soon to follow and in the past years several single-amplified genomes were published, inhabiting various environments like coastal water (Stepanauskas and Sieracki 2007), hot spring sediments (Dodsworth et al. 2013), anoxic springs (Youssef et al. 2011) or biofilm in hospital sink (McLean et al. 2013).

The cases where amplification of single chromosomes were used to address a biological issue seem to be rarer sight, although some of the fundamental questions can be answered only by utilizing this method. There is an increasing number of evidence, that phenotypic effect of various sequence variations in human are best understood in form of haplotypes rather than isolated SNPs (Kirkness et al. 2013). Sequencing six billion bases of human genome is currently almost a routine task, as is identifying polymorphisms on homologous chromosomes. The problem is to phase these variants along chromosomes and thus creating haplotypes on a genome-wide scale. Fan et al. (2011) used a custom-built microfluidic device to separate and amplify homologous copy of each chromosome from a human metaphase cell. Analysing SNPs on array allowed the creation of personal haplotypes of four individuals by completely phasing 90% of the 900,000 SNPs, including for example human leukocyte antigen loci (HLA). The highly polymorphic HLA system was also genotyped in single individual by deep sequencing of flow-sorted chromosome 6 and phasing both allelic variants, thus defining the two haplotypes (Murphy et al. 2016). Flow-sorted individual chromosomes were used even earlier to completely phase SNPs along chromosome 19 in human in order to infer long-range haplotypes (Yang et al. 2011). Single chromosome approach was used to create whole-chromosome painting probes in mouse by glass-needle microdissecting the metaphase cell and amplification of all individual chromosomes (Kosyakova et al. 2013). It is a technique that can be readily utilized in characterizing intrachromosomal rearrangements or identifying chromosomal breakpoints, also for karyotyping in

species where G-banding is not able to resolve between the chromosomes due to their similar morphology and banding patterns. Chromosome painting probes were likewise constructed from single flow-sorted chromosomes of human B-cell line and used for detection of single aberrant chromosome in a patient with this earlier described translocation (Gribble et al. 2005).

4.3 Emergence of single-cell/chromosome approaches in plants

As indicated above, this will be a short one. The only truly single-cell study performed in plants to date has been the analysis of meiotic recombination based on tetrad analysis in maize (Li et al. 2015). The authors describe the method for isolation of four haploid microspores from a tetrad by micromanipulated extraction from the anthers, their subsequent whole genome amplification and sequencing. This approach allows to directly study the recombination events during the process of meiosis by observing genomic exchanges between the four chromatids of the same tetrad. From the total of 24 tetrads a high-resolution recombination map was constructed, comprising almost 600,000 SNPs with average distance of 235 bp between two neighbouring markers, allowing for very fine recombination mapping. Similar approach was executed with flow-sorted nuclei of barley, amplifying and genotyping WGA products of 50 single pollen nuclei allowed to estimate the number of meiotic crossovers per chromosome, count recombination frequencies along the barley chromosome 3H and compare the segregation distortions with doubled haploid mapping population produced by cross of the same genotypes (Dreissig et al. 2014). Both published studies hold the promise of utilizing this approach as a complement to genetic mapping using segregating populations.

The utilization of single amplified chromosomes in plant genomics has been pioneered in Doležel's lab. We have reported successful flow-sorting, amplification and sequencing of the largest of bread wheat chromosomes – 3B.

The merged sequences from three independent amplifications resulted in 60% coverage of the chromosome, entirely covering 2204 genes annotated in 3B pseudomolecule (Cápal et al. 2015); a result that reveals the potential use for obtaining chromosome-specific sequences in plant species where individual chromosomes cannot be resolved by conventional flow cytometry. The number of successfully detected PCR markers led to the application of this method for localization of single copy transgene in three transgenic bread wheat lines (Cápal et al. 2016). Surprisingly, in each of these similarly performing lines, the transgene has integrated into a different chromosome. The protocol comprised FISHIS labelling of chromosomes prior the flow-sorting, FISH identification of candidate chromosomes and single chromosome amplification to pinpoint the chromosomal affiliation of the transgene. A detailed view of both studies is listed in the supplement section of this thesis. Moreover, the data from sequencing of single amplified chromosomes of pea and lentil are being used as a tool to validate whole genome assemblies of the two legumes by resolving sequence contigs to their respective chromosomes as well as for anchoring linkage groups to chromosomes (unpublished results).

5 Aims

Present thesis aims at broadening the spectrum of applications of flow-sorted chromosomes in plant genomics. It summarizes the current knowledge about plant genomes, briefly describes the principles of flow cytometry and its uses in plant genome research, namely the chromosome-based strategy in plant genomics. A particular focus is laid upon utilizing single flow-sorted chromosomes. The experimental part concentrated on single chromosome genomics with the following aims:

1) Development of a protocol for single flow-sorted chromosome amplification and sequencing
2) Characterization of resulting sequences using chromosome 3B of bread wheat as a model
3) Localization of single copy transgene in three distinct wheat transgenic lines
4) Measuring recombination rate along chromosome 3H of barley in pollen of *Morex-Barke* F1 hybrid

6 General conclusion

The thesis focused on expanding the spectrum of applications of chromosome genomics in plant sciences. It couples flow cytometric sorting plant chromosomes with the deployment of single-cell and single-chromosome approaches. This methodology simplifies genome analysis even in plant species where chromosomes cannot be discriminated one from each other by flow cytometric analysis. Its feasibility was demonstrated on individual chromosomes as well as on plant flow-sorted nuclei in three original studies. Theoretically, this protocol enables obtaining chromosome-specific DNA in every plant species and cultivar where a liquid suspension of chromosomes can be prepared. This achievement provides new opportunities for the analysis of chromosome heterozygosity, allel phasing alleles and complete haplotypes along the chromosomes and studies of meiotic recombination in plants.

7 Publications

7.1 Multiple displacement amplification of the DNA from single flow–sorted plant chromosome

Cápal P, Blavet N, Vrána J, Kubaláková M, Doležel J (2016) Multiple displacement amplification of the DNA from single flow–sorted plant chromosome. *Plant J* **84**:838-844

Abstract:
A protocol is described for production of micrograms of DNA from single copies of flow-sorted plant chromosomes. Of 183 single copies of wheat chromosome 3B, 118 (64%) were successfully amplified. Sequencing DNA amplification products using an Illumina HiSeq 2000 system to 10x coverage and merging sequences from three separate amplifications resulted in 60% coverage of the chromosome 3B reference, entirely covering 30% of its genes. The merged sequences permitted de novo assembly of 19% of chromosome 3B genes, with 10% of genes contained in a single contig, and 39% of genes covered for at least 80% of their length. The chromosome-derived sequences allowed identification of missing genic sequences in the chromosome 3B reference and short sequences similar to 3B in survey sequences of other wheat chromosomes. These observations indicate that single-chromosome sequencing is suitable to identify genic sequences on particular chromosomes, to develop chromosome-specific DNA markers, to verify assignment of DNA sequence contigs to individual pseudomolecules, and to validate whole-genome assemblies. The protocol expands the potential of chromosome genomics, which may now be applied to any plant species from which chromosome samples suitable for flow cytometry can be prepared, and opens new avenues for studies on chromosome structural heterozygosity and haplotype phasing in plants.

DOI: 10.1111/tpj.13035

7.2 The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat

Cápal P, Endo TR, Vrána J, Kubaláková M, Karafiátová M, Komínková E, Mora-Ramírez I, Weshke W, Doležel J (2016) The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat. *Plant Methods* **12**:24

Abstract:
Background: Identification of transgene insertion sites in plant genomes has practical implications for crop breeding and is a stepping stone to analyze transgene function. However, single copy sequences are not always easy to localize in large plant genomes by standard approaches.

Results: We employed flow cytometric chromosome sorting to determine chromosomal location of barley sucrose transporter construct in three transgenic lines of common wheat. Flow-sorted chromosomes were used as template for PCR and fluorescence in situ hybridization to identify chromosomes with transgenes. The chromosomes carrying the transgenes were then confirmed by PCR using DNA amplified from single flow-sorted chromosomes as template.

Conclusions: Insertion sites of the transgene were unambiguously localized to chromosomes 4A, 7A and 5D in three wheat transgenic lines. The procedure presented in this study is applicable for localization of any single-copy sequence not only in wheat, but in any plant species where suspension of intact mitotic chromosomes suitable for flow cytometric sorting can be prepared.

DOI: 10.1186/s13007-016-0124-8

7.3 Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley

Dreissig S, Fuchs J, Capal P, Kettles N, Byrne E, Houben A (2015) Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley. *PLoS ONE* **10**:e0137677

Abstract:
The detection of meiotic crossovers in crop plants currently relies on scoring DNA markers in a segregating population or cytological visualization. We investigated the feasibility of using flow-sorted haploid nuclei, Phi29 DNA polymerase-based whole-genome-amplification (WGA) and multi-locus KASP-genotyping to measure meiotic crossovers in individual barley pollen grains. To demonstrate the proof of concept, we used 24 gene-based physically mapped single nucleotide polymorphisms to genotype the WGA products of 50 single pollen nuclei. The number of crossovers per chromosome, recombination frequencies along chromosome 3H and segregation distortion were analysed and compared to a doubled haploid (DH) population of the same genotype. The number of crossovers and chromosome wide recombination frequencies show that this approach is able to produce results that resemble those obtained from other methods in a biologically meaningful way. Only the segregation distortion was found to be lower in the pollen population than in DH plants.

DOI: 10.1371/journal.pone.0137677

7.4 Advances in plant chromosome genomics

Doležel J, Vrána J, Cápal P, Kubaláková M, Burešová V, Šimková H (2014) Advances in plant chromosome genomics. *Biotechnol Adv* **32**:122-136

Abstract:

Next generation sequencing (NGS) is revolutionizing genomics and is providing novel insights into genome organization, evolution and function. The number of plant genomes targeted for sequencing is rising. For the moment, however, the acquisition of full genome sequences in large genome species remains difficult, largely because the short reads produced by NGS platforms are inadequate to cope with repeat-rich DNA, which forms a large part of these genomes. The problem of sequence redundancy is compounded in polyploids, which dominate the plant kingdom. An approach to overcoming some of these difficulties is to reduce the full nuclear genome to its individual chromosomes using flow-sorting. The DNA acquired in this way has proven to be suitable for many applications, including PCR-based physical mapping, in situ hybridization, forming DNA arrays, the development of DNA markers, the construction of BAC libraries and positional cloning. Coupling chromosome sorting with NGS offers opportunities for the study of genome organization at the single chromosomal level, for comparative analyses between related species and for the validation of whole genome assemblies. Apart from the primary aim of reducing the complexity of the template, taking a chromosome-based approach enables independent teams to work in parallel, each tasked with the analysis of a different chromosome(s). Given that the number of plant species tractable for chromosome sorting is increasing, the likelihood is that chromosome genomics - the marriage of cytology and genomics - will make a significant contribution to the field of plant genetics.

DOI: 10.1016/j.biotechadv.2013.12.011

7.5 Flow cytometry in plant research. In Applied plant cell biology

Vrána J, Cápal P, Bednářová M, Doležel J (2014) Flow cytometry in plant research. In Applied plant cell biology, Editors: Peter Nick, Zdeněk Opatrný, pp 395-430. Springer-Verlag Berlin Heidelberg 2014

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

DOI: 10.1007/978-3-642-41787-0_13

8 References

Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol **56**:1919-1925

Amborella Genome Project (2013) The *Amborella* genome and the evolution of flowering plants. Science **342**:1241089

Barteneva NS, Fasler-Kan E, Vorobjev IA (2012) Imaging Flow Cytometry: Coping with Heterogeneity in Biological Systems. J Histochem Cytochem **60**:723-733

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

Bartoš J, Paux E, Kofler R, Havránková M, Kopecký D, Suchánková P, et al. (2008) A first survey of the rye (*Secale cereale*) genome composition through BAC end sequencing of the short arm of chromosome 1R. BMC Plant Biol **8**:95

Bengtsson M, Ståhlberg A, Rorsman P, Kubista M (2005) Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals $\log_{normal}$ distribution of mRNA levels. Genome Res **15**:1388–1392

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

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

Bennetzen JL, Wang H (2014) The contribution of transposable elements to the structure, function and evolution of plant genomes. Annu Rev Plant Biol **65**:505-530

Bennetzen JL (1993): The contributions of retroelements to plant genome organization, function and evolution. Trends Microbiol **4**:347–353

Berkman PJ, Skarshewski A, Lorenc MT, Lai K, Duran Ch, Ling EYS, et al. (2011) Sequencing and assembly of low copy and genic regions of isolated Triticum aestivum chromosome arm 7DS. Plant Biotechnol J **9**:768775

Berkman PJ, Skarshewski A, Manoli S, Lorenc MT, Stiller J, Smits L, et al. (2012) Sequencing wheat chromosome arm 7BS delimits the 7BS/4AL translocation and reveals homoeologous gene conservation. Theor Appl Genet **124**:423–32

Berkman PJ, Visendi P, Lee HC, Stiller J, Manoli S, Lorenc MT, et al. (2013) Dispersion and domestication shaped the genome of bread wheat. Plant Biotechnol J **11**:564–71

Blainey PC (2013). The future is now: single-cell genomics of bacteria and archaea. FEMS Microbiol Rev **37**:407–427

Blanc G, Hokamp K, Wolfe KH (2003) A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. Genome Res **13**:137-144

Böck G, Steinlein P, Huber LA (1997) Cell biologists sort things out: Analysis and purification of intracellular organelles by flow cytometry. Trends Cell Biol **7**:499-503

Bureau TE, Wessler SE (1994) Mobile inverted-repeat elements of the *Tourist* family are associated with the genes of many plant genomes. Proc Natl Acad Sci USA **91**:1411–1415

Cápal P, Blavet N, Vrána J, Kubaláková M, Doležel J (2016) Multiple displacement amplification of the DNA from single flow–sorted plant chromosome. Plant J **84**:838-844

Cápal P, Endo TR, Vrána J, Kubaláková M, Karafiátová M, Komínková E, Mora-Ramírez I, Weshke W, Doležel J (2016) The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat. Plant Methods **12**:24

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

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

Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J et al. (2014) Structural and functional partitioning of bread wheat chromosome 3B. Science **345**:article no 1249721

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**:795–804

Cronn R, Wendel JF (1998) Simple methods for isolating homoeologous loci from allopolyploid genomes. Genome **41**:756-761

De Bodt S, Maere S, Van de Peer Y (2005) Genome duplication and the origin of angiosperms. Trends Ecol Evol **20**:591-7

De Bourcy CFA, De Vlaminck I, Kanbar JN, Wang J, Gawad C, Quake SR (2014). A Quantitative comparison of Single-Cell Whole Genome Amplification Methods. Plos ONE **9**: e105585

De Laat AMM, Blaas J (1984) Flow-cytometric characterization and sorting of plant chromosomes. Theor Appl Genet **67**:463-467

Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P et al. (2002) Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci USA **99**:5261-5266

D'Hondt A, Denoeud F, Aury JM, Baurens FC, Carreel F, Garsmeur O (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. Nature **488**:213-219

Dodsworth JA, Blainey PC, Murugapiran SK, Swingley WD, Ross CA, Tringe SG et al. (2013) Single-cell and metagenomic analyses indicate a fermentative and saccharolytic lifestyle for members of the OP9 lineage. Nature Commun **4**:1854

Doležel J, Číhalíková J, Lucretti S (1989) Analysis of Nuclear DNA content in plant cells by Flow cytometry. Biol Plantarum **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 (1997) Application of flow cytometry for the study of plant genomes. J Appl Genet **38**:285– 302

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

Doležel J, Vrána J, Šafář J, Bartoš J, Kubaláková M, šimková H (2012) Chromosomes in the flow to simplify genome analysis. Funct Integr Genomics **12**:397-416

Doležel J, Vrána J, Cápal P, Kubaláková M, Burešová V, Šimková H (2014) Advances in plant chromosome genomics. Biotechnol Adv **32**:122-136

Dreissig S, Fuchs J, Cápal P, Kettles N, Byrne E, Houben A (2015) Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley. PLoS ONE **10**:e0137677

Dudin G, Steegmayer EW, Vogt P, Schnitzer H, Diaz E, Howell KE, Cremer T, Cremer C (1988) Sorting of chromosomes by magnetic separation. Hum Genet **80**:111–116

Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, Lehmann HS et al. (2012) Single-Neuron Sequencing Analysis of L1 Retrotransposition and Somatic Mutation in the Human Brain Cell **151**: 483–496

Fan HC, Wang J, Potanina A, Quake SR (2011) Whole-genome molecular haplotyping of single cells. Nat Biotechnol **29**:51-60

Feuillet C, Eversole K (2007) Physical mapping of the wheat genome: a coordinated effort to lay the foundation for genome sequencing and develop tools for breeders. Isr J Plant Sci **55**:307–13

Fiorentino F (2012) Molecular Genetic Analysis of Single Cells. Semin Reprod Med **30**:268–283

Flavell AJ, Pearce SR, Kumar A (1994) Plant transposable elements and the genome. Curr Opin Genet Dev **4**:838–844

Fluch S, Kopecky D, Burg K, Šimková H, Taudien S, Petzold A, et al. (2012) Sequence composition and gene content of the short arm of rye (*Secale cereale*) chromosome 1. PLoS One **7**:e30784

Frumkin D, Wasserstrom A, Kaplan S, Feige U, Shapiro E (2005) Genomic variability within an organism exposes its cell lineage tree. PLoS Comput Biol **1**:e50

Galbraith DW (1994) Flow cytometry and sorting of plant protoplasts and cells. Methods Cell Biol **42B**: 539–561

Galbraith DW (2007) Protoplast analysis using flow cytometry. In: Doležel J, Greilhuber J, Suda J [eds], Flow cytometry with plant cells: Analysis of genes, chromosomes and genomes, p. 231–250, Wiley–VCH, Weinheim

Gawad C, Koh W, Quake SR (2016) Single-cell genome sequencing: current state of the science. Nat Reviews **17**:175-188

Gill KS, Arumuganathan K, Le JH (1999) Isolating individual wheat (*Triticum aestivum*) chromosome arm by flow cytometric analysis of ditelosomic lines. Theor Appl Genet **98**:1248-1252

Giorgi D, Farina A, Grosso V, Gennara A, Ceoloni C, Lucretti S (2013) FISHIS: fluorescence *in situ* hybridization in suspension and chromosome flow sorting made easy. PLoS ONE **8**:e57994

Glower NM, Redestig H, Dessimoz C (2016) Homoeologs: What are they and how do we infer them? Trends Plant Sci **21**:609-621

Grandbastien M-A (1992) Retroelements in higher plants. Trends Genet **8**:103–108

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

Gribble S, Ng BL, Prigmore E, Burford DC, Carter NP (2004) Chromosome paints from single copies of chromosomes. Chrom Res **12**:143–151

Grindberg RV, Yee-Greenbaum JL, McConnell MJ, Novotny M, O'Shaughnessy AL, Lambert GM et al. (2013) RNA-sequencing from single nuclei. Proc Natl Acad Sci USA **110**: 19802–19807

Gygi MP, Ferguson, MD, Mefford HC, Lund KP, O'Day C, Zhou P et al. (2002) Use of fluorescent sequence-specific polyamides to discriminate human chromosomes by microscopy and flow cytometry. Nucleic Acids Res **30**:2790-2799

Han L, Zi X, Garmire LX, Wu Y, Weissman SM, Pan X, Fan R (2014). Co-detection and sequencing of genes and transcripts from the same single cells facilitated by a microfluidics platform. Sci Rep **4**:6485

Handyside AH, Robinson MD, Simpson RJ, Omar MB, Shaw M-A, Grudzinkas jG, Rutherford A (2004) Isothermal whole genome amplification from single and small numbers of cells: a new era for preimplantation genetic diagnosis of inherited disease. Mol Hum Reprod **10**:767–772

Hawkins JS, Kim H, Nason JD, Wing RA, Wendel JF (2006) Differential lineage-specific amplification of transposable elements is responsible for genome size variation in *Gossypium*. Genome Res **16**:1252–1261

Hernandez P, Martis M, Dorado G, Pfeifer M, Gálvez S, Schaaf S, 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**:377–86.

Hollister JD, Smith LM, Guo Y-L, Ott F, Weigel D, Gaut BS (2011) Transposable elements and small RNAs contribute gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Proc Natl Acad Sci USA **108**:2322–2327

Hou Y, Fan W, Yan L, Lian Y, Huang J, Li J et al. (2013) Genome analyses of single human oocytes. Cell **155**:1492–1506

Huskins CL (1931) A cytological study of Vilmorin's unfixable dwarf wheat. J Genet **25**:113-124

International Wheat Genome Sequencing Consortium (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science **345**:article no. 1251788

Jakobson I, Reis D, Tiidema A, Peusha H, Timofejeva L, Valárik M, et al. (2012) Fine mapping, phenotypic characterization and validation of non-race-specific resistance to powdery mildew in a wheat–*Triticum militinae* introgression line. Theor Appl Genet **125**:609–23

Janda J, Šafář J, Kubaláková M, Bartoš J, Kovářová P, Suchánková P, et al. (2006) Advanced resources for plant genomics: BAC library specific for the short arm of wheat chromosome 1B. Plant J **47**:977–86

Jiang Z, Zhang X, Deka R, Jin L (2005) Genome amplification of single sperm using multiple displacement amplification. Nucleic Acids Res **33**:e91

Kalisky T, Blainey P, Quake SR (2011) Genomic analysis at the single-cell level. Annu Rev Genet **45**: 431–445

Karafiátová M, Bartoš J,Kopecký D, Ma L, Sato K, Houben A, et al. (2013) Mapping nonrecombining regions in barley using multicolor FISH. Chromosome Res **21**:739–51

Kejnovský E, Vrána J, Matsunaga S, Souček P, Široký J, Doležel J, et al. (2001) Localization of male-specifically expressed MROS genes of *Silene latifolia* by PCR on flow-sorted sex chromosomes and autosomes. Genetics **158**:1269–77

Kejnovský E, Hawkins JS, Feschotte C (2012) Plant Transposable Elements: Biology and Evolution. In Plant Genome Diversity Volume 1, Eds: Wendel JF, Greilhuber J, Doležel J, Leitch IJ. Springer International Publishing AG, Cham (Switzerland)

Kirkness EF, Grindberg RV, Yee-Greenbaum J (2013) Sequencing of isolated sperm cells for direct haplotyping of a human genome. Genome Res **23**:826-832

Kejnovský E, Hawkins JS, Feschotte C (2012) Plant Transposable Elements: Biology and Evolution. In Plant Genome Diversity Volume 1, eds: Wendel JF, Greilhuber J, Doležel J, Leitch IJ, Springer 2012

Knight RL (1949) Dictionary of genetics, Chronica Botanica

Kofler R, Bartoš J, Gong L, Stift G, Suchánková P, Šimková H, et al. (2008) Development of microsatellite markers specific for the short arm of rye (Secale cereale L.) chromosome 1. Theor Appl Genet **117**:915–26

Kosyakova N, Hamid BA, Chaveerach A, Pinthong K, Siripiyasing P, Supiwong W (2013) Generation of multicolor banding probes for chromosomes of different species. Mol Cytogenetics **6**:6

Kubaláková M, Valárik M, Bartoš J, Vrána J, Číhalíková J, Molnár-Lang M, Doležel J (2003) Analysis and sorting of rye (*Secale cereale* L) chromosomes using flow cytometry. Genome **46**:893-905

Langmore JP (2002) Rubicon Genomics. Pharmacogenomics **3**:557-560

Lasken RS (2013) Single-cell sequencing in its prime. Nat Biotech **31**:211-212

Lasken RS, McLean JS (2014). Recent advances in genomic DNA sequencing of microbial species from single cells. Nat Genet **15**:577-584

Le Caignec C, Spitz C, Sermon K, De Rycke M Thienpont B, Debrock S et al. (2006) Single-cell chromosomal imbalances detection by array CGH. Nucleic Acids Res **34**:e68

Lee JH, Arumuganathan K, Chung YS, Kim KY, Chung WB, Bae KS et al. (2000) Flow cytometric analysis and chromosome sorting of barley (*Hordeum vulgare*). Moll Cells **10**:619-625

Li IJ, Arumuganathan K, Rines HW, Phillips RL, Riera-Lizarazu O, Sandhu D, Zhou Y, Gill KS (2001) Flow cytometric sorting of maize chromosome 9 from an oat–maize chromosome addition line. Theor Appl Genet **102**:658–663

Li Y, Xu X, Song L, Hou Y, Li Z, et al. (2012) Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer. Gigascience **1**: 12

Li A, Liu D, Wu J, Zhao X, Hao M, Geng S et al. (2014) mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. Plant Cell Online **26**:1878-1900

Li X, Li L, Yan J (2015) Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize. Nat Commun **6**:6648

Little SE, Vuononvirta R, Reis-Filho JS, Natrajan R, Iravani M, Fenwick K et al. (2006) Array CGH using whole genome amplification of fresh-frozen and formalin-fixed, paraffin-embedded tumor DNA. Genomics **87**:298 – 306

Lucas SJ, Šimková H, Šafář J, Jurman I, Cattonaro F, Vautrin S, et al. (2012) Functional features of a single chromosome arm in wheat (1AL) determined from its structure. Funct Integr Genomics **12**:173–82

Lucas SJ, Akpınar BA, Kantar M,Weinstein Z, Aydınoglu F, Šafář J, et al. (2013) Physical mapping integrated with syntenic analysis to characterize the gene space of the long arm of wheat chromosome 1A. PLoS One **8**:e59542

Lucretti S, Doležel J, Schubert I, Fuchs J (1993) Flow karyotyping and sorting of *Vicia faba* chromosomes. Theor Appl Genet **85**:665-672

Macas J, Doležel J, Lucretti S, Pich U, Meister A, Fuchs J, et al. (1993) Localization of seed protein genes on flow-sorted field bean chromosomes. Chromosome Res **1**:107–15

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

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M et al. (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell **161**:1202–1214

Maeshima K, Janssen S, Laemmli UK (2001) Specific targeting of insect and vertebrate telomeres with pyrrole and imidazole polyamides. EMBO J **20**:3218-3228

Marcy Y, Ishoey T, Lasken RS, Stockwell TB, Walenz BP, Halpern AL et al. (2007a) Nanoliter reactors improve multiple displacement amplification of genomes from single cells. PLoS Genet **3**:e155

Marcy Y, Ouverney C, Bik EM, Losekann T, Ivanova N, Martin HC et al. (2007b) Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. Proc Natl Acad Sci USA **104**:11889–11894

Martis MM, Zhou R, Haseneyer G, Schmutzer T, Vrána Jan, Kubaláková M, et al. (2013) Reticulate Evolution of the Rye Genome. Plant Cell **25**:3685–98

Mayer KFX, Taudien S, Martis M, Šimková H, Suchánková P, Gundlach H, et al. (2009) Gene content and virtual gene order of barley chromosome 1H. Plant Physiol **151**:496–505

Mayer KFX, Martis M, Hedley PE, Šimková H, Liu H, Morris JA, et al. (2011) Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell **23**:1249–63

McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron Cet al. (2013) Mosaic Copy Number Variation in Human Neurons. Science **342**:632-637

McLean JS, Lombardo M-J, Badger JH, Edlund A, Novotny M, Yee-Greenbaum Jet al. (2013) Candidate phylum TM6 genome recovered from a hospital sink biofilm provides genomic insights into this uncultivated phylum. Proc Natl Acad Sci USA **110**:E2390–E2399

Muñoz-Amatriaín M, Moscou MJ, Bhat PR, Svensson JT, Bartoš J, Suchánková P, et al. (2011) An improved consensus linkage map of barley based on flow-sorted chromosomes and single nucleotide polymorphism markers. Plant Genome **4**:238–49

Murphy NM, Burton M, Powell DR, Rossello FJ, Cooper D, Chopra A et al. (2016) Haplotyping the human leukocyte antigen system from single chromosomes. Sci Rep **6**:30381

Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature **472**:90-96

Nie X, Li B,Wang L, Liu P, Biradar SS, Li T, et al. (2012) Development of chromosome-arm-specific microsatellite markers in *Triticum aestivum* (*Poaceae*) using NGS technology. Am J Bot **99**:369–71

Neumann P, Lysák M, Doležel J, Macas J (1998) Isolation of chromosomes from *Pisum sativum* L. hairy root cultures and their analysis by flow cytometry. Plant Sci **137**:205–215

Neumann P, Požárková D, Vrána J, Doležel J, Macas J. (2002) Chromosome sorting and PCR-based physical mapping in pea (Pisum sativum L.). Chromosome Res **10**:63–71

Osborn TC, Pires JC, Birchler JA, Auger DL, Chen J, Lee H-S et al. (2003) Understanding mechanisms of novel gene expression in polyploids. Trend Genet **19**:141-147

Papadakis AK, Roubelakis-Angelakis A (2002) Oxidative stress could be responsible for the recalcitrance of plant protoplasts. Plant Physiol **40**:549-559

Park SY, Gonen M, Kim HJ, Michor F, Polyak K (2010) Cellular andgenetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. J Clin Invest **120**:636–644

Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, et al. (2008) A physical map of the 1-gigabase bread wheat chromosome 3B. Science **322**:101–4

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

Petit J-M, Denis-Gay M, Ratinaud M-H (1993) Assessment of fluorochromes for cellular structure and function studies by flow cytometry. Biol Cell **78**:1-13

Pfosser M, Heberle-Bors E, Amon A, Lelley T (1995) Evaluation of sensitivity of flow cytometry in detecting aneuploidy in wheat using disomic and ditelosomic wheat–rye addition lines. Cytometry **21A**:387–393

Philippe R, Paux E, Bertin I, Sourdille P, Choulet F, Laugier C, et al. (2013) A high density physical map of chromosome 1BL supports evolutionary studies, map-based cloning and sequencing in wheat. Genome Biol **14**:R64

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

Piegu B, Guyot R, Picault N, Roulin A, Saniyal A, Kim H, Collura K, Brar DS, Jackson S, Wing RA et al. (2006) Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. Genome Res **16**:1262–1269

Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JPT, et al. (2013) Advances in plant gene-targeted and functional markers: a review. Plant Methods **9**:6

Požárková D, Koblížková A, Román B, Torres AM, Lucretti S, Lysák M, et al. (2002) Development and characterization of microsatellite markers from chromosome 1-specific DNA libraries of *Vicia faba*. Biol Plant **45**:337–45

Ramsey J, Schemske DW (1998) Pathways, mechanisms and rates of polyploid formation in flowering plants. Annu Rev Ecol Syst **29**:467-501

Riera-Lizarazu O, Rines HW, Phillips RL (1996) Cytological and molecular characterization of oat x maize partial hybrids. Theor Appl Genet **93**:123–135

Román B, Satovic Z, Požárková D, Macas J, Doležel J, Cubero JI, et al. (2004) Development of a composite map in Vicia faba, breeding applications and future prospects. Theor Appl Genet **108**:1079–88

Roux N, Toloza A, Radecki Z, Zapata-Arias FJ, Doležel J (2003) Rapid detection of aneuploidy in *Musa* using flow cytometry. Plant Cell Rep **21**:483–490

Šafář J, Bartoš J, Janda J, Bellec A, Kubaláková M, Valárik M, et al. (2004) Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J **39**:960–968

Šafář J, Šimková H, Kubaláková M, Číhalíková J, Suchánková P, Bartoš J, et al. (2010) Development of chromosome-specific BAC resources for genomics of bread wheat. Cytogenet Genome Res **129**:211–223

Salmon A, Udall JA, Jeddeloh JA, Wendel J (2012) Targeted capture of homoeologous coding and noncoding sequence in polyploid cotton. G3 (Bethesda) **2**:921-930

SanMiguel P, Bennetzen JL (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann Bot **82**:37–44

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nat Genet **20**:43–45

Schatz MC, Witkowski J, McCombie WR (2012) Current challenges in *de novo* plant genome sequencing and assembly. Genome Biol **13**:243

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S et al. (2009) The B73 maize genome: complexity, diversity, and dynamics, Science **326**:1112–1115

Schreiber AW, Hayden MJ, Forrest KL, Kong SL, Langridge P, Baumann U (2012) Transcriptome-scale homoeolog-specific transcript assemblies of bread wheat. BMC Genomics **13**:492

Šesek P, Šuštar-Vozlič J, Bohanec B (2000) Determination of aneuploids in hop (*Humulus lupulus* L.) using flow cytometry. Pflugers Arch **439**:16–18

Shangguan L, Han J, Kayesh E, Sun X, Zhang C, Pervaiz T, et al. (2013) Evaluation of genome sequencing quality in selected plant species using expressed sequence tags. PLoS One **8**:e69890

Shatalina M, Wicker T, Buchmann JP, Oberhaensli S, Šimková H, Doležel J, et al. (2013) Genotype-specific SNP map based on whole chromosome 3B sequence information from wheat cultivars Arina and Forno. Plant Biotechnol J **11**:23–32

Shapiro H (2003) Practical flow cytometry, 4th edition, Wiley, Hoboken

Simillion C, Vandepoele K, Van Montagu MCE, Zabeau M, Van de Peer Y (2002) The hidden duplication past of Arabidopsis thaliana. Proc Natl Acad Sci USA **99**:13627-13632

Šimková H, Číhalíková J, Vrána J, Lysák MA, Doležel J. (2003) Preparation of HMW DNA from plant nuclei and chromosomes isolated from root tips. Biol Plant **46**:369–73.

Šimková H, Svensson JT, Condamine P, Hřibová E, Suchánková P, Bhat PR, et al. (2008) Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genomics **9**:294.

Šimková H, Šafář J, Kubaláková M, Suchánková P, Číhalíková J, Robert-Quatre H, et al. (2011) BAC libraries from wheat chromosome 7D: efficient tool for positional cloning of aphid resistance genes. J Biomed Biotechnol **2011**:302543.

Small RL, Ryburn JA, Wendell JF (1999) Low levels of nucleotide diversity at homoeologous *Adh* loci in allotetraploid cotton (*Gossypium* L.). Mol Biol Evol **16**:491-501

Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. Annu Rev Plant Bio **60**:561-588

Stepanauskas R, Sieracki ME (2007) Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. Proc Natl Acad Sci **104**:9052-9057

Suchánková P, Kubaláková M, Kovářová 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

Swift H (1950) The constancy of deoxyribose nucleic acid in plant nuclei. Proc Natl Acad Sci USA **36**:643-654

The *Arabidopsis* genome initiative (2002) Analysis of the genome sequence of flowering plant *Arabidopsis thaliana*. Nature **408**:796-815

The French-Italian Public Consortium for Grapevine Genome Characterization (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature **449**:463- 468

Trevors JT (1996) Genome size in bacteria. A Van Leeuw J Microb **69**:293-30

Trask BJ, Van den Engh G, Landegent J, In de Wal NJ, Van der Ploegh M (1985) Detection of DNA sequences in nuclei in suspension by in situ hybridization and dual beam flow cytometry. Science **230**:1401–1403

Udall JA, Swanson JM, Haller K, Rapp, RA, Sparks ME, Hatfield J et al. (2006) A global assembly of cotton ESTs. Genome Res **16**:441-450

Vaijayanthi T, Bando T, Pandian GN, Sugiyama H (2012) Progress and prospects of pyrrole-imidazole polyamide-fluorophore conjugates as sequence-selective DNA probes. ChemBioChem **13**:2170-2185

Valárik M, Bartoš J, Kovářová P, Kubaláková M, de Jong H, Doležel J (2004) High-resolution FISH on super-stretched flow-sorted plant chromosomes. Plant J **37**:940–50

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

Vrána J, Kubaláková M, Šimková H, Číhalíková J, Lysák MA, Doležel J (2000) Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). Genetics **156**:2033–2041

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics Nature Reviews Genetics **10**:57-63

Wang J, Fan CH, Behr B, Quake SR (2012) Genome-wide Single-Cell Analysis of Recombination Activity and De Novo Mutation Rates in Human Sperm. Cell **150**:402–412

Wenzl P, Li H, Carling J, Zhou M, Raman H, Paul E et al. (2006) A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics **7**:206

Wenzl P, Suchánková P, Carling J, Šimková H, Huttner E, Kubaláková M, et al. (2010) Isolated chromosomes as a new and efficient source of DArTmarkers for the saturation of genetic maps. Theor Appl Genet **121**:465–74

Wu AR, Neff NF,Kalisky T, Dalerba P, Treutlein B, Rothenberg ME et al. (2014) Quantitative assessment of single-cell RNA-sequencing methods. Nat Methods **11**:41-46

Yang H, Chen X, Hung Wong W (2011) Completely phased genome sequencing through chromosome sorting. Proc Natl Acad Sci USA **109**:12-17

Youssef NH, Blainey PC, Quake SR, Elshahed MS (2011) Partial genome assembly for a candidate division OP11 single cell from an anoxic spring (Zodletone Spring, Oklahoma). Appl Environ Microbiol **77**:7804–7814

Zhang X, Wessler SR (2004) Genome-wide comparative analysis of the transposable elements in the related species *Arabidopsis thaliana* and *Brassica oleracea*. Proc Natl Acad Sci USA **101**:5589–5594

Zong C, Lu S, Chapman AR, Xie XS (2012) Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. Science **338**:1622

9 Abbreviations

BAC – bacterial artificial chromosome

CNV – copy number variation

DArT- diversity arrays technology

FACS – fluorescence activated cell sorting

HMW – high molecular weight (DNA)

ISBP - insertion site based polymorphism

IWGSC – international wheat genome sequencing consortium

MDA – multiple displacement amplification

SNP – single nucleotide polymorphism

SSR – simple sequence repeats

WGA – whole genome amplification

10 Supplements

Appendix I

Multiple displacement amplification of the DNA from single flow–sorted plant chromosome

DOI: 10.1111/tpj.13035

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2015) **84,** 838–844 doi: 10.1111/tpj.13035

TECHNICAL ADVANCE

Multiple displacement amplification of the DNA from single flow–sorted plant chromosome

Petr Cápal[1], Nicolas Blavet[1,2], Jan Vrána[1], Marie Kubaláková[1] and Jaroslav Doležel[1,*]

[1]*Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, CZ-78371,Olomouc, Czech Republic, and*

[2]*Palacký University Olomouc, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 27, CZ-78371,Olomouc, Czech Republic*

Received 11 August 2015; revised 13 September 2015; accepted 17 September 2015; published online 24 September 2015.
*For correspondence (e-mail dolezel@ueb.cas.cz).

SUMMARY

A protocol is described for production of micrograms of DNA from single copies of flow-sorted plant chromosomes. Of 183 single copies of wheat chromosome 3B, 118 (64%) were successfully amplified. Sequencing DNA amplification products using an Illumina HiSeq 2000 system to 10× coverage and merging sequences from three separate amplifications resulted in 60% coverage of the chromosome 3B reference, entirely covering 30% of its genes. The merged sequences permitted *de novo* assembly of 19% of chromosome 3B genes, with 10% of genes contained in a single contig, and 39% of genes covered for at least 80% of their length. The chromosome-derived sequences allowed identification of missing genic sequences in the chromosome 3B reference and short sequences similar to 3B in survey sequences of other wheat chromosomes. These observations indicate that single-chromosome sequencing is suitable to identify genic sequences on particular chromosomes, to develop chromosome-specific DNA markers, to verify assignment of DNA sequence contigs to individual pseudomolecules, and to validate whole-genome assemblies. The protocol expands the potential of chromosome genomics, which may now be applied to any plant species from which chromosome samples suitable for flow cytometry can be prepared, and opens new avenues for studies on chromosome structural heterozygosity and haplotype phasing in plants.

Keywords: whole-genome amplification, single-chromosome genomics, next-generation sequencing, flow cytometry, chromosome sorting, sequence assembly, *Triticum aestivum* L. (bread wheat).

INTRODUCTION

Higher plants exhibit a broad range of genome sizes, with a number of agronomically important species having large genomes exceeding 10 Gb (Bennet and Leitch, 1995). The huge genomes consist mostly of various classes of repetitive DNA sequences, and are characterized by enormous sequence redundancy. This hampers genome mapping and sequencing as well as positional gene cloning. The complexity is further increased in polyploids due to the presence of homoeologs. As nuclear genomes are divided into chromosomes, which represent natural subunits, a lossless way to reduce their complexity is to dissect the genomes into separate chromosomes (Doležel *et al.*, 2014).

The utility of chromosome-based approaches in plant genomics has already been well demonstrated. Sequencing mitotic metaphase chromosomes isolated by flow cytometric sorting provided insights into the organization and evolution of the complex genomes of barley, rye and bread wheat (Mayer *et al.*, 2011; Martis *et al.*, 2013; International Wheat Genome Sequencing Consortium, 2014), facilitated development of markers linked to important traits (Shatalina *et al.*, 2013; Tiwari *et al.*, 2014), revealed the molecular organization and origin of supernumerary plant chromosomes (Martis *et al.*, 2012), and was instrumental in improving whole-genome shotgun assemblies (Ruperao *et al.*, 2014).

The application of flow-sorted chromosomes in genomics (chromosome genomics) depends on the ability to discriminate individual chromosomes by flow cytometry,

and this may be hampered by similarities in size and AT/GC content among the chromosomes (Doležel *et al.*, 2012). Thus, in a majority of cases, chromosome-specific DNA has been obtained after sorting chromosomes from cytogenetic stocks such as chromosome deletion, translocation and alien addition lines (Doležel *et al.*, 2014). Recently, Giorgi *et al.* (2013) developed a method termed FISHIS to label repetitive DNA on chromosomes prior to flow cytometric analysis. This approach permits discrimination of chromosomes that have the same or very similar relative DNA content and size. However, the method is suitable for labeling only some DNA sequences, and works only in some species (Doležel *et al.*, 2014).

A radical solution may be to sort single copies of chromosomes and use their DNA as template for whole-genome amplification to obtain chromosome-derived DNA. This approach would provide DNA samples that always originate from only one chromosome, and are thus free of contamination by other chromosomes, even if the chromosome of interest cannot be discriminated from other chromosomes. By producing a series of single-chromosome DNA samples, it is possible to obtain chromosome-derived DNA samples covering the whole genome.

In contrast to flow cytometry, chromosome microdissection allows the researcher to select chromosomes to be isolated and thus avoids problems with chromosome identification and contamination by other chromosomes (Hobza and Vyskot, 2007). A further advantage of chromosome microdissection is that sample preparation is more straightforward compared to flow cytometry. On the other hand, microdissection has much lower throughput than flow sorting. Chromosome microdissection has a number of uses; for example, DNA amplified from single microdissected chromosomes has been used to prepare chromosome painting probes and determine molecular haplotypes in humans (Thalhammer *et al.*, 2004; Ma *et al.*, 2010). Nevertheless, to our knowledge, there has been no report on successful whole-genome amplification of chromosomes microdissected from preparations made from fixed plant tissues. This may indicate problems with DNA accessibility.

Single cell- and single-chromosome genomics is a rapidly developing field, stimulated by the improvement of whole-genome amplification protocols and reduced sequencing costs. It has enabled the study of recombination rates along chromosomes (Jiang *et al.*, 2005) and generation of chromosome-specific painting probes (Kosyakova *et al.*, 2013), and opened new possibilities for prenatal genomic diagnostics (Handyside *et al.*, 2004). Other important applications include the exploration of genomes of uncultivated bacteria (Stepanauskas and Sieracki, 2007), phasing of alleles across whole chromosomes (Fan *et al.*, 2011), and assessment of genomic heterogeneity in tumors (Navin *et al.*, 2011) as well as in healthy tissues (Evrony *et al.*, 2012). However, despite the obvious potential, none of these applications have involved plants.

Here we report on the production of DNA from single copies of plant chromosomes isolated by flow cytometric sorting. We chose chromosome 3B of bread wheat as a model, because its reference sequence is available (Choulet *et al.*, 2014).

RESULTS AND DISCUSSION

A total of 183 single copies of wheat chromosome 3B were amplified using a GenomiPhi V2 kit (GE Healthcare) according to the manufacturer's instructions with modifications. The kit, which relies on multiple displacement amplification (MDA) using polymerase phi29, has been designed for whole-genome amplification, and the manufacturer recommends using 10 ng of template DNA to achieve representative amplification. Our modification of the protocol (see Experimental procedures) permitted amplification of DNA from single copies of mitotic metaphase chromosome 3B of wheat, which corresponds to only 0.002 ng template DNA, i.e. four orders of magnitude less than recommended. Successful DNA amplification was observed in 118 of 183 samples (64% amplification success). The mean yield was 1.4 μg DNA per single chromosome reaction, and the size of DNA fragments ranged from 300 bp to over 20 kb, with a majority of fragments around 10 kb. In contrast, the mean amount of DNA obtained in 18 negative control reactions was only 80 ng (Figure S1).

Each amplification product was evaluated for the presence of 24 unique ISBP (insertion site-based polymorphism) markers that are known to be evenly distributed along the whole length of 3B. On average, 54% of the 24 ISBP markers were present in each chromosome MDA product (range 8–71%). However, in a pool of three single chromosome MDA products, a mean of 92% ISBP markers were present. This success rate is comparable to studies on single sperms (Kirkness *et al.*, 2013) and single human neurons (Evrony *et al.*, 2012). Three samples of amplified chromosome 3B DNA were then randomly selected and sequenced individually using an Illumina HiSeq 2000 system at 10× chromosome coverage. In total, 87, 96 and 85 million cleaned 100 bp paired-end reads were obtained from each sample. Over 74% of them could be mapped onto the chromosome 3B pseudomolecule and scaffolds (Choulet *et al.*, 2014), and more than 60% of them were properly paired, i.e. with both 100 bp ends separated by a 100 bp stretch (Figure 1a). Due to the difficulty in unambiguous determination of the position of reads mapping to more than one locus on the 3B reference, only uniquely mapped reads (60% of all reads) were considered subsequently. Of these uniquely mapped paired-end reads, 83% had both ends mapped properly (Figure 1b).

Mapping sequence reads from three single-chromosome MDA products highlighted the random character of MDA

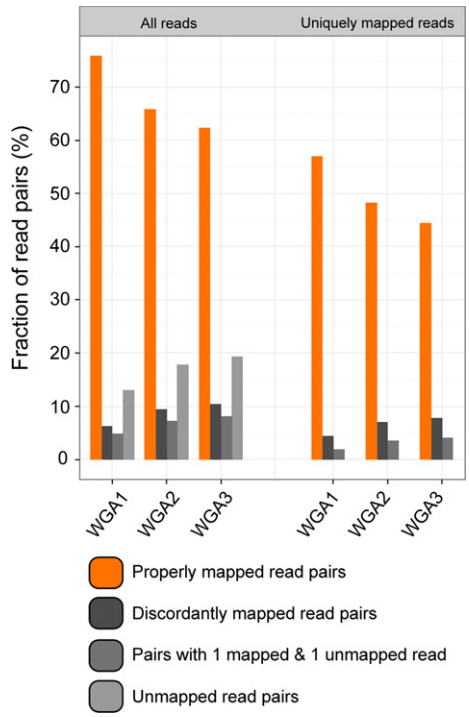

Figure 1. Mapping Illumina pair-end reads obtained after sequencing three independent chromosome 3B MDA products (WGA1–3).
After mapping the reads on the reference sequence of chromosome 3B, four classes of reads were observed: (i) properly mapped paired-end reads, (ii) one read mapped but its mate not mapped, (iii) read pairs discordantly mapped because of incorrect insert size and/or mapped to another contig of the 3B reference sequence, and (iv) reads that were not mapped. Data are shown for all reads and for uniquely mapped reads.

amplification due to the extremely low amount of template DNA, and indicated the need to combine several independent MDA products (Figure 2a). Indeed, some regions of the 3B reference that were highly covered in one amplification reaction were poorly covered or not amplified at all in another reaction (Figure 2a). A close-up view representing one of the scaffolds that constitute the 3B pseudomolecule is shown in Figure 2(b). Each of the three amplifications covered up to 30% of the 3B pseudomolecule, with 10% of its length covered at 10×. Merging sequence reads from three MDA products increased the pseudomolecule coverage to 60%, with more than 30% of its length covered at 10× (Figure 2c). Merging the reads also improved gene coverage, doubling the number of genes covered completely by uniquely mapped reads (Figure 2d). Each of the three sequenced chromosome MDA products completely covered only 7–15% genes of the 7234 genes on the pseudomolecule (Choulet *et al.*, 2014), with 35% of the genes missing. However, after merging reads from the three MDA products, approximately 30% of genes were fully covered and only 16% remained not mapped (Figure 2d). Moreover, 3833 genes, i.e. more than half of the genes on the 3B reference, were covered for 80% of their length (Figure S2). In order to independently assess the quality of the single chromosome-derived sequences, the pooled reads were assembled *de novo* using the SOAP program (Luo *et al.*, 2012), and the resulting sequence contigs were assigned to the 3B reference using BLAST. This *de novo* assembly entirely rebuilt 19% of genes, with 10% of genes contained in a single contig. Moreover, 39% of genes were covered for at least 80% of their length (Figure S2).

As mentioned above, more than 74% of all reads mapped to the 3B reference. The remaining 26% of reads that did not map onto the reference (unmapped reads) may be classified into two groups. The first group of reads consists of approximately 40% of the sequences, and did not have one of their paired reads mapped to the reference (Figure 1). This observation is not surprising, as it has been estimated that up to 20% (approximately 200 Mb) of chromosome 3B was not assembled (Choulet *et al.*, 2014). The remaining 60% of unmapped reads consisted of paired reads. In order to explore their nature, the reads were assembled into 105 798 contigs (Table S1). Most of the contigs contained parts of repetitive DNA elements, with 75% of nucleotides repeat-masked, indicating that most of these sequences originated from non-genic regions. Only 19 121 contigs contained <40% masked nucleotides and were evaluated in more detail. Of them, 15 621 contigs matched wheat chromosome survey sequences (International Wheat Genome Sequencing Consortium, 2014), which were obtained by sequencing DNA amplified from 25 000 copies of flow-sorted chromosome 3B, and 50 000–100 000 copies of the arms of the remaining 20 wheat chromosomes (each sample was equivalent to approximately 25 ng DNA after purification and prior to MDA). Of the 15 621 contigs, 79% matched chromosome 3B, and these sequences may represent loci that are missing in the current reference sequence of 3B (Choulet *et al.*, 2014). The remaining 21% of contigs mapped to survey sequences of the arms of wheat chromosomes other than 3B (Figure S3).

The presence of sequences on other chromosomes of wheat that matched chromosome 3B-derived sequences obtained in the present study may be explained by contamination of the chromosome arm fractions used to produce survey sequences (International Wheat Genome Sequencing Consortium, 2014). The sequences were obtained by sequencing DNA amplified from large numbers of chromosomes, and the fractions may be contaminated by other chromosomes, including 3B and its fragments, during sorting. This is in line with the previously reported 2–16% contamination by other chromosomes (International Wheat Genome Sequencing Consortium, 2014). However, we cannot exclude the possibility that some of the sequences did not originate as 3B contamination, but represent low-copy sequences that were similar to their 3B counterparts.

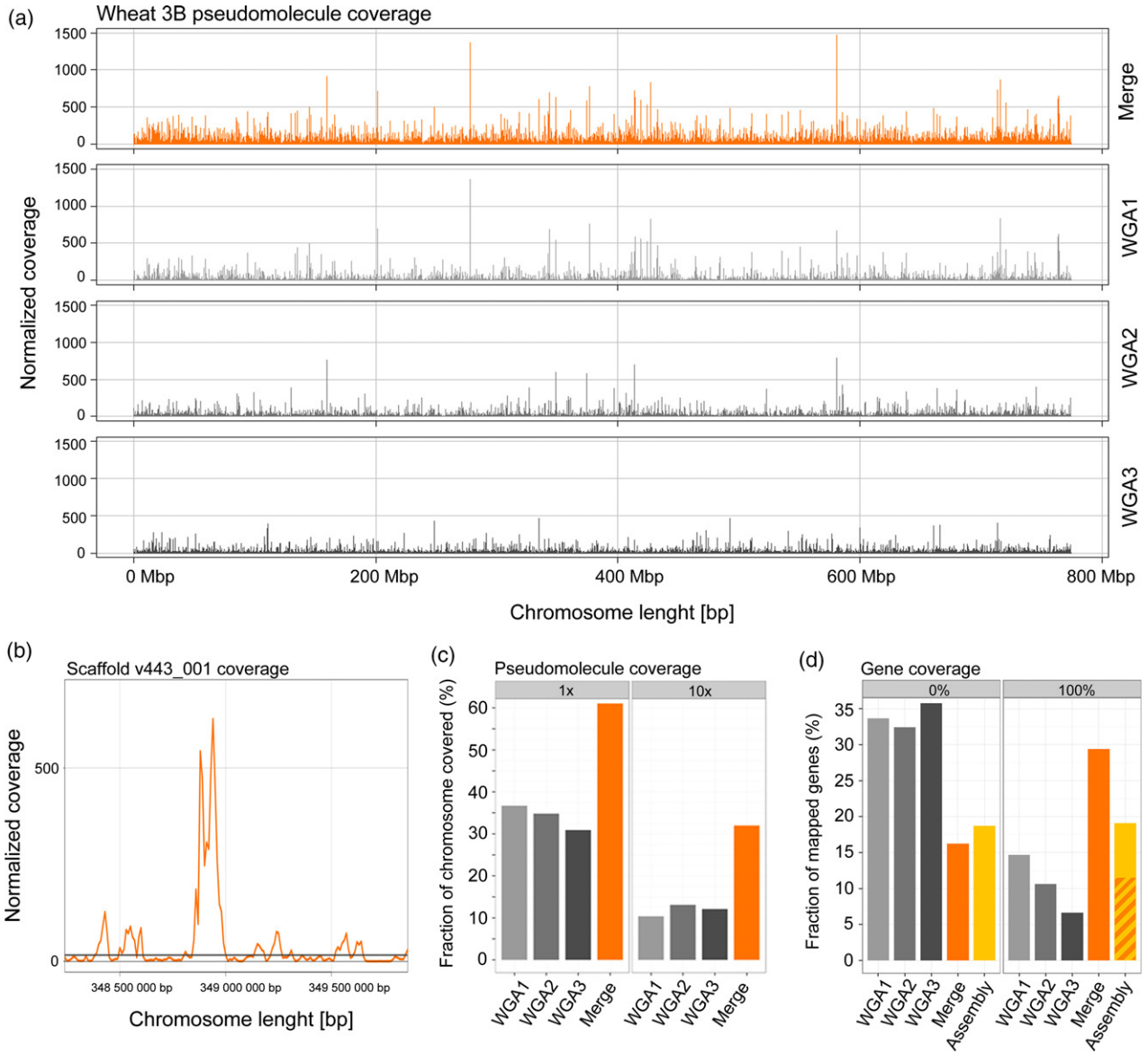

Figure 2. Sequence coverage of the wheat chromosome 3B reference sequence by Illumina reads obtained from three single-chromosome MDA products. The results are given for each MDA product separately and pooled.
(a) Normalized read coverage along chromosome 3B.
(b) Detailed view of the coverage of pseudomolecule 3B spanning the scaffold v443_0001 by sequences from the pool of three chromosome MDA products.
(c) The length of the chromosome 3B pseudomolecule covered at 1× and 10×.
(d) Fraction of genes that are missing or completely covered by read mapping. The last column summarizes the results for the assembly; the dashed part of the bar represents the fraction of genes covered by one *de novo* assembled contig.

To further explore the nature of the sequence contigs that did not map to the chromosome 3B reference, the contigs were compared with barley high-confidence proteins (Mayer *et al.*, 2011). Of the 19 121 unmapped 'non-repetetive' contigs, 172 (0.9%) showed a good match (>80% identity, >80% gene length) with 62 barley genes (Table S2), indicating that a few genes may be missing in the reference chromosome 3B sequence.

The results obtained in this study indicate that single-chromosome sequencing may become an important tool to identify putative contaminating sequences from chromosome shotgun assemblies obtained by sequencing DNA amplified from thousands of chromosomes (International Wheat Genome Sequencing Consortium, 2014). These assemblies are more representative than the single-chromosome ones. However, as a higher number of chromosomes are sorted, it is difficult to avoid contamination by other chromosomes. Ideally, both approaches should be combined to achieve representative chromosome coverage and identify contamination from other chromosomes.

The possibility of amplifying DNA from single chromosomes makes the present method suitable for production of chromosome-derived DNA even in species where individual chromosomes cannot be discriminated and flow-sorted in large numbers as required for standard MDA amplification (Šimková *et al.*, 2008). Thus, the protocol is suitable for any species for which liquid suspensions of intact chromosomes suitable for flow cytometry may be prepared. Although the DNA thus produced is not generally sufficient for *de novo* sequencing without pooling multiple reactions, the protocol provides unique opportunities to identify genic sequences on particular chromosomes, develop chromosome-specific DNA markers, verify the assignment of DNA sequence contigs to individual pseudomolecules, and validate whole-genome assemblies. The new method will be particularly useful in species with large and (allo)polyploid genomes, in which it may simplify sequencing efforts by creating pools of chromosome-specific sequences suitable for assigning shotgun sequence reads to one of the homoeologous chromosomes. The fact that an extremely low amount of template DNA may be used for MDA to produce microgram amounts of chromosome-specific DNA opens new possibilities for the application of chromosome genomics in plants, which include analysis of meiotic recombination in heterogamous plants, analysis of structural chromosome heterozygosity, haplotyping and allele phasing.

Recently, progress has been achieved in reducing DNA amplification bias, either by using an improved amplification method (Zong *et al.*, 2012), or by performing MDA in microfluidic chips (Blainey, 2013). MDA performed in small volumes results in improved genome coverage and reduced bias (Marcy *et al.*, 2007). Based on these results and those of the present work, we suggest coupling of chromosome flow sorting with MDA in microfluidic chips to develop a high-throughput pipeline for preparation of chromosome-derived DNA samples. Flow sorting is an efficient approach to prepare chromosome fractions enriched for particular chromosomes or chromosome groups (Vrána *et al.*, 2015). The samples with a reduced proportion or absence of chromosomes that are not targeted in the particular study, and are free of nuclear and cellular remnants, would be ideal samples for MDA in microfluidic chips, where the chromosomes are distributed one by one into separate chambers and amplified with significantly reduced bias.

EXPERIMENTAL PROCEDURES

Chromosome sorting

Liquid suspensions of intact mitotic metaphase chromosomes were prepared from synchronized root tips of young seedlings of bread wheat (*Triticum aestivum* L.) cv. Chinese Spring as described previously (Vrána *et al.*, 2000). Chromosome samples were stained using DAPI (4′,6–diamidino-2–phenylindole) at a final concentration of 7 μM, and analyzed at rate of 300 chromosomes per second on a FACSAria SORP flow cytometer and sorter (BD Biosciences, http://m.bdbiosciences.com/us/home). In order to sort chromosome 3B, the instrument was set to single-cell one-drop mode, and the sort window was set as shown in Figure S4. GenomiPhi V2 sample buffer (GE Healthcare, http://www3.gehealthcare.com) was mixed with proteinase K (10 mg ml^{-1}; Sigma-Aldrich, https://www.sigmaaldrich.com) at ratio of 10:1 prior to use. Then 3 μl of the mixture was pipetted into a 0.2 ml PCR tube, and chromosomes (one chromosome per tube) were flow-sorted directly into the tubes. One thousand copies of chromosome 3B were sorted into one tube and served as a positive control.

Chromosome DNA amplification

Each sorted chromosome was immediately spun using MiniStar silverline microcentrifuge (VWR, https://vwr.com/) down and incubated in GenomiPhi V2 sample buffer supplemented with proteinase at 50°C overnight in order to digest proteins. The proteinase was then inactivated by heating to 85°C for 15 min in a thermocycler, and deproteinized chromosomes were stored at −20°C until use. Lysis buffer (600 mM KOH, 100 mM dithiothreitol, 10 mM EDTA, pH 8) and neutralization buffer (Tris/HCl, pH 8, 300 mM HCl) were prepared, filter-sterilized through a Millex 0.22 μm syringe-driven filter unit (Merck Millipore, http://www.emdmillipore.com) and UV-irradiated in a Stratalinker 2400 crosslinker (Agilent, http://www.agilent.com) set to 1200 mJ. Then 1.5 μl of lysis buffer was added to each tube containing a deproteinized chromosome, and the samples were incubated at 30°C for 15 min. The reactions were then neutralized using 1.5 μl neutralization buffer, spun down and kept on ice until use. A DNA amplification master mix was prepared by mixing 4 μl sample buffer, 9 μl reaction buffer and 1 μl enzyme (all reagents from the GenomiPhi V2 kit) and added to each tube. The amplification was performed at 30°C for 4 h, and the enzyme was subsequently inactivated at 65°C for 10 min. The negative control was processed exactly the same way, except that no chromosome was sorted into the tube. All pipetting steps were performed in a UV-irradiated biohazard cabinet using Axygen sterile filter pipette tips (Corning, https://www.corning.com).

Evaluation of MDA products and sequencing

The amplification products were evaluated on a 1.5% agarose gel (Figure S1), and the products of successful amplifications were purified using Agencourt Ampure XP beads (Beckman Coulter, www.beckmancoulter.com) according to the manufacturer's instructions. The DNA concentration was estimated using a Nanodrop 1000D spectrophotometer (NanoDrop Products, http://www.nanodrop.com/) and a Modulus fluorimeter (Turner BioSystems/Promega, https://worldwide.promega.com) using a Quant–iT™ PicoGreen® dsDNA assay kit (Life Technologies, https://www.thermofisher.com). The amplification products were then evaluated by PCR using a set of 24 single loci 3B-specific ISBP markers (Paux *et al.*, 2008) as follows: initial denaturation at 95°C for 3 min, 30 cycles of denaturation at 95°C for 30 sec, annealing at 55–67°C for 30 sec, extension at 72°C for 30 sec, and a final extension at 72°C for 5 min, followed by storage at 4°C. ISBP marker primer sequences, annealing temperatures and amplicon sizes are listed in Table S3.

Three MDA products were randomly selected, sheared using Covaris S2 (Covaris, http://covarisinc.com/) using a protocol for a 300 bp insert size, and DNA sequencing libraries were prepared

using a NEBNext Ultra library preparation kit using NEB index primers 1, 10 and 11 (New England Biolabs, http://www.neb.uk.com/). Each library was sequenced on an Illumina (https://www.illumina.com/) HiSeq 2000 system at 10× coverage. Sequences were deposited in the National Center for Biotechnology Information Sequence Read Archive under accession number SRP062037.

Read trimming, mapping and assembly

DNA sequence reads were checked for quality using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), filtered for reads corresponding to bacteria using ERNE version 1.2 (Del Fabbro *et al.*, 2013) and *E. coli* K12 DH10B sequence as reference, and then trimmed for quality using Trimmomatic version 0.30 (Bolger *et al.*, 2014) using the list of adapters provided by the software and the following parameters: ILLUMINACLIP:adapters.fasta:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:80 (Table S4). Cleaned reads were mapped on the assembled pseudomolecule and scaffolds of wheat chromosome 3B (Choulet *et al.*, 2014) using Burrows–Wheeler alignment software (BWA) version 0.75a (Li and Durbin, 2010) with parameters controlling differences between the reads and the reference of *–n* 2 *–k* 0. Statistical analyses were performed using SAMtools version 0.1.19 (Li *et al.*, 2009) and BEDTools version 2.17 (Quinlan and Hall, 2010).

Two approaches were used to determine the number of genes in chromosome 3B MDA products. The first was based on combination of the mapping results with known gene positions to directly determine the gene coverage. The second employed assembly of read sequences into scaffolds using SOAP version 2 (Luo *et al.*, 2012) with a *k*–mer size of 59 (Table S1). Gene coverage was then determined from the results obtained by mapping the assembled contigs using BLAST version 2.2.26 (Altschul *et al.*, 1997) (E–value cut-off 1E–5). Paired reads that did not map onto the 3B reference were assembled using SOAP with similar parameters. Repetitive DNA elements were masked in the resulting contigs using Vmatch version 2.2.0 (http://www.vmatch.de) against the MIPS Poaceae repeat element database version 9.0 http://pgsb.helmholtz-muenchen.de/plant/recat/, and sequences containing more than 40% repeats were filtered out using an in-house Perl script. The remaining sequences were analyzed using BLAST against wheat chromosome survey sequences (International Wheat Genome Sequencing Consortium, 2014) and barley protein sequences (Mayer *et al.*, 2011).

ACCESSION NUMBER

Sequences were deposited in the National Center for Biotechnology Information Sequence Read Archive under accession number SRP062037.

ACKNOWLEDGEMENTS

We thank Vladimír Beneš and his team (Genecore/European Molecular Biology Laboratory, Heidelberg, Germany) for DNA sequencing. We are grateful to Michael Abrouk for useful suggestions, and Jarmila Čihalíková and Zdenka Dubská for chromosome sample preparations and excellent technical assistance. This work was supported by the Czech Science Foundation (award P501/12/G090), and the National Program of Sustainability I (award LO1204). N.B. was supported by the Operational Program Education for Competitiveness – European Social Fund (award CZ.1.07/2.3.00/20.0165). Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the program 'Projects of Large Infrastructure for Research, Development, and Innovations' (award LM2010005), is greatly appreciated.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Evaluation of MDA products.

Figure S2. Sequence coverage of genes located on chromosome 3B.

Figure S3. Assignment of unmapped contigs to other wheat chromosomes.

Figure S4. Flow cytometric analysis and sorting.

Table S1. Assembly statistics of the pooled library.

Table S2. Assignment of unmapped reads to barley genes.

Table S3. ISBP markers list.

Table S4. Trimming statistics for separate sequencing libraries.

REFERENCES

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.

Bennet, M.D. and Leitch, I.J. (1995) Nuclear DNA amounts in angiosperms. *Ann. Bot.* **76**, 113–176.

Blainey, P.C. (2013) The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol. Rev.* **37**, 407–427.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Choulet, F., Alberti, A., Theil, S. *et al.* (2014) Structural and functional partitioning of bread wheat chromosome 3B. *Science*, **345**, 1249721.

Del Fabbro, C., Scalabrin, S., Morgante, M. and Giorgi, F.M. (2013) An extensive evaluation of read trimming effects on Illumina NGS data analysis. *PLoS ONE*, **8**, e85024.

Doležel, J., Vrána, J., Šafář, J., Bartoš, J., Kubaláková, M. and Šimková, H. (2012) Chromosomes in the flow to simplify genome analysis. *Funct. Integr. Genomics* **12**, 397–416.

Doležel, J., Vrána, J., Cápal, P., Kubaláková, M., Burešová, V. and Šimková, H. (2014) Advances in plant chromosome genomics. *Biotechnol. Adv.* **32**, 122–136.

Evrony, G.D., Cai, X., Lee, E. *et al.* (2012) Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell*, **151**, 483–496.

Fan, H.C., Wang, J., Potanina, A. and Quake, S.R. (2011) Whole-genome molecular haplotyping of single cells. *Nat. Biotechnol.* **29**, 51–57.

Giorgi, D., Farina, A., Grosso, V., Gennara, A., Ceoloni, C. and Lucretti, S. (2013) FISHIS: fluorescence *in situ* hybridization in suspension and chromosome flow sorting made easy. *PLoS ONE*, **8**, e57994.

Handyside, A.H., Robinson, M.D., Simpson, R.J., Omar, M.B., Shaw, M.A., Grudzinskas, J.G. and Rutherford, A. (2004) Isothermal whole genome amplification from single and small numbers of cells: a new era for preimplantation genetic diagnosis of inherited disease. *Mol. Hum. Reprod.* **10**, 767–772.

Hobza, R. and Vyskot, B. (2007) Laser microdissection-based analysis of plant sex chromosomes. *Methods Cell Biol.* **82**, 433–453.

International Wheat Genome Sequencing Consortium (2014) A chromosome-based draft sequence of the hexaploid wheat (*Triticum aestivum*) genome. *Science*, **345**, 1251788.

Jiang, Z., Zhang, X., Deka, R. and Jin, L. (2005) Genome amplification of single sperm using multiple displacement amplification. *Nucleic Acids Res.* **33**, e91.

Kirkness, E.F., Grindberg, R.V., Yee-Greenbaum, J., Marshall, C.R., Scherer, S.W., Lasken, R.S. and Venter, J.C. (2013) Sequencing of isolated sperm cells for direct haplotyping of a human genome. *Genome Res.* **23**, 826–832.

Kosyakova, N., Hamid, A.B., Chaveerach, A., Pinthong, K., Siripiyasing, P., Supiwong, W., Romanenko, S., Trifonov, V. and Fan, X. (2013) Genera-

tion of multicolor banding probes for chromosomes of different species. *Mol Cytogenetics*, **6**, 6.

Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennel, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Luo, R., Liu, B., Xie, Y. *et al.* (2012) SOAPdenovo2: an empirically improved memory-efficient short read *de novo* assembly. *GigaScience*, **1**, 18.

Ma, L., Xiao, Y., Huang, H., Qingwei Wang, Q., Rao, W., Feng, Y., Kui Zhang, K. and Son, Q. (2010) Direct determination of molecular haplotypes by chromosome microdissection. *Nat. Methods*, **7**, 299–302.

Marcy, Y., Ishoey, T., Lasken, R., Stockwell, T.B., Walenz, B.P., Halpern, A.L., Beeson, K.Y., Goldberg, S.M.D. and Quake, S.R. (2007) Nanoliter reactors improve multiple displacement amplification of genomes from single cells. *PLoS Genet.* **3**, 1702–1708.

Martis, M.M., Klemme, S., Banaei-Moghaddam, A.M. *et al.* (2012) Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proc. Natl Acad. Sci. USA*, **109**, 13343–13346.

Martis, M.M., Zhou, R., Haseneyer, G. *et al.* (2013) Reticulate evolution of the rye genome. *Plant Cell*, **25**, 3685–3698.

Mayer, K.F.X., Martis, M., Hedley, P.E. *et al.* (2011) Unlocking the barley genome by chromosomal and comparative genomics. *Plant Cell*, **23**, 1249–1263.

Navin, N., Kendall, J., Troge, J. *et al.* (2011) Tumour evolution inferred by single-cell sequencing. *Nature*, **427**, 90–94.

Paux, E., Sourdille, P., Salse, J. *et al.* (2008) A physical map of the 1-gigabase bread wheat chromosome 3B. *Science*, **322**, 101–104.

Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

Ruperao, P., Chan, C.K.K., Azam, S. *et al.* (2014) A chromosomal genomics approach to assess and validate the desi and kabuli draft chickpea genome assemblies. *Plant Biotechnol. J.* **12**, 778–786.

Shatalina, M., Wicker, T., Buchmann, J.P., Oberhaensli, S., Simkova, H., Doležel, J. and Keller, B. (2013) Genotype-specific SNP map based on whole chromosome 3B sequence information from wheat cultivars Arina and Forno. *Plant Biotechnol. J.* **11**, 23–32.

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 Genom.* **9**, 294.

Stepanauskas, R. and Sieracki, M.E. (2007) Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proc. Natl Acad. Sci. USA*, **104**, 9052–9057.

Thalhammer, S., Langer, S., Speicher, M.R., Wolfgang, M., Heckl, W.M. and Geigl, J.B. (2004) Generation of chromosome painting probes from single chromosomes by laser microdissection and linker-adaptor PCR. *Chrom. Res.* **12**, 337–343.

Tiwari, V.K., Wang, S., Sehgal, S. *et al.* (2014) SNP discovery for mapping alien introgressions in wheat. *BMC Genom.* **15**, 273.

Vrána, J., Kubaláková, M., Šimková, H., Číhalíková, J., Lysák, M.A. and Doležel, J. (2000) Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). *Genetics*, **156**, 2033–2041.

Vrána, J., Kubaláková, M., Číhalíková, J., Valárik, M. and Doležel, J. (2015) Preparation of sub-genomic fractions enriched for particular chromosomes in polyploid wheat. *Biol. Plant.* **59**, 445–455.

Zong, C., Lu, S., Chapman, A.R. and Xie, X.S. (2012) Genome-wide detection of single nucleotide and copy-number variations of a single human cell. *Science*, **338**, 1622–1626.

Appendix II

The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat

DOI: 10.1186/s13007-016-0124-8

Cápal *et al. Plant Methods (2016) 12:24*
DOI 10.1186/s13007-016-0124-8

Plant Methods

METHODOLOGY

Open Access

The utility of flow sorting to identify chromosomes carrying a single copy transgene in wheat

Petr Cápal[1], Takashi R. Endo[1,2], Jan Vrána[1], Marie Kubaláková[1], Miroslava Karafiátová[1], Eva Komínková[1], Isabel Mora-Ramírez[3], Winfriede Weschke[3] and Jaroslav Doležel[1*]

Abstract

Background: Identification of transgene insertion sites in plant genomes has practical implications for crop breeding and is a stepping stone to analyze transgene function. However, single copy sequences are not always easy to localize in large plant genomes by standard approaches.

Results: We employed flow cytometric chromosome sorting to determine chromosomal location of barley sucrose transporter construct in three transgenic lines of common wheat. Flow-sorted chromosomes were used as template for PCR and fluorescence in situ hybridization to identify chromosomes with transgenes. The chromosomes carrying the transgenes were then confirmed by PCR using DNA amplified from single flow-sorted chromosomes as template.

Conclusions: Insertion sites of the transgene were unambiguously localized to chromosomes 4A, 7A and 5D in three wheat transgenic lines. The procedure presented in this study is applicable for localization of any single-copy sequence not only in wheat, but in any plant species where suspension of intact mitotic chromosomes suitable for flow cytometric sorting can be prepared.

Keywords: Transgene localization, Flow cytometric sorting, Single chromosome amplification, *Triticum aestivum*, *Hordeum vulgare*, HvSUT1

Background

During the past 30 years, many cultivars of agricultural crops beneficial to humankind have been developed by means of genetic engineering, including plants resistant to herbicides, pests or viruses, bearing fruits with prolonged shelf life and products more suited for industrial processing [for review see 1]. Wheat ranks 5th in the commodities produced worldwide and is the second most-produced food crop occupying more than 50 % of the world crop area (http://faostat3.fao.org/). In the light of climate change and world population growth, future challenges for the increase of crop production have constantly been discussed. However, FAO statistics show that the wheat production is reaching a plateau and is severely affected by climate change. This is a consequence of a slowdown in wheat yield increase, accounting for only 0.5 % per year in the last decade [2].

Breeding improved cultivars with increased tolerance to adverse climatic conditions and with increased yield and quality could be facilitated by genetic engineering and introduction of beneficial genes from other organisms. The insertion site of a transgene is of great importance for the transgene function [3, 4] which is also influenced by its position on the chromosome, including the flanking DNA sequences [5]. However, transgene localization is not easy by routine approaches, like fluorescence in situ hybridization (FISH), or Southern blotting. A prevalent method for detection of transgenes in animals and plants is FISH, which has its pros and cons [6]. In barley and common wheat, FISH enables cytological localization of cDNAs, as short as 1.5 kb, on a

*Correspondence: dolezel@ueb.cas.cz
[1] Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, 78371 Olomouc, Czech Republic
Full list of author information is available at the end of the article

chromosome or chromosomes that had already been known to carry the cDNAs [7, 8]. Although some authors succeeded in localizing transgenes on plant chromosomes using FISH [9–11], this approach has not become a routine application.

Weichert et al. [12] obtained transgenic lines (HOSUT) of hexaploid wheat carrying barley (*Hordeum vulgare*) sucrose transporter HvSUT1 (SUT) gene that is overexpressed under the control of the endosperm-specific Hordein B1 promoter (HO). The HOSUT lines were found to increase grain yield significantly as compared to control non-transformed plants [13]. However, the genomic location of the transgene in these lines was not known. In the present work we employed a novel approach for unambiguous identification of chromosomes carrying the transgene in three HOSUT lines. The protocol takes the advantage of the availability of a procedure for flow cytometric chromosome sorting in wheat and the fact that flow-sorted chromosomes are suitable as templates for PCR and FISH [14]. Moreover, a protocol has been developed recently for representative DNA amplification from single copies of chromosomes [15]. By combining these approaches we could assign the transgene to particular chromosomes in three HOSUT lines of wheat.

Results and discussion

The experimental workflow is shown on Fig. 1. As the first step, we prepared liquid suspensions of intact mitotic chromosomes from all five lines of wheat (see "Methods" section) and analyzed them by flow cytometry. Monovariate flow karyotypes (histograms of relative fluorescence intensity) were obtained after the analysis of DAPI-stained chromosomes, and bivariate flow karyotypes obtained after the analysis of DAPI-stained chromosomes with FITC-labelled GAA microsatellites. We observed differences between flow karyotypes of the HOSUT lines and the model hexaploid wheat cultivar Chinese Spring. The alterations concerned the profiles of major composite peaks on monovariate flow karyotypes (Additional file 1: Figure S1) and the distribution of chromosome populations on bivariate flow karyotypes (Fig. 2). This observation reflected the differences in karyotypes (chromosome polymorphism) between the cultivar Certo, used to produce the HOSUT lines (data not shown), and Chinese Spring. On the other hand, flow karyotypes of the three HOSUT lines were indistinguishable from each other.

In order to identify chromosomes carrying the transgenes, we first used the approach described by Vrána et al. [16]. Fractions of 200 chromosomes were sorted from different regions of monovariate flow karyotypes as shown on Additional file 1: Figure S1, and DNA of the sorted chromosomes was used as template for PCR. This analysis identified one region (sort gate) in each line as representing chromosomes bearing a transgene. As each region (sort gate) on a monovariate flow karyotype may represent more than one chromosome type, in the next step we sorted chromosomes from regions delineated on bivariate flow karyotypes (Fig. 2). The sort gates were designed to include chromosome populations corresponding to the positive sort gates on monovariate flow karyotypes. From these regions, and also from nearby regions, chromosomes were sorted into PCR tubes (100 chromosomes per tube) and immediately afterwards also onto microscopic slides (ca. 1000 chromosomes per slide). The results obtained by PCR with primers amplifying HvSUT-RT sequence (Fig. 3) and identification of chromosomes from sort gates for each transgenic line by FISH with probes targeting Afa-repeat family and GAA-microsatellites (Fig. 4) are summarized in Table 1.

FISH analysis showed that more than 90 % of chromosomes flow-sorted from the region defined by the green rectangle consisted of one type of chromosome in each of the HOSUT lines. This fact together with the results of PCR suggested that the transgene was located on chromosome 7A in HOSUT 12/44, on chromosome 5D in HOSUT 20/6 and on chromosome 4A in HOSUT 24/31. In the former two lines, the critical type of chromosome was not found among the chromosomes flow-sorted from the region defined by red rectangles. However, chromosome 4A was found to represent 12.39 % of chromosomes flow-sorted from the red region in HOSUT 24/31. This was probably due to the similarities in size and the amount of GAA-FITC fluorescence of chromosomes 4A and 7A. Due to this similarity, mixture of the two chromosomes 4A and 7A was also observed in the chromosome fraction sorted from the green region in HOSUT 12/44.

To confirm chromosomal locations of the transgene and avoid ambiguous results due to possible contamination of flow-sorted fraction by other chromosomes, PCR was done on DNA amplified from single flow-sorted chromosomes. As each time only one copy of chromosome is sorted, the DNA cannot be contaminated by other chromosomes. Five single chromosomes were sorted from the green sort regions of the HOSUT lines and their DNA was separately amplified using multiple displacement amplification (MDA). Out of the five sorted chromosomes, whole genome amplification was successful with three chromosomes in HOSUT 12/44, two chromosomes in HOSUT 20/6 and four chromosomes in HOSUT 24/31. The successful amplification was defined by the production of measurable amount of DNA after MDA and by the presence of at least one marker for the

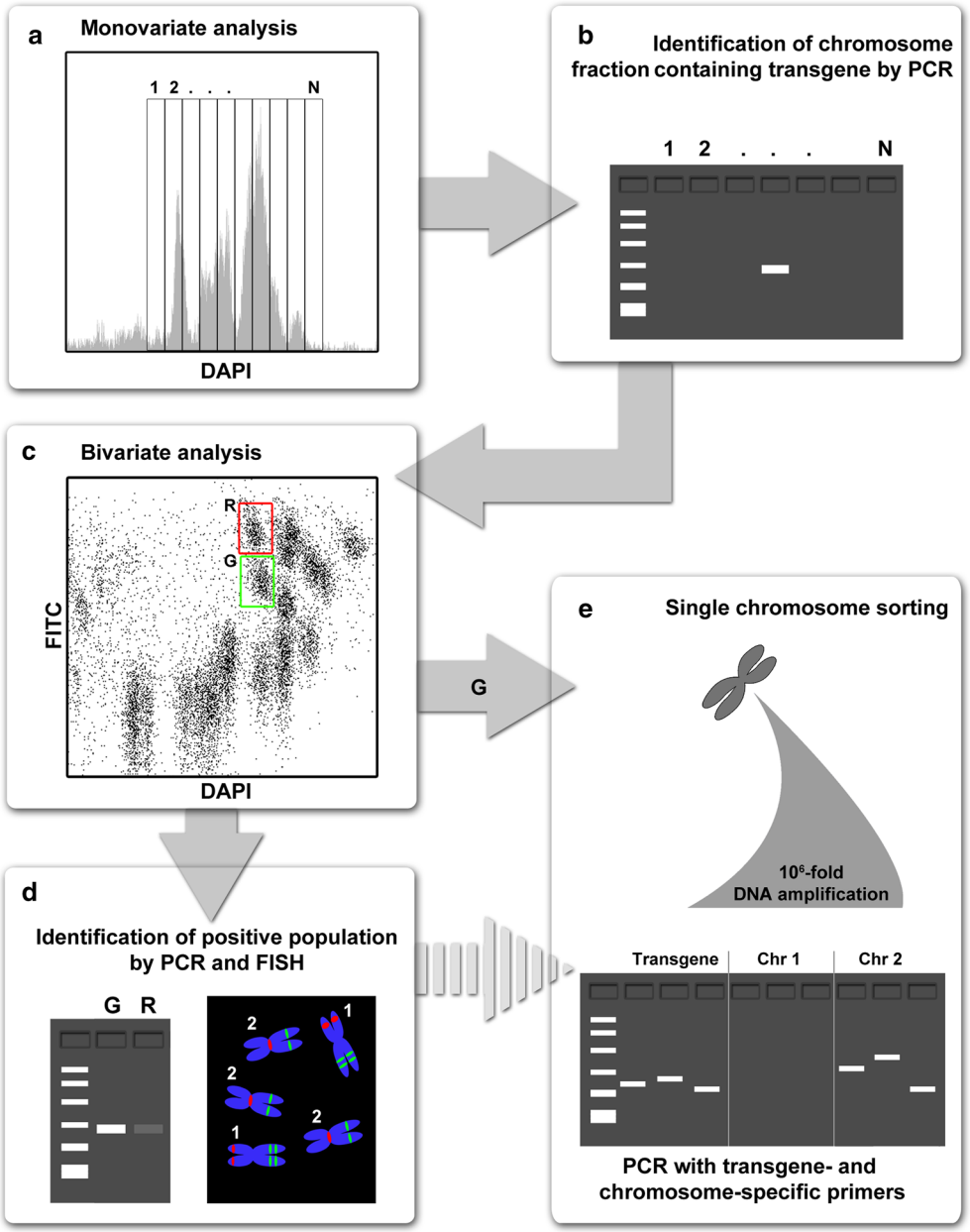

Fig. 1 Experimental workflow. **a** Monovariate flow karyotype is dissected into small regions. From each region, 200 chromosomes are sorted. **b** Flow-sorted chromosomes are used as template for PCR with transgene-specific marker. **c** The region representing chromosome with the transgene identified on monovariate flow karyotype is dissected by sorting chromosomes from sort regions on bivariate flow karyotype. **d** From the sort gates on the bivariate flow karyotype, 100 chromosomes are sorted for PCR with transgene-specific marker. Moreover, 1000 chromosomes are sorted immediately afterwards onto the microscopic slides to identify flow-sorted chromosomes by FISH. **e** From the sort gate most enriched for the transgene-bearing chromosome (region *G* in this example), single chromosomes are sorted individually into PCR tubes. DNA of single-flow sorted chromosomes is amplified and resulting DNA is used as template for PCR to identify the presence of multiple transgene- and chromosome-specific sequences. This step unambiguously confirms the chromosome with integrated transgene

transgene and one marker for the wheat chromosome. The reason for occasional failure to amplify DNA from single chromosomes, which was observed previously [15] is not clear. One explanation is that a droplet with sorted chromosome lands on side wall of PCR tube and the chromosome is excluded from the MDA reaction. The amount of chromosomal DNA in successfully amplified samples ranged from 0.3 to 1.7 µg DNA.

Chromosome specificity of sequence tagged site (STS) markers used in this work to identify individual

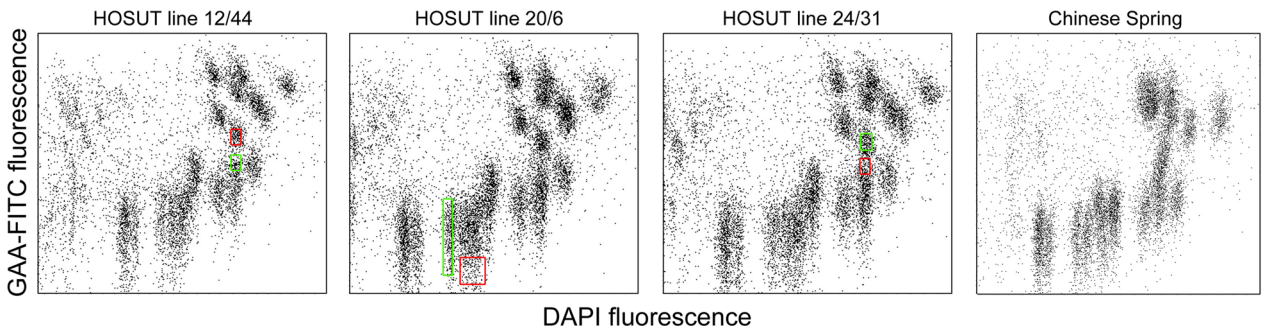

Fig. 2 Bivariate flow karyotypes of three transgenic HOSUT lines of wheat obtained after the analysis of chromosomes with FITC-labelled $(GAA)_n$ microsatellites and stained by DAPI. The position of *red* and *green* regions used to sort particular chromosomes is indicated. The *green* sort gate was found to represent chromosomes carrying transgene. Chromosomes were flow-sorted also from the neighboring population delineated by *red* gate and were used as a control. Although the transgene-bearing chromosome should not be included in this region, the sorted population could potentially be contaminated with transgene-bearing chromosomes due to similarity in chromosome size and DNA content

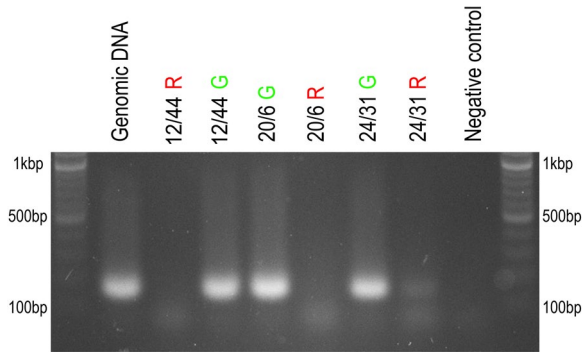

Fig. 3 Agarose gel electrophoresis of PCR products obtained with primers for the transgene and DNA of chromosomes flow-sorted from three HOSUT lines using the *green* and *red* sort regions as shown in Fig. 1. The amplicon of HvSUT-RT (169 bp) was obtained with chromosomes sorted from the *green* sort region in all three HOSUT lines. When chromosomes were sorted from the *red* sort regions, no PCR amplification occurred for HOSUT 12/44 and HOSUT 20/6. However, a weak band was observed for HOSUT 24/31. Genomic DNA of the transgenic lines served as positive control

chromosomes was first tested using the euploid and corresponding nulli-tetrasomic lines of Chinese Spring (Additional file 2: Figure S2). The results confirmed that the markers were suitable for unambiguous identification of wheat chromosomes 1A, 4A, 5D and 7A. PCR analysis using both transgene- and chromosome-specific markers clearly confirmed chromosome location of transgenes as determined in the first part of this study. In case of HOSUT 24/31, where the location of the transgene was ambiguous, all four transgene markers were detected in DNA amplified from single chromosomes sorted from the green region (Fig. 5), and all four 4A-specific markers were also amplified in the same amplicons. None of the four 7A-specific markers was found in the same amplicons.

Conclusions

Coupling PCR and FISH mapping using flow-sorted mitotic chromosomes as templates narrowed down the list of candidate chromosomes harboring the transgene to one or two chromosomes. PCR on DNA amplified

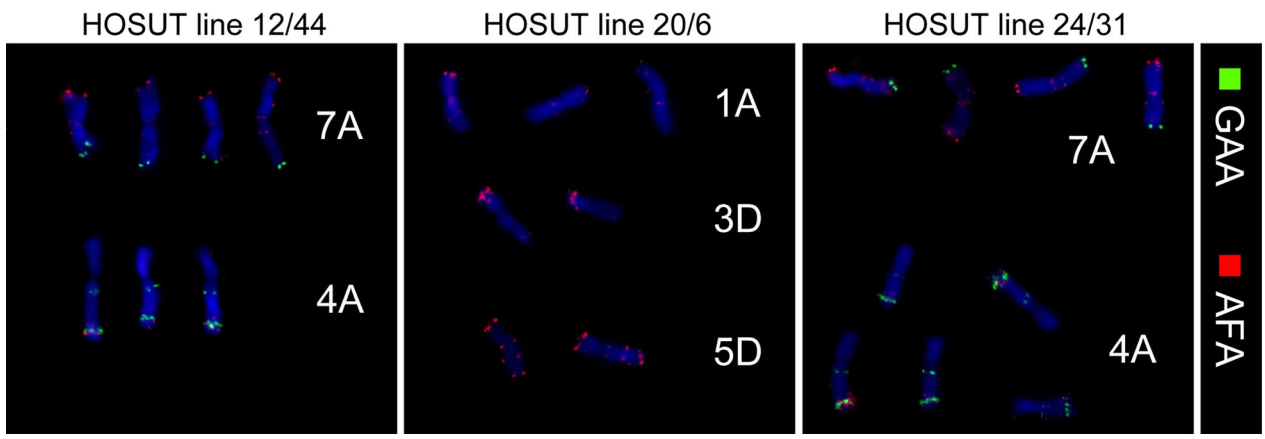

Fig. 4 Representative images of chromosomes flow-sorted from three HOSUT lines using the *green* and *red* sort regions on bivariate flow karyotypes as shown in Fig. 2. FISH was done using probes for Afa-family (*red* signals) and GAA microsatellites (*green* signals). Chromosomes were counterstained with DAPI (*blue*)

Table 1 PCR and FISH analysis of chromosomes sorted from each of the sort gates in three HOSUT lines

Transgenic line	Sort gate[a]	PCR result	Chromosomes identified FISH[c]
HOSUT 12/44	Red	Negative	4A (92.65 %)
	Green	Positive	7A (90.90 %)
			4A (4.45 %)
HOSUT 20/6	Red	Negative	1A (63.75 %)
			3D (36.25 %)
	Green	Positive	5D (94.66 %)
			1A (5.33 %)
HOSUT 24/31	Red	Semi-positive[b]	7A (83.19 %)
			4A (12.39 %)
			2A (4.42 %)
	Green	Positive	4A (97.30 %)

[a] Sort gates delineated with green and red rectangles in Fig. 2

[b] A faint band was visible after agarose gel electrophoresis of PCR product

[c] More than 1000 chromosomes were examined in each sorted fraction in each line

from single flow-sorted chromosomes then unambiguously identified the chromosomes with the integrated transgene. If chromosome-specific PCR-based markers are available, mapping on single copy chromosomes could be an ultimate approach to assign single copy DNA sequences, including transgenes, to particular chromosomes. Moreover, the sequence assembly of amplicons from the chromosome could allow detecting the position of transgene insertion, if enough sequence information on the chromosome is available. However the main purpose of this work was to assign a transgene to particular chromosomes. The approach presented here is currently applicable to more than 25 plant species, which include important cereals and legumes [14] where liquid suspensions of mitotic chromosomes suitable for flow cytometric sorting can be prepared.

Methods

Plant material

We used German winter wheat cultivar Certo (*Triticum aestivum* L., 2n = 6x = 42, genome formula AABBDD) and its three transgenic lines, HOSUT 12/44, HOSUT 20/6 and HOSUT 24/31. The transgenic lines contain a single copy of the HvSUT1-cDNA (1894 bp) fused to the barley HorB1 promoter (550 bp) and the barley HorB1 terminator (1663 bp) [12]. We also used euploid and nullisomic–tetrasomic (Nt1A1B, Nt7A7B, Nt5D5B) lines of hexaploid wheat cultivar Chinese Spring (obtained from NBRP-wheat) to confirm the specificity of PCR markers to particular wheat chromosomes.

Flow cytometric chromosome sorting

Cell cycle synchronization and metaphase accumulation of root tip meristem cells was performed as described

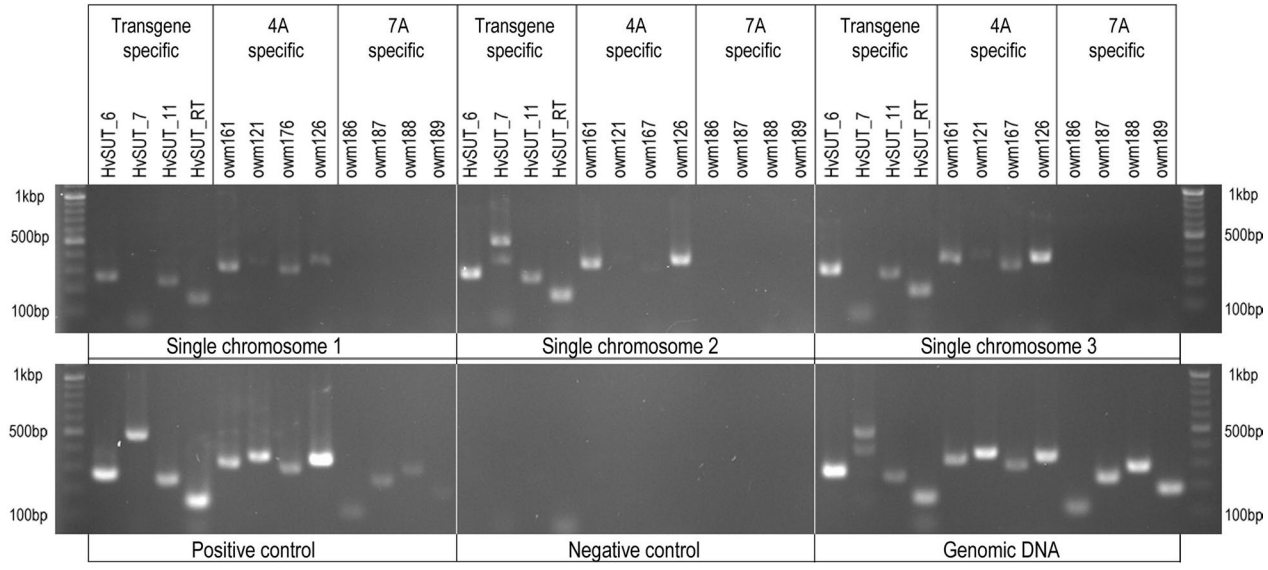

Fig. 5 Agarose gel electrophoresis of PCR products obtained using DNA produced by multiple displacement amplification of three single chromosomes flow-sorted from the sort region representing chromosome 4A in HOSUT 24/31 line. PCR with primers for the four transgenes resulted in products of expected length. The same was true for the chromosome 4A-specific STS markers. Note that none of the chromosome 7A-specific markers was detected in the samples of single chromosome DNA. PCR with genomic DNA of HOSUT 24/31 as template detected both 4A and 7A chromosome-specific markers. PCR with the positive control (represented by 1000 chromosomes sorted from *green* sorting region and amplified) showed slight PCR bands of chromosome 7A, which reflects a minor contamination of the sorted chromosome 4A by chromosome 7A

previously [17], except for the formaldehyde fixation, which was shortened to 15 min. Isolated chromosomes were labelled by FISHIS (fluorescence in situ hybridization in suspension) using FITC-labeled GAA probe following the protocol of Giorgi et al. [18]. Flow cytometric analysis and sorting was done on BD FACSAria II high speed flow sorter equipped with 390 nm laser for DAPI excitation and 488 nm laser for FITC excitation. Sort gates were initially drawn on monovariate flow karyotypes of DAPI fluorescence (not shown) and subsequently on bivariate flow karyotypes of DAPI fluorescence versus GAA-FITC fluorescence as shown in Fig. 1.

Fluorescence in situ hybridization (FISH)

For microscopic observations, 1000 chromosomes were sorted onto microscope slides from each of the sort regions. The slides were left to air-dry in the dark overnight. Then the preparations were used for FISH following the protocol of Kubaláková et al. [19] using a Cy5-labeled probe targeting Afa-family repeats, the chromosomes were already labeled by a GAA microsatellite probe during the FISHIS procedure.

PCR

PCR was done using primers specific for the HOSUT transgene and for markers specific for candidate wheat chromosomes (Table 2). Of the four HOSUT primers, three were designed in the HvSUT1 region (accession no. AJ272309) and one in the HorB1 terminator region (accession no. FN643080). Wheat chromosome-specific markers were designed by Primer3 based on the chromosome sequences from the International Wheat Genome Sequencing Consortium (IWGSC), while preventing the primers from amplifying the sequence from the homoeologous chromosomes. PCR conditions were set as follows: initial denaturation 95 °C for 3 min, 35 cycles of 30 s denaturation at 95 °C, annealing at 58–62 °C (see Table 2 for T_a of the primer pairs) for 30 s and extension at 72 °C for 30 s, followed by final extension at 72 °C for 5 min. The amount of template DNA was 5 ng for each reaction. PCR

Table 2 List of PCR primers for the HOSUT transgene construct and PCR primers for wheat STS markers on chromosomes 1A, 4A, 7A and 5D

Name	Target	Forward primer sequence	Reverse primer sequence	Amplicon size (bp)	Annealing temperature (°C)
HvSUT_6	HOSUT1	AGCGGCGGCGGTCACTGACTG	CCAAAGGACGACACCCCAGCC	265	62
HvSUT_7	HorB1 terminator	ATTAATTCCTCCCCGACCCTGC	CAATGGAGACGGCGCGTGCAA	471	62
HvSUT_11	HOSUT1	GGCGGAACCCGCCGTGCAG	CCTGCGTCTTCCCCATCTGGAAGTA	241	62
HvSUT_RT	HOSUT1	CGGGCGGTCGCAGCTCGCGTCTATT	CATACAGTGACTCTGACCGGCACACA	169	62
Owm121	Chromosome 4A	ATTGCCGTCGCGAACTAGA	CGGGACGAGCTTGACGAT	351	60
Owm126	Chromosome 4A	CCAGTCAGAAATTATTATGAACCTATC	CGCTGTCTCGAGATTGGAGT	342	60
Owm161	Chromosome 4A	TTTTCAAGCAGGTTTTGTGC	TCACTTCTCTTCTTTGCGTTCA	324	60
Owm167	Chromosome 4A	TTTTCTTGGTCAGTATAACCTGTTTTT	TGAGCAGAGAAAAATTTCCAAG	285	60
Owm174	Chromosome 1A	GCATCCTAGTTTCTCTCTCAAGT	AACAAGATCACGAGCGAATTG	157	58
Owm175	Chromosome 1A	AAACCCCTGATACTCATGCG	GTTTCTTGTCATTCATGTCACTTGT	530	58
Owm176	Chromosome 1A	TTCCTGTCTGACTCCGCG	AACCACAACCGTCAACCG	104	58
Owm177	Chromosome 1A	GTAGTCTGCTCCCGAGGAAT	GTCTCTAACCATACATCCATGAAGT	192	58
Owm178	Chromosome 1A	CAACTTCTTCACATCCCGGAA	ATTTGGCCCTATGAGATATAATTACG	306	58
Owm179	Chromosome 1A	ACACTGTGATACCTCTAGATGTATG	CACATTGCCTATAAATTCTAAAAGGTC	425	58
Owm180	Chromosome 5D	CGGACGAGCAGCAGTACC	GCAGATCGGCATAAATTGAATGT	292	58
Owm181	Chromosome 5D	GGAGGTGTTCTAGGTGTACTTACT	AGAGCAATGTCAGAAGTCATCG	240	58
Owm182	Chromosome 5D	TCTCCACCTGCAGAGTCG	CATCAGGCCACAGTGTCAAT	119	58
Owm183	Chromosome 5D	TGTCCACACATTTCCCGTATG	AGTGGTGGATGTGGTTGCT	196	58
Owm184	Chromosome 5D	AGCATGCTCCCAAAGACTATTAC	GTTATGATGGTGGTAGCAATTTGA	400	58
Owm185	Chromosome 5D	GTGAACCTATATGACATCTTACCGG	GGGGCAGTTGTCAAGTATTGC	421	58
Owm186	Chromosome 7A	CTCTCTGTGGCCAATAGTGC	TCTATACCTCAACCCTACATCCA	112	58
Owm187	Chromosome 7A	GGCCACGAATTCCACAAGTA	CTATCGATCAACCAACCATCCA	229	58
Owm188	Chromosome 7A	GTACGAGTGCAGACAGTGTG	ACAATTAATTATACGCCCAGTTAAGC	282	58
Owm189	Chromosome 7A	CGTGCTTTCTTCTTCCTCCG	GCAGGTTAGTTTCTTGTGGTTG	185	58
Owm190	Chromosome 7A	CGCATGGACATTGTTCTAGTCA	GCACTTAGGCACGCTTGAG	517	58
Owm191	Chromosome 7A	CGACGACATTAGGAATATGGGAT	TGCGTGTGGGTGTGCTTA	402	58

products were run on 1.5 % agarose gels. PCR using 100 sorted chromosomes as template was conducted after a few freeze–thaw cycles to disintegrate the chromosomes and the initial denaturation step was prolonged for 7 min.

Whole genome amplification of single chromosomes

DNA amplification of single chromosomes was performed by MDA using a GE Healthcare GenomiPhi V2 kit (GE Healthcare Life Sciences, Little Chalfont, UK) according to Cápal et al. [15]. Five individual chromosomes were flow-sorted into five 0.2 ml PCR tubes from green sort gates from each HOSUT line and their DNA amplified. The amplified DNA was evaluated on 1.5 % agarose gel, purified using magnetic beads (AMPure XP system, Beckman Coulter, Inc., Brea, CA, USA) and the concentration was measured by a spectrophotometer (NanoDrop, Thermo Fisher Scientific Inc., Waltham, MA, USA).

Additional files

Additional file 1: Figure S1. Flow karyotypes (histograms of fluorescence intensity) obtained after the analysis of DAPI-stained chromosomes isolated from three transgenic lines and cv. Chinese Spring of common wheat. Flow karyotypes of the transgenic lines are indistinguishable from each other, and slightly differ in profiles of the major composite peaks from those of Chinese Spring.

Additional file 2: Figure S2. Verification of marker specificity. PCR with a full set of chromosome-specific wheat STS markers was performed using genomic DNA of cv. Chinese Spring and corresponding nullitetrasomic lines for chromosomes 1A, 4A, 5D and 7A. The markers, which resulted in amplification products only in Chinese Spring and not in the nullitetrasomic lines, were used in this study.

Authors' contributions

The study was conceived and designed by TE, PC and JD, experiments were performed by PC, TE, M Ka, M Ku, JV, IMR and EK, manuscript was written by PC, TE, JD and WW. All authors read and approved the final manuscript.

Author details

[1] Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, 78371 Olomouc, Czech Republic. [2] Faculty of Agriculture, Ryukoku University, 1-5 Yokotani, Seta Oe-cho, Otsu, Shiga 520-2194, Japan. [3] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstrasse 3, 06466 Stadt Seeland, Germany.

Acknowledgements

We thank Zdeňka Dubská for technical assistance with flow cytometric chromosome sorting. This work was supported by the National Program of Sustainability (Award No. LO 2014) and the Czech Science Foundation (Award No. P501-12-G090).

Competing interests

The authors declare that they have no competing interests.

Received: 20 February 2016 Accepted: 19 April 2016
Published online: 25 April 2016

References

1. Ahmad P, Ashraf M, Younis M, Hu X, Kumar A, Akram NA, Al-Qurainy F. Role of transgenic plants in agriculture and biopharming. Biotechnol Adv. 2012;30:524–40.
2. Fischer RA. The importance of grain or kernel number in wheat: a reply to Sinclair and Jamieson. Field Crops Res. 2008;105:15–21.
3. Svitashev S, Anaiev E, Pawlowski WP, Somers DA. Association of transgene integration sites with chromosome rearrangements in hexaploid oat. Theor Appl Genet. 2000;100:872–80.
4. Kumar S, Fladung M. Gene stability in transgenic aspen (Populus). II. Molecular characterization of variable expression of transgene in wild and hybrid aspen. Planta. 2001;213:731–40.
5. Kooter JM, Matzke A, Meyer P. Listening to the silent genes: transgene silencing, gene regulation and pathogen control. Trends Plant Sci. 1999;4:340–7.
6. Svitashev KS, Somers DA. Characterization of transgene loci in plants using FISH: a picture is worth a thousand words. Plant Cell Tiss Org Cult. 2002;69:205–14.
7. Danilova TV, Friebe B, Gill BS. Development of a wheat single gene FISH map for analyzing homoeologous relationship and chromosomal rearrangements within the Triticeae. Theor Appl Genet. 2014;127(3):715–30.
8. Karafiátová M, Bartoš J, Kopecký D, Ma L, Sato K, Houben A, Doležel J. Mapping nonrecombining regions in barley using multicolor FISH. Chromosome Res. 2013;21:739–51.
9. Pedersen C, Zimny J, Becker D, Jähne-Gärtner A, Lörz H. Localization of introduced genes on the chromosomes of transgenic barley, wheat and triticale by fluorescence in situ hybridization. Theor Appl Genet. 1997;94:749–57.
10. Dong J, Pushpa K, Cervera M, Hall TC. The use of FISH in chromosomal localization of transgenes in rice. Methods Cell Sci. 2001;23:105–13.
11. Khrustaleva LI, Kik C. Localization of single-copy T-DNA insertion in transgenic shallots (*Allium cepa*) by using ultra-sensitive FISH with tyramide signal amplification. Plant J. 2001;25:699–707.
12. Weichert N, Saalbach I, Weichert H, Kohl S, Erban A, Kopka J, Hause B, Varshney A, Sreenivasulu N, Strickert M, Kumlehn J, Weschke W, Weber H. Increasing sucrose uptake capacity of wheat grains stimulates storage protein synthesis. Plant Physiol. 2010;152:698–710.
13. Saalbach I, Mora-Ramírez I, Weichert N, Andersch F, Guild G, Wieser H, Koehler P, Stangoulis J, Kumlehn J, Weschke W, Weber H. Increased grain yield and micronutrient concentration in transgenic winter wheat by ectopic expression of a barley sucrose transporter. J Cereal Sci. 2014;60:75–81.
14. Doležel J, Vrána J, Cápal P, Kubaláková M, Burešová V, Šimková H. Advances in plant chromosome genomics. Biotechnol Adv. 2014;32:122–36.
15. Cápal P, Blavet N, Vrána J, Kubaláková M, Doležel J. Multiple displacement amplification of the DNA from single flow-sorted plant chromosome. Plant J. 2015. doi:10.1111/tpj.13035.
16. Vrána J, Kubaláková M, Číhalíková J, Valárik M, Doležel J. Preparation of sub-genomic fractions enriched for particular chromosomes in polyploid wheat. Biol Plant. 2015;59:445–55.
17. Vrána J, Kubaláková M, Šimková H, Číhalíková J, Lysák MA, Doležel J. Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). Genetics. 2000;156:2033–41.
18. Giorgi D, Farina A, Grosso V, Gennara A, Ceoloni C, Lucretti S. FISHIS: fluorescence in situ hybridization in suspension and chromosome flow sorting made easy. PLoS One. 2013;8:e57994.
19. Kubaláková M, Kovářová P, Suchánková P, Číhalíková J, Bartoš J, Lucretti S, Watanabe N, Kianian SF, Doležel J. Chromosome sorting in tetraploid wheat and its potential for genome analysis. Genetics. 2005;170:823–9.

Appendix III

Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley

DOI: 10.1371/journal.pone.0137677

PLOS | ONE

RESEARCH ARTICLE

Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley

Steven Dreissig[1], Jörg Fuchs[1], Petr Cápal[2], Nicola Kettles[3], Ed Byrne[3], Andreas Houben[1]*

1 Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstrasse 3, 06466 Stadt Seeland, Germany, **2** Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, Olomouc CZ-78371, Czech Republic, **3** KWS-UK Ltd, 56 Church Street, Thriplow, Hertfordshire, SG8 7RE, United Kingdom

* houben@ipk-gatersleben.de

Citation: Dreissig S, Fuchs J, Cápal P, Kettles N, Byrne E, Houben A (2015) Measuring Meiotic Crossovers via Multi-Locus Genotyping of Single Pollen Grains in Barley. PLoS ONE 10(9): e0137677. doi:10.1371/journal.pone.0137677

Editor: Michael Lichten, National Cancer Institute, UNITED STATES

Received: April 20, 2015

Accepted: August 19, 2015

Published: September 10, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the IPK, P. Cápal was supported by the Ministry of Education, Youth and Sports of the Czech Republic (grant LO1204 from the National Program of Sustainability I). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

The detection of meiotic crossovers in crop plants currently relies on scoring DNA markers in a segregating population or cytological visualization. We investigated the feasibility of using flow-sorted haploid nuclei, Phi29 DNA polymerase-based whole-genome-amplification (WGA) and multi-locus KASP-genotyping to measure meiotic crossovers in individual barley pollen grains. To demonstrate the proof of concept, we used 24 gene-based physically mapped single nucleotide polymorphisms to genotype the WGA products of 50 single pollen nuclei. The number of crossovers per chromosome, recombination frequencies along chromosome 3H and segregation distortion were analysed and compared to a doubled haploid (DH) population of the same genotype. The number of crossovers and chromosome wide recombination frequencies show that this approach is able to produce results that resemble those obtained from other methods in a biologically meaningful way. Only the segregation distortion was found to be lower in the pollen population than in DH plants.

Introduction

Meiotic recombination is the primary mechanism of generating novel allelic combinations and introducing genetic diversity. In barley (*Hordeum vulgare* L.), as well as in many other crops, recombination frequencies are elevated in distal gene-rich chromosomal regions. Nevertheless, 24.7% of the total barley gene content is located in low recombining regions [1] representing an untapped resource which is unavailable for plant breeding [2]. Hence, strategies to modulate the recombination frequency along chromosomes are needed. The ability to induce an increase in meiotic recombination is so far limited to the model species *Arabidopsis thaliana* via a mutation of the FANCM helicase [3]. In barley, Higgins et al. [4] demonstrated a shift of meiotic crossovers towards interstitial and proximal regions at higher temperatures during meiosis.

There are different ways of monitoring meiotic crossovers in plants. They can be identified using molecular markers in a segregating population [5], or alternatively the frequency and distribution of crossovers can be visualized by cytological means like analysis of pairing configurations [6] or immunolabeling of proteins involved in meiotic recombination such as the barley

MutL homologue (HvMLH3) [7]. One limitation of using microscopy-based methods is that the sites of recombination events can only be resolved at the chromosomal level. Another limitation is an uncertainty about perfect agreement between protein localization and crossover. Other tools for efficient determination of recombination events such as the tetrad analysis based on the quartet (qrt) mutation are currently only available for *Arabidopsis thaliana* [8].

In human and livestock genetics, recombination analysis using meiotic gametophytes was developed more than 20 years ago for the high-resolution mapping of recombination sites. In plants, the idea of analysing pollen grains has already brought forward a number of studies. Petersen et al. [9] extracted DNA from single barley and rye pollen grains for PCR amplification and subsequent sequencing of high and single copy genes. Chen et al. [10] developed a method using pollen grains of several plant species for molecular analysis utilizing randomly amplified polymorphic DNA and simple sequence repeat markers. The introduction of whole-genome-amplification (WGA) methods, such as primer extension pre-amplification, enabled Aziz and Sauve [11] to further increase the amount of information gained from single pollen grains. However, other WGA methods based on isothermal amplification via the Phi29 DNA polymerase hold the potential to enable the analysis of hundreds to thousands of markers in a single cell [12].

In the current study, we describe a strategy to perform a parallel analysis of individual haploid nuclei derived from pollen grains by utilizing fluorescence activated cell sorting (FACS) coupled with Phi29 DNA polymerase-based whole-genome-amplification (WGA) and multi-locus KASP genotyping. The meiotic crossover measurements were compared to data obtained by different methods in comparable genetic environments.

Materials and Methods

Plant material and isolation of pollen nuclei and genomic DNA

Pollen grains were collected from an F_1 plant of the barley cultivars Morex x Barke (*Hordeum vulgare* L.). Mature anthers of 20 flowers were collected in a 1.5 ml Eppendorf tube using forceps. Afterwards, 300 μl of ddH_2O were added and the suspension was vortexed for approximately 30 sec. The suspension was shaken at 1500 rpm for 10 min at room temperature to release all pollen grains. Afterwards, the pollen suspension was centrifuged for 5 min at 13,000 rpm and all empty anthers were manually removed using forceps. After centrifugation for 5 min at 13,000 rpm the supernatant was removed and the pollen pellet was resuspended in 100 μl Galbraith buffer (45 mM $MgCl_2$, 30 mM sodium citrate, 20 mM MOPS, 0.1% Triton-X100, pH to 7.0; [13]) and transferred into a 2.0 ml Eppendorf tube containing two metallic beads of 6 mm in diameter (Intec GmbH) as described in [14]. The pollen-bead mixture was centrifuged for 5 min at 13,000 rpm prior to homogenization at 30 Hz for 40 seconds using a MM 400 ball mill (Retsch). After homogenization, another 500 μl Galbraith buffer were added and the suspension was filtered through a 30 μm filter (Sysmex-Partec). For the purpose of providing genomic DNA for marker testing, genomic DNA was extracted from leaf tissue using the DNeasy Plant Mini kit (Qiagen) and measured using nanodrop (Peqlab).

FACS-based purification of single haploid nuclei and whole-genome-amplification

The nuclei suspension was stained with 4',6-diamidino-2-phenylindole (DAPI; 1.5 μg/ml) and single 1C nuclei were sorted using a BD FACSAria IIu (BD Biosciences) flow-sorter into individual wells of a 384-microwell plate containing 2 μl lysis solution (0.5 μl lysis buffer composed of 400 mM KOH, 100 mM DTT, 10 mM EDTA; [15], 0.5 μl ddH_2O, 1 μl sample buffer

(Genomiphi V2, GE Healthcare)) for whole-genome-amplification. Note, in contrast to the manufacturer's protocol the sample buffer containing random primers for whole-genome-amplification was added to lysis solution. Whole-genome-amplification was carried out using the Genomiphi V2 kit (GE Healthcare) according to the manufacturer's protocol with the following modifications: Nuclei lysis and DNA denaturation was conducted by incubation at 65°C for 3 min in 2 μl lysis solution. The lysis solution was neutralized by adding 0.5 μl neutralisation buffer (666 mM Tris-HCl, 250 mM HCl; [15]). Afterwards, a master mix composed of 3.5 μl sample buffer, 4.5 μl reaction buffer and 0.5 μl enzyme mix (all Genomiphi V2, GE Healthcare) per reaction was added and samples were incubated at 30°C for 8 hours followed by inactivation of the enzyme at 65°C for 10 minutes. Subsequently, each sample was diluted with 500 μl ddH_2O. The DNA concentration of the WGA products of single pollen nuclei was measured by fluorometric quantitation (Qubit, Life Technologies).

Each sample was subjected to a PCR using primers for the *Ty3/gypsy*-like retroelement *cereba* in order to validate the successful sorting of pollen nuclei into the microwells. The reaction volume of the *cereba* amplification was 10 μl containing 5 μl WGA product, 1x PCR buffer (Qiagen), 0.2 mM dNTPs (Bioline), 1x Q-solution (Qiagen), 0.6 μM of each primer and 0.02 units Taq DNA polymerase (Qiagen). The following thermal cycling conditions were used: DNA polymerase activation: 3 min at 95°C; denaturation: 30 sec at 95°C; annealing: 30 sec at 60°C; extension: 30 sec at 72°C; final extension: 10 min at 72°C; 30 cycles in total. The *cereba*-positive samples were further analysed with 8 chromosome 3H-specific primers to quantify the efficiency of the whole-genome-amplification (S1 Table). These primer pairs were targeting single copy sequences to test if the WGA was able to amplify unique sequences. The reaction volume of the 3H-specific amplification was 10 μl containing 5 μl WGA product, 1x PCR buffer (Qiagen), 0.2 mM dNTPs (Bioline), 1x Q-solution, 0.3 μM of each primer and 0.02 units Taq DNA polymerase (Qiagen). The following thermal cycling conditions were used: DNA polymerase activation: 3 min at 95°C; denaturation: 30 sec at 95°C; annealing: 30 sec at 65°C, reduced by 1°C for 9 cycles; extension: 30 sec at 72°C; 25 cycles at final annealing temperature.

KASP-genotyping

A set of 24 chromosome 3H-specific single nucleotide polymorphisms (SNPs) [16] based on the current barley genome sequence assembly [17] was chosen and converted into KASP markers (LGC Genomics, S2 Table). Thermal cycling conditions were adopted from Mirouze et al. [18] and end-point signals were read out on a BioRad iQ5 cycler at 30°C. Genomic DNA from cv. Morex, cv. Barke and Morex x Barke F_1 plants were genotyped in parallel. Additionally, replicate single pollen nuclei of the cultivar Morex were subjected to whole-genome-amplification and subsequent genotyping to act as a positive control against amplification errors. This was done in order to test the possibility of false allele calling due to inaccurate WGA or contamination. Allele calling was done manually by plotting relative fluorescence values of FAM and HEX against each other. Heterozygous signals were discarded as genotyping errors since we expect haploid nuclei to only provide homozygous signals.

Analysis of segregation distortion loci and crossovers

To test for segregation distortion we conducted a χ^2-test assuming an expected segregation ratio of 1:1 for each marker. Segregation distortion loci (SDL) were identified by significant deviation from the expected ratio of 1:1 ($P < 0.05$). Crossovers were detected by visualizing our marker data using flapjack [19] by identifying allele calls for which there was a switch from allele A (Morex) to allele B (Barke) and vice versa. The physical position of each marker on barley chromosome 3H was derived from the barley genome sequence assembly [17] thus enabling

us to count crossovers without constructing a linkage map. The recombination frequency between two adjacent marker pairs was measured as the proportion of crossovers to no-crossovers. Marker pairs within a sample involving missing data points were omitted from the analysis. The number of crossovers was normalized according to the following calculation:

$$\mathbf{CO\ ratio} = \frac{\boldsymbol{n}(\boldsymbol{CO})}{\boldsymbol{n}(\boldsymbol{no}-\boldsymbol{CO}) + \boldsymbol{n}(\boldsymbol{CO})} \quad (1)$$

where n(CO) is the number of crossovers and n(no-CO) is the number of no-crossovers.

Similarly, the number of no-crossovers was normalized according to the following calculation:

$$\mathbf{no-CO\ ratio} = \frac{\boldsymbol{n}(\boldsymbol{no}-\boldsymbol{CO})}{\boldsymbol{n}(\boldsymbol{no}-\boldsymbol{CO}) + \boldsymbol{n}(\boldsymbol{CO})} \quad (2)$$

The ratio of (1) to (2) enabled us to calculate recombination frequencies for each marker pair in a comparable manner. To compare our data to the Morex x Barke DH population data, we used raw genotyping-by-sequencing data (wheat.pw.usda.gov, S3 Table) and extracted the physical map position for each marker [17]. Segregation distortion and crossover analysis of the Morex x Barke DH population was performed as described above. A two-tailed unpaired Student's t-test was performed to compare the average number of crossovers in both populations and the distribution of the number of crossovers was compared using a χ^2-test assuming the Morex x Barke DH crossover data as expected values.

Results and Discussion

Whole-genome-amplification of single haploid nuclei

In order to develop a strategy for the high-throughput analysis of meiotic recombination events in barley pollen, we first needed to prove the successful extraction, amplification and genotyping of single pollen DNA. To overcome problems associated with the rigid cell wall of pollen grains described by Chen et al. [10], we selected a novel approach to isolate haploid nuclei suitable for flow-sorting.

Isolated haploid nuclei from pollen grains were individually sorted via a FACS-based approach into individual wells of a 384-microwell plate (Fig 1). After nuclei lysis and DNA denaturation, whole-genome-amplification was performed using Phi29 DNA polymerase. Each reaction yielded 1 to 3 μg of DNA consisting of barley-specific products and likely also unspecific products as expected from whole-genome-amplification via Phi29 DNA polymerase [20]. The products of 192 single-nuclei whole genome amplification (WGA) reactions were analysed by PCR for the presence of the barley high copy *Ty3/gypsy*-like retroelement *cereba* to confirm the successful sorting of nuclei into the individual microwells. From a total of 192 samples, 168 contained PCR amplified barley DNA giving an accuracy of our FACS approach of 87.5%. To preselect samples for further genotyping, we used PCR to amplify eight 3H-specific single copy sequences located across both arms of chromosome 3H. Out of the 168 single-nuclei amplifications positively tested for *cereba*, we selected 50 samples which showed successful amplifications of at least 3 single copy sequences from both arms for further genotyping (S1 Fig). To measure meiotic crossovers on chromosome 3H, we selected 24 KASP markers. The suitability of the 24 selected chromosome 3H-specific KASP markers was confirmed using genomic DNA isolated from leaves of both genotypes. Next, the same set of markers was used to genotype WGA-amplified DNA derived from individual haploid nuclei. The selected 50 nuclei samples revealed an average marker call rate of 70%, indicating that the majority of samples were effectively amplified (Fig 2A). Considering preselection via PCR, we found a weak

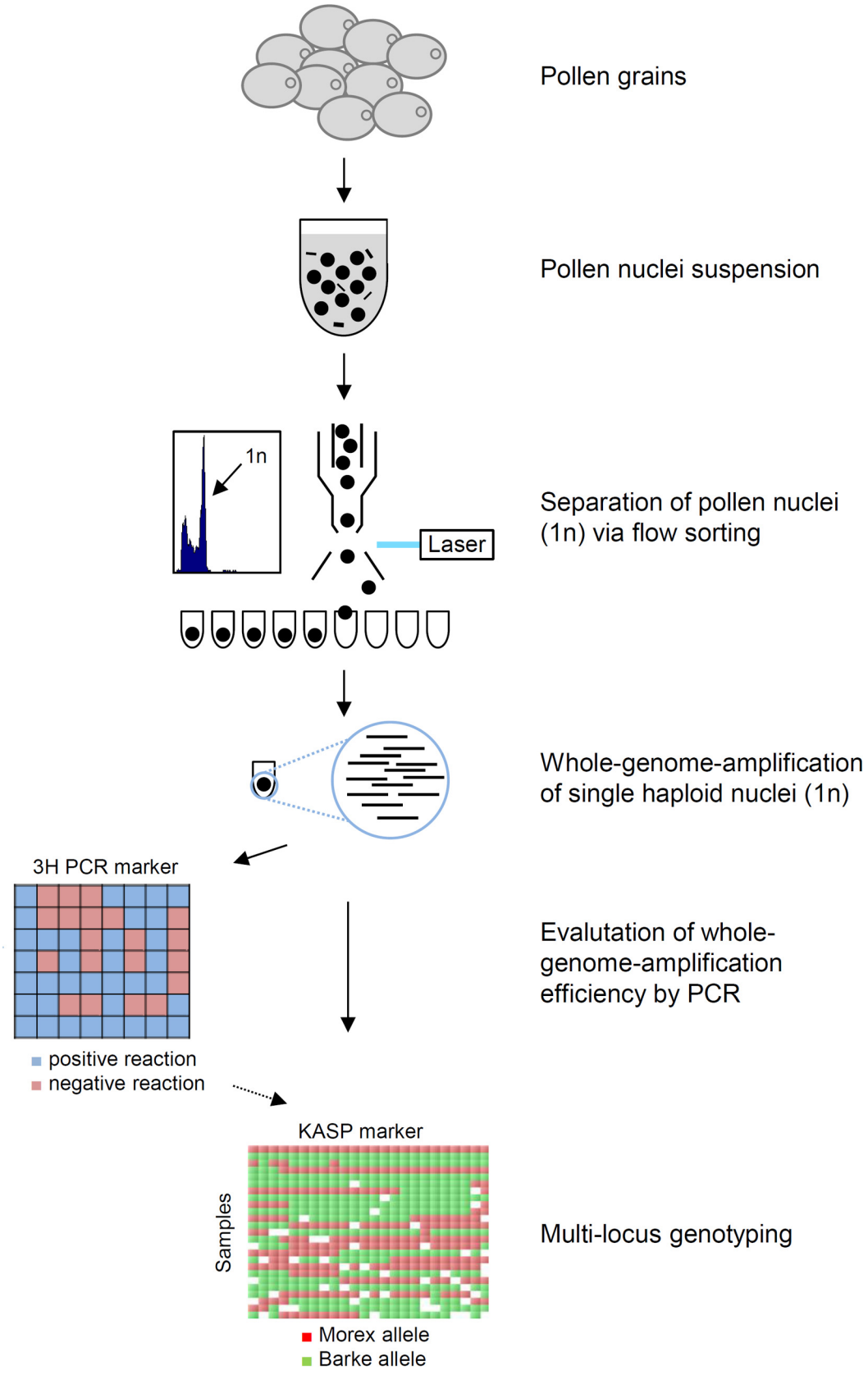
Pollen grains
Pollen nuclei suspension
1n
Laser
Separation of pollen nuclei (1n) via flow sorting
Whole-genome-amplification of single haploid nuclei (1n)
3H PCR marker
positive reaction
negative reaction
Evalutation of whole-genome-amplification efficiency by PCR
KASP marker
Samples
Morex allele
Barke allele
Multi-locus genotyping

Fig 1. Scheme of experimental workflow developed in the current study. Haploid nuclei were extracted from pollen grains, separated via flow-sorting and individually subjected to whole-genome-amplification (WGA). High quality samples, evaluated by PCR with chromosome 3H-specific primers, were genotyped using 25 KASP markers to measure crossover frequency and distribution.

doi:10.1371/journal.pone.0137677.g001

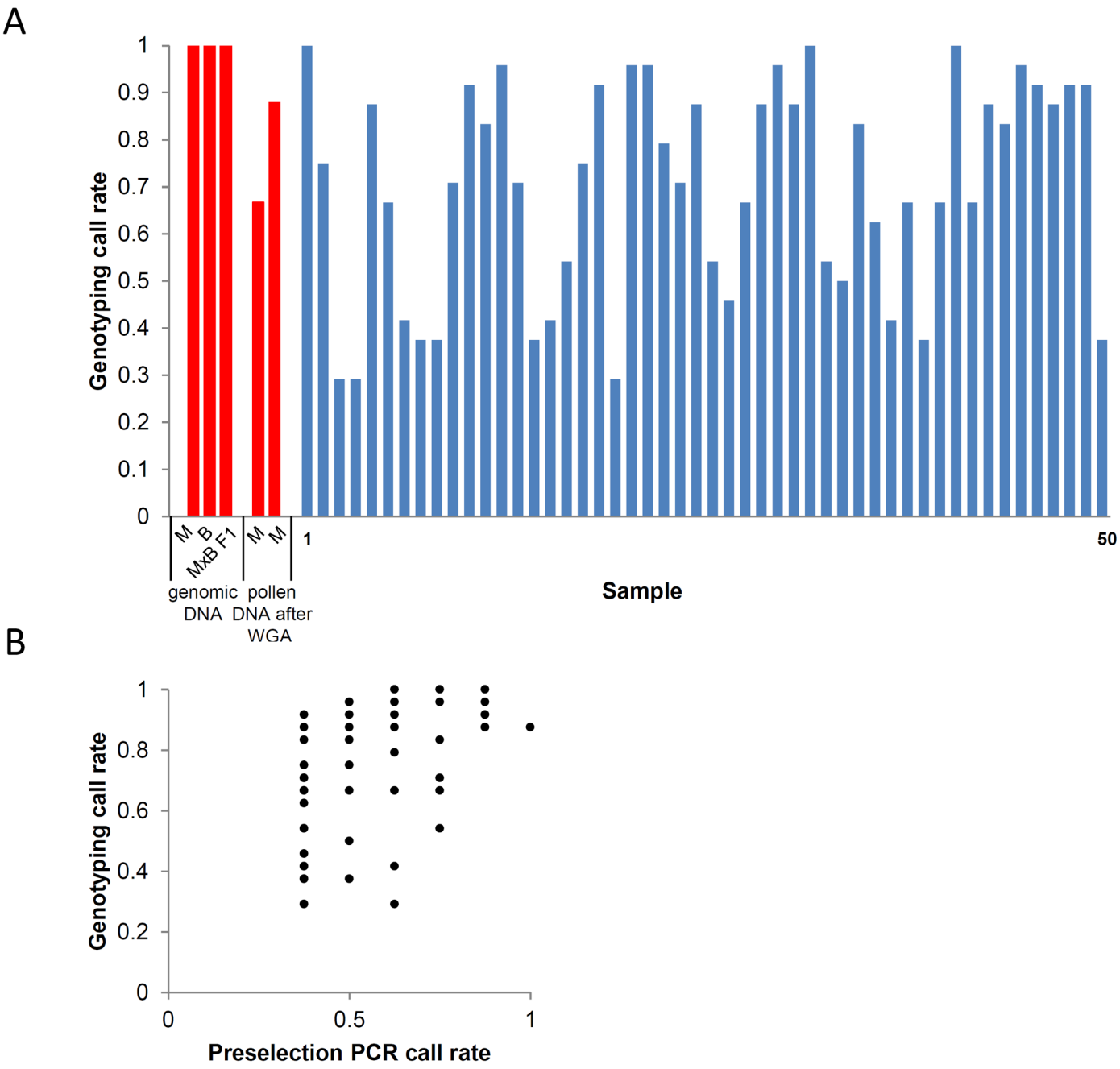

Fig 2. KASP genotyping performance. (A) Genotyping call rate of the positive controls across 24 KASP markers is indicated in red. Genomic DNA from Morex, Barke and Morex x Barke F1 plants was used. Two individual nuclei derived from Morex pollen grains were used to test for false allele calling due to whole-genome-amplification. The selected 50 Morex x Barke pollen nuclei samples are indicated in blue showing an average genotyping call rate of 71%. **(B)** Correlation between KASP genotyping call rate and preselection PCR call rate. $r = 0.51$, $r^2 = 0.26$.

doi:10.1371/journal.pone.0137677.g002

but yet positive correlation between final genotyping call rate and preselection PCR call rate ($r = 0.51$, $r^2 = 0.26$, Fig 2B) which indicates the advantage of preselecting samples after WGA before conducting downstream analyses. Only three samples resulted in products with less than 35% of the markers, probably due to inefficient amplification of the Phi29 polymerase-based whole-genome-amplification system, as described previously [12, 20]. 98% of the positive KASP reactions (1226 out of 1250) showed clear homozygous signals in agreement to the positive-control. 24 heterozygous calls (1.92%) were observed. These were randomly distributed across all samples and markers and therefore unlikely to be caused by an erroneous sorting of two nuclei into one microwell, so they were discarded as genotyping errors. Furthermore, no false allele calling, e.g. Barke allele instead of Morex, due to WGA errors was found in the positive controls using haploid nuclei of Morex. We conclude that multi-locus KASP-based genotyping on WGA-amplified DNA derived from single haploid nuclei is feasible.

Monitoring meiotic recombination by genotyping single pollen grains

Meiotic crossovers along chromosome 3H were measured with a mean inter-marker distance of 14.35 and 12.32 mega base-pairs (Mbp) for the short and long arm of chromosome 3H, respectively [17]. The average number of crossovers for chromosome 3H in our pollen population was 1.92, while the corresponding number for a corresponding DH population was 1.84. These two values are not significantly different ($P = 0.72$). Looking at the distribution of the total number of crossovers, we did not find a significantly different pattern for both pollen and DH population (Fig 3A) indicating the reliability of our approach (χ^2-test, $P = 0.99$). Although one individual nucleus showing 6 crossovers on chromosome 3H was found to have a low genotyping call rate of 0.38, there was no significant correlation between the number of crossovers and genotyping call rate ($r = -0.16$, $r^2 = 0.03$, Fig 3B). We further determined recombination

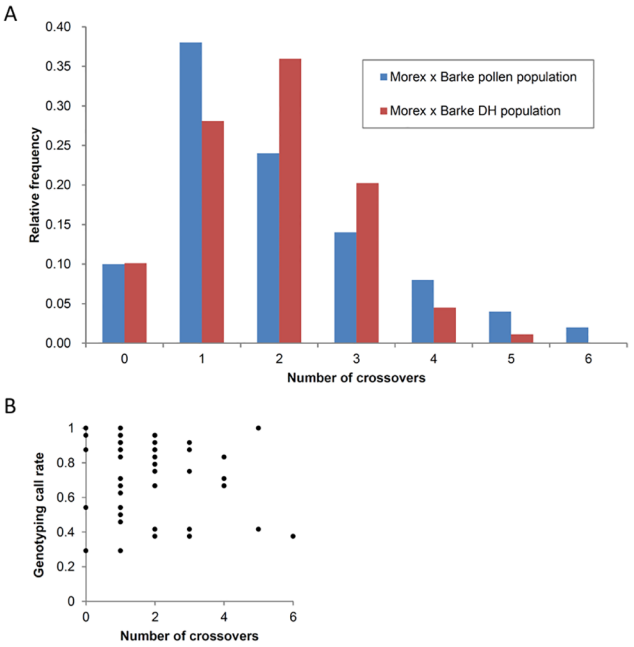

Fig 3. Comparison of the distribution of the number of crossovers. (A) The relative frequency of the total number of crossovers per chromosome 3H grouped into classes ranging from 0 to 6 of the Morex x Barke pollen population (blue) in comparison to the Morex x Barke DH population data (red) [17]. **(B)** Correlation between KASP genotyping call rate and the number of crossovers found for each sample ($r = -0.16$, $r^2 = 0.03$).

doi:10.1371/journal.pone.0137677.g003

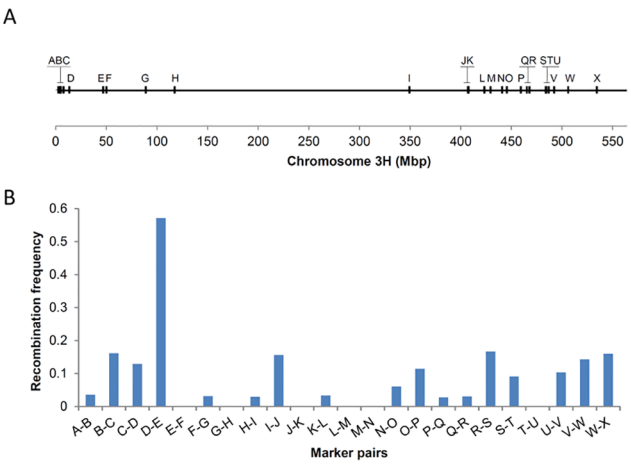

Fig 4. Recombination frequency along chromosome 3H determined by pollen genotyping. (A) KASP marker positions (A to X) are shown as vertical bars along chromosome 3H. The physical length of the chromosome (Mbp) is shown on the x-axis. **(B)** The recombination frequency along chromosome 3H of a given physical interval measured as the proportion of crossovers to no-crossovers for each marker pair. A distal bias is shown by higher recombination frequencies towards the chromosome ends and low recombination frequencies between markers H and I.

doi:10.1371/journal.pone.0137677.g004

frequencies along chromosome 3H by counting the number of crossovers in neighbouring marker intervals (Fig 4). A typical pattern of elevated recombination frequencies towards the distal regions of the chromosome was found which is in agreement with previous reports for barley based on molecular marker data [1, 21] and cytological visualizations of crossovers [7].

To assess the extent of segregation distortion in pollen grains, we investigated the number of loci showing segregation distortion in our pollen population and compared it to equivalent data derived from a doubled haploid (DH) population of the same genotype. In our pollen population, 24 loci on chromosome 3H were scored for presence or absence of each allele. Segregation distortion was found for 8.3% (2 of 24) of the markers (S3 Table). This proportion appears to be lower compared to the Morex x Barke DH population which showed 25.7% (25 of 97) of all loci on chromosome 3H having distorted segregation ratios (S3 Table). However, the difference in sample size, which is 50 pollen grains compared to 89 DH individuals, allows only major effects to be detected. This difference might be explained by selection against particular genotype combinations during anther culture of the DH lines or by absence of selection in pollen grains for pollination and fertilization success. However, this tendency is in agreement with Sayed et al. [22] who compared segregation distortion of a barley DH population versus an F_2 population, finding a difference of 44.2% versus 16.3%, respectively.

We conclude that it is feasible to genotype single pollen grains using our amplification approach combined with KASP. It offers the opportunity to efficiently monitor meiotic recombination in individual pollen nuclei and avoids the necessity to generate segregating populations. Due to the high amount of DNA obtained from a single haploid nucleus via WGA, we suggest that our approach might be used for genome wide analyses. This will be particularly useful in plant breeding to monitor the recombination landscape of any genotype of interest.

Supporting Information

S1 Fig. Distribution of the number of positive PCR markers to assess WGA efficiency.
(TIF)

S2 Fig. Distribution of KASP marker call rate and frequency of double crossover. The marker call rate of each KASP marker (A to X) is shown (blue) as well as the frequency of presumptive double crossover (red), e.g. a crossover on both sides of a given marker.
(TIF)

S1 Table. List of primers used to evaluate WGA performance.
(DOCX)

S2 Table. Gene-based KASP markers.
(DOCX)

S3 Table. Genotypic data and segregation statistics of the pollen population and DH population. Sample number is shown in the first column and the first row indicates the marker number. Genotypic values are shown as A for Morex and B for Barke. Missing values are indicated by a minus.
(XLSX)

S4 Table. Supplementary information to Figs 2, 3, 4 and S5, S6 Figs. Actual data referring to Figs 2, 3, 4 and S5, S6 Figs are summarized in separate tables.
(XLSX)

Acknowledgments

We thankfuly acknowledge Tim Sharbel and Lala Aliyeva-Schnorr for technical advice. This work was supported by the IPK, P. Cápal was supported by the Ministry of Education, Youth and Sports of the Czech Republic (grant LO1204 from the National Program of Sustainability I).

Author Contributions

Conceived and designed the experiments: SD AH. Performed the experiments: SD JF. Analyzed the data: SD NK. Contributed reagents/materials/analysis tools: PC. Wrote the paper: SD JF AH EB.

References

1. Baker K, Bayer M, Cook N, Dreissig S, Dhillon T, Russell J, et al. The low-recombining pericentromeric region of barley restricts gene diversity and evolution but not gene expression. Plant J. 2014; 79 (6):981–92. doi: 10.1111/tpj.12600 PMID: 24947331.
2. Tondelli A, Xu X, Moragues M, Sharma R, Schnaithmann F, Ingvardsen C, et al. Structural and Temporal Variation in Genetic Diversity of European Spring Two-Row Barley Cultivars and Association Mapping of Quantitative Traits. Plant Genome-Us. 2013; 6(2). doi: 10.3835/plantgenome2013.03.0007 PMID: WOS:000322887700009.
3. Crismani W, Girard C, Froger N, Pradillo M, Santos JL, Chelysheva L, et al. FANCM limits meiotic crossovers. Science. 2012; 336(6088):1588–90. doi: 10.1126/science.1220381 PMID: 22723424.
4. Higgins JD, Perry RM, Barakate A, Ramsay L, Waugh R, Halpin C, et al. Spatiotemporal Asymmetry of the Meiotic Program Underlies the Predominantly Distal Distribution of Meiotic Crossovers in Barley. Plant Cell. 2012; 24(10):4096–109. doi: 10.1105/tpc.112.102483 PMID: ISI:000312378300018.
5. Salome PA, Bomblies K, Fitz J, Laitinen RA, Warthmann N, Yant L, et al. The recombination landscape in Arabidopsis thaliana F2 populations. Heredity. 2012; 108(4):447–55. doi: 10.1038/hdy.2011.95 PMID: 22072068; PubMed Central PMCID: PMC3313057.
6. Sybenga J. Quantitative Analysis of Chromosome Pairing and Chiasma Formation Based on Relative Frequencies of M I Configurations .4. Interchange Heterozygotes. Genetica. 1966; 37(2):199–&. doi: 10.1007/Bf01547131 PMID: WOS:A19668461900007.
7. Phillips D, Wnetrzak J, Nibau C, Barakate A, Ramsay L, Wright F, et al. Quantitative high resolution mapping of HvMLH3 foci in barley pachytene nuclei reveals a strong distal bias and weak interference.

J Exp Bot. 2013; 64(8):2139–54. doi: 10.1093/jxb/ert079 PMID: 23554258; PubMed Central PMCID: PMC3654414.

8. Copenhaver GP, Keith KC, Preuss D. Tetrad analysis in higher plants. A budding technology. Plant Physiol. 2000; 124(1):7–16. PMID: 10982416.

9. Petersen G, Johansen B, Seberg O. PCR and sequencing from a single pollen grain. Plant Mol Biol. 1996; 31:189–91. PMID: 8704154

10. Chen PH, Pan YB, Chen RK. High-throughput procedure for single pollen grain collection and polymerase chain reaction in plants. J Integr Plant Biol. 2008; 50(3):375–83. doi: 10.1111/j.1744-7909.2007.00624.x PMID: ISI:000253711300015.

11. Aziz AN, Sauve RJ. Genetic mapping of Echinacea purpurea via individual pollen DNA fingerprinting. Mol Breeding. 2008; 21(2):227–32. doi: 10.1007/s11032-007-9123-9 PMID: ISI:000251868400009.

12. Wang JB, Fan HC, Behr B, Quake SR. Genome-wide Single-Cell Analysis of Recombination Activity and De Novo Mutation Rates in Human Sperm. Cell. 2012; 150(2):402–12. doi: 10.1016/j.cell.2012.06.030 PMID: ISI:000306595700018.

13. 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. 1983; 220(4601):1049–51. doi: 10.1126/science.220.4601.1049 PMID: 17754551.

14. De Storme N, Geelen D. The Arabidopsis mutant jason produces unreduced first division restitution male gametes through a parallel/fused spindle mechanism in meiosis II. Plant Physiol. 2011; 155(3):1403–15. doi: 10.1104/pp.110.170415 PMID: 21257792; PubMed Central PMCID: PMC3046594.

15. Gole J, Gore A, Richards A, Chiu YJ, Fung HL, Bushman D, et al. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. Nat Biotechnol. 2013; 31(12):1126–32. doi: 10.1038/nbt.2720 PMID: 24213699; PubMed Central PMCID: PMC3875318.

16. Comadran J, Kilian B, Russell J, Ramsay L, Stein N, Ganal M, et al. Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nature genetics. 2012; 44(12):1388–92. doi: 10.1038/ng.2447 PMID: 23160098.

17. International Barley Genome Sequencing Consortium (IBGSC). A physical, genetic and functional sequence assembly of the barley genome. Nature. 2012; 491(7426):711–+. doi: 10.1038/Nature11543 PMID: ISI:000311606000035.

18. Mirouze M, Lieberman-Lazarovich M, Aversano R, Bucher E, Nicolet J, Reinders J, et al. Loss of DNA methylation affects the recombination landscape in Arabidopsis. Proc Natl Acad Sci U S A. 2012; 109(15):5880–5. doi: 10.1073/pnas.1120841109 PMID: 22451936; PubMed Central PMCID: PMC3326504.

19. Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas WT, et al. Flapjack—graphical genotype visualization. Bioinformatics. 2010; 26(24):3133–4. doi: 10.1093/bioinformatics/btq580 PMID: 20956241; PubMed Central PMCID: PMC2995120.

20. de Bourcy CF, De Vlaminck I, Kanbar JN, Wang J, Gawad C, Quake SR. A quantitative comparison of single-cell whole genome amplification methods. Plos One. 2014; 9(8):e105585. doi: 10.1371/journal.pone.0105585 PMID: 25136831; PubMed Central PMCID: PMC4138190.

21. Kunzel G, Korzun L, Meister A. Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics. 2000; 154(1):397–412. PMID: 10628998; PubMed Central PMCID: PMC1460903.

22. Sayed H, Kayyal H, Ramsey L, Ceccarelli S, Baum M. Segregation distortion in doubled haploid lines of barley (Hordeum vulgare L.) detected by simple sequence repeat (SSR) markers. Euphytica. 2002; 125(2):265–72. doi: 10.1023/A:1015861610226 PMID: WOS:000176184000014.

Appendix IV

Advances in plant chromosome genomics

DOI: 10.1016/j.biotechadv.2013.12.011

Biotechnology Advances 32 (2014) 122–136

Contents lists available at ScienceDirect

Biotechnology Advances

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

Research review paper

Advances in plant chromosome genomics ☆

Jaroslav Doležel *, Jan Vrána, Petr Cápal, Marie Kubaláková, Veronika Burešová, Hana Šimková

Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany, Šlechtitelů 31, Olomouc CZ-78371, Czech Republic

ARTICLE INFO

Available online 7 January 2014

Keywords:
BAC library
Chromosome sorting
Cytogenetics
DNA arrays
DNA marker development
Flow cytometry
Genome assembly
Mitotic metaphase chromosomes
Next generation sequencing
Physical mapping

ABSTRACT

Next generation sequencing (NGS) is revolutionizing genomics and is providing novel insights into genome organization, evolution and function. The number of plant genomes targeted for sequencing is rising. For the moment, however, the acquisition of full genome sequences in large genome species remains difficult, largely because the short reads produced by NGS platforms are inadequate to cope with repeat-rich DNA, which forms a large part of these genomes. The problem of sequence redundancy is compounded in polyploids, which dominate the plant kingdom. An approach to overcoming some of these difficulties is to reduce the full nuclear genome to its individual chromosomes using flow-sorting. The DNA acquired in this way has proven to be suitable for many applications, including PCR-based physical mapping, *in situ* hybridization, forming DNA arrays, the development of DNA markers, the construction of BAC libraries and positional cloning. Coupling chromosome sorting with NGS offers opportunities for the study of genome organization at the single chromosomal level, for comparative analyses between related species and for the validation of whole genome assemblies. Apart from the primary aim of reducing the complexity of the template, taking a chromosome-based approach enables independent teams to work in parallel, each tasked with the analysis of a different chromosome(s). Given that the number of plant species tractable for chromosome sorting is increasing, the likelihood is that chromosome genomics – the marriage of cytology and genomics – will make a significant contribution to the field of plant genetics.

Contents

1. Sequencing of plant genomes . . . 123
2. Reducing the complexity of the sequencing template . . . 123
3. Flow cytometry . . . 124
   3.1. Sample preparation . . . 125
   3.2. Analysis and sorting . . . 126
4. Uses of flow-sorted chromosomes . . . 128
   4.1. Physical mapping . . . 128
      4.1.1. Mapping by PCR . . . 128
      4.1.2. Construction of clone-based physical maps . . . 129
      4.1.3. Cytogenetic mapping . . . 130
   4.2. Genetic marker development . . . 130
      4.2.1. SSRs and ISBPs . . . 130
      4.2.2. DArT markers . . . 130
      4.2.3. Marker development from chromosome-specific shotgun sequences . . . 130
      4.2.4. Marker specificity . . . 131
   4.3. Sequencing . . . 131
      4.3.1. BAC clones . . . 131
      4.3.2. Whole chromosome sequencing using 454 technology . . . 131
      4.3.3. Whole chromosome sequencing using Illumina technology . . . 132
      4.3.4. Validation of whole genome assemblies . . . 133

* Corresponding author. Tel.: +420 585 238 703.
E-mail addresses: dolezel@ueb.cas.cz (J. Doležel), vrana@ueb.cas.cz (J. Vrána), capal@ueb.cas.cz (P. Cápal), kubalakova@ueb.cas.cz (M. Kubaláková), buresova@ueb.cas.cz (V. Burešová), simkovah@ueb.cas.cz (H. Šimková).

0734-9750/$ – see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.biotechadv.2013.12.011

5. Conclusions . . . 133
Acknowledgements . . . 133
References . . . 133

1. Sequencing of plant genomes

The last decade has seen a major leap in our understanding of plant genome structure, function and evolutionary dynamics. The main driver of this advance has been the elaboration of next generation sequencing (NGS) platforms, which allow for the parallel acquisition of huge numbers of reads, representing hundreds of billions of nucleotides; in concert, advances in bioinformatics have been necessary to enable this ocean of DNA sequence to be analyzed. The first plant genome to be fully sequenced was that of *Arabidopsis thaliana*, chosen for its small genome of ~150 Mb; although this represented a logistical challenge in the context of 1990s sequencing technology, it would no longer do so, given the capacity of modern instruments, which can generate up to 60 Gb of sequence per run. The *A. thaliana* genome was acquired using a clone-by-clone (CBC) strategy (The Arabidopsis Genome Initiative, 2000). The minimum set of clones to be sequenced, termed the "minimum tiling path" (MTP), is elaborated from the physical map, which is constructed on the basis of overlapping large-insert DNA clones. The second plant species to be sequenced was rice, using a similar strategy (Matsumoto et al., 2005). Apart from its importance as a crop species, rice was selected also because of its relatively small genome size (~400 Mb). The acquisition of these two whole genome sequences marked a new departure for plant genetics, allowing, for the first time, a holistic view of the entire genome. Since the beginning of the present century, the pace of sequencing has accelerated, so that by 2010, a number of important plant species had been sequenced.

A gradual shift in sequencing strategy, moving away from the CBC approach to a whole genome shotgun (WGS) one was already underway during the first phase of plant genome sequencing. The shotgun method was used for acquiring the genome sequences of poplar (Tuskan et al., 2006), grapevine (Jaillon et al., 2007) and sorghum (Paterson et al., 2009). The 2.5 Gb maize genome was published in 2009, but exceptionally relied on the CBC approach (Schnable et al., 2009). Since 2010, NGS technologies have become routine, and have greatly driven down both the price and effort required of genome sequencing. In this second phase of plant genome sequencing, already some 40 plant species have been sequenced, and the expectation is that not only reference genome sequences will be acquired for most of the economically and scientifically important plant species, but that the scale of re-sequencing will grow by orders of magnitude (The million plant and animal genomes project, 2013). Unlike *de novo* sequencing, which requires the assembly of the genome from short reads, re-sequencing is technically simpler, as the reads can be referenced to an available complete genome sequence. The quality of re-sequenced genomes is therefore determined by the quality of the reference genome sequence; the fuller the coverage of the reference sequence, the more correctly the re-sequenced contigs will be ordered. The feasibility of sequencing many individuals from the same species offers opportunities for population genetics analysis and genotype-based breeding (Long et al., 2013).

High quality reference genome sequences are particularly important for the analysis of the functional organization of DNA. The function of the nuclear genome cannot be understood without an understanding of its various components, as exemplified by the human genome ENCODE project (Gerstein et al., 2012). An unfortunate consequence of the widespread use of NGS shotgun sequencing is a drop in assembly quality, so that the highest quality genome sequences remain those of *A. thaliana*, rice and maize, which were acquired by the CBC method (Feuillet et al., 2011; Shangguan et al., 2013). Assembly is particularly problematical for large genome species such as Norway spruce (1C: ~20 Gb), where only some 25% of the genome was assemblable into scaffolds longer than 10 Kb (Nystedt et al., 2013); such issues can arise in smaller genomes too, for example in chickpea (1C: ~0.9 Gb), where the genome sequence presently comprises over 180,000 scaffolds (Jain et al., 2013). Of course, it is not always necessary to generate a gold standard sequence, since for some applications a rough genome draft is sufficient for the purpose. The difficulty arises when such draft genome assemblies are presented as reference sequences (Sierro et al., 2013). In some cases, projects relying on incomplete genome sequences may fail, and there are examples where funding proposals aimed at the acquisition of a high quality reference sequence have been declined as the donors believed that the work had already been done.

The power of NGS lies in its capacity to generate a huge volume of reads, but its weakness is that these reads are rather short. Plant genomes are populated by many families of repetitive DNA elements (Schmidt and Heslop-Harrison, 1998), and these can be impossible to resolve when only short reads are available. The problem of sequence redundancy is compounded in polyploids, which dominate the plant kingdom. Genome assembly from shotgun reads may not be straightforward even in compact genomes having a small content of repetitive DNA. A good example is the bladderwort *Utricularia gibba*, with a genome size of just 77 Mb, of which only 3% is repetitive; nevertheless an attempt at shotgun sequencing resulted in a set of >3800 sequence contigs arranged in over 1200 scaffolds (Ibarra-Laclette et al., 2013). Technical improvements in read length and/or the algorithms used for sequence assembly should in time, however, enable reference genome sequences to be produced by NGS shotgun methods (Roberts et al., 2013). NGS shotgun sequencing may be at present be of limited utility in acquiring gold standard reference sequences (Marx, 2013), but the technology is very powerful for simpler templates such as bacterial artificial chromosomes (BACs), which form the backbone of many physical maps (Feuillet et al., 2011). Incomplete sequence assembly is then limited to at most 100 Kb, the genomic location of which is known. BAC clones are commonly sequenced in pools to reduce cost (Sato et al., 2011; Steuernagel et al., 2009), and this requires a bar-coding strategy to attribute the resulting contigs to their specific BAC. The sequence redundancy typical of large and particularly of polyploid genomes, makes the construction of a physical map based on BAC clones difficult (Meyers et al., 2004; Paux et al., 2008); it is a task which would be greatly simplified if the template complexity could be reduced.

2. Reducing the complexity of the sequencing template

As both the CBC and the NGS shotgun sequencing strategies are compromised by sequence redundancy, any reduction in template complexity would be helpful. Breaking down the genome into its individual chromosomes represents an attractive option, especially for polyploid genomes, as this would abolish the problem of redundancy due to the presence of homoeologs (Fig. 1). Flow-sorting has been developed to achieve exactly this result, and this review outlines its potential for plant genome analysis and sequencing. Methods designed to simplify the assembly of shotgun sequence reads and to construct ready-to-sequence clone-based physical maps are described. Chromosome sorting is not, of course, the sole option available for reducing template complexity prior to DNA sequencing. The selection of DNA based on either its renaturation kinetics ("Cot filtration") (Peterson et al., 2002)

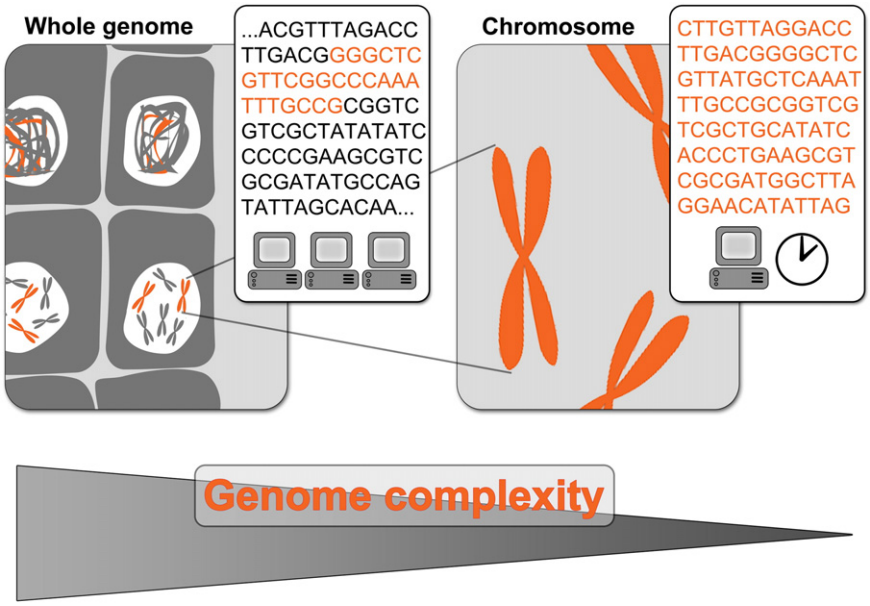

Fig. 1. Chromosome genomics: instead of treating the whole nuclear genome as a unit, single chromosomes are isolated and their DNA used as the template for genomic analyses. The reduction in template complexity achieved speeds mapping, sequencing and sequence analysis, and simplifies the necessary bioinformatics. In polyploids, interference from homoeologs is minimized.

or its methylation status (Rabinowicz et al., 2003) both were designed to eliminate much of the repetitive DNA component, leaving mainly low copy sequences. A complexity reduction step has also been incorporated into genotyping-by-sequencing, based on the use of methylation-sensitive restriction enzymes to eliminate the highly methylated repetitive component prior to sequencing (Elshire et al., 2011), and several other target-enrichment strategies have been developed (Mamanova et al., 2010). Inevitably, this sort of strategy, unlike one based on individual chromosomes, cannot deliver a complete genome sequence. Chromosome number is variable from species to species, but is typically in the range 5–20. Thus, complexity can in principle, be reduced by around an order of magnitude. For example, each barley or bread wheat chromosome harbors, on average, respectively about 14% and 5% of the full genome complement.

During most of an organism's life cycle, its chromosomes are extended and intimately intertwined with one another in interphase nuclei. The exceptions are during cell division, when the chromosomes become very much shortened and are physically separated from one another. Attempts have been made to isolate mitotic chromosomes using microdissection (Matsunaga et al., 1999; Stein et al., 1998). A clear advantage of this approach is that the chromosomes have already been attached to a fixed surface, where they can be optically identified prior to their mechanical isolation. However the process is highly labor-intensive, so only small populations of individual chromosomes can be isolated; while the resulting DNA can be amplified to provide template sufficient for sequencing, the required amplification imposes such a restriction on the length of the DNA recovered (Schondelmaier et al., 1993; Stein et al., 1998) that it become unsuitable for constructing the large insert libraries required to assemble a physical map. Moreover, extensive amplification inevitably introduces a bias. The alternative to micro-dissection is to isolate large populations of intact mitotic metaphase chromosomes in suspension. The methods required to achieve this necessitate not just the ability to prepare such suspensions, but also the means to physically separate a specific chromosome from the mass of non-homologs present. Attempts have been made to achieve this separation using gradient centrifugation (Stubblefield and Oro, 1982) or by capture on magnetic beads following hybridization with a labeled chromosome-specific probe (Dudin et al., 1988; Vitharana and Wilson, 2006); however, to date, the most successful method is flow-sorting (Doležel et al., 1994, 2007a, 2011). In what follows, we first explain the methodology involved in flow cytometric chromosome analysis and sorting (termed "flow cytogenetics") and then discuss current and potential applications of flow-sorted chromosomes in plant genomics ("chromosome genomics").

3. Flow cytometry

Flow cytometry was initially developed as an alternative to microscopy for counting blood cells; its advantage is its high throughput and potential for automation. The capacity to handle large numbers of individual cells enables the detection of rare mutants, and can deliver meaningful statistical data with respect to frequency. A typical flow cytometer does not capture images of the cells; rather the aim is to analyze light scatter and fluorescence. Flow cytometers need to be capable of measuring these properties simultaneously in real time, as they combine to provide a wealth of information (Rieseberg et al., 2001), specifically regarding cell viability, physiological status, apoptosis, ploidy and cell cycle status. Supported by a variety of fluorescent probes and antibodies, flow cytometry has developed into a ubiquitous tool in immunology, pathology, oncology and other areas of biomedical research (Shapiro, 2003). Although less commonly exploited in plant biology, these devices have found a number of fundamental research and industrial uses, the main ones being the estimation of genome size and ploidy level (Doležel et al., 2007b). The salient feature of flow cytometry is that the target particles are suspended in a narrow stream of liquid (typically saline); they are forced to move in a single file, where they can be made to interact one-by-one with an orthogonally oriented light beam (Fig. 2). Solid state lasers provide the most commonly used light source, and it is not unusual to install more than one laser, with each set up to excite a different fluorochrome incorporated into the particles. The flow rate is typically several thousand per second. To sort the particles into discrete sub-populations, the stream is broken into ~1 nL droplets. Those carrying a target particle are electrically charged and then deflected from the main stream of non-target particles by passage through an electrical field. Because the rate of droplet generation exceeds the particle flow rate, the majority of droplets are empty and very few droplets contain more than one particle. Clumps of particles tend to block the narrow orifice (typically < 100 μm) of the flow chamber, thereby disrupting laminar flow and compromising the analysis (Shapiro, 2003). Poor results are also obtained if the particles are

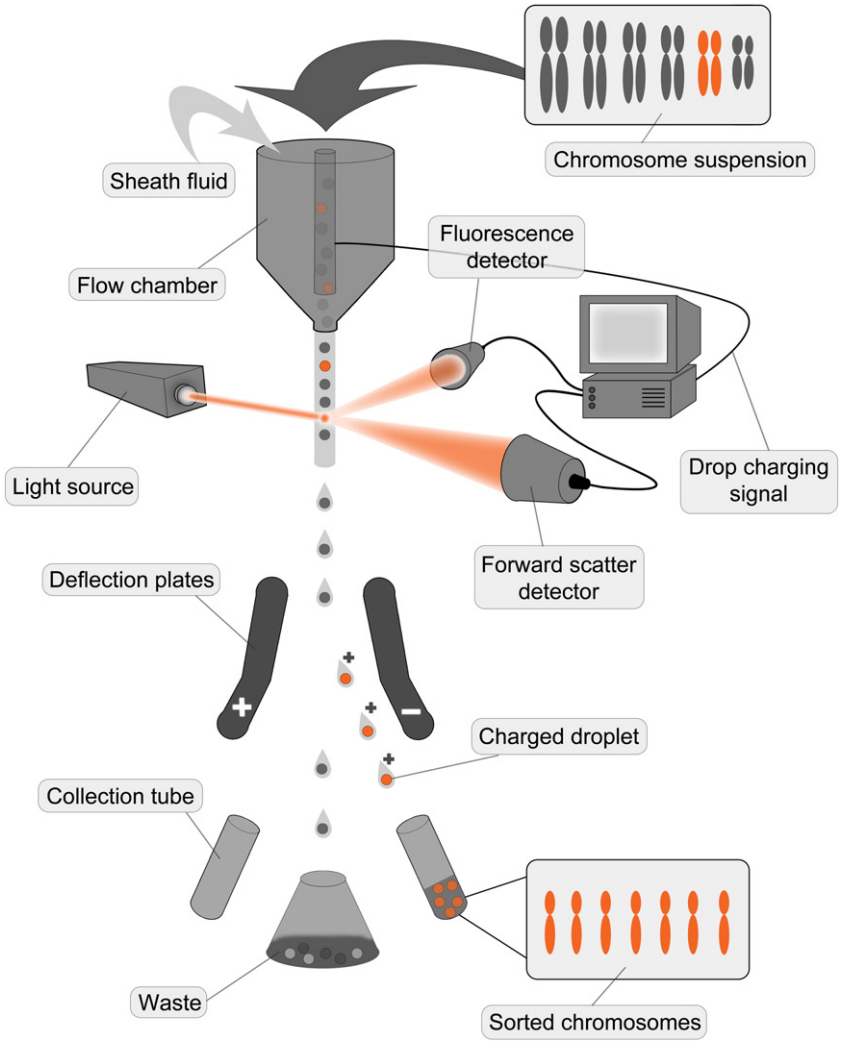

Fig. 2. The mechanics of flow-sorting. Chromosomes held in liquid suspension are stained by a fluorochrome and passed into a flow chamber containing sheath fluid. The geometry of the chamber forces the chromosome suspension into a narrow stream in which the chromosomes become aligned in a single file, and so are able to interact individually with an orthogonally directed laser beam(s). Pulses of scattered light and emitted fluorescence are detected and converted to electric pulses. If the chromosome of interest differs in fluorescence intensity from other chromosomes, it is identified and sorted. The sorting is achieved by breaking the stream into droplets and by electrically charging droplets carrying chromosomes of interest. The droplets are deflected during passage through electrostatic field between defection plates and collected in suitable containers.

mechanically damaged. Thus sample quality is of prime importance, and this is especially the case for chromosome analysis and sorting, where any reduction in resolution will produce unwanted contamination of a sorted chromosome by other chromosomes, chromosome fragments or aggregates. The elaboration of a robust method for preparing sufficiently high quality chromosome suspensions has been the most serious barrier to the development of flow cytogenetics in humans and animals, but particularly in plants.

3.1. Sample preparation

Since in general somatic tissue is easier to obtain than reproductive tissue, flow cytogenetics has largely concentrated on isolating mitotic metaphase chromosomes. At any given time, the majority of plant and animal cells in non-reproductive tissue are in interphase, so accessing mitotic metaphase chromosomes requires a pre-treatment to first encourage cell division, and then to arrest cells at mitotic metaphase. (Note that targeting meiotic chromosomes in microspores is in principle highly attractive, since cell division is well synchronized in these cells. However there are practical difficulties associated with the acquisition of a sufficient number of dividing cells.) Some technical issues surround the release of metaphase chromosomes into the liquid medium. Current protocols designed to prepare chromosome suspensions from human or animal cells are based on either synchronized cell lines or stimulated peripheral blood; the chromosomes are released by hypotonic lysis (Chen et al., 2008; Yang et al., 2011). Plant cells are less tractable, mainly because of their rigid cell wall. Synchronizing mitosis is also less straightforward than in animal cells. De Laat and Blaas (1984), who were the first to demonstrate the sorting of plant chromosomes, used hydroxyurea for synchronization and colchicine to arrest cells at metaphase. A similar approach was taken by Arumuganathan et al. (1991) in tomato and by Schwarzacher et al. (1997) in wheat. Although plant cells, like animal ones, can be cultured *in vitro,* such cultures are often karyologically unstable (Leitch et al., 1993; Schwarzacher et al., 1997), and their cell cycle is not well synchronized (typically not exceeding 35%, see Arumuganathan et al. (1994) and de Laat and Blaas (1984)). Following the animal cell protocols, hypotonic lysis was used in early experiments to release plant chromosomes, but this was only feasible if the cell walls were first digested enzymatically. While this step provides a non-disruptive means of releasing the chromosomes, it also introduces a time delay between metaphase arrest and the chromosome release, which lowers the chromosome yield due to the premature separation of sister chromatids and/or chromosome decondensation.

The release of chromosomes from leaf-derived protoplasts was described by Conia et al. (1987). The strategy adopted was to force arrest of the cells in the G1 phase, and then to transfer the cells into a

medium formulated to initiate cell cycling, so that they would enter mitosis in synchrony. Unfortunately the induction was not sufficiently effective, since only 10% of the cells divided. Additionally, chromosome release was hampered by a partial regeneration of the cell wall. A major advantage of sourcing chromosomes from live plant tissue (such as the leaf) as opposed to *in vitro* cultured cells is that their karyotype is normal. The choice of root tip meristems as a source of mitotic chromosomes is based on a naturally high rate of cell division, and (unlike leaf-derived cells), the ease of synchronizing mitosis, with rates above 50% being attainable (Doležel et al., 1992). A productive method of chromosome release from root tips, avoiding the need to digest the cell wall, was elaborated by Doležel et al. (1992). The material was first fixed in formaldehyde to render the chromosomes mechanically stable and resistant to shearing forces, and then homogenized. Apart from karyological stability, the advantage of using root tips is that seedlings can be obtained in a majority of plants and roots can be exposed to various treatments using a hydroponic system. The procedure can be extended to species which produce few (or no) seeds by inducing hairy root cultures (Neumann et al., 1998; Veuskens et al., 1995).

A typical root tip-based protocol (e.g., Vrána et al., 2012) involves seed germination, the exposure of roots of young seedlings to hydoxyurea (a DNA synthesis inhibitor) to arrest the cells at the G1/S interface, followed by recovery to synchronize the cell cycle through the S and G2 phases and into mitosis. Dividing cells are arrested at mitotic metaphase by treating with a mitotic spindle poison such as the herbicides amiprophos-methyl, oryzalin or trifluralin (Doležel et al., 1992; Guo et al., 2006; Vláčilová et al., 2002). In species where these compounds induce chromosome stickiness, a treatment with nitrous oxide (Kato, 1999) has proven to be efficacious (unpublished data). An option is an overnight exposure to ice water prior to fixation, a treatment which can improve the dispersion of chromosomes in the cytoplasm and thereby increase the chromosome yield (Vrána et al., 2000). The treated roots are then fixed in formaldehyde and the chromosomes released into the isolation buffer by chopping using a sharp scalpel or razor blade (Doležel et al., 1992). When working with small root tips, homogenization using a handheld homogenizer is both rapid and convenient (Gualberti et al., 1996). Of especial importance is the composition of the isolation buffer, as this ensures the maintenance of chromosome morphology and DNA integrity, as well as providing a compatible environment for DNA staining.

3.2. Analysis and sorting

To date, flow cytometry has been used to sort chromosomes in 24 plant species, belonging to 18 genera (Table 1). Staining chromosomal DNA with a fluorochrome (commonly either ethidium bromide (Li et al., 2004), Hoechst 333242 (Conia et al., 1987) or DAPI (Kubaláková et al., 2005)) results in a distribution of fluorescence signal intensity (the "flow karyotype"), in which, ideally, each chromosome can be recognized by a different peak. Formaldehyde fixation has been found to interfere with the stoichiometric binding of some fluorochromes to chromosomal DNA, and DAPI has been found to be the least sensitive of the fluorochromes in this respect (Doležel and Lucretti, 1995). The size of the peak is dependent on the DNA content, and it is common to find that the DNA content of two (or more) of the chromosomes is so similar that they are represented in the flow karyotype as a single, broad peak. Thus, for example, in the flow karyotype of chickpea, six of the eight chromosomes can be separated, while the other two form a single peak (Fig. 3A). In contrast, the bread wheat ($n = 21$) flow karyotype comprises only one single chromosome peak (chromosome 3B), with the other 20 chromosomes forming three composite peaks (Fig. 3B). Karyotype variation within wheat has allowed some additional chromosomes to be discriminated (Kubaláková et al., 2002), and the same is the case for chickpea (Vláčilová et al., 2002; Zatloukalová et al., 2011). The pattern of light scatter can be used to differentiate between chromosomes and cell detritus (Conia et al., 1987), while the width of the fluorescence pulse aids in the discrimination of chromosomes doublets (Lucretti et al., 1993).

The inability to discriminate each chromosome in the flow karyotype presents a serious limitation to the utility of flow cytometry, so substantial effort has been devoted to overcoming this problem. An early strategy was to simultaneously stain the material with two fluorochromes differing in their base pair preference (for instance Hoechst 33258 which binds preferentially to AT rich sequence and Chromomycin A3, which targets GC rich sequence). In the human karyotype, this method effectively discriminates almost every chromosome (Ferguson-Smith,

Table 1
List of plant species for which a flow cytometric analysis of mitotic chromosomes has been published.

Genus	Species	Common name	n	References
Aegilops	*biuncialis*	Goatgrasses	14	Molnár et al. (2011)
	comosa		7	Molnár et al. (2011)
	geniculata		14	Molnár et al. (2011)
	umbellulata		7	Molnár et al. (2011)
Avena	*sativa*	Oat	21	Li et al. (2001)
Cicer	*arietinum*	Chickpea	8	Vláčilová et al. (2002), Zatloukalová et al. (2011)
Dasypyrum	*villosum*		7	Grosso et al., 2012; Giorgi et al., 2013
Festuca	*pratensis*	Meadow fescue	7	Kopecký et al., 2013
Haplopappus	*gracilis*		2	de Laat and Blaas (1984), de Laat and Schel (1986)
Hordeum	*vulgare*	Barley	7	Lysák et al. (1999), Lee et al. (2000), Suchánková et al. (2006)
Lycopersicon	*esculentum*	Tomato	12	Arumuganathan et al. (1991)
	pennellii	Tomato	12	Arumuganathan et al. (1991, 1994)
Nicotiana	*plumbaginifolia*	Tobacco	10	Conia et al. (1989)
Oryza	*sativa*	Rice	12	Lee and Arumuganathan (1999)
Petunia	*hybrida*	Petunia	7	Conia et al. (1987)
Picea	*abies*	Norway spruce	12	Überall et al. (2004)
Pisum	*sativum*	Pea	7	Gualberti et al. (1996), Neumann et al. (1998, 2002)
Secale	*cereale*	Rye	7	Kubaláková et al. (2003)
Silene	*latifolia*	White campion	12	Veuskens et al. (1995), Kejnovský et al. (2001)
Triticum	*aestivum*	Bread wheat	21	Wang et al. (1992), Schwarzacher et al. (1997), Lee et al. (1997), Gill et al. (1999), Vrána et al. (2000), Kubaláková et al. (2002), Giorgi et al., 2013
	durum	Durum wheat	14	Kubaláková et al. (2005), Giorgi et al., 2013
Vicia	*faba*	Field bean	6	Lucretti et al. (1993), Doležel and Lucretti (1995), Lucretti and Doležel (1997)
	sativa	Common vetch	6	Kovářová et al. (2007)
Zea	*mays*	Maize	10	Lee et al. (1996, 2002), Li et al. (2001, 2004)

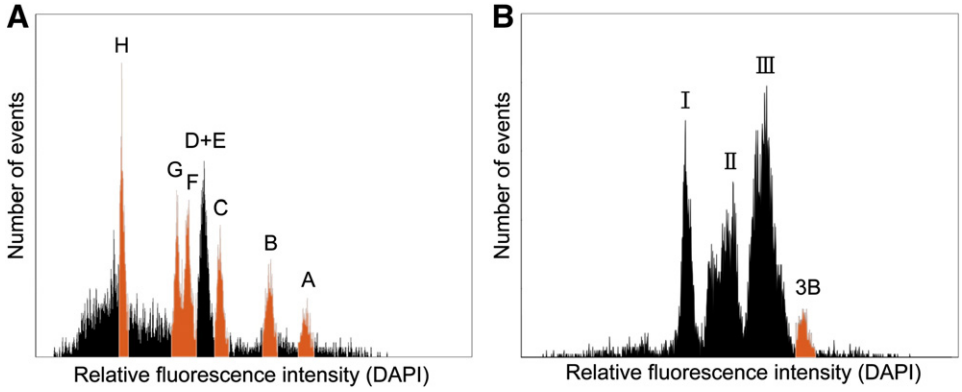

Fig. 3. Flow karyotyping in chickpea and bread wheat. The fluorescence intensity histograms (flow karyotypes) were obtained from DAPI-stained suspensions of mitotic chromosomes. (A) Chickpea cv. Frontier (2n = 2x = 16) forms seven peaks, six of which each represent a single chromosome (A–C and F–H). The seventh peak harbors both chromosomes D and E. (B) In the wheat cv. Chinese Spring (2n = 6x = 42) flow karyotype, only chromosome 3B forms a discrete peak. The remaining 20 chromosomes are dispersed into the composite peaks I–III.

1997; Langlois et al., 1982), but it has not been successful in plant genomes (Lee et al., 1997, 2000; Lucretti and Doležel, 1997; Schwarzacher et al., 1997), presumably because global variation in AT/GC ratio among the chromosomes is masked by the ubiquitous presence of repetitive DNA (Fuchs et al., 1996; Schubert et al., 2001). The approach taken attempted to exploit polymorphism in chromosome length resulting from deletions and translocations. Lucretti et al. (1993) were the first to show that reciprocal translocations in field bean could be used to identify a number of its chromosomes, and a similar success was recorded by Neumann et al. (1998) working with garden pea. In some cases, chromosome sorting has been facilitated by cryptic structural features (Kubaláková et al., 2002; Lee et al., 2002). The tolerance of polyploids to aneuploidy has been used to develop a plethora of true-breeding cytogenetic materials, especially in bread wheat. Of particular interest in the context of flow karyotyping are telocentric chromosomes (telosomes), in which an entire arm has been lost; a collection of these, involving most of the 42 chromosome arms of wheat was generated by Sears (Sears and Sears, 1978). The small size of telosomes means that their peaks become well separated from the rest of the flow karyotype, allowing them to be readily sorted (Gill et al., 1999; Guo et al., 2006; Kubaláková et al., 2002) (Fig. 4A).

The tolerance of polyploids to aneuploidy has also allowed for the production of stable lines in which a single chromosome pair from a related species can be maintained in isolation from the others. If this "alien" chromosome differs in DNA content from those of the host species, its peak should be recognizable, and can therefore be sorted. For example, in cereal rye, the only chromosome which can be successfully sorted from the other six is 1R, but the other six proved to be sortable when represented in a single chromosome addition line (Kubaláková et al., 2003) (Fig. 4B). In the case of barley, the peaks overlap with those of wheat, so the chromosome addition line approach is not fruitful. However, it has proved possible to discriminate and sort barley telosomes present as a single pair in a wheat background (Suchánková et al., 2006). The availability of such addition lines has been a boon for chromosome sorting in the wild relatives of wheat, which otherwise have proven difficult to purify (Grosso et al., 2012; Molnár et al., 2011). Like bread wheat, oat is also a hexaploid able to tolerate the addition of an alien chromosome pair, and this property has been used to sort certain maize chromosomes (Li et al., 2001). Some plant species possess so called B chromosomes, whose evolution, function and molecular organization have long been controversial (Jones, 1995; Jones and Houben, 2003). They are typically much smaller than the standard chromosomes, and therefore are amenable to sorting (Kubaláková et al., 2003; Martis et al., 2012). A further example is represented by the dioecious species white campion, which carries a sex chromosome which differs in size from the rest of the chromosome complement, and can thus be sorted (Kejnovský et al., 2001; Veuskens et al., 1995).

Sorting specific chromosomes using an addition line is a convenient means of isolating a portion of the donor genome. However, the development of these lines is very laborious, so they can only ever be generated from a limited number of donors. A similar consideration relates to translocation and deletion lines. Many applications, however, focus on a

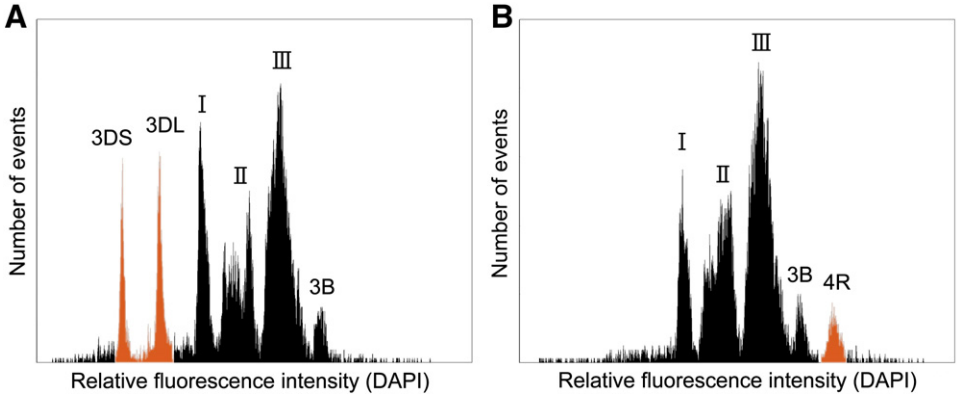

Fig. 4. The use of cytogenetic stocks to isolate particular wheat chromosomes. Flow karyotypes were obtained from DAPI-stained suspensions of mitotic chromosomes. (A) The double ditelosomic line dDt3D (20″ + t″3DS + t″3DL) carries the two arms of chromosome 3D in the form of two distinct telosomes, each of which is smaller than any of the 20 entire wheat chromosomes; these form discrete, sortable peaks. (B) The wheat-rye (Chinese Spring/Imperial) disomic addition line 4R (2n = 44; 21_W″ + 1_R″) forms peaks I–III and 3B, and a discrete, sortable peak harboring rye chromosome 4R.

specific accession of the donor species, for example because it harbors a specific gene or allele. Currently, two approaches have been elaborated to discriminate chromosomes without recourse to specialized cytogenetic stocks. In the first, composite peaks are divided into sections and those which are enriched for the chromosome of interest are retained (Vrána et al., submitted for publication). Although the purity level attained is necessarily lower than is achievable from well discriminated peaks, fractions with a contamination level as low as 20% can be prepared from composite peaks in wheat. Importantly, a majority of wheat chromosomes sorted in this way have proven to be free of contamination by homoeologs, which greatly simplifies sequence analysis. The second approach relies on the differential labeling of chromosomes, based on the presence of repetitive sequences. The earliest attempts to achieve this goal, as described by Macas et al. (1995), involved a modification of the PRINS (primed *in situ* DNA labeling) technique. While this did lead to some successful results (Pich et al., 1995), it was plagued by poor reproducibility and by non-quantitative labeling DNA repeats (unpublished data). Both suspended rye and barley chromosomes were labeled with fluorochromes by Ma et al. (2005), but no attempt was made to apply flow cytometry to these preparations. Finally, Giorgi et al. (2013) developed a reproducible method termed FISHIS (FISH in suspension), which differentially labels chromosomes by hybridizing with oligonucleotide probes targeting specific microsatellite sequences (Fig. 5). The successful binding of these probes may well be related either to their ability, as small molecules, to easily invade the chromosomes, or be the result of the formation of alternative B-DNA structures (Cuadrado and Jouve, 2010). As yet, it has not been established to what extent (if any) the FISHIS procedure damages chromosomal DNA and proteins, and hence with which downstream applications FISHIS-labeled chromosomes will be compatible.

Given a sorting speed 5–40 chromosomes per second (Doležel and Lucretti, 1995; Vrána et al., 2012), it is feasible to recover some 200,000 chromosomes per working day using a commercial flow-sorter (Šafář et al., 2010), a number sufficient to acquire microgram quantities of chromosome-specific DNA. The two major factors influencing the yield of sorted chromosomes are the level of resolution achievable and the quality of the initial sample (specifically, the overall number of intact chromosomes present and the amount of debris). Where aneuploid material is the source, yields can be reduced because the target chromosome is not represented in the disomic state in every seedling. The assignation of chromosome identity to flow karyotype peak is most conveniently achieved using a chromosome-specific PCR assay (Lysák et al., 1999; Vrána et al., 2000), particularly as such assays only require a small amount of DNA as a template. PCR assays are not, however, capable of estimating peak purity; in principle, this could be achieved using a quantitative PCR assay based on a set of primers designed to specifically recognize each chromosome in the genome. More straightforwardly, the chromosomal content of a given peak can be inspected by conventional microscopy following a PRINS or FISH labeling protocol (Kubaláková et al., 2000, 2005). Such an analysis of course requires a prior characterization of the karyotype.

4. Uses of flow-sorted chromosomes

Because the morphology of flow-sorted chromosomes isolated from formaldehyde-fixed root tips is well preserved (Doležel et al., 1992), high molecular weight DNA is readily derivable. As a result, flow-sorted plant chromosomes have proven valuable for a range of applications, including cytogenetic analysis, physical and genetic mapping and whole genome sequencing (Fig. 6).

4.1. Physical mapping

4.1.1. Mapping by PCR

As the template requirement for PCR is small, sorted chromosomes have proven to represent an elegant means of chromosomally assigning a given DNA sequence. This approach was adopted to map vicillin genes in field bean, since these genes were difficult to map genetically due to a paucity of allelic variation (Macas et al., 1993); similarly, genes mapping to the sex chromosome in white campion were successfully identified (Kejnovský et al., 2001; Matsunaga et al., 2003, 2005), and the genetic and physical maps of both garden pea (Neumann et al., 2002) and chickpea (Vláčilová et al., 2002; Zatloukalová et al., 2011) were successfully integrated. Macas et al. (1993) and Neumann et al. (2002) exploited sorted reciprocal translocation chromosomes in field bean and garden pea to locate a number of DNA sequences to their sub-chromosomal region. More recently, PCR amplification of template consisting of flow-sorted chromosomes has been used to develop DNA markers to support positional cloning (Šimková et al., 2011a, 2011b). Such an approach is particularly useful in allopolyploid species, where the development (and subsequent mapping) of low copy sequences can be complicated by the presence of three homoeologs.

Physical mapping applications which require a larger quantity of DNA of course require a more prolonged chromosome sorting effort, although where high molecular weight DNA is not needed, this can be avoided by the amplification of template derived from a modest number of sorted chromosomes. Šimková et al. (2008a) showed that microgram quantities of chromosomal DNA with a majority of fragments between 5 and 30 Kb can be produced using a multiple displacement amplification (MDA) protocol based on ϕ29 DNA polymerase. Starting with a 10 ng aliquot of DNA derived from a population of 10,000 barley

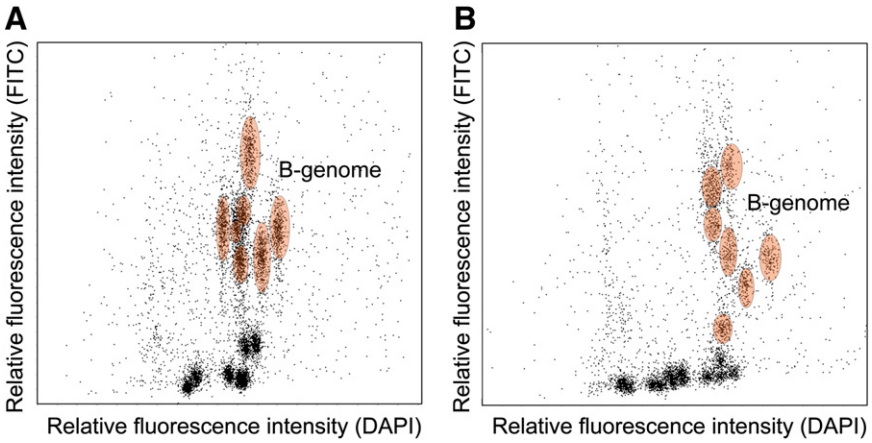

Fig. 5. Flow karyotypes of (A) tetraploid (cv. Creso, 2n = 28) and (B) hexaploid (cv. Chinese Spring, 2n = 42) wheat after the joint fluorescent labeling of GAA_n microsatellites and DAPI staining. The former was achieved by hybridization with a GAA_7-FITC probe, following the FISHIS procedure. The B genome chromosomes have a higher GAA content than either the A or D genome ones, and so can be discriminated on the basis of their higher FITC fluorescence (highlighted in orange).

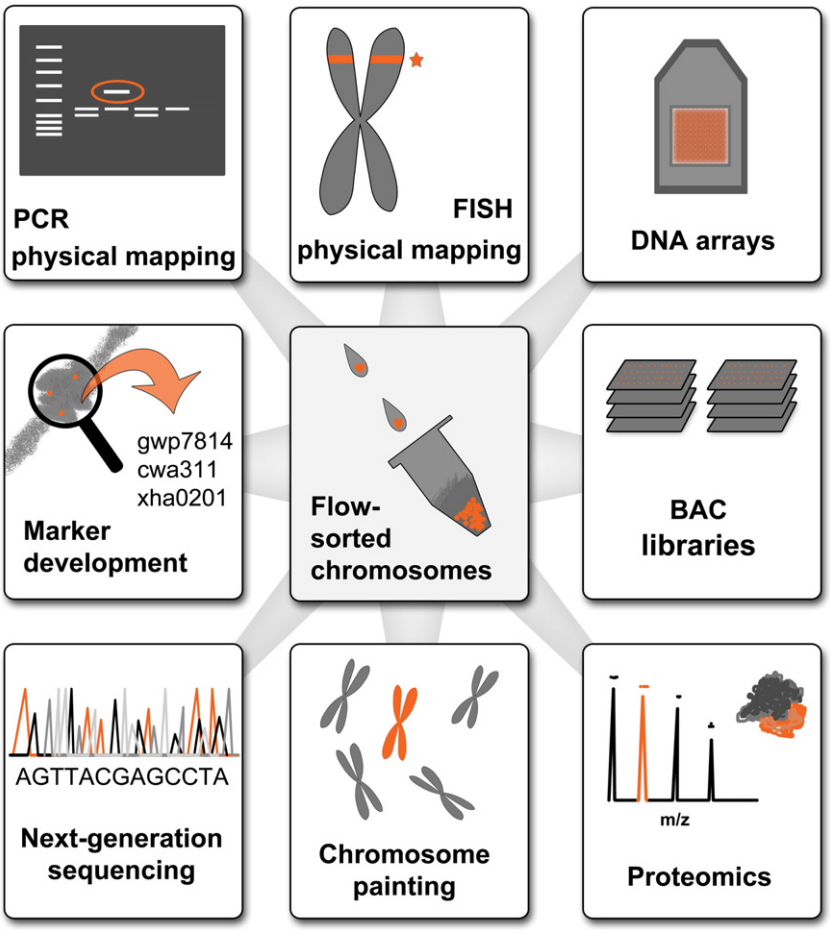

Fig. 6. Major current and potential uses of flow-sorted chromosomes.

chromosomes 1H, the amplification product proved to be very representative of the whole chromosome, since only 1.9% of SNP loci known to map to this chromosome failed to be recovered. On this basis, the chromosome 1H DNA pool was used infer a 1H location to 40 SNP loci which had hitherto not been mapped. When DNA was amplified in this way from each of the 12 individual arms of chromosomes 2H–7H maintained individually in wheat–barley telosome addition lines, 370 SNP loci which had not hitherto been genetically mapped were allocated a chromosome arm (Muñoz-Amatriaín et al., 2011). Prior to using 7H-specific simple sequence repeat (SSR) markers to characterize a spontaneous wheat–barley Robertsonian translocation, Cseh et al. (2011) were able to verify their chromosomal arm location by testing against a template of flow-sorted chromosome arms 7HS and 7HL. Chromosome sorting was also exploited for the positional cloning of a powdery mildew resistance gene located on wheat chromosome arm 4AL (Jakobson et al., 2012).

4.1.2. Construction of clone-based physical maps

The construction of a physical map as a template for either CBC sequencing or positional cloning requires large insert genomic DNA libraries, most commonly generated in the form of bacterial artificial chromosomes (BACs), which are able to accommodate an insert of up to several hundred Kb in length (Shizuya et al., 1992). The quantity of high molecular weight DNA required for this purpose is in the microgram range, so achieving this from flow-sorted material involved the elaboration of a customized protocol (Šimková et al., 2003). Using this protocol, Šafář et al. (2004) succeeded in constructing the first documented chromosome-specific BAC library of a eukaryotic organism; the chromosome involved was wheat 3B, and was sourced from a set of two million sorted chromosomes, prepared over 18 working days. The library comprised about 68,000 clones with a mean insert size of 103 Kb, and represented more than 6 x coverage of the chromosome. Further improvements to the protocol extended the coverage to >15x and the mean insert size to >120 Kb (IEB genomic resources database, 2013; Šafář et al., 2010). In addition to a number of wheat whole chromosome- and chromosome arm-specific BAC libraries, a library has also been constructed from the short arm of cereal rye chromosome 1R (Šimková et al., 2008b).

The International Wheat Genome Sequencing Consortium (IWGSC) has chosen a CBC chromosome-based strategy to produce a reference sequence of the wheat genome (Feuillet and Eversole, 2007) and so a chromosome-specific BAC library has been generated for each of the 21 chromosomes of the model cultivar Chinese Spring (IEB genomic resources database, 2013). The feasibility of constructing a physical map of each wheat chromosome based on such libraries was confirmed by the successful contig map of chromosome 3B produced by fingerprinting the 3B BAC library (Paux et al., 2008). The initial version of the map comprised just over 1000 contigs anchored with nearly 1500 molecular markers, and represented 82% of the chromosome. The lessons drawn from this exercise have been incorporated into the ongoing effort to establish a physical map for each of the remaining 20 wheat chromosomes (Lucas et al., 2013; Philippe et al., 2013; Sehgal et al., 2012). The sequencing of 13 of the 3B contigs involved over 150 BACs (Paux et al., 2008), and led to the annotation of >18 Mb of sequence. While the global gene density was found to be about one per 104 Kb, some 75% of the genes clustered into small groups (each containing on average three genes), and the density increased by two fold in regions close to the telomere, largely as a consequence of tandem and interchromosomal duplications. Using the same physical map, Rustenholz et al. (2011) were able to locate some 3,000 genes, distributed along the

whole chromosome, and a similar pattern of gene islands and greater gene density at the chromosome ends emerged. Most of the gene islands resulted from interchromosomal duplications specific to polyploid wheat and are enriched in genes sharing the same function or expression profile. Gene space organization and evolution proved to be similar on chromosome arm 1BL (Philippe et al., 2013). The definition of an MTP for both chromosome 3B and chromosome arm 3DS enabled Bartoš et al. (2012) to attempt a comparison of the molecular organization of these two homoeologs. What was revealed was a similar rate of non-collinear gene insertion, with the majority of duplications occurring prior to the divergence of the B and D genomes some 30 Mya. One third of insertions occurred during the past 2.5–4.5 My, leading to the suggestion that gene insertion was accelerated by allopolyploidisation. Pseudogenes appear to represent only a small fraction of the non-collinear genic sequence; for the most part, they seem to have arisen during the evolution of the polyploid wheat genome and not from insertion of non-functional genes.

Beyond their utility for acquiring the genome sequence of wheat, the chromosome-specific BAC libraries have found a number of other uses. The chromosome 7DL and 7DS libraries have been queried with markers linked to the aphid resistance genes *Dn2401* and *Gb3* (Šimková et al., 2011a, 2011b). Both PCR- and hybridization-based screening has demonstrated the gain in efficiency brought about by the reduction in complexity of the template. Thus, just three rounds of screening on three high density filters were sufficient to build a BAC contig spanning *Gb3*. To achieve the positional cloning of genes not present in cv. Chinese Spring (such as *Gb3*), other cultivars have been targeted for making chromosome-specific BAC libraries (IEB genomic resources database, 2013; Janda et al., 2006). Chromosome-specific BAC libraries are especially valuable in polyploids as they avoid the problem of homoeology. Additionally the necessary size of such libraries is an order of magnitude lower and so are more straightforward to store, handle and screen (Šimková et al., 2011a). Finally, the dissection of a large genome into its constituent chromosomes parts helps to structure collaborative projects where each of the various partners can be made responsible for the management of a specific chromosome(s), even though the BAC libraries have been generated centrally.

4.1.3. Cytogenetic mapping

Ordering and orienting BAC and sequence contigs is an important step in, respectively, building a clone-based physical map and assembling a shotgun sequence. Genetic markers are seldom helpful in proximal chromosome regions because these are associated with a low frequency of recombination. An alternative means of ordering is to apply FISH to mitotic or meiotic chromosomes (Karafiátová et al., 2013; Tang et al., 2009). A development of this idea is to apply FISH to mechanically stretched (by 100 fold), flow-sorted mitotic metaphase chromosomes (Valárik et al., 2004). The stretching greatly improves the achievable level of spatial resolution, to an extent where the individual probes can be ordered.

4.2. Genetic marker development

The reduction in template complexity achieved by targeting individual chromosomes has been beneficial for genetic marker development. Markers are a critical resource for the construction of genetic linkage maps, the understanding of trait inheritance, the assembly of physical maps and DNA shotgun sequences, and positional cloning. An array of marker types have been developed, the most ubiquitously used of which in plant genetics are SSRs, DArTs (diversity array technology), ISBPs (insertion site based polymorphisms) and SNPs (single nucleotide polymorphisms) (Poczai et al., 2013). Various strategies have been employed to base marker development on chromosome-specific libraries.

4.2.1. SSRs and ISBPs

An initial approach began by cloning the DNA derived from sorted chromosomes (generally following an amplification step) to generate small insert DNA libraries (Macas et al., 1996). The earliest significant marker type was hybridization-based, such as the RFLPs (restriction fragment length polymorphisms) generated on tomato chromosome 2 by Arumuganathan et al. (1994). With the advent of PCR, attention switched to SSRs. An enrichment was carried out on the initial library to bias the recovery of a target microsatellite motif (Koblížková et al., 1998). Požárková et al. (2002) developed a set of SSR markers from chromosome 1 of field bean, and some of these were later used to elaborate a genetic map of the species (Román et al., 2004). Kofler et al. (2008) employed a similar approach to develop 57 SSR markers from MDA-amplified DNA of rye chromosome arm 1RS, a source of a number of agronomically important genes for wheat (Lukaszewski, 1990). In addition to developing SSR markers from amplified 1RS DNA, Kofler et al. (2008) also developed 138 SSR assays from 2778 BAC end sequences (BES) obtained from the 1RS-specific BAC library The same set of BES included 249 transposable element junctions which could be exploited to produce 64 ISBP markers, of which 12 were 1RS specific (Bartoš et al., 2008). BES derived from chromosome-specific libraries have proven informative for marker development in wheat itself as well. For example, Lucas et al. (2012) identified 433 potential SSRs and 9,338 potential ISBP sequences from ~13,500 BES generated from chromosome arm 1AL. About one half of the putative ISBP markers tested proved to be functional. Similarly, among ~10,000 3AS BES, Sehgal et al. (2012) identified over 1,000 potential SSR and nearly 8,000 potential ISBP sequences, of which an estimated 18% and 29%, respectively, marked loci on 3AS.

4.2.2. DArT markers

Wenzl et al. (2010) demonstrated how useful chromosome sorting is to develop DArT markers to significantly increase saturation of linkage maps at specific genome regions. Using DNA of chromosome 3B and chromosome arm 1BS of wheat, the authors developed DArT arrays with 2,688 and 384 clones, respectively. Out of 711 polymorphic 3B-derived markers, 553 (78%) mapped to chromosome 3B, while 59 of 68 polymorphic 1BS-derived markers (87%) mapped to chromosome arm 1BS. Hence a majority of markers were specific to target chromosomes. The 3B DArT array was used in development of a new consensus genetic map of the chromosome, leading to doubling the number of genetically distinct loci on 3B. The efficiency of chromosome targeting can be estimated by comparing the 510 polymorphic 3B markers obtained by screening 2,688 3B-derived clones with 269 polymorphic markers identified by screening approximately 70,000 whole genome-derived clones (Wenzl et al., 2010). Coupling chromosome sorting with the DArT platform is straightforward, as the DNA requirement is only ~5 ng, a quantity which can be recovered in less than one hour of flow-sorting.

4.2.3. Marker development from chromosome-specific shotgun sequences

The combination of MDA-generated chromosome-specific DNA and high throughput sequencing platforms offers an efficient route towards whole genome shotgun sequencing and the *in silico* identification of genetic markers. The development of a SNP map of wheat chromosome 3B serves as a good example of the power of this approach (Shatalina et al., 2013). A set of 737 gene-containing contigs harboring chromosome 3B SNPs between the two cultivars Arina and Forno was selected, and a subset of 96 of these SNPs used to genotype an Arina x Forno recombinant inbred line population; of these, 70 mapped to the expected chromosome. The 454-derived sequence of rye chromosome arm 1RS allowed Fluch et al. (2012) to identify >4000 potential SSR loci, and similarly Nie et al. (2012) used Illumina-derived sequence of wheat chromosome arm 7DL to identify >16,000 putative SSR loci. When a random set of 33 of the latter was tested by PCR, 18 proved to be informative across a panel of 20 cultivars. Similarly, the 454-derived sequence produced from wheat chromosome arm 1BL (Wicker et al., 2011) was used by Philippe et al. (2013) to identify nearly 19,000 putative ISBPs and 200

SSRs. Finally, a comparison of homoeologous group 7 sequences across four Australian wheat cultivars located some 900,000 informative SNP loci (Berkman et al., 2013).

4.2.4. Marker specificity

A feature of the chromosome-based strategy is that it can save a substantial volume of screening effort, particularly in polyploid species. Thus, for example, Požárková et al. (2002) were able to use flow-sorted fractions as a PCR template to verify the chromosome specificity of SSR markers in filed bean. Michalak de Jimenez et al. (2013) used a radiation hybrid approach to map wheat chromosome 1D, exploiting DNA amplified from the homoeologous group 1 chromosomes as a source of 1D-specific markers. Shotgun sequences of each chromosome of barley (Mayer et al., 2011), rye (Martis et al., 2013) and bread wheat (K Eversole, pers. comm.) have now been acquired using either the Illumina or the 454 platform; thus it should be in future possible to rapidly verify chromosome-specificity *in silico* in these species.

4.3. Sequencing

4.3.1. BAC clones

NGS technology has the capacity to shotgun-sequence whole genomes, but the quality of genome assembly in large genome species is poor compared to that obtained using the CBC method, as used to derive the reference sequences of *A. thaliana*, rice and maize (Shangguan et al., 2013). Handling a genomic BAC library of a large genome species is cumbersome, because of the number of clones involved. Particular problems are associated with the presence of homoeology in polyploid genomes. A chromosome-based strategy at present represents the most promising one in these cases, and has been adopted for the acquisition of the hexaploid wheat genome sequence (The International Wheat Genome Sequencing Consortium, 2013); so far it has generated a 1 Gb reference sequence of chromosome 3B after sequencing its MTP using a combination of Roche 454 and Illumina technologies (Choulet et al., submitted for publication). The project of the International Wheat Genome Sequencing Consortium involves the construction of a full set of chromosome-specific BAC libraries, the definition of an MTP for each, and the CBC-sequencing of the MTP using NGS.

The availability of a number of chromosome-specific BAC libraries has already provided some interesting research opportunities. Bartoš et al. (2008) end-sequenced a random set of 1,536 clones from a BAC library specific for the short arm of rye chromosome 1R (1RS). The analysis of repeat content indicated a similar fraction of repeats as in the B genome of wheat (84%). However, as the rye genome is much larger (almost 8 Gb/1C vs. ~5.6 Gb/1C), a lower than expected proportion of repeats was probably due to insufficient representation of rye repeats in DNA sequence databases that were searched to identify repeats. Since only 0.9% of the 1RS derived BES were classified as genic sequences, it was estimated that the arm harbored about 2000 genes. A similar analysis of ~10,000 3AS BACs led to an estimate that the proportion of repetitive DNA present was 79% (Sehgal et al., 2012). About 1.4% of the DNA was estimated to represent coding sequence, producing an estimated 2,850 genes as present on the arm, the length of which is just 0.8 times the size of the entire rice genome, which is estimated to harbor over 45,000 genes (Yu et al., 2002). An increase in gene density towards the telomere was noted, and for up to 30% of the genes, synteny was not maintained with the rice, sorghum and *B. distachyon* genomes. Similarly, Lucas et al. (2012) used >13,000 1AL BES to characterize the composition of this chromosome arm, producing an estimate of ~1.0% for the proportion of the arm's DNA which represented coding sequence and a gene number of 4700. The analysis confirmed the presence of two known major synteny blocks (Mayer et al., 2009), as well as three smaller blocks not previously identified.

4.3.2. Whole chromosome sequencing using 454 technology

The combination of NGS technology and chromosome sorting currently represents the most affordable means of obtaining the sequence composition of single chromosomes. Generally, MDA-amplified DNA, which typically generates fragments in the size range 5–30 Kb (Šimková et al., 2008a), is suitable for this purpose. However, it is unsuitable for constructing paired-end and mate-pair libraries with insert sizes >3 Kb (Belova et al., 2013). If longer insert sequencing libraries are needed, the amplification step should be avoided and a larger number of chromosomes need to be sorted. Amplified chromosomal DNA from barley chromosome 1H was sequenced using the 454 technology by Mayer et al. (2009). Comparison of the sequences with genes of rice and sorghum and with EST datasets of barley and wheat identified 5400 genes. Based on the integration with synteny data from the two grass model species, the authors proposed a virtually ordered inventory of 1987 genes and their work increased the number of 1H anchored genes by 6-fold compared to previous map resources. Mayer et al. (2011) exploited the same approach by adding low-pass 454-acquired sequence from the other barley chromosomes, incorporating at the same time all available full length cDNA sequence and DNA microarray hybridization data. The result was a sequence-based gene map of barley capturing an estimated 86% of the total gene content. This so-called "Genome Zipper" approach is illustrated in Fig. 7, and has succeeded in precisely localizing six of the seven barley centromeres, and established gene order in the poorly recombining proximal chromosome regions. Due to its relative simplicity, Genome Zipper is an attractive approach for all species, whose genomes have not been sequenced and in which chromosomes can be isolated by flow-sorting.

The possibility of sequencing all six arms of the wheat group 1 homoeologs allowed Wicker et al. (2011) to make structural comparisons at the single chromosomal level. Analysis of sequences from low-pass sequencing with Roche 454 technology (1.3- to 2.2x chromosome coverage) indicated that all three wheat subgenomes have similar sets of genes that are syntenic with model grass genomes. However, the number of genic sequences that have their homologs outside the group 1 syntenic region in the grass models outnumbers the syntenic ones. Further analysis indicated that a large proportion of the genes that are found in only one of the three homoeologous wheat chromosomes were most probably pseudogenes resulting from transposon activity and double strand break repair. The 1A sequences were later used by Lucas et al. (2013) to produce a virtual gene order along chromosome arm 1AL, adopting the Genome Zipper approach, and this was readily integrated into a physical map of the arm. The analysis confirmed the presence of non-syntenic genes and identified some putative translocations.

Vitulo et al. (2011) characterized the content of wheat chromosome 5A by acquiring 454-derived sequence from each arm. Their estimate was that coding sequence represented 1.1% of 5AS and 1.3% of 5AL, leading to the prediction that the whole chromosome harbors just over 5,000 genes. Similarly, Hernandez et al. (2012) analyzed chromosome 4A, a chromosome which has undergone a major series of evolutionary re-arrangements (Devos et al., 1995). Application of the Genome Zipper method produced a virtual gene map capturing at least 85% of the chromosome's estimated gene content. A comparison with the maps of barley chromosomes 4H, 5H and 7H identified and ordered five distinct regions (Fig. 8), the gene content and order within each of which being inferred from synteny. A 454-derived sequence of both arms of chromosome 3A recognized over 3500 contigs (Akhunov et al., 2013). A comparison with the equivalent sequences of the model grass genomes detected that some 35% of genes had experienced structural rearrangements leading to a variety of mis-sense and non-sense mutations. In particular, 38% of these genes were affected by a premature stop codon, which is on line with other studies indicating ongoing pseudogenization of the wheat genome. Alternative splicing patterns were diverse between homoeologs, perhaps an effect of the genetic redundancy resulting from polyploidy.

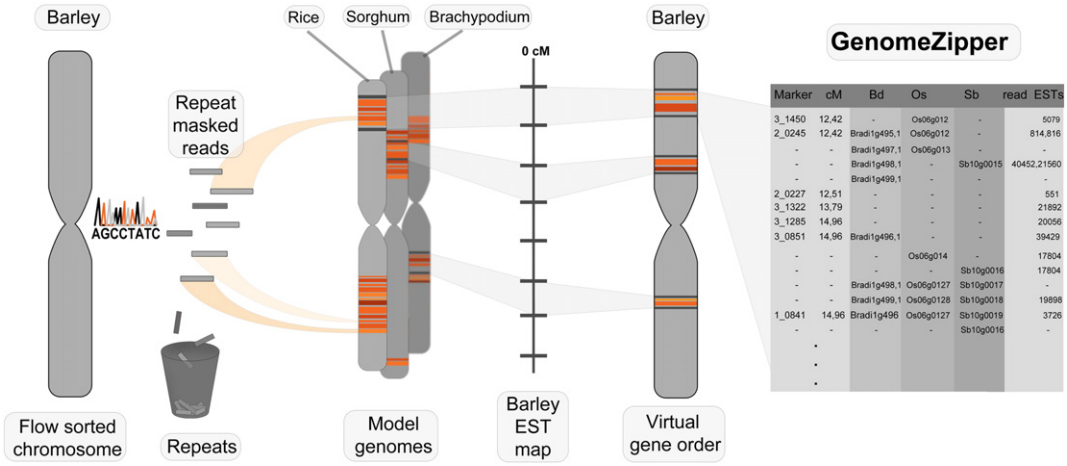

Fig. 7. Genome Zipper analysis in barley chromosome. MDA-amplified DNA of flow-sorted barley chromosomes was sequenced by 454 technology. Repeat-masked sequence reads identified syntenic regions in the rice, sorghum and *B. distachyon* genomes. Genes located in these regions were then aligned with a EST-based barley map of barley, which served as a scaffold to anchor collinear segments derived from the non-barley genomes. Genic sequence reads of barley were integrated and ordered by assuming collinearity within syntenic regions, leading to the derivation of a virtual gene map of barley.

The 454 technology has also been applied to sequence 1RS chromosome arm of rye, revealing the presence of just over 3000 gene loci and identifying syntenic regions in model genomes of rice and brachypodium, and in barley chromosome 1H (Fluch et al., 2012). The subsequent 454-based sequencing of all chromosomes of rye established their virtual linear gene order models (genome zippers) comprising over 22,000 or 72% of the detected set of ~31,000 rye genes (Martis et al., 2013). The study revealed six major translocations that shaped the modern rye genome in comparison to a putative Triticeae ancestral genome. Moreover, the results indicated that introgressive hybridizations and/or a series of whole-genome or chromosome duplications played a role in rye speciation and genome evolution.

A very attractive application of flow cytometric sorting is to isolate specialized chromosomes such as sex chromosomes and supernumerary B chromosomes. Since B chromosomes act as a selfish genetic element, they have been proposed as a vehicle for chromosome-mediated gene transfer (Birchler et al., 2008). The structure of rye B chromosomes has been elucidated by sequencing flow-sorted material using the 454 platform (Martis et al., 2012). Although they have long been considered to be gene poor (Jones, 1995; Jones and Houben, 2003), a sequence alignment with rice, *B. distachyon*, sorghum and barley genomic sequence identified the presence of almost 5000 putative gene fragments. A strong indication was that their DNA probably originated from both chromosome arm 3RS and chromosome 7R, although the sequence appears to have been subjected to complex rearrangement. Molecular clock-based dating of the rye B chromosomes' origin places it at 1.1–1.3 Mya, which is not long after the formation of the genus *Secale* (1.7 Mya).

4.3.3. Whole chromosome sequencing using Illumina technology

The initial attempts at shotgun sequencing of flow-sorted plant chromosomes were based on the 454 platform, which generates read lengths of several hundred nucleotides. With the development of the Illumina platform, Berkman et al. (2011) were able to demonstrate that short read sequencing technology could equally be used for chromosome shotgun sequencing and subsequent assembly. Thus, a coverage of >30× was achieved for chromosome arm 7DS, and the subsequent assembly comprised over 550,000 contigs (up to 32.6 Kb in length) with an N50 of 1159 bp. The coverage represented approximately 40% of the whole arm, since much of the repetitive DNA collapsed into a single contig. A comparison with the *B. distachyon* sequence identified nearly 1,500 genes, of which about one in three were non-syntenous. A comparison with bin-mapped wheat ESTs (Qi et al., 2004) highlighted possible erroneous allocations, with the result that the 7DS assembly probably captured all or nearly all of the arm's gene content. The same approach was used to sequence and assemble chromosome arm 7BS (Berkman et al., 2012). A comparison between the assemblies of 7DS, 7BS and 4AL recognized the known evolutionary translocation between chromosomes 7B and 4A and closely defined its break-point. The level of collinearity between 7BS and 7DS was 84%, while that between the wheat and *B. distachyon* was 60%. Extending the approach to cover the whole of the group 7 homoeologs showed that there has been more gene loss in 7A and 7B than in 7D (Berkman et al., 2013).

Micro RNAs (miRNAs) are an important component of post-transcriptional gene regulation, so their distribution at the chromosome

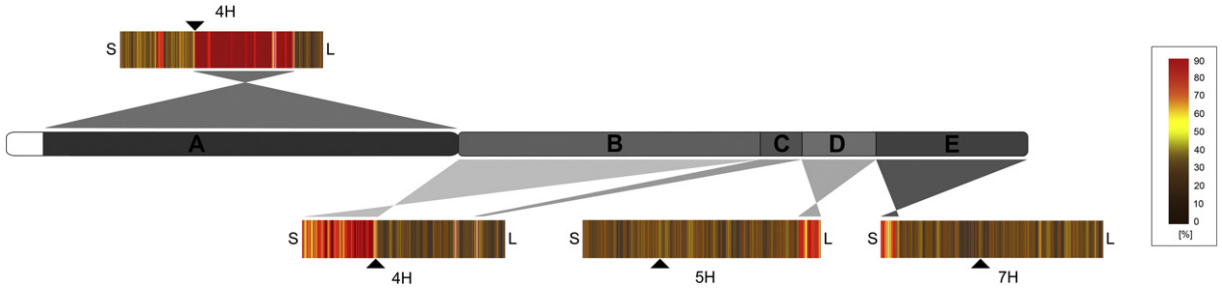

Fig. 8. The 4A shotgun sequence of barley. Repeat-masked 4AS and 4AL shotgun sequence reads were compared with the sequence of virtual barley chromosomes (Mayer et al., 2011). Syntenic regions on chromosomes 4H, 7H and 5H are shown in red, and non-syntenic regions in brown. The centromeres are indicated by black triangles and the chromosome arms are labeled S and L. Connectors indicate corresponding segments and the orientation of the individual segments. Taken with permission from Hernandez et al., Plant Journal 2012;69:377–86, John Wiley & Sons Ltd. Modified.

level is expected to provide novel insights into genome organization and function. Vitulo et al. (2011) used chromosome shotgun sequences to identify 195 candidate miRNA precursors belonging to 16 miRNA families on chromosome 5A, while Kantar et al. (2012), focusing on chromosome 4A, found 68 different miRNAs of which 37 had not been observed previously in wheat. The two chromosome arms differed with respect to both the variety and representation of miRNAs. Among the 62 putative targets identified, 24 were found to give hits to expressed sequences.

4.3.4. Validation of whole genome assemblies

Many genomes have already been sequenced using NGS shotgun approach, and it is not realistic to expect that they will be sequenced again following CBC strategy. Additional approaches are therefore needed to improve the assemblies. These may include improved bioinformatics tools for whole genome assembly, incorporation of sequences obtained using methods resulting in longer reads (Roberts et al., 2013), optical mapping (Dong et al., 2013) and mapping on nanochannel arrays (Hastie et al., 2013; Lam et al., 2012). Cytogenetic mapping has a role to play in the verification of sequence assemblies (Febrer et al., 2010; Islam-Faridi et al., 2009). However, a powerful option is to sequence isolated chromosomes using NGS and compare chromosome-derived sequences with whole genome assemblies. Preliminary results obtained with genome assemblies of two types of chickpea (Jain et al., 2013; Varshney et al., 2013) highlighted regions that appear to have been mis-assembled and provided the basis for genome assembly improvement (R. Varshney and D. Edwards, pers. comm.). Thus, chromosome genomics can be employed in genome sequencing projects to validate and assist in the accurate sequence assemblies obtained by NGS shotgun.

5. Conclusions

The recent past few years have witnessed marked progress in chromosome genomics, a technology which has rapidly established itself as a facilitator of mapping and sequencing of plant genomes. The number of species tractable to flow-sorting has expanded, confirming the broad applicability of suspensions of intact chromosomes obtained from synchronized root tips (Doležel et al., 1992). The development of the FISHIS technique (Giorgi et al., 2013) should expand the reach of flow-sorting, since it provides a powerful means of discriminating between chromosomes which are similar in size, thereby easing the dissection of complete genomes into their individual chromosome components. There has also been a notable increase in the number and variety of applications using flow-sorted chromosomes, driven most importantly by the step change in sequencing power achieved by NGS technologies, but also by the possibility of producing microgram quantities of chromosomal DNA via MDA. Chromosome genomics has been especially useful in species lacking a reference genome sequence. The analysis of sequence at the single chromosome level has provided new insights into the structure of complex, and particularly polyploid genomes, where comparisons between homoeologs has informed the process of genome evolution in a polyploid setting. Sequencing single chromosomes has been highly productive in the context of marker development and validation. Finally, chromosome-specific shotgun sequences are proving to represent a convenient means of verifying genome sequence assemblies, of identifying candidate genes and of analyzing the organization and evolution of specialized chromosomes such as sex chromosomes and supernumerary B chromosomes.

The chromosome genomics approach has been particularly fruitful in the wheat genome, the analysis of which using a whole genome approach is hampered by the size of the genome and the presence of homoeologs. The current international effort coordinated by IWGSC to sequence the wheat genome has therefore been largely based on the construction of ready-to-sequence chromosome arm-specific BAC libraries. The experience gained in this task already suggests that chromosome genomics can contribute materially to the analysis of genomes lacking a high quality reference sequence. A number of potential applications still remain to be addressed. A prime example is chromosome mapping on nanochannel arrays (Lam et al., 2012), the availability of which would ease the initial assembly and validation of genome sequences. The organization of the chromosomes during interphase and their behavior during most of both mitosis and meiosis are difficult to unravel in large genome species in the absence of chromosome painting probes; isolated single chromosomes would certainly offer an excellent opportunity to develop these. As the function of the nuclear genome is intimately linked to DNA organization and the architecture of the interphase nucleus, there is also a need to study chromatin proteins and their dynamics. A proteomic analysis of flow-sorted chromosomes should represent an attractive approach to study chromatin free of contaminating cytoplasmic components.

Acknowledgements

We thank Drs. Eva Hřibová and David Kopecký for their help with FISHIS. We appreciate the useful comments made by the three anonymous reviewers. This research was supported by the Czech Science Foundation (awards P501/10/1740, P501/10/1778 and P501/12/G090) and by the Ministry of Education, Youth and Sports of the Czech Republic and the European Regional Development Fund (Operational Programme Research and Development for Innovations No. ED0007/01/01).

References

Akhunov ED, Sehgal S, Liang H, Wang S, Akhunova AR, Kaur G, et al. Comparative analysis of syntenic genes in grass genomes reveals accelerated rates of gene structure and coding sequence evolution in polyploid wheat. Plant Physiol 2013;161:252–65.
Arumuganathan K, Slattery JP, Tanksley SD, Earle ED. Preparation and flow cytometric analysis of metaphase chromosomes of tomato. Theor Appl Genet 1991;82:101–11.
Arumuganathan K, Martin GB, Telenius H, Tanksley SD, Earle ED. Chromosome 2-specific DNA clones from flow-sorted chromosomes of tomato. Mol Gen Genet 1994;242:551–8.
Bartoš J, Paux E, Kofler R, Havránková M, Kopecký D, Suchánková P, et al. A first survey of the rye (*Secale cereale*) genome composition through BAC end sequencing of the short arm of chromosome 1R. BMC Plant Biol 2008;8:95.
Bartoš J, Vlček Č, Choulet F, Džunková M, Cviková K, Šafář J, et al. Intraspecific sequence comparisons reveal similar rates of non-collinear gene insertion in the B and D genomes of bread wheat. BMC Plant Biol 2012;12:155.
Belova T, Zhan B, Wright J, Caccamo M, Asp T, Šimková H, et al. Integration of mate pair sequences to improve shotgun assemblies of flow-sorted chromosome arms of hexaploid wheat. BMC Genomics 2013;14:222.
Berkman PJ, Skarshewski A, Lorenc MT, Lai K, Duran Ch, Ling EYS, et al. Sequencing and assembly of low copy and genic regions of isolated *Triticum aestivum* chromosome arm 7DS. Plant Biotechnol J 2011;9:768775.
Berkman PJ, Skarshewski A, Manoli S, Lorenc MT, Stiller J, Smits L, et al. Sequencing wheat chromosome arm 7BS delimits the 7BS/4AL translocation and reveals homoeologous gene conservation. Theor Appl Genet 2012;124:423–32.
Berkman PJ, Visendi P, Lee HC, Stiller J, Manoli S, Lorenc MT, et al. Dispersion and domestication shaped the genome of bread wheat. Plant Biotechnol J 2013;11:564–71.
Birchler JA, Yu W, Han F. Plant engineered minichromosomes and artificial chromosome platform. Cytogenet Genome Res 2008;120:228–32.
Chen W, Kalscheuer V, Tzschach A, Menzel C, Ullmann R, Schulz MH, et al. Mapping translocation breakpoints by next-generation sequencing. Genome Res 2008;18:1143–9.
Choulet J, Alberti A, Theil S, Glover N, Barbe V, Daron J, et al. A reference sequence of wheat chromosome 3B reveals structural and functional compartmentalization. Science 2014. (submitted for publication).
Conia J, Bergounioux C, Perennes C, Muller P, Brown S, Gadal P. Flow cytometric analysis and sorting of plant chromosomes from *Petunia hybrida* protoplasts. Cytometry 1987;8:500–8.
Conia J, Muller P, Brown S, Bergounioux C, Gadal P. Monoparametric models of flow cytometric karyotypes with spreadsheet software. Theor Appl Genet 1989;77:295–303.
Cseh A, Kruppa K, Molnár I, Rakszegi M, Doležel J, Molnár-Láng M. 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 2011;54:795–804.
Cuadrado A, Jouve N. Chromosomal detection of simple sequence repeats (SSRs) using nondenaturating FISH (ND-FISH). Chromosoma 2010;119:495–503.
de Laat AMM, Blaas J. Flow-cytometric characterization and sorting of plant chromosomes. Theor Appl Genet 1984;67:463–7.
de Laat AMM, Schel JHN. The integrity of metaphase chromosomes of *Haplopappus gracilis* (Nutt) Gray isolated by flow cytometry. Plant Sci 1986;47:145–51.

Devos KM, Dubcovsky J, Dvorak J, Chinoy CN, Gale MD. Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor Appl Genet 1995;91:282–8.
Doležel J, Lucretti S. High-resolution flow karyotyping and chromosome sorting in *Vicia faba* lines with standard and reconstructed karyotypes. Theor Appl Genet 1995;90: 797–802.
Doležel J, Číhalíková J, Lucretti S. A high-yield procedure for isolation of metaphase chromosomes from root tips of *Vicia faba* L. Planta 1992;188:93–8.
Doležel J, Lucretti S, Schubert I. Plant chromosome analysis and sorting by flow cytometry. Crit Rev Plant Sci 1994;13:275–309.
Doležel J, Kubaláková M, Paux E, Bartoš J, Feuillet C. Chromosome-based genomics in cereals. Chromosome Res 2007a;15:51–66.
Doležel J, Greilhuber J, Suda J. Flow cytometry with plants: an overview. In: Doležel J, Greilhuber J, Suda J, editors. Flow cytometry with plant cells. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2007b. p. 41–65.
Doležel J, Kubaláková M, Číhalíková J, Suchánková P, Šimková H. Chromosome analysis and sorting using flow cytometry. In: Birchler JA, editor. Plant chromosome engineering. Methods and protocols. New York: Springer; 2011. p. 221–38.
Dong Y, Xie M, Jiang Y, Xiao N, Du X, Zhang W, et al. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). Nat Biotechnol 2013;31:135–41.
Dudin G, Steegmayer EW, Vogt P, Schnitzer H, Diaz E, Howell KE, et al. Sorting of chromosomes by magnetic separation. Hum Genet 1988;80:111–6.
Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Edward S, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 2011;6:e19379.
Febrer M, Goicoechea JL, Wright J, McKenzie N, Song X, Lin J, et al. An integrated physical, genetic and cytogenetic map of *Brachypodium distachyon*, a model system for grass research. PLoS One 2010;5:e13461.
Feuillet C, Eversole K. Physical mapping of the wheat genome: a coordinated effort to lay the foundation for genome sequencing and develop tools for breeders. Isr J Plant Sci 2007;55:307–13.
Feuillet C, Leach JE, Rogers J, Schnable PS, Eversole K. Crop genome sequencing: lessons and rationales. Trends Plant Sci 2011;16:77–88.
Ferguson-Smith MA. Genetic analysis by chromosome sorting and painting: phylogenetic and diagnostic applications. Eur J Hum Genet 1997;5:253–65.
Fluch S, Kopecky D, Burg K, Šimková H, Taudien S, Petzold A, et al. Sequence composition and gene content of the short arm of rye (*Secale cereale*) chromosome 1. PLoS One 2012;7:e30784.
Fuchs J, Houben A, Brandes A, Schubert I. Chromosome 'painting' in plants—A feasible technique? Chromosoma 1996;104:315–20.
Gerstein MB, Kundaje A, Hariharan M, Landt SG, Yan KK, Cheng C, et al. Architecture of the human regulatory network derived from ENCODE data. Nature 2012;489:91–100.
Gill KS, Arumuganathan K, Le JH. Isolating individual wheat (*Triticum aestivum*) chromosome arm by flow cytometric analysis of ditelosomic lines. Theor Appl Genet 1999;98:1248–52.
Giorgi D, Farina A, Grosso V, Gennaro A, Ceoloni C, Lucretti S. FISHIS: fluorescence *in situ* hybridization in suspension and chromosome flow sorting made easy. PLoS One 2013;8:e57994.
Grosso V, Farina A, Gennaro A, Giorgi D, Lucretti S. Flow sorting and molecular cytogenetic identification of individual chromosomes of *Dasypyrum villosum* L. (*H. villosa*) by a single DNA probe. PLoS One 2012;7:e50151.
Gualberti G, Doležel J, Macas J, Lucretti S. Preparation of pea (*Pisum sativum* L.) chromosome and nucleus suspensions from single root tips. Theor Appl Genet 1996;92: 744–51.
Guo DW, Ma YZ, Li LC, Chen YF. Flow sorting of wheat chromosome arms from the ditelosomic line 7BL. Plant Mol Biol Rep 2006;24:23–31.
Hastie AR, Dong L, Smith A, Finklestein J, Lam ET, Naxin Huo N, et al. Rapid genome mapping in nanochannel arrays for highly complete and accurate *de novo* sequence assembly of the complex *Aegilops tauschii* genome. PLoS One 2013;8:e55864.
Hernandez P, Martis M, Dorado G, Pfeifer M, Gálvez S, Schaaf S, et al. Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. Plant J 2012;69:377–86.
Ibarra-Laclette E, Lyons E, Hernández-Guzmán G, Pérez-Torres CA, Carretero-Paulet L, Chang TH, et al. Architecture and evolution of a minute plant genome. Nature 2013;498:94–8.
Islam-Faridi MN, Nelson CD, DiFazio SP, Gunter LE, Tuskan GA. Cytogenetic analysis of Populus trichocarpa–ribosomal DNA, telomere repeat sequence, and marker-selected BACs. Cytogenet Genome Res 2009;125:74–80.
Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 2007;449:463–7.
Jain M, Misra G, Patel RK, Priya P, Jhanwar S, Khan AW, et al. A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). Plant J 2013;74:715–29.
Jakobson I, Reis D, Tiidema A, Peusha H, Timofejeva L, Valárik M, et al. Fine mapping, phenotypic characterization and validation of non-race-specific resistance to powdery mildew in a wheat–*Triticum militinae* introgression line. Theor Appl Genet 2012;125:609–23.
Janda J, Šafář J, Kubaláková M, Bartoš J, Kovářová P, Suchánková P, et al. Advanced resources for plant genomics: BAC library specific for the short arm of wheat chromosome 1B. Plant J 2006;47:977–86.
Jones RN. Tansley review no. 85. B chromosomes in plants. New Phytol 1995;131:411–34.
Jones N, Houben A. B chromosomes in plants: escapees from the A chromosome genome? Trends Plant Sci 2003;8:417–23.
Kantar M, Akpınar BA, Valárik M, Lucas SJ, Doležel J, Hernández P, et al. Subgenomic analysis of microRNAs in polyploid wheat. Funct Integr Genomics 2012;12:465–79.
Karafiátová M, Bartoš J, Kopecký D, Ma L, Sato K, Houben A, et al. Mapping nonrecombining regions in barley using multicolor FISH. Chromosome Res 2013;21:739–51.
Kato A. Air drying method using nitrous oxide for chromosome counting in maize. Biotech Histochem 1999;74:160–6.
Kejnovský E, Vrána J, Matsunaga S, Souček P, Široký J, Doležel J, et al. Localization of male-specifically expressed MROS genes of *Silene latifolia* by PCR on flow-sorted sex chromosomes and autosomes. Genetics 2001;158:1269–77.
Koblížková A, Doležel J, Macas J. Subtraction with 3′ modified oligonucleotides eliminates amplification artifacts in DNA libraries enriched for microsatellites. Biotechniques 1998;25:32–8.
Kofler R, Bartoš J, Gong L, Stift G, Suchánková P, Šimková H, et al. Development of microsatellite markers specific for the short arm of rye (*Secale cereale* L.) chromosome 1. Theor Appl Genet 2008;117:915–26.
Kopecký D, Martis M, Číhalíková J, Hřibová E, Vrána J, Bartoš J, et al. Genomics of meadow fescue chromosome 4 F. Plant Physiol 2013;163:1323–37.
Kovářová P, Navrátilová A, Macas J, Doležel J. Chromosome analysis and sorting in *Vicia sativa* using flow cytometry. Biol Plant 2007;51:43–8.
Kubaláková M, Lysák MA, Vrána J, Šimková H, Číhalíková J, Doležel J. Rapid identification and determination of purity of flow-sorted plant chromosomes using C-PRINS. Cytometry 2000;41:102–8.
Kubaláková M, Vrána J, Číhalíková J, Šimková H, Doležel J. Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 2002;104: 1362–72.
Kubaláková M, Valárik M, Bartoš J, Vrána J, Číhalíková J, Molnár-Láng M, et al. Analysis and sorting of rye (*Secale cereale* L.) chromosomes using flow cytometry. Genome 2003;46:893–905.
Kubaláková M, Kovářová P, Suchánková P, Číhalíková J, Bartoš J, Lucretti S, et al. Chromosome sorting in tetraploid wheat and its potential for genome analysis. Genetics 2005;170:823–9.
Lam ET, Hastie A, Lin C, Ehrlich D, Das SK, Austin MD, et al. Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. Nat Biotechnol 2012;30:771–6.
Langlois RG, Yu LC, Gray JW, Carrano AV. Quantitative karyotyping of human chromosomes by dual beam flow cytometry. Proc Natl Acad Sci U S A 1982;79:7876–80.
Lee JH, Arumuganathan K. Metaphase chromosome accumulation and flow karyotypes in rice (*Oryza sativa* L.) root tip meristem cells. Mol Cells 1999;9:436–9.
Lee JH, Arumuganathan K, Kaeppler SM, Papa CM, Kaeppler HF. Cell synchronization and isolation of metaphase chromosomes from maize (*Zea mays* L.) root tips for flow cytometric analysis and sorting. Genome 1996;39:697–703.
Lee JH, Arumuganathan K, Yen Y, Kaeppler S, Kaeppler H, Baenziger PS. Root tip cell cycle synchronization and metaphase-chromosome isolation suitable for flow sorting in common wheat (*Triticum aestivum* L.). Genome 1997;40:633–8.
Lee JH, Arumuganathan K, Chung YS, Kim KY, Chung WB, Bae KS, et al. Flow cytometric analysis and chromosome sorting of barley (*Hordeum vulgare* L.). Mol Cells 2000;10:619–25.
Lee JH, Arumuganathan K, Kaeppler SM, Park SW, Kim KY, Chung YS, et al. Variability of chromosomal DNA contents in maize (*Zea mays* L.) inbred and hybrid lines. Planta 2002;215:666–71.
Leitch AR, Schwarzacher T, Wang ML, Leitch IJ, Surlan-Momirovich G, Moore G, et al. Molecular cytogenetic analysis of repeated sequences in a long term wheat suspension culture. Plant Cell Tissue Organ Cult 1993;33:287–96.
Li IJ, Arumuganathan K, Rines HW, Phillips RL, Riera-Lizarazu O, Sandhu D, et al. Flow cytometric sorting of maize chromosome 9 from an oat-maize chromosome addition line. Theor Appl Genet 2001;102:658–63.
Li L, Arumuganathan K, Gill KS, Song Y. Flow sorting and microcloning of maize chromosome 1. Hereditas 2004;141:55–60.
Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A, et al. Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. Nat Genet 2013;45:884–90.
Lucas SJ, Šimková H, Šafář J, Jurman I, Cattonaro F, Vautrin S, et al. Functional features of a single chromosome arm in wheat (1AL) determined from its structure. Funct Integr Genomics 2012;12:173–82.
Lucas SJ, Akpınar BA, Kantar M, Weinstein Z, Aydınoglu F, Šafář J, et al. Physical mapping integrated with syntenic analysis to characterize the gene space of the long arm of wheat chromosome 1A. PLoS One 2013;8:e59542.
Lucretti S, Doležel J. Bivariate flow karyotyping in broad bean (*Vicia faba*). Cytometry 1997;28:236–42.
Lucretti S, Doležel J, Schubert I, Fuchs J. Flow karyotyping and sorting of *Vicia faba* chromosomes. Theor Appl Genet 1993;85:665–72.
Lukaszewski AJ. Frequency of 1RS. 1AL and 1RS. 1BL translocations in United States wheats. Crop Sci 1990;30:1151–3.
Lysák MA, Číhalíková J, Kubaláková M, Šimková H, Künzel G, Doležel J. Flow karyotyping and sorting of mitotic chromosomes of barley (*Hordeum vulgare* L.). Chromosome Res 1999;7:431–44.
Ma YZ, Lee JH, Li LC, Uchiyama S, Ohmido N, Fukui K. Fluorescent labeling of plant chromosomes in suspension by FISH. Genes Genet Syst 2005;80:35–9.
Macas J, Doležel J, Lucretti S, Pich U, Meister A, Fuchs J, et al. Localization of seed protein genes on flow-sorted field bean chromosomes. Chromosome Res 1993;1:107–15.
Macas J, Doležel J, Gualberti G, Pich U, Schubert I, Lucretti S. Primer-induced labelling of pea and field bean chromosomes in situ and in suspension. Biotechniques 1995;19:402–8.
Macas J, Gualberti G, Nouzová M, Samec P, Lucretti S, Doležel J. Construction of chromosome-specific DNA libraries covering the whole genome of field bean (*Vicia faba* L.). Chromosome Res 1996;4:531–9.
Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, et al. Target-enrichment strategies for next-generation sequencing. Nat Methods 2010;7: 111–8.

Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, et al. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. Proc Natl Acad Sci U S A 2012;109:13343–6.
Martis MM, Zhou R, Haseneyer G, Schmutzer T, Vrána Jan, Kubaláková M, et al. Reticulate Evolution of the Rye Genome. Plant Cell 2013;25:3685–98.
Marx V. The genome jigsaw. Nature 2013;501:263–8.
Matsumoto T, Wu JZ, Kanamori H, et al. The map-based sequence of the rice genome. Nature 2005;436:793–800.
Matsunaga S, Kawano S, Michimoto T, Higashiyama T, Nakao S, Sakai A, et al. Semi-automatic laser beam microdissection of the Y chromosome and analysis of Y chromosome DNA in a dioecious plant, *Silene latifolia*. Plant Cell Physiol 1999;40:60–8.
Matsunaga S, Isono E, Kejnovský E, Vyskot B, Doležel J, Kawano S, et al. Duplicative transfer of a MADS box gene to a plant Y chromosome. Mol Biol Evol 2003;20:1062–9.
Matsunaga S, Lebel-Hardenack S, Kejnovský E, Vyskot B, Grant SR, Kawano S. An anther- and petal-specific gene SlMF1 is a multicopy gene with homologous sequences on sex chromosomes. Genes Genet Syst 2005;80:395–401.
Mayer KFX, Taudien S, Martis M, Šimková H, Suchánková P, Gundlach H, et al. Gene content and virtual gene order of barley chromosome 1H. Plant Physiol 2009;151:496–505.
Mayer KFX, Martis M, Hedley PE, Šimková H, Liu H, Morris JA, et al. Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell 2011;23:1249–63.
Meyers BC, Scalabrin S, Morgante M. Mapping and sequencing complex genomes: let's get physical. Nat Rev Genet 2004;5:578–88.
Michalak de Jimenez MK, Bassi FM, Ghavami F, Simons K, Dizon R, Seetan RI, et al. A radiation hybrid map of chromosome 1D reveals synteny conservation at a wheat speciation locus. Funct Integr Genomics 2013;13:19–32.
Molnár I, Kubaláková M, Šimková H, Cseh A, Molnár-Láng M, Doležel J. Chromosome isolation by flow sorting in Aegilops umbellulata and Ae. comosa and their allotetraploid hybrids Ae. biuncialis and Ae. geniculata. PLoS One 2011;6:e27708.
Muñoz-Amatriaín M, Moscou MJ, Bhat PR, Svensson JT, Bartoš J, Suchánková P, et al. An improved consensus linkage map of barley based on flow-sorted chromosomes and single nucleotide polymorphism markers. Plant Genome 2011;4:238–49.
Neumann P, Lysák M, Doležel J, Macas J. Isolation of chromosomes from *Pisum sativum* L. hairy root cultures and their analysis by flow cytometry. Plant Sci 1998;137:205–15.
Neumann P, Požárková D, Vrána J, Doležel J, Macas J. Chromosome sorting and PCR-based physical mapping in pea (*Pisum sativum* L.). Chromosome Res 2002;10:63–71.
Nie X, Li B, Wang L, Liu P, Biradar SS, Li T, et al. Development of chromosome-arm-specific microsatellite markers in *Triticum aestivum* (Poaceae) using NGS technology. Am J Bot 2012;99:369–71.
Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, et al. The Norway spruce genome sequence and conifer genome evolution. Nature 2013;497:579–84.
Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, et al. The *Sorghum bicolor* genome and the diversification of grasses. Nature 2009;457:551–6.
Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, et al. A physical map of the 1-gigabase bread wheat chromosome 3B. Science 2008;322:101–4.
Philippe R, Paux E, Bertin I, Sourdille P, Choulet F, Laugier C, et al. A high density physical map of chromosome 1BL supports evolutionary studies, map-based cloning and sequencing in wheat. Genome Biol 2013;14:R64.
Pich U, Meister A, Macas J, Doležel J, Lucretti S, Schubert I. Primed in situ labelling facilitates flow sorting of similar sized chromosomes. Plant J 1995;7:1039–44.
Peterson DG, Wessler SR, Paterson AH. Efficient capture of unique sequences from eukaryotic genomes. Trends Genet 2002;18:547–50.
Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JPT, et al. Advances in plant gene-targeted and functional markers: a review. Plant Methods 2013;9:6.
Požárková D, Koblížková A, Román B, Torres AM, Lucretti S, Lysák M, et al. Development and characterization of microsatellite markers from chromosome 1-specific DNA libraries of *Vicia faba*. Biol Plant 2002;45:337–45.
Qi LL, Echalier B, Chao S, Lazo GR, Butler GE, Anderson OD, et al. A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. Genetics 2004;168:701–12.
Rabinowicz PD, McCombie WR, Martienssen RA. Gene enrichment in plant genomic shotgun libraries. Curr Opin Plant Biol 2003;6:150–6.
Rieseberg M, Kasper C, Reardon KF, Scheper T. Flow cytometry in biotechnology. Appl Microbiol Biotechnol 2001;56:350–60.
Roberts RJ, Carneiro MO, Schatz MC. The advantages of SMRT sequencing. Genome Biol 2013;14:405.
Román B, Satovic Z, Požárková D, Macas J, Doležel J, Cubero JI, et al. Development of a composite map in *Vicia faba*, breeding applications and future prospects. Theor Appl Genet 2004;108:1079–88.
Rustenholz C, Choulet F, Laugier C, Šafář J, Šimková H, Doležel J, et al. A 3000-loci transcription map of chromosome 3B unravels the structural and functional features of gene islands in hexaploid wheat. Plant Physiol 2011;157:1596–608.
Šafář J, Bartoš J, Janda J, Bellec A, Kubaláková M, Valárik M, et al. Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J 2004;39:960–8.
Šafář J, Šimková H, Kubaláková M, Číhalíková J, Suchánková P, Bartoš J, et al. Development of chromosome-specific BAC resources for genomics of bread wheat. Cytogenet Genome Res 2010;129:211–23.
Sato K, Motoi Y, Yamaji N, Yoshida H. 454 sequencing of pooled BAC clones on chromosome 3H of barley. BMC Genomics 2011;12:246.
Schnable PS, Ware Doreen, Fulton Robert S, Stein Joshua C, Wei Fusheng, Pasternak Shiran, et al. The B73 maize genome: complexity, diversity, and dynamics. Science 2009;326:1112–5.
Schmidt T, Heslop-Harrison JS. Genomes, genes and junk: the large-scale organization of plant chromosomes. Trends Plant Sci 1998;3:195–9.
Schondelmaier J, Martin R, Jahoor A, Houben A, Graner A, Koop HU, et al. Microdissection and microcloning of the barley (*Hordeum vulgare* L.) chromosome 1HS. Theor Appl Genet 1993;86:629–36.
Schubert I, Fransz PF, Fuchs J, de Jong JH. Chromosome painting in plants. Methods Cell Sci 2001;23:57–69.
Schwarzacher T, Wang ML, Leitch AR, Miller N, Moore G, Heslop-Harrison JS. Flow cytometric analysis of the chromosomes and stability of a wheat cell-culture line. Theor Appl Genet 1997;94:91–7.
Sears ER, Sears LMS. The telocentric chromosomes of common wheat. In: Ramanujams S, editor. Proc. 5th Int. Wheat Genetics Symp. New Delhi: Indian Agricultural Research Institute; 1978. p. 389–407.
Sehgal SK, Li W, Rabinowicz PD, Chan A, Šimková H, Doležel J, et al. Chromosome arm-specific BAC end sequences permit comparative analysis of homoeologous chromosomes and genomes of polyploid wheat. BMC Plant Biol 2012;12:64.
Shangguan L, Han J, Kayesh E, Sun X, Zhang C, Pervaiz T, et al. Evaluation of genome sequencing quality in selected plant species using expressed sequence tags. PLoS One 2013;8:e69890.
Shapiro HM. Practical Flow Cytometry. 4th ed. Inc.: John Wiley & Sons; 2003
Shatalina M, Wicker T, Buchmann JP, Oberhaensli S, Šimková H, Doležel J, et al. Genotype-specific SNP map based on whole chromosome 3B sequence information from wheat cultivars Arina and Forno. Plant Biotechnol J 2013;11:23–32.
Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, et al. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci U S A 1992;89:8794–7.
Sierro N, Battey JND, Ouadi S, Bovet L, Goepfert S, Bakaher N, et al. Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. Genome Biol 2013;14:R60.
Šimková H, Číhalíková J, Vrána J, Lysák MA, Doležel J. Preparation of HMW DNA from plant nuclei and chromosomes isolated from root tips. Biol Plant 2003;46:369–73.
Šimková H, Svensson JT, Condamine P, Hřibová E, Suchánková P, Bhat PR, et al. Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genomics 2008a;9:294.
Šimková H, Šafář J, Suchánková P, Kovářová P, Bartoš J, Kubaláková M, et al. A novel resource for genomics of Triticeae: BAC library specific for the short arm of rye (*Secale cereale* L.) chromosome 1R (1RS). BMC Genomics 2008b;9:237.
Šimková H, Šafář J, Kubaláková M, Suchánková P, Číhalíková J, Robert-Quatre H, et al. BAC libraries from wheat chromosome 7D: efficient tool for positional cloning of aphid resistance genes. J Biomed Biotechnol 2011a;2011:302543.
Šimková H, Staňková H, Valárik M, Šafář J, Lapitan N, Luo M-Ch, et al. Chromosome genomics facilitates cloning of a Russian wheat aphid resistance gene. Abstracts of the 21st International Triticeae Mapping Initiative Workshop; 2011b. p. 17. [Mexico City].
Stein N, Ponelies N, Musket T, McMullen M, Weber G. Chromosome micro-dissection and region-specific libraries from pachytene chromosomes of maize (*Zea mays* L.). Plant J 1998;13:281–9.
Steuernagel B, Taudien S, Gundlach H, Seidel M, Ariyadasa R, Schulte D, et al. De novo 454 sequencing of barcoded BAC pools for comprehensive gene survey and genome analysis in the complex genome of barley. BMC Genomics 2009;10:547.
Stubblefield E, Oro J. The isolation of specific chicken macrochromosomes by zonal centrifugation and flow sorting. Cytometry 1982;5:273–81.
Suchánková P, Kubaláková M, Kovářová P, Bartoš J, Číhalíková J, Molnár-Láng M, et al. Dissection of the nuclear genome of barley by chromosome flow sorting. Theor Appl Genet 2006;113:651–9.
Tang X, De Boer JM, Van Eck HJ, Bachem C, Visser RG, De Jong H. Assignment of genetic linkage maps to diploid *Solanum tuberosum* pachytene chromosomes by BAC-FISH technology. Chromosome Res 2009;17:899–915.
The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 2000;408:796–815.
. [Internet]The IEB genomic resources database. The Centre of Plant Structural and Functional Genomics of the Institute of Experimental Botany, Olomouc, Czech Republic. [cited 2013 Dec 20. Available from]http://olomouc.ueb.cas.cz/dna-libraries/cereals.
. [Internet]The International Wheat Genome Sequencing Consortium (IWGSC). [cited 2013 Dec 20. Available from]http://www.wheatgenome.org/.
. [Internet]The million plant and animal genomes project. BGI. [cited 2013 Dec 20. Available from]http://www.genomics.cn/en/navigation/show_navigation?nid=5657.
Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 2006;15:1596–604.
Überall I, Vrána J, Bartoš J, Šmerda J, Doležel J, Havel L. Isolation of chromosomes from *Picea abies* L. and their analysis by flow cytometry. Biol Plant 2004;48:199–203.
Valárik M, Bartoš J, Kovářová P, Kubaláková M, de Jong H, Doležel J. High-resolution FISH on super-stretched flow-sorted plant chromosomes. Plant J 2004;37:940–50.
Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, et al. Draft genome sequence of kabuli chickpea (*Cicer arietinum*): genetic structure and breeding constraints for crop improvement. Nat Biotechnol 2013;31:240–6.
Veuskens J, Marie D, Brown SC, Jacobs M, Negrutiu I. Flow sorting of the Y sex chromosome in the dioecious plant *Melandrium album*. Cytometry 1995;21:363–73.
Vitharana SN, Wilson GS. Fractionation of chromosome 15 with an affinity-based approach using magnetic beads. Genomics 2006;87:158–64.
Vitulo N, Albiero A, Forcato C, Campagna D, Dal Pero F, Bagnaresi P, et al. First survey of the wheat chromosome 5A composition through a next generation sequencing approach. PLoS One 2011;6:e26421.
Vláčilová K, Ohri D, Vrána J, Číhalíková J, Kubaláková M, Kahl G, et al. Development of flow cytogenetics and physical genome mapping in chickpea (*Cicer arietinum* L.). Chromosome Res 2002;10:695–706.
Vrána J, Kubaláková M, Šimková H, Číhalíková J, Lysák MA, Doležel J. Flow-sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). Genetics 2000;156:2033–41.

Vrána J, Šimková H, Kubaláková M, Číhalíková J, Doležel J. Flow cytometric chromosome sorting in plants: the next generation. Methods 2012;57:331–7.
Vrána J, Kubaláková M, Číhalíková J, Valárik M, Doležel J. Preparation of subgenomic fractions enriched for particular chromosomes in plants. Cytometry 2014. submitted for publication.
Wang ML, Leitch AR, Schwarzacher T, Heslop-Harrison JS, Moore G. Construction of a chromosome-enriched HpaII library from flow-sorted wheat chromosomes. Nucleic Acids Res 1992;20:1897–901.
Wenzl P, Suchánková P, Carling J, Šimková H, Huttner E, Kubaláková M, et al. Isolated chromosomes as a new and efficient source of DArT markers for the saturation of genetic maps. Theor Appl Genet 2010;121:465–74.
Wicker T, Mayer KFX, Gundlach H, Martis M, Steuernagel B, Scholz U, et al. Frequent gene movement and pseudogene evolution is common to the large and complex genomes of wheat, barley, and their relatives. Plant Cell 2011;23:1706–18.
Yang H, Chen X, Wong WH. Completely phased genome sequencing through chromosome sorting. Proc Natl Acad Sci U S A 2011;108:12–7.
Yu J, Hu S, Wang J, Wong KGS, Li S, Liu B, et al. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 2002;296:79–92.
Zatloukalová P, Hřibová E, Kubaláková M, Suchánková P, Šimková H, Adoración C, et al. Integration of genetic and physical maps of the chickpea (*Cicer arietinum* L.) genome using flow-sorted chromosomes. Chromosome Res 2011;19:729–39.

Appendix V

Flow cytometry in plant research. In Applied plant cell biology

DOI: 10.1007/978-3-642-41787-0_13

Flow Cytometry in Plant Research: A Success Story

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

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

1 Introduction

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

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

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

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

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

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

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

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

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

2 Principles of Flow Cytometry

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

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

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

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

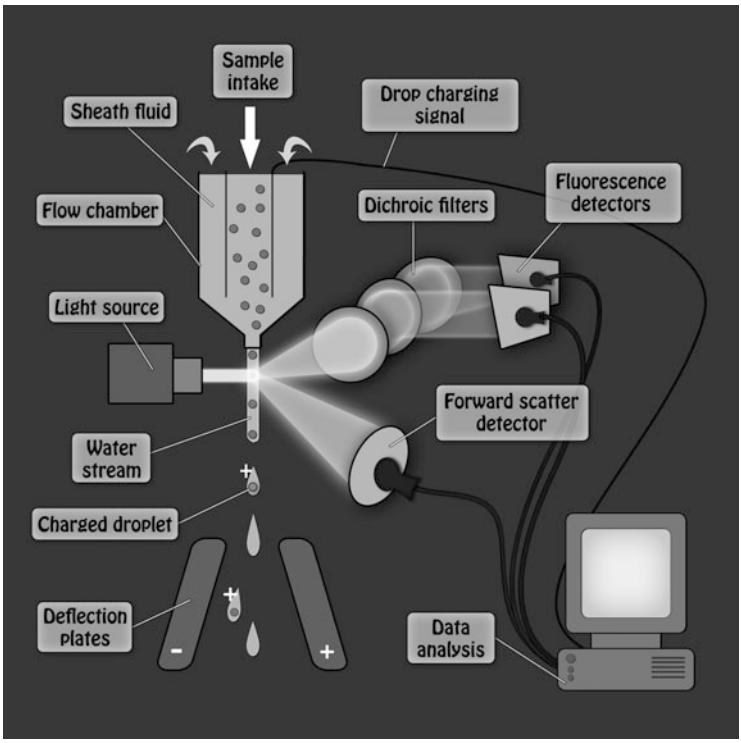

Fig. 1 Schematic view of a flow cytometer and sorter with one excitation light source, one detector of forward light scatter and two detectors of fluorescence signals. This type of instrument can sort two different populations of particles simultaneously

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

3 Plants and Flow Cytometry: An Uneasy but Fruitful Marriage

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

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

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

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

4 Plant Nuclei in Focus

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

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

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

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

5 Principles and Methodology

5.1 Sample Preparation

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

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

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

5.2 Nuclear Isolation Buffers

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

5.3 Parameters Analysed on Plant Nuclei

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

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

5.4 Nuclear Staining

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

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

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

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

5.5 Secondary Metabolites

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

5.6 Materials for Nuclear Isolation

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

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

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

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

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

5.7 DNA Flow Cytometry

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

6 Biological Aspects: Analysis of Nuclei

6.1 Genome Size

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

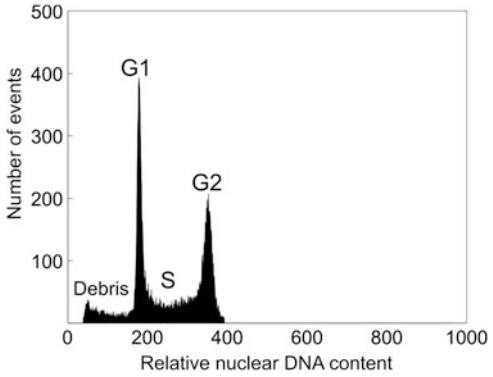

Fig. 2 Distribution of nuclear DNA content in a population of cycling cells. Histogram of relative DNA content obtained after the analysis of DAPI-stained nuclei isolated from actively growing root tips of bread wheat. As cells progress from G_1 to G_2 phase of cell cycle, their DNA content doubles.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

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

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

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

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

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

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

$$\begin{aligned}&\text{2C value of unknown sample} = \\ &[\text{G}_0/\text{G}_1 \text{ peak mean of unknown sample}/\text{G}_0/\text{G}_1 \text{ peak mean of reference standard}] \\ &\times \text{2C value of reference standard (pg)}.\end{aligned}$$

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

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

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

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

6.2 Ploidy

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

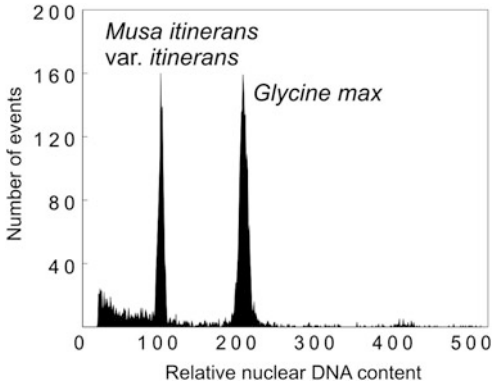

Fig. 3 Estimation of nuclear DNA amount in absolute units. Histogram of relative DNA content was obtained after simultaneous analysis of propidium iodide-stained nuclei of a banana species *Musa itinerans* var. *itinerans* and soybean (*Glycine max*). The soybean with 2C = 2.50 pg DNA was used as internal reference standard to estimate DNA amount of the banana species according to the formula:
2C value of *Musa* = (2.5 × G_1 peak mean of *Musa*)/(G_1 peak mean of *Glycine*)
The estimated 2C amount of *Musa* was 1.217 pg DNA. The mass of DNA in pg can be converted to the number of base pairs considering 1 pg DNA = 0.978 × 10^9 bp (Doležel et al. 2003)

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

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

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

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

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

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

6.3 Base Composition

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

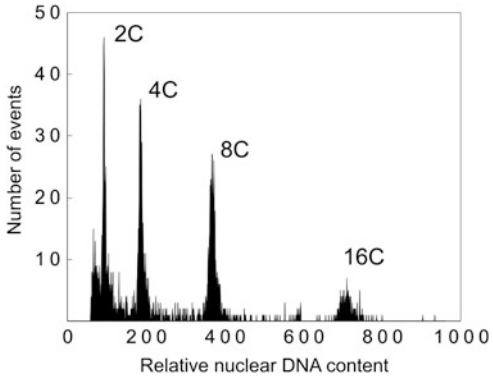

Fig. 4 Analysis of endoreduplication. Histogram of relative DNA amount was obtained after the analysis of DAPI-stained nuclei isolated from 14-day-old plants of *Arabidopsis thaliana*. The histogram comprises three peaks – one representing G_1 phase cells with 2C DNA content, one representing G_2 phase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

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

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

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

6.4 Cell Cycle

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

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

6.5 Gene Expression

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

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

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

7.1 Taxonomy

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

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

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

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

7.2 Evolution of Plant Genomes

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

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

7.3 Ecology and Plant Population Biology

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

7.4 Plant Sex and Reproduction

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

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

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

7.5 Cell Biology

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

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

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

7.6 Plant Breeding

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

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

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

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

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

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Bennett MD, Leitch IJ (1995) Nuclear DNA amounts in angiosperms. Ann Bot 76:113–176
Bennett MD, Leitch IJ (1997) Nuclear DNA amounts in angiosperms – 583 new estimates. Ann Bot 80:169–196
Bennett MD, Leitch IJ (2005) Nuclear DNA amounts in angiosperms – progress, problems and prospects. Ann Bot 95:45–90
Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Philos Trans R Soc Lond Ser B Biol Sci 274:227–274
Bennett MD, Smith JB, Smith RIL (1982) DNA amounts of angiosperms from the Antarctic and South Georgia. Environ Exp Bot 22:307–318
Bennett MD, Leitch IJ, Hanson L (1998) DNA amounts in two samples of angiosperm weeds. Ann Bot 82:121–134
Bennett MD, Bhandol P, Leitch IJ (2000) Nuclear DNA amounts in angiosperms and their modern uses – 807 new estimates. Ann Bot 86:859–909
Bennett MD, Leitch IJ, Price HJ, Johnston JP (2003) Comparisons with *Caenorhabditis* (100 Mb) and Drosophila (175 Mb) using flow cytometry show genome size in Arabidopsis to be 157 Mb and thus 25 % larger than the Arabidopsis genome initiative estimate of 125 Mb. Ann Bot 91:547–557
Bergounioux C, Perennes C, Brown SC, Gadal P (1988) Nuclear RNA quantification in protoplast cell cycle phases. Cytometry 9:84–87
Binarová P, Doležel J (1993) Effect of anti-microtubular drug amiprophos-methyl on somatic embryogenesis and DNA ploidy levels in alfalfa and carrot cell suspension cultures. Biol Plant 35:329–339
Binarová P, Číhalíková J, Doležel J (1993) Localization of MPM–2 recognized phosphoproteins and tubulin during cell cycle progression in synchronized *Vicia faba* root meristem cells. Cell Biol Int 17:847–856
Binarová P, Doležel J, Dráber P, Heberle-Bors E, Strnad M, Bogre L (1998a) Treatment of *Vicia faba* root tip cells with specific inhibitors to cyclin–dependent kinases leads to abnormal spindle formation. Plant J 16:697–707
Binarová P, Hause B, Doležel J, Dráber P (1998b) Association of γ–tubulin with kinetochore/centromeric region of plant chromosomes. Plant J 14:751–757
Binarová P, Cenklová V, Hause B, Kubátová E, Lysák M, Doležel J, Bögre L, Dráber P (2000) Nuclear γ-tubulin during acentriolar plant mitosis. Plant Cell 12:433–442
Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN (2003) A gene expression map of the *Arabidopsis* root. Science 302:1956–1960
Boivin A, Vendrely R, Vendrely C (1948) L'acide désoxyribonucléique du noyau cellulaire dépositaire des caractéres héréditaires; arguments d'ordre analytique. C R Acad Sci 226:1061–1063
Bourdon M, Coriton O, Pirrello J, Cheniclet C, Brown SC, Poujol C, Chevalier C, Renaudin JP, Frangne N (2011) In planta quantification of endoreduplication using fluorescent in situ hybridization (FISH). Plant J 66:1089–1099
Bourdon M, Pirrello J, Cheniclet C, Coriton O, Bourge M, Brown S, Moïse A, Peypelut M, Rouyère V, Renaudin JP, Chevalier C, Frangne N (2012) Evidence for karyoplasmic homeostasis during endoreduplication and a ploidy-dependent increase in gene transcription during tomato fruit growth. Development 139:3817–3826
Buck SB, Bradford J, Gee KR, Agnew BJ, Clarke ST, Salic A (2008) Detection of S–phase cell cycle progression using 5–ethynyl–2′–deoxyuridine incorporation with click chemistry, an alternative to using 5–bromo–2′–deoxyuridine antibodies. Biotechniques 44:927–929
Cardoso DC, Carvalho CR, Cristiano MP, Soares FAF, Tavares MG (2012) Estimation of nuclear genome size of the genus *Mycetophylax emery*, 1913: evidence of no whole–genome duplication in Neoattini. C R Biol 335:619–624
Ceccarelli M, Sanantonio E, Marmottini F, Amzallag GN, Cionini PG (2006) Chromosome endoreduplication as a factor of salt adaptation in *Sorghum bicolour*. Protoplasma 227:113–118

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Doležel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc 2:2233–2244

Doležel J, Vrána J, Šafář J, Bartoš J, Kubaláková M, Šimková H (2012) Chromosomes in the flow to simplify genome analysis. Funct Integr Genomics 12:397–416

Doleželová M, Doležel J, Van den Houwe I, Roux N, Swennen R (2005) Focus on the *Musa* collection: ploidy levels revealed. InfoMusa 14:34–36

Doskočilová A, Kohoutová L, Volc J, Kourová H, Benada O, Chumová J, Plíhal O, Petrovská B, Halada P, Bögre L, Binarová P (2013) Nitrilase1 regulates the exit from proliferation, genome stability and plant development. New Phytol 198:685–698

Dubelaar GBJ, Gerritzen PL, Beeker AER, Jonker RR, Tangen K (1999) Design and first results of CytoBuoy: a wireless flow cytometer for in situ analysis of marine and fresh waters. Cytometry 37:247–254

Dubelaar GBJ, Casotti R, Tarran GA, Biegala IC (2007) Phytoplankton and their analysis by flow cytometry. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 287–322

Eeckhaut T, Leus L, van Huylenbroeck J (2005) Exploitation of flow cytometry for plant breeding. Acta Phys Plant 27:743–750

Egesi CN, Pillay M, Asiedu R, Egunjobi JK (2002) Ploidy analysis in water yam, *Dioscorea alata* L., germplasm. Euphytica 128:225–230

Favoreto FC, Carvalho CR, Lima ABP, Ferreira A, Clarindo WR (2012) Genome size and base composition of Bromeliaceae species assessed by flow cytometry. Plant Syst Evol 298:1185–1193

Fox MH, Galbraith DW (1990) The application of flow cytometry and sorting to higher plant systems. In: Melamed MR, Lindmo T, Mendelsohn ML (eds) Flow cytometry and sorting, 2nd edn. Wiley-Liss, New York, pp 633–650

Fujikura Y, Doležel J, Číhalíková J, Bögre L, Heberle-Bors E, Hirt H, Binarová P (1999) Vicia faba germination: synchronized cell growth and localization of nucleolin and α–tubulin. Seed Sci Res 9:297–304

Galbraith DW (1990) Flow cytometric analysis of plant genomes. Methods Cell Biol 33:549–562

Galbraith DW (1994) Flow cytometry and sorting of plant protoplasts and cells. Methods Cell Biol 42B:539–561

Galbraith DW (2007) Protoplast analysis using flow cytometry. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 231–250

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

Gaurav V, Kolewe ME, Roberts SC (2010) Flow cytometric methods to investigate culture heterogeneities for plant metabolic engineering. In: Germano A (ed) Plant secondary metabolism engineering. Methods in molecular biology. Humana Press, New York, pp 243–262, Springer protocols

Gendreau E, Hofte H, Grandjean O, Brown S, Traas J (1998) Phytochrome controls the number of endoreduplication cycles in the *Arabidopsis thaliana* hypocotyls. Plant J 13:221–230

Giglioli-Guivarc'h N, Pierre JN, Vidal J, Brown S (1996) Flow cytometric analysis of cytosolic pH of mesophyll cell protoplasts from the crabgrass *Digitaria sanguinalis*. Cytometry 23:241–249

Glab N, Labidi B, Qin LX, Trehin C, Bergounioux C, Meijer L (1994) Olomoucine, an inhibitor of the cdc2/cdk2 kinases activity, blocks plant cells at the G1 to S and G2 to M cell cycle transitions. FEBS Lett 353:207–211

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Loureiro J, Pinto G, Lopes T, Doležel J, Santos C (2005) Assessment of ploidy stability of the somatic embryogenesis process in *Quercus suber* L. using flow cytometry. Planta 221:815–822
Loureiro J, Rodriguez E, Doležel J, Santos C (2006a) Comparison of four nuclear isolation buffers for plant DNA flow cytometry. Ann Bot 98:679–689
Loureiro J, Rodriguez E, Doležel J, Santos C (2006b) Flow cytometric and microscopic analysis of the effect of tannic acid on plant nuclei and estimation of DNA content. Ann Bot 98:515–527
Loureiro J, Rodriguez E, Doležel J, Santos C (2007a) Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. Ann Bot 100:875–888
Loureiro J, Kopecký D, Castro S, Santos C, Silveira P (2007b) Flow cytometric and cytogenetic analyses of Iberian Peninsula *Festuca* spp. Plant Syst Evol 269:89–105
Loureiro J, Trávníček P, Rauchová J, Urfus T, Vít P, Štech M, Castro S, Suda J (2010) The use of flow cytometry in the biosystematics, ecology and population biology of homoploid plants. Preslia 82:3–21
Lucretti S, Nardi L, Nisini PT, Moretti F, Gualberti G, Doležel J (1999) Bivariate flow cytometry DNA/BrdUrd analysis of plant cell cycle. Methods Cell Sci 21:155–166
Lysák MA, Doleželová M, Horry JP, Swennen R, Doležel J (1999) Flow cytometric analysis of nuclear DNA content in *Musa*. Theor Appl Genet 98:1344–1350
Macas J, Lambert GM, Doležel D, Galbraith DW (1998) Nuclear expressed sequence tag (NEST) analysis: a novel means to study transcription through amplification of nuclear RNA. Cytometry 33:460–468
Marhold K, Kudoh H, Pak JH, Watanabe K, Španiel S, Lihová J (2010) Cytotype diversity and genome size variation in eastern Asian polyploid *Cardamine* (Brassicaceae) species. Ann Bot 105:249–264
Marie D, Brown SC (1993) A cytometric exercise in plant DNA histograms, with 2C values for 70 species. Biol Cell 78:41–51
Marum L, Rocheta M, Maroco J, Oliveira MM, Muguel C (2009) Analysis of genetic stability at SSR loci during somatic embryogenesis in maritime pine (*Pinus pinaster*). Plant Cell Rep 28:673–682
Matzk F (2007) Reproduction mode screening. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 131–152
Matzk F, Meister A, Schubert I (2000) An efficient screen for reproductive pathways using mature seeds of monocots and dicots. Plant J 21:97–108
Meister A, Barow M (2007) DNA base composition of plant genomes. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells. analysis of genes, chromosomes, and genomes. Viley-VCH, Weinheim, pp 177–215
Melaragno JE, Mehrotra B, Coleman AW (1993) Relationship between endopolyploidy and cell size in epidermal tissue of *Arabidopsis*. Plant Cell 5:1661–1668
Mishra MK (1997) Stomatal characteristics at different ploidy levels in *Coffea* L. Ann Bot 80:689–692
Morgan-Richards M, Trewick SA, Chapman HM, Krahulcova A (2004) Interspecific hybridization among *Hieracium* species in New Zealand: evidence from flow cytometry. Heredity 93:34–42
Naill MC, Roberts SC (2005) Flow cytometric analysis of protein content in *Taxus* protoplasts and single cells as compared to aggregated suspension cultures. Plant Cell Rep 23:528–533
Neelakandan AK, Wang K (2012) Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications. Plant Cell Rep 31:597–620
Noirot M, Barre P, Louarn J, Duperray C, Hamon S (2000) Nucleus-cytosol interactions – A source of stoichiometric error in flow cytometric estimation of nuclear DNA content in plants. Ann Bot 86:309–316
Noirot M, Barre P, Louarn J, Duperray C, Hamon S (2002) Consequences of stoichiometric error on nuclear DNA content evaluation in *Coffea liberica* var. *dewevrei* using DAPI and propidium iodide. Ann Bot 89:385–389

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

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

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

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

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

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

Ochatt SJ, Patat-Ochatt EM, Moessner A (2011) Ploidy level determination within the context of in vitro breeding. Plant Cell Tissue Organ Cult 104:329–341

Overton WR, McCoy JP (1994) Reversing the effect of formalin on the binding of propidium iodide to DNA. Cytometry 16:351–356

Ozaki Y, Kumiko N, Michikazu H, Kenji U, Hiroshi O (1998) Application of flow cytometry for rapid determination of ploidy levels in Asparagus (*Asparagus officinalis* L.). J Fac Agric Kyushu Univ 43:83–88

Palomino G, Doležel J, Mendez I, Rubluo A (2003) Nuclear genome size analysis of *Agave tequilana*. Caryologia 56:37–46

Pawlowski TA, Bergervoet JHW, Bino RJ, Groot SPC (2004) Cell cycle activity and β-Tubulin accumulation during dormancy breaking of *Acer platanoides* L. seeds. Biol Plant 48:211–218

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

Perfetto SP, Chattopadhyay PK, Roederer M (2004) Seventeen–colour flow cytometry: unravelling the immune system. Nat Rev Immunol 4:648–655

Petit PX (1992) Flow Cytometric Analysis of rhodamine 123 fluorescence during modulation of the membrane potential in plant mitochondria. Plant Physiol 98:279–286

Petit PX, Diolez P, Muller P, Brown SC (1986) Binding of concanavalin A to the outer membrane of potato tuber mitochondria detected by flow cytometry. Febs Lett 196:65–70

Petrovská B, Cenklová V, Pochylová Z, Kourová H, Doskočilová A, Plíhal O, Binarová L, Binarová P (2012) Plant Aurora kinases play a role in maintenance of primary meristems and control of endoreduplication. New Phytol 193:590–604

Pfosser M (1989) Improved method for critical comparison of cell cycle data of asynchronously dividing and synchronized cell cultures of *Nicotiana tabacum*. J Plant Physiol 134:741–745

Pfosser M, Heberle-Bors E, Amon A, Lelley T (1995) Evaluation of sensitivity of flow cytometry in detecting aneuploidy in wheat using disomic and ditelosomic wheat–rye addition lines. Cytometry 21A:387–393

Pfündel E, Meister A (1996) Flow cytometry of mesophyll and bundle sheath chloroplast thylakoids of maize (*Zea mays* L.). Cytometry 23A:97–105

Picot J, Guerin CL, Le Van KC, Boulanger CM (2012) Flow cytometry: retrospective, fundamentals and recent instrumentation. Cytotechnology 64:109–130

Pillay M, Ogundiwin E, Tenkouano A, Doležel J (2006) Ploidy and genome composition of Musa germplasm at the International Institute of Tropical Agriculture (IITA). Afr J Biotechnol 5:1224–1232

Praça-Fontes MM, Carvalho CR, Clarindo WR, Cruz CD (2011) Revisiting the DNA C-values of the genome size-standards used in plant flow cytometry to choose the "best primary standards". Plant Cell Rep 30:1183–1191

Price HJ, Hodnett G, Johnston JS (2000) Sunflower (*Helianthus annuus*) leaves contain compounds that reduce nuclear propidium iodide fluorescence. Ann Bot 86:929–934

Rani V, Raina SN (2000) Genetic fidelity of organized meristem–derived micropropagated plants: a critical reappraisal. In Vitro Cell Dev Biol Plant 36:319–330

Reichheld JP, Vernoux T, Lardon F, van Montagu M, Inzé D (1999) Specific checkpoints regulate plant cell cycle progression in response to oxidative stress. Plant J 17:647–656

Rewers M, Sliwinska E (2012) Endoreduplication Intensity as a marker of seed developmental stage in the Fabaceae. Cytometry 81A:1067–1075

Ricroch A, Yockteng R, Brown SC, Nadot S (2005) Evolution of genome size across some cultivated Allium species. Genome 48:511–520

Roberts AV (2007) The use of bead beating to prepare suspensions of nuclei for flow cytometry from fresh leaves, herbarium leaves, petals and pollen. Cytometry 71A:1039–1044

Roux N, Doležel J, Swennen R, Zapata-Arias FJ (2001) Effectiveness of three micropropagation techniques to dissociate cytochimeras in *Musa* spp. Plant Cell Tissue Organ Cult 66:189–197

Roux N, Toloza A, Radecki Z, Zapata-Arias FJ, Doležel J (2003) Rapid detection of aneuploidy in *Musa* using flow cytometry. Plant Cell Rep 21:483–490

Rowan BA, Oldenburg DJ, Bendich AJ (2007) A high–throughput method for detection of DNA in chloroplasts using flow cytometry. Plant Methods 3:5

Šafářová L, Duchoslav M, Jandová M, Krahulec F (2011) *Allium oleraceum* in Slovakia: cytotype distribution and ecology. Preslia 83:513–527

Schmidt G, Thannhauser SJ (1945) A method for the determination of deoxyribonucleic acid, ribonucleic acid and phosphoproteins in animal tissues. J Biol Chem 161:83–89

Schönswetter P, Suda J, Popp M, Weis-Schneeweiss H, Brochmann C (2007) Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. Mol Phylogen Evol 42:92–103

Schulze D, Pauls KP (1998) Flow cytometric characterization of embryogenic and gametophytic development in *Brassica napus* microspore cultures. Plant Cell Physiol 39:226–234

Šesek P, Šuštar-Vozlič J, Bohanec B (2000) Determination of aneuploids in hop (*Humulus lupulus* L.) using flow cytometry. Pflugers Arch 439:16–18

Sgorbati S, Levi M, Sparvoli E, Trezzi F, Lucchini G (1986) Cytometry and flow cytometry of 4′,6-diamidino-2-phenylindole (DAPI)-stained suspensions of nuclei released from fresh and fixed tissues of plants. Physiol Plant 68:471–476

Sgorbati S, Sparvoli E, Levi M, Galli MG, Citterio S, Chiatante D (1991) Cell cycle kinetic analysis with flow cytometry in pea root meristem synchronized with aphidicolin. Physiol Plant 81:507–512

Shapiro HM (1985) Practical flow cytometry. Wiley, New York

Shapiro HM (1988) Practical flow cytometry, 2nd edn. Wiley-Liss, New York

Shapiro HM (1995) Practical flow cytometry, 3rd edn. Wiley-Liss, New York

Shapiro HM (2003) Practical flow cytometry, 4th edn. Wiley, Hoboken

Siljak-Yakovlev S, Benmale S, Cerbah M, Coba de la Pena T, Bounaga N, Brown SC, Sarr A (1996) Chromosomal sex determination and heterochromatin structure in date palm. Sex Plant Reprod 9:127–132

Silva TCR, Abreu IS, Carvalho CR (2010) Improved and reproducible flow cytometry methodology for nuclei isolation from single root meristem. J Bot 2010, Article ID 320609

Śliwinska E, Thiem B (2007) Genome size stability in six medicinal plant species propagated in vitro. Biol Plant 51:556–558

Śliwinska E, Zielinska E, Jedrzejcyk I (2005) Are seeds suitable for flow cytometric estimation of plant genome size? Cytometry 64A:72–79

Śliwinska E, Bassel GW, Bewley D (2009) Germination of *Arabidopsis thaliana* seeds is not completed as a result of elongation of the radicle but of the adjacent transition zone and lower hypocotyl. J Exp Bot 60:3578–3594

Šmarda P (2006) DNA ploidy levels and intraspecific DNA content variability in Romanian fescues (*Festuca*, Poaceae) measured in fresh and herbarium material. Folia Geobot 41:417–432

Šmarda P, Stančík D (2006) Ploidy level variability in South American fescues (*Festuca* L., Poaceae): use of flow cytometry in up to 5 1/2–year–old caryopses and herbarium specimens. Plant Biol 8:73–80

Šmarda P, Müller J, Vrána J, Kočí K (2005) Ploidy level variability of some Central European fescues (*Festuca* subg. Festuca, Poaceae). Biologia (Bratislava) 60:25–36

Šmarda P, Bureš P, Horová L, Foggi B, Rossi G (2008) Genome size and GC content evolution of *Festuca*: ancestral expansion and subsequent reduction. Ann Bot 101:421–433

Šmarda P, Bureš P, Šmerda J, Horová L (2011) Measurements of genomic GC content in plant genomes with flow cytometry: a test for reliability. New Phytol 193:513–521

Soltis DE, Soltis PS, Bennett MD, Leitch IJ (2003) Evolution of genome size in the angiosperms. Am J Bot 90:1596–1603

Stehlík I, Barrett SCH (2005) Mechanisms governing sex-ratio variation in dioecious *Rumex nivalis*. Evolution 59:814–825

Stehlík I, Kron P, Barrett SCH, Husband BC (2007) Sexing pollen reveals female bias in a dioecious plant. New Phytol 175:185–194

Suda J, Leitch IJ (2010) The quest for suitable reference standards in genome size research. Cytometry 77A:717–720

Suda J, Trávníček P (2006) Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry: new prospects for plant research. Cytometry 69A:273–280

Suda J, Kyncl T, Jarolímová V (2005) Genome size variation in Macaronesian angiosperms: forty percent of the Canarian endemic flora completed. Plant Syst Evol 252:215–238

Suda J, Krahulcová A, Trávníček P, Krahulec F (2006) Ploidy level versus DNA ploidy level: an appeal for consistent terminology. Taxon 55:447–450

Suda J, Kron P, Husband BC, Trávníček P (2007a) Flow cytometry and ploidy: applications in plant systematics, ecology and evolutionary biology. In: Doležel J, Greilhuber J, Suda J (eds) Flow cytometry with plant cells: analysis of genes, chromosomes and genomes. Wiley-VCH, Weinheim, pp 103–130

Suda J, Krahulcová A, Trávníček P, Rosenbaumová R, Peckert T, Krahulec F (2007b) Genome size variation and species relationships in *Hieracium* subgen. *Pilosella* (Asteraceae) as inferred by flow cytometry. Ann Bot 100:1323–1335

Suda J, Trávníček P, Mandák B, Berchová-Bímová K (2010) Genome size as a marker for identifying the invasive alien taxa in *Fallopia* section *Reynoutria*. Preslia 82:97–106

Sugimoto-Shirasu K, Roberts K (2003) "Big it up": endoreduplication and cell-size control in plants. Curr Opin Plant Biol 6:544–553

Sun YL, Sun Y, Lin GG, Zhang R, Zhang K, Xie JH, Wang LN, Li JM (2012) Multicolor flow cytometry analysis of the proliferations of T-lymphocyte subsets in vitro by EdU incorporation. Cytometry 81A:901–909

Swift H (1950) The constancy of desoxyribose nucleic acid in plant nuclei. Proc Natl Acad Sci U S A 36:643–654

Taylor IW, Milthorpe BK (1980) An evaluation of DNA fluorochromes, staining techniques, and analysis for flow cytometry. I. Unperturbed cell populations. J Histochem Cytochem 28:1224–1232

Temsch EM, Greilhuber J (2000) Genome size variation in *Arachis hypogaea* and *A. monticola* re-evaluated. Genome 43:449–451

Temsch EM, Greilhuber J (2001) Genome size in *Arachis duranensis*: a critical study. Genome 44:826–830

Temsch EM, Temsch W, Ehrendorfer–Schratt L, Greilhuber J (2010) Heavy metal pollution, selection, and genome size: the species of the Žerjav study revisited with flow cytometry. J Bot 2010, Article ID 596542

Tiersch TR, Chandler RW, Wachtel SS, Elias S (1989) Reference standards for flow cytometry and application in comparative studies of nuclear DNA content. Cytometry 10A:706–710

Typas MA, Heale JB (1980) DNA content of germinating sores, individual hyphal cells and resting structure cells of *Verticillium* spp. measured by microdensitometry. Microbiology 121:231–242

Ulrich I, Ulrich W (1991) High-resolution flow cytometry of nuclear DNA in higher plants. Protoplasma 165:212–215

Ulrich I, Fritz B, Ulrich W (1988) Application of DNA fluorochromes for flow cytometric DNA analysis of plant protoplasts. Plant Sci 55:151–158

van Duren M, Morpurgo R, Doležel J, Afza R (1996) induction and verification of autotetraploids in diploid banana (*Musa acuminata*) by in vitro techniques. Euphytica 88:25–34

Vidic T, Greilhuber J, Vilhar B, Dermastia M (2009) Selective significance of genome size in a plant community with heavy metal pollution. Ecol Appl 19:1515–1521

Vilhar B, Greilhuber J, Koce JD, Temsch EM, Dermastia M (2001) Plant genome size measurement with DNA image cytometry. Ann Bot 87:719–728

Vilhar B, Kladnik A, Blejec A, Chourey PS, Dermastia M (2002) Cytometrical evidence that the loss of seed weight in the miniature seed mutant of maize is associated with reduced mitotic activity in the developing endosperm. Plant Physiol 129:23–30

Vinogradov AE (1994) Measurement by flow cytometry of genomic AT/GC ratio and genome size. Cytometry 16A:34–40

Voglmayr H (2000) Nuclear DNA amounts in mosses (Musci). Ann Bot 85:531–546

Voglmayr H, Greilhuber J (1998) Genome size determination in Peronosporales (Oomycota) by Feulgen image analysis. Fungal Genet Biol 25:181–195

Watanabe M, Setoguchi D, Uehara K, Ohtsuka W, Watanabe Y (2002) Apoptosis–like cell death of *Brassica napus* leaf protoplasts. New Phytol 156:417–426

Weber S, Unker F, Friedt W (2005) Improved doubled haploid production protocol for *Brassica napus* using microspore colchicine treatment in vitro and ploidy determination by flow cytometry. Plant Breed 124:511–513

Whittemore AT, Olsen RT (2011) *Ulmus americana* (Ulmaceae) is a polyploid complex. Am J Bot 98:754–760

Yanpaisan W, King NJC, Doran PM (1998) Analysis of cell cycle activity and population dynamics in heterogeneous plant cell suspensions using flow cytometry. Biotechnol Bioeng 58:515–528

Yao N, Eisfelder BJ, Marvin J, Greenberg JT (2004) The mitochondrion – an organelle commonly involved in programmed cell death in *Arabidopsis thaliana*. Plant J 40:596–610

Yokoya K, Roberts AV, Mottley J, Lewis R, Brandham PE (2000) Nuclear DNA amounts in roses. Ann Bot 85:557–561

Záveský L, Jarolímová V, Štěpánek J (2005) Nuclear DNA content variation within the genus *Taraxacum* (Asteraceae). Folia Geobotanica 40:91–104

Zhang CQ, Gong FC, Lambert GM, Galbraith DW (2005) Cell type-specific characterization of nuclear DNA contents within complex tissues and organs. Plant Methods 1:7

Zhang CQ, Barthelson RA, Lambert GM, Galbraith DW (2008) Characterization of cell-specific gene expression through fluorescence–activated sorting of nuclei. Plant Physiol 147:30–40

Zhao J, Cui J, Liu J, Liao F, Henny RJ, Chen J (2012) Direct somatic embryogenesis from leaf and petiole explants of *Spathiphyllum* 'Supreme' and analysis of regenerants using flow cytometry. Plant Cell Tissue Organ Cult 110:239–249

Zonneveld BJM, Leitch IJ, Bennett MD (2005) First nuclear DNA amounts in more than 300 angiosperms. Ann Bot 96:229–244