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

* 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
http://dx.doi.org/10.1016/j.biotechadv.2013.12.011

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)

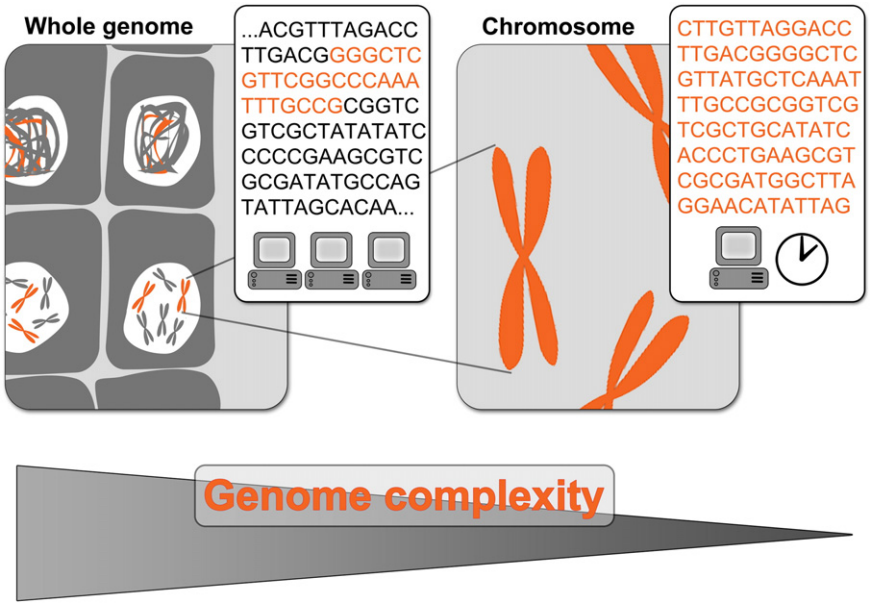

Fig. 1. Chromosome genomics: instead of treating the whole nuclear genome as a unit, single chromosomes are isolated and their DNA used as the template for genomic analyses. The reduction in template complexity achieved speeds mapping, sequencing and sequence analysis, and simplifies the necessary bioinformatics.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

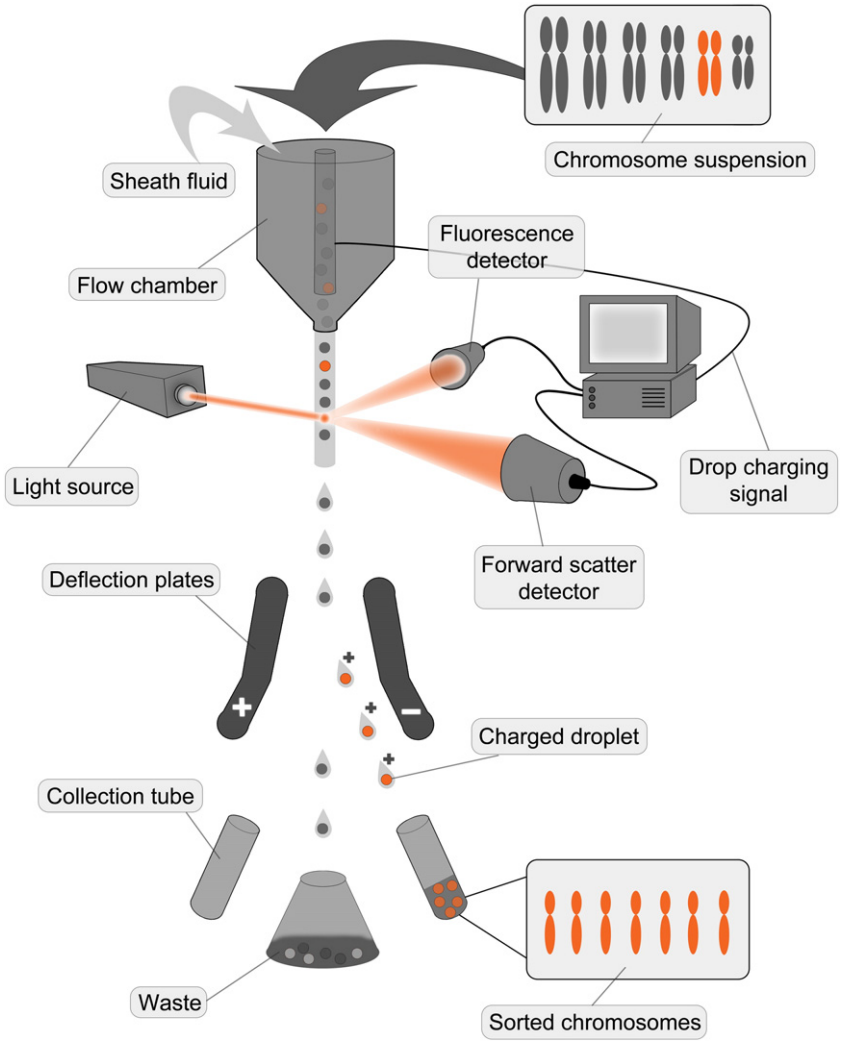

Fig. 2. The mechanics of flow-sorting. Chromosomes held in liquid suspension are stained by a fluorochrome and passed into a flow chamber containing sheath fluid. The geometry of the chamber forces the chromosome suspension into a narrow stream in which the chromosomes become aligned in a single file, 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)

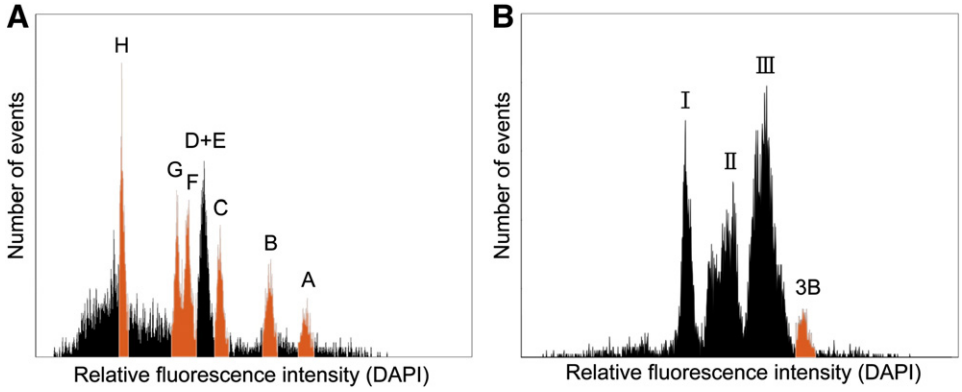

Fig. 3. Flow karyotyping in chickpea and bread wheat. The fluorescence intensity histograms (flow karyotypes) were obtained from DAPI-stained suspensions of mitotic chromosomes. (A) Chickpea cv. Frontier (2n = 2x = 16) forms seven peaks, six of which each represent a single chromosome (A–C and F–H). 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

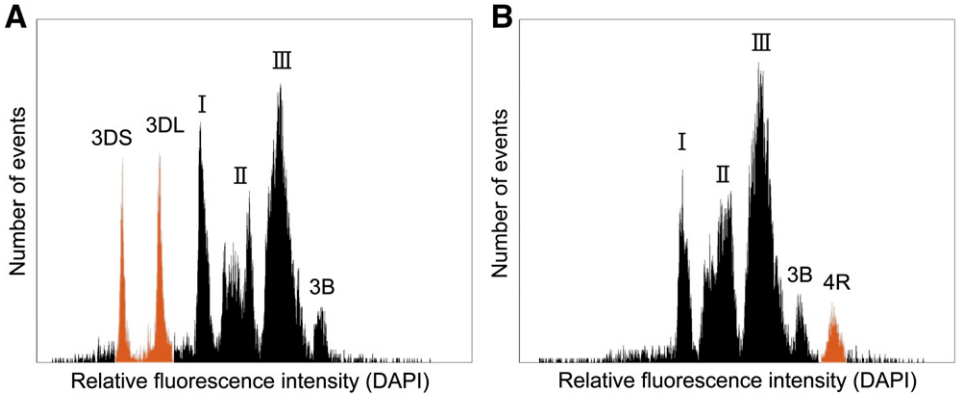

Fig. 4. The use of cytogenetic stocks to isolate particular wheat chromosomes. Flow karyotypes were obtained from DAPI-stained suspensions of mitotic chromosomes. (A) The double ditelosomic line dDt3D (20″ + t″3DS + t″3DL) carries the two arms of chromosome 3D in the form of two distinct telosomes, each 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

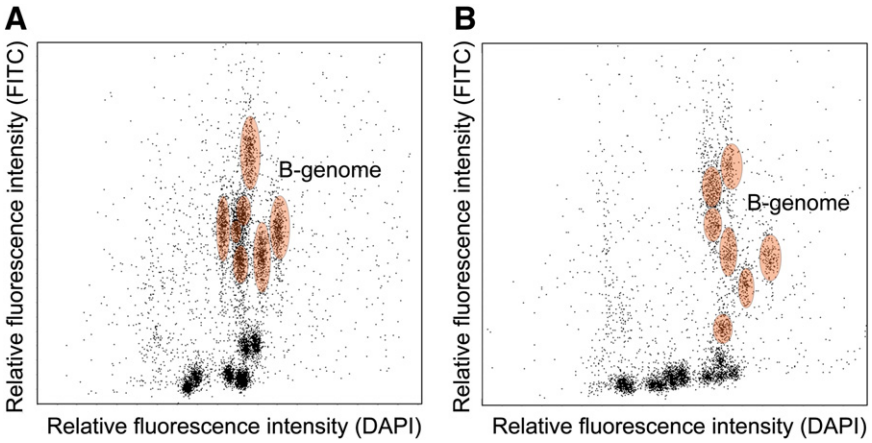

Fig. 5. Flow karyotypes of (A) tetraploid (cv. Creso, 2n = 28) and (B) hexaploid (cv. Chinese Spring, 2n = 42) wheat after the joint fluorescent labeling of GAA_n microsatellites and DAPI staining. The former was achieved by hybridization with a GAA_7-FITC probe, following the FISHIS procedure. The B 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).

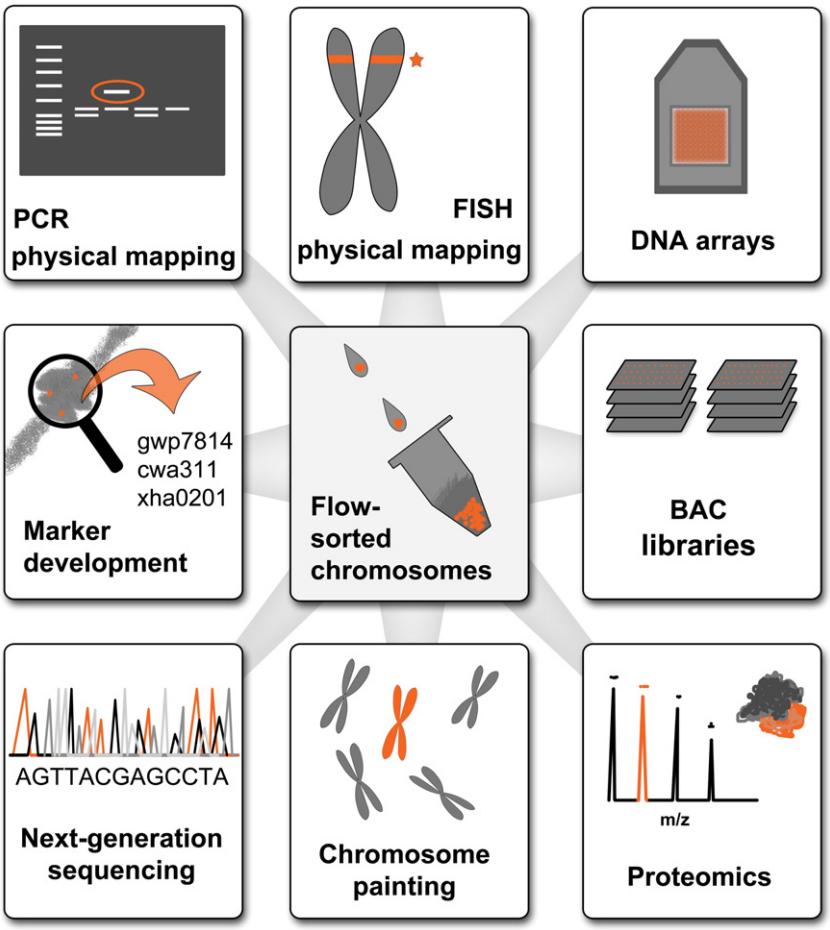

Fig. 6. Major current and potential uses of flow-sorted chromosomes.

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.

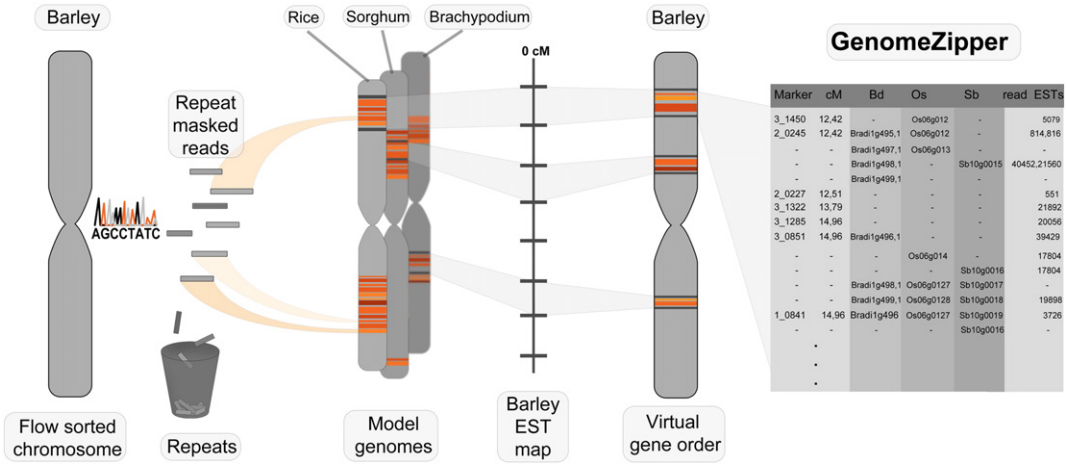

Fig. 7. Genome Zipper analysis in barley chromosome. MDA-amplified DNA of flow-sorted barley chromosomes was sequenced by 454 technology. Repeat-masked sequence reads identified syntenic regions in the rice, sorghum and *B. distachyon* genomes. Genes located in these regions were then aligned with a 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

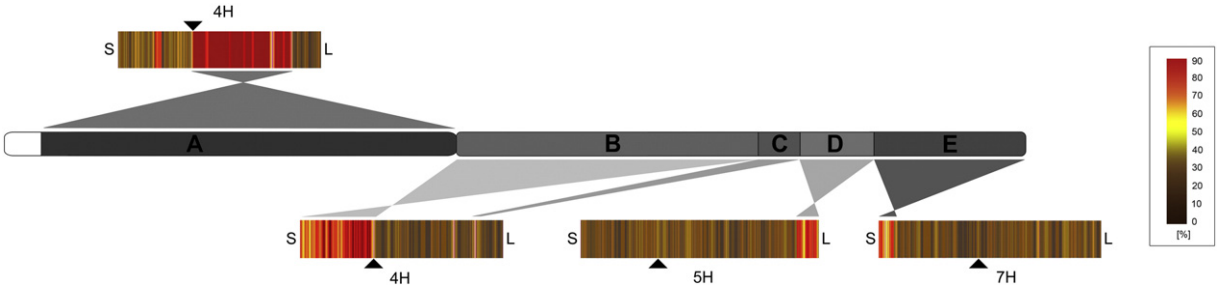

Fig. 8. The 4A shotgun sequence of barley. Repeat-masked 4AS and 4AL shotgun sequence reads were compared with the sequence of virtual barley chromosomes (Mayer et al., 2011). Syntenic regions on chromosomes 4H, 7H and 5H are shown in red, and non-syntenic regions in brown. The centromeres are indicated 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.