

Integrated consensus genetic and physical maps of flax (*Linum usitatissimum* L.)

Sylvie Cloutier · Raja Ragupathy · Evelyn Miranda ·
Natasa Radovanovic · Elsa Reimer · Andrzej Walichnowski ·
Kerry Ward · Gordon Rowland · Scott Duguid · Mitali Banik

Received: 28 March 2012 / Accepted: 21 July 2012 / Published online: 14 August 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract Three linkage maps of flax (*Linum usitatissimum* L.) were constructed from populations CDC Bethune/Macbeth, E1747/Viking and SP2047/UGG5-5 containing between 385 and 469 mapped markers each. The first consensus map of flax was constructed incorporating 770 markers based on 371 shared markers including 114 that were shared by all three populations and 257 shared between any two populations. The 15 linkage group map corresponds to the haploid number of chromosomes of this species. The marker order of the consensus map was largely collinear in all three individual maps but a few local inversions and marker rearrangements spanning short intervals were observed. Segregation distortion was present in all linkage groups which contained 1–52 markers

displaying non-Mendelian segregation. The total length of the consensus genetic map is 1,551 cM with a mean marker density of 2.0 cM. A total of 670 markers were anchored to 204 of the 416 fingerprinted contigs of the physical map corresponding to ~274 Mb or 74 % of the estimated flax genome size of 370 Mb. This high resolution consensus map will be a resource for comparative genomics, genome organization, evolution studies and anchoring of the whole genome shotgun sequence.

Introduction

Flax (*Linum usitatissimum* L., $2n = 2x = 30$), is an annual self-pollinated crop that is commercially grown as a source of stem fibre and seed oil. Flax seed oil is utilized for the fabrication of various biodegradable products such as high quality drying oil, paints, varnishes and linoleum flooring. Flax oil is a rich source of omega-3 fatty acids used as nutraceuticals and also as a functional food for both humans and animals. Fibre and oilseed flax belong to the same species but are morphologically different. Oilseed type flax plants (linseed) are more branched and shorter than the fibre type (Gill 1987). Fibre flax is grown mainly in Northern Europe, Russia and China but linseed is the primary type grown in Canada, USA, Argentina and India as well as Russia and China (Gill 1987; Marchenkov et al. 2003).

Development and characterization of flax genetic resources and assessment of genetic variability are essential for germplasm conservation and breeding. Flax germplasm collections contain thousands of accessions of *L. usitatissimum* and related species, of which, subsets were assessed for the extent of genetic diversity for morphological characteristics (Diederichsen and Hammer 1995; Diederichsen 2001; Diederichsen and Raney 2006;

Communicated by R. Visser.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-012-1953-0) contains supplementary material, which is available to authorized users.

S. Cloutier (✉) · R. Ragupathy · E. Miranda ·
N. Radovanovic · E. Reimer · A. Walichnowski ·
K. Ward · M. Banik
Cereal Research Centre, Agriculture and Agri-Food Canada,
195 Dafoe Road, Winnipeg, MB R3T 2M9, Canada
e-mail: sylvie.j.cloutier@agr.gc.ca

S. Cloutier
Department of Plant Science, University of Manitoba,
66 Dafoe Road, Winnipeg, MB R3T 2N2, Canada

G. Rowland
Crop Development Centre, University of Saskatchewan,
51 Campus Drive, Saskatoon, SK S7N 5A8, Canada

S. Duguid
Morden Research Station, Agriculture and Agri-Food Canada,
101 Route 100, Unit 100, Morden, MB R6M 1Y5, Canada

Saeidi 2008). A variety of molecular markers including random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) have been developed and used in assessing flax genetic diversity (Spielmeyer et al. 1998; Oh et al. 2000; Wiesner et al. 2001; Fu et al. 2003; Adugna et al. 2006; Fu 2006; Roose-Amsaleg et al. 2006; Cloutier et al. 2009, 2012; Uysal et al. 2010; Deng et al. 2010, 2011; Bickel et al. 2011; Kale et al. 2012; Rachinskaya et al. 2011; Soto-Cerda et al. 2011a, 2011b). While the reports are numerous, the number of informative markers in each of the studies is somewhat limited with the majority reporting between 9 and 60 markers only (Cloutier et al. 2012).

SSRs are stretches of DNA consisting of a variable number of short tandem repeats that are generally co-dominant, highly polymorphic, multi-allelic, relatively abundant, heritable, reproducible and reliable (Powell et al. 1996; Hwang et al. 2009). They also show cross-species usefulness and can be used in closely related species (Powell et al. 1996; Collard et al. 2005; Varshney et al. 2005). SSRs have been developed from genomic sequences or Expressed Sequence Tags (ESTs). In flax, Ragupathy et al. (2011) identified 4,064 putative SSRs from bacterial artificial chromosome (BAC) end sequences (BES). SSR markers have also been developed from various flax EST libraries (Cloutier et al. 2009; Soto-Cerda et al. 2011b) and from SSR-enriched genomic libraries or other genomic sequences (Roose-Amsaleg et al. 2006; Deng et al. 2010, 2011; Bickel et al. 2011; Kale et al. 2012; Rachinskaya et al. 2011). There are currently 1,326 SSR markers published in flax (Cloutier et al. 2012). SSR markers have been used for the construction of genetic maps of many plant species and provide dependable landmarks throughout the genome (Cheng et al. 2009; Studer et al. 2010). In flax, genetic maps and genetic diversity assessment were achieved with this type of marker (Fu and Peterson 2010; Cloutier et al. 2011; Soto-Cerda et al. 2011b). Genetic maps are useful for evolutionary and comparative studies as they provide both intra- and inter-species genome wide insights on recombination rates and gene rearrangements within or across chromosomes (Ball et al. 2010; Wang et al. 2011).

Only three individual flax linkage maps (Spielmeyer et al. 1998; Oh et al. 2000; Cloutier et al. 2011) have been published to date. The linkage map developed by Cloutier et al. (2011) had 113 markers, mostly SSRs, grouped into 24 linkage groups, while those of Spielmeyer et al. (1998) and Oh et al. (2000) were based on 213 AFLP markers forming 18 linkage groups and 94 RFLP/RAPD markers grouped into 15 linkage groups, respectively. The limitations of these maps reside in either or both the type and limited number of markers. Hence the need exists for a reliable, high density genetic map of flax that would serve as reference for a wide

variety of applications such as QTL mapping, map based gene cloning, marker assisted crop improvement, linkage disequilibrium (LD) mapping, phylogenetic analysis and anchoring of the whole genome shotgun sequence assembly. Consensus genetic linkage maps have been constructed for various plant species including *Arabidopsis* (Hauge et al. 1993), *Brassica* (Xu et al. 2010, Wang et al. 2011), barley (Langridge et al. 1995; Varshney et al. 2007), sorghum (Mace et al. 2009), wheat (Somers et al. 2004), rice (Antonio et al. 1996), maize (Cone et al. 2002), red clover (Isobe et al. 2009), lettuce (Truco et al. 2007), rye (Gustafson et al. 2009), soybean (Hwang et al. 2009), melon (Diaz et al. 2011), grapevine (Vezzulli et al. 2008), cowpea (Muchero et al. 2009), chickpea (Millan et al. 2010), potato (Danan et al. 2011), eucalyptus (Brondani et al. 2006), *Cucurbita pepo* (Zraidi et al. 2007) and *Zoysia* species (Li et al. 2010). High density consensus maps are well suited as references for the incorporation of information from genetically diverse individuals or multiple populations thus facilitating comparative analyses across germplasm.

Whole genome physical maps have been constructed for maize (Messing et al. 2004), *Brachypodium* (Gu et al. 2009), melon (González et al. 2010), grapevine (Scalabrini et al. 2010), *Arabidopsis* (Mozo et al. 1999), *Brassica rapa* (Mun et al. 2008), soybean (Wu et al. 2004), apple (Han et al. 2011) and flax (Ragupathy et al. 2011). BAC-based physical maps have been anchored to genetic maps in a number of plants such as rice (Chen et al. 2002), maize (Wei et al. 2009), papaya (Yu et al. 2009), *Medicago* (Mun et al. 2006), bean (Córdoba et al. 2010), poplar (Kelleher et al. 2007), grapevine (Scalabrini et al. 2010) and melon (González et al. 2010), where they were used to order physical maps and provide a framework for genome sequence assemblies.

The physical map of the flax genome cv. CDC Bethune, an oilseed flax variety, consists of 416 fingerprinted contigs (FPC) spanning ~368 Mb, very close to the estimated genome size of CDC Bethune of approximately 370 Mb (Ragupathy et al. 2011). The present study was intended to: (1) construct three independent genetic maps, (2) create an integrated consensus genetic map and (3) anchor the consensus genetic and physical maps of flax to provide the backbone information for ordering the whole genome shotgun (WGS) sequence assembly.

Materials and methods

Plant material, DNA extraction and marker amplification

Three segregating populations were used for mapping. CDC Bethune/Macbeth (BM) is comprised of 243 F₆-

derived recombinant inbred lines (RILs). The two parents are current varieties termed ‘conventional’ oilseed types because they contain 55–57 % linolenic acid, a “standard” amount for oilseed flax varieties. E1747/Viking (EV) received from S. Knapp (University of Georgia, USA) consists of 90 F₆-derived RILs generated from a cross between the low linolenic acid line E1747 and the European fibre flax variety Viking. SP2047/UGG5-5 (SU) is an F₁-derived doubled haploid (DH) population of 78 individuals. SP2047 is a solin breeding line characterized by its 2–4 % linolenic acid content and yellow seeds while UGG5-5 is a “high-lin” line with 65–70 % linolenic acid (Banik et al. 2011; Cloutier et al. 2011).

Genomic DNA was extracted from lyophilized leaf tissue (~100 mg fresh) of individual seedlings of all the segregating and parental lines of the three mapping populations using the DNeasy 96 plant kit according to manufacturer’s instructions (Qiagen Inc, Toronto, ON, Canada). The genomic DNA was quantified by fluorometer and re-suspended to a final concentration of 6 ng/μl. Amplification of template DNA with SSR primers was performed in 384-well plates in a final volume of 10 μl. The amplification products were resolved on an ABI 3130xl Genetic analyzer (Applied Biosystems, Foster City, CA, USA) and scored for segregation of parental alleles at each SSR locus. A total of five SNPs and seven genes (*fad2A*, *fad2B*, *fad3A*, *fad3B*, *dgatA*, *dgatB* and *ysc1*) were also positioned on the maps (Cloutier et al. 2011). Protocols and primer information for SSR markers Lu4 to Lu1193, Lu2097 to Lu2300 and Lu2331 to Lu3291 were previously described (Cloutier et al. 2009, 2011, 2012). In addition, SSR markers Lu2001 to Lu2096 were from previously published reports (Roose-Amsaleg et al. 2006; Deng et al. 2010, 2011) while Lu2301 to Lu2330 were designed from scaffold 505 of the flax WGS sequence assembly (<http://www.phytozome.net>), as previously described (Cloutier et al. 2009, 2012). References for individual markers of the consensus map are listed (Supplementary Table S1).

Anchoring the genetic and physical maps

Genetic and physical map anchoring was performed using complementary strategies. First, the CDC Bethune BAC library was screened with a subset of the SSR primers positioned on the genetic maps to identify BAC addresses and their corresponding FPC contigs (Ragupathy et al. 2011). SSR markers Lu2097-Lu2300 and Lu2331-Lu3291 were derived directly from BESs and, as such, were directly assigned BAC addresses and corresponding FPC contigs. These BES anchors were confirmed by performing BLASTn searches of the SSR primer sequences against the flax WGS sequence assembly (<http://www.phytozome.net>). Only perfect matches of both primer sequences to the same

scaffold were considered for anchoring. BLASTn searches were also performed using the BESs from which the SSRs were derived and matches with an *e* value <1e-25 were considered scaffold anchors. Similarly, primer BLAST and BLASTn were performed for SSR markers Lu4 to Lu1193 derived from ESTs (Cloutier et al. 2009, 2012). The marker name convention “marker name_FPC contig number” (e.g. Lu3156_405) was adopted to indicate positioning of the markers on the physical map. Markers not anchored to an FPC contig were labelled ‘0’ (e.g. Lu2312_0). Markers amplifying two or three loci were labelled with a single contig number but likely belong to paralogous contigs.

Map construction and linkage analysis

Linkage analysis was carried out independently for each mapping population using JoinMap 4.0 (Van Ooijen 2006) with a LOD of 4.0 and a maximum recombination frequency of 40. Marker segregation was tested against the expected Mendelian ratio of 1:1 using the Chi-square goodness-of-fit test. For conversion of recombination frequency into map distances expressed in centiMorgans (cM), the Kosambi mapping function which accounts for genetic interference from double cross over was used (Kosambi 1944).

The consensus linkage map was constructed based on the principle illustrated by Stam (1993) using JoinMap 4.0 (Van Ooijen 2006). LOD scores and pairwise recombination frequencies were computed for all linkage groups (LGs) of individual populations. They were then combined into a single group node in the navigation tree using the ‘grouping node’ command. Consensus LGs were obtained using the ‘combine groups for map integration’ function that is based on the presence of a minimum of two marker loci common to at least two populations on the basis of the mean recombination frequencies and combined LOD scores of pairwise data from the three segregating populations. The consensus map was constructed according to Alheit et al. (2011) and Gong et al. (2008) using the following parameters: Kosambi mapping function, regression mapping option, maximum recombination frequency of 40, LOD >1.0, ripple = 1, third round = yes and goodness-of-fit jump threshold for removal of loci = 5.0. Comparative analyses of marker distance and marker order were performed across individual maps and with the consensus map by visualization of the four final maps obtained as described above. The individual homologous LGs from the three populations were integrated using commonly mapped markers based on 411 segregating lines (243 RILs from BM, 90 RILs from EV and 78 DHs from SU). The 15 consensus LGs were numbered LG1 to LG15, in decreasing size order (cM). The population specific LGs were described using the population acronyms and the same

number as in the consensus map (BMLG1-BMLG16, EVLG1-EVLG18 and SULG1-SULG15).

Results

Individual genetic linkage maps

CDC Bethune/Macbeth

A total of 389 segregating marker loci were polymorphic in the BM population, of which, 385 marker loci assembled into 16 LGs spanning 2,007 cM, for an average density of one marker locus every 5.2 cM, and 4 marker loci remained unlinked (Table 1). LGs ranged from 27 to 187 cM and contained 9–40 markers. The *dgatA* and *dgatB* genes were mapped in this population. The population showed segregation distortion for 56 loci ($P < 0.05$) with equal numbers of loci skewed towards each parent (Supplementary Table S2).

E1747/Viking

The EV population was assayed with 443 polymorphic marker loci, 442 of which grouped into 18 LGs leaving a single marker unlinked (Table 1). The length of the LGs and number of marker loci per LG varied from 10 to 168 cM and 2–46, respectively. The total length of the map was 1,731 cM with a mean marker density of 3.9 cM between loci. 77 of the 442 loci diverged significantly from the expected 1:1 segregation ratio with 33 skewed towards E1747 and 44 skewed towards Viking (Supplementary Table S2).

SP2047/UGG5-5

The previously published SU map was based on 125 marker loci assembled in 24 linkage groups spanning

834 cM (Cloutier et al. 2011). The SU map constructed herein is more saturated and comprehensive with 477 polymorphic marker loci, of which 8 remained unlinked and the remaining 469 formed 15 LGs totalling 3,044 cM (Table 1). The length of the LGs and the number of loci per LG varied from 136 to 362 cM and 20–47, respectively. The approximate average marker density was one every 6.5 cM. A total of 168 loci showed segregation distortion (Supplementary Table S2). All LGs contained distorted markers except SULG13 which contained a single polymorphic marker in the SU population and SULG14 (Supplementary Table S2). A total of seven genes including six from fatty acid biosynthetic pathways (*fad2A*, *fad2B*, *fad3A*, *fad3B*, *dgatA* and *dgatB*) were also mapped.

Consensus map

A total of 795 markers generated 821 loci for a locus per marker ratio of 1.03 because 18 markers identified two loci and four markers identified three loci. Of the 821 loci scored in the three populations, 114 were common to all three populations and another 257 were common to two of the three populations (Supplementary Table S3). A total of 770 marker loci were assembled into the 15 LGs constituting the consensus map (Fig. 1; Table 1). Four additional LGs contained 19 marker loci ordered based on a single mapping population and 32 markers remained unlinked (Table 1; Supplementary Table S2). The total length of the consensus genetic linkage map was 1,551 cM and LGs ranged from 60 to 170 cM. The consensus map had an average marker density of one per 2.0 cM. Assuming a genome size of 370 Mb for CDC Bethune (Ragupathy et al. 2011), the genome wide ratio of physical to genetic distance averaged 239 Kb/cM, equivalent to an average of one marker per 481 Kb.

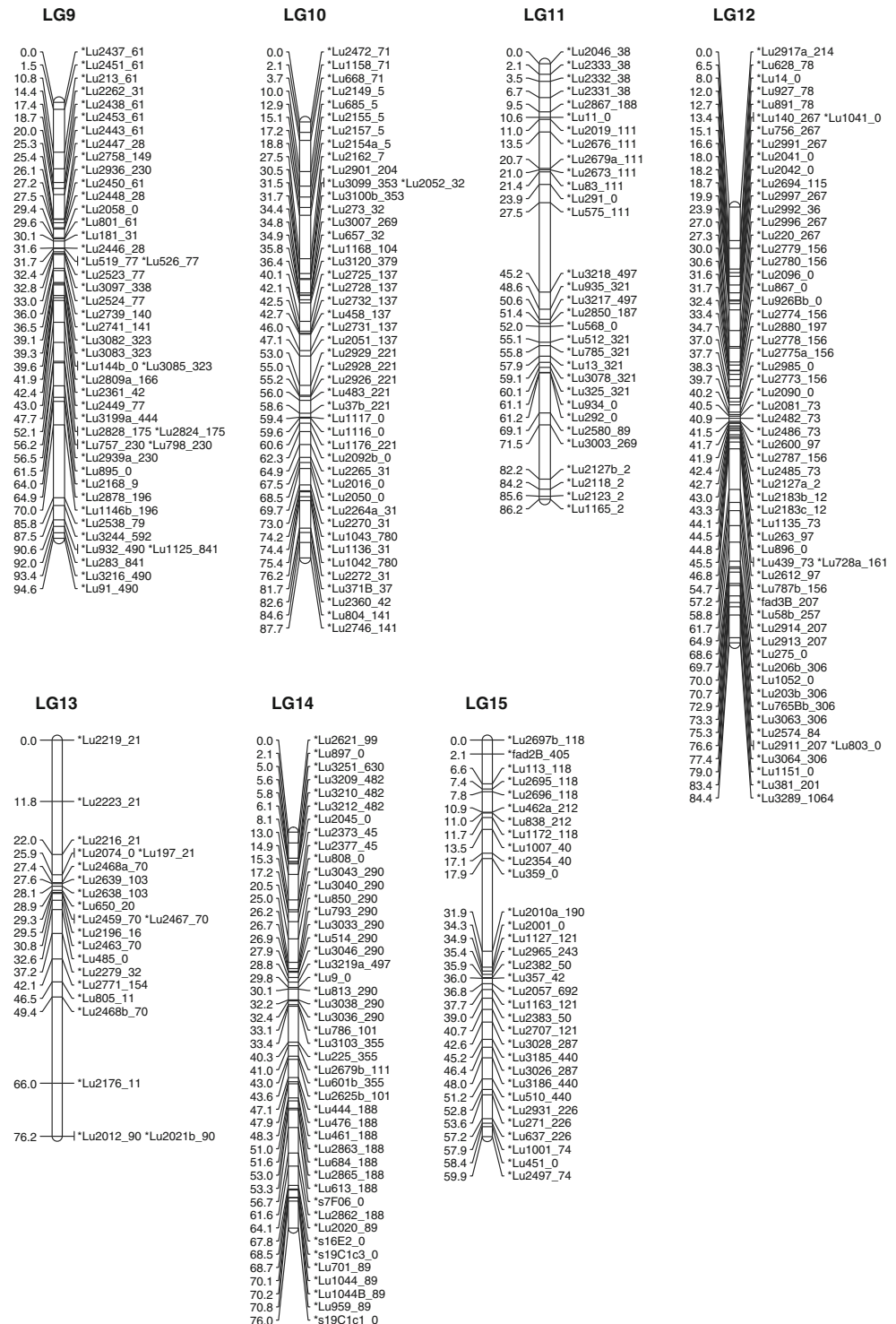
The length of the consensus map was shorter than the individual population specific maps (Table 1). All LGs of

Table 1 Mapping statistics for the three individual and the consensus genetic maps of flax

Populations	No. individuals	Total no. marker loci	No. marker loci in LGs ^a	Length (cM)	No. LGs ^a	No. unlinked marker loci	No. marker loci in the consensus map	Percent marker loci in the consensus map
CDC Bethune/Macbeth (BM)	243	389	385	2,007	16	4	373	95.9
E1747/Viking (EV)	90	443	442	1,731	18	1	419	94.6
SP2047/UGG5-5 (SU)	78	477	469	3,044	15	8	463	97.1
Consensus	411	821	770 (19)	1,551	15 (4)	32	770	93.8

^a Numbers in brackets represent marker loci and LGs belonging to a single population and that did not incorporate in the 15 LGs of the consensus map

Fig. 1 continued



the consensus map were constructed based on markers shared among the three populations except LG13 which was constructed mostly with markers from the BM and EV populations because all markers, with the exception of Lu850, were monomorphic in the SU population. The

consensus map displayed a few gaps that were mostly smaller than 10 cM. The largest gap of 20.8 cM is in LG1 (Fig. 1; Supplementary Table S1). The marker orders were consistent between the three independent genetic maps with some local inversions.

The percentages of distorted loci in BM, EV and SU populations were 15, 17 and 36 %, respectively (Supplementary Table S2). Out of the 770 loci of the consensus map, 292 (38 %) loci exhibited segregation distortion in at least one of the three mapping populations (Supplementary Table S2). The presence of a large number of distorted loci seems to have caused the large gap in LG11. Although LG8 carried the most distorted loci (76 %), the length of that LG was not affected as compared to the bi-parental populations where marker segregation was Mendelian, although some local rearrangements of closely linked markers were observed. Overall, only a few ambiguities were identified with respect to marker position compared to individual maps. Markers Lu359_0 and Lu2354_40, common to all three populations, mapped at the proximal end of each individual population LG15. However, these markers were placed at internal positions (41.977 and 42.829 cM) in the consensus LG15 (Supplementary Table S1).

Anchoring genetic and physical maps

Of the 770 loci in the consensus genetic map, 670 were anchored to 204 of the 416 FPC contigs of the physical map (Ragupathy et al. 2011) corresponding to 274 Mb or 74 % of the flax genome (Fig. 2). Twenty-one of the 204 FPC contigs were anchored at more than one map location because some markers amplified two or three polymorphic loci (labelled with a small a, b or c suffix in Fig. 1). FPC contigs were anchored with 1–14 markers. Examples of FPC contigs anchored with 14 markers include FPC contigs 79 and 82 estimated at 2.836 Mb with 205 BAC clones and 3.192 Mb with 265 BAC clones, respectively. The largest FPC contig (21), estimated at 5.562 Mb, consisted of 437 BAC clones and was anchored with four markers. Sixty-seven FPC contigs contained a single marker.

Discussion

Comparison of individual and consensus maps

Three genetic maps of flax have been published to date: a 213 AFLP marker-based map of 18 LGs covering 1,400 cM (Spielmeyer et al. 1998), a 1,000 cM RFLP/RAPD map with 94 markers assembled into 15 LGs (Oh et al. 2000) and a 113 marker map containing EST-SSRs, SNPs, genes and one phenotypic trait grouped into 24 LGs and spanning 834 cM (Cloutier et al. 2011). Major QTL for *fusarium* wilt (Spielmeyer et al. 1998), for fatty acid composition and seed coat colour (Cloutier et al. 2011) were identified using these genetic maps. The increased marker density of the SU map from 113 (Cloutier et al. 2011) to 469 (this publication) significantly improved the map by bridging gaps and thereby reducing the number of LGs from 24 to 15 while increasing the map coverage from 834 to 3,044 cM which promises to enhance QTL detection. Collinearity with two additional genetic maps (BM and EV) further confirmed the accuracy of the groupings. The three maps were largely collinear with few marker inversions. Although some LGs or portions thereof were fixed in some of the individual populations (e.g. LG4, LG13 and LG15 in SU; LG7, LG9 and LG15 in BM; LG15 in EV), the consensus map successfully bridged LGs and resulted in good coverage across the genome with few gaps. Some local inconsistencies of marker order such as small inversions or local rearrangements between individual and consensus maps were observed, particularly in closely linked markers and markers located at the distal ends of LGs as previously reported for rye (Studer et al. 2010), cotton (Xu et al. 2008), *Zoysia* (Li et al. 2010), grapevine (Vezzulli et al. 2008) and *Eucalyptus* (Brondani et al. 2006). The single most striking discrepancy resided in the total length of the SU map which exceeded the size of the BM and EV maps by more than 1,000 cM. The higher

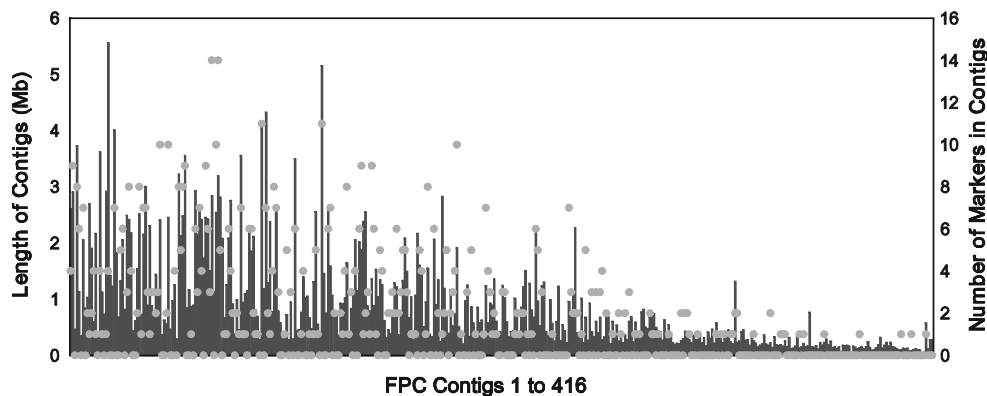


Fig. 2 Distribution of the genetic markers of the consensus map across the FPC contigs of the physical map (Ragupathy et al. 2011). A total of 204 of the 416 FPC contigs (x axis) were anchored by 670

marker loci (dots). The length of the contigs (Mb) is on the left y axis. From 1 to 14 marker loci (dots) were anchored onto each FPC contig represented on the right y axis

percentage (36 %) of distorted markers of this DH population, i.e. at least twice as high as the other two populations, may be responsible for the artifactual expansion of the genetic map length (Garcia-Dorado and Gallego 1992; Zhu et al. 2007; Li et al. 2011).

Comparative and consensus mapping are advantageous to obtain an unbiased linkage map representing the genome under investigation. As discussed above, mapping of markers employing multiple populations provides increased genome coverage because it is unlikely that multiple parents would all be fixed (monomorphic) in the same genomic regions. Also, overall population size afforded by multiple populations increases the chances of capturing recombination events, the foundation of genetic mapping. Reports of overrepresentation of localized crossover promoting 13mer motifs (Myers et al. 2008) in recombination hot spots of 1–2 kb (Ptak et al. 2005), and the influence of highly polymorphic trans-acting loci such as PRDM9 on the activation of those recombination hot-spots in human (Baudat et al. 2010; Paigen and Petkov 2010) indicates the importance of the genomic background in crossover frequencies. Considering crossing over as a fundamental cellular process conserved across eukaryotes, variability for distribution of recombination hot spots and its genetic determinants can be determined using multiple populations. Comparative mapping can also offer evidence for duplications or chromosomal rearrangements (Sewell et al. 1999). As a consequence of merging of datasets from three populations, the consensus map had fewer and smaller gaps compared to the individual genetic maps, hence it was more comprehensive. Fatty acid desaturase genes *fad2A*, *fad2B*, *fad3A* and *fad3B*, diacyl glycerol transferase genes *dgatA* and *dgatB* and seed coat color gene *ysc1* were positioned to seven different LGs of the consensus map. The majority were polymorphic in a single population but common neighbouring polymorphic markers permitted their integration in the consensus map, illustrating another advantage of consensus mapping. Flax has a relatively low level of genetic polymorphism, indicating a lower degree of genome divergence (Deng et al. 2010; Cloutier et al. 2011; Kale et al. 2012), unlike crops like maize where extensive molecular variation has been reported, primarily due to the activity of transposable elements (Llaca et al. 2011). The use of multiple populations followed by consensus mapping greatly increases marker saturation, a valuable feature for all mapping applications, for understanding the LD structure across genomes and germplasm characterization by association mapping (Soto-Cerda and Cloutier 2012).

The present consensus map of 770 SSR markers represents a major improvement over the low resolution phylogenetic analyses published to date with other marker types (McDill et al. 2009; Fu and Allaby 2010) and those

published with few SSR markers within *Linum usitatissimum* (Wiesner et al. 2001; Cloutier et al. 2009) and across *Linum* species (Fu and Peterson 2010; Soto-Cerda et al. 2011b). Pale flax (*Linum bienne* Mill, *L. angustifolia* Huds) is the wild progenitor of cultivated flax. Both have similar karyotypes bearing equal numbers of chromosomes ($2n = 2x = 30$) (Muravenko et al. 2003) and interspecific crosses between them produce fertile progeny (Gill and Yermanos 1967; Diederichsen and Hammer 1995). Pale and cultivated flax have been inferred to differ by a single translocation event (Gill and Yermanos 1967). The exceptionally high transferability (97 %) of EST-SSRs from cultivated flax to *L. bienne* supports the assignment of pale flax to the primary gene pool (Diederichsen 2007; Fu and Peterson 2010; Soto-Cerda et al. 2011b). A genetic map for pale flax or an interspecific cross map has yet to be produced. The current availability of SSR markers combined with their cross applicability should allow for an in-depth analysis of genetic diversity in *L. bienne* which should be useful to explore its potential to widen the gene pool of cultivated flax to meet breeding objectives.

Linum usitatissimum is a self-pollinated diploid species which, like a number of crop genomes, is an ancient polyploid (Blanc and Wolfe 2004; Paterson et al. 2004; Pfeil et al. 2005; Sterck et al. 2005; Gong et al. 2008; Soltis et al. 2009; Schmutz et al. 2010; Jiao et al. 2011; Lin and Paterson 2011). The remnant of ancestral whole genome duplication is reflected by the fact that a subset of SSR markers amplified two paralogous loci, although in most cases, only one was polymorphic (Cloutier et al. 2009, 2012). Mapping of the markers that amplified multiple polymorphic loci revealed ancestral chromosomal rearrangements resulting from paleopolyploidization events as noticed in LG6 and LG8. The existence of duplicated regions in consensus linkage groups LG6 and LG8 delimited by Lu2561 and Lu3057 markers indicate signatures of ancient duplication (Supplementary Figure 1). Analyses of a large collection of flax ESTs also corroborate the ancient duplication of flax (Venglat et al. 2011), as exemplified by the duplicate nature of the genes of the fatty acid biosynthetic pathway (Cloutier et al. 2011). Global comparative analysis of the scaffolds of the WGS sequence assembly promises a more comprehensive picture of the events that have shaped the flax genome through evolution.

Distorted markers

The segregation distortion in the three populations ranged from 15 to 36 %, an intermediate level, comparable to extent of distorted markers reported in common bean (37.3 %, de Campos et al. 2011), maize (19–36 %, Lu et al. 2002), red clover (5.8–45 %, Isobe et al. 2009), *Medicago truncatula* (27 %, Thoquet et al. 2002) and peanut

(8.5–22.8 %, Hong et al. 2010) but higher than *C. pepo* (3.7 %, Zraidi et al. 2007; Gong et al. 2008), *Brassica rapa* (2.6 %, Song et al. 1991), grapevine (7–11 %, Doligez et al. 2006) and globe artichoke (13 %, Portis et al. 2009). Species such as *Arabidopsis* (43.0 %, Reiter et al. 1992), cotton (71 %, Lacape et al. 2009), tomato (68 %, Paterson et al. 1988), perennial ryegrass (32–63 %, Anhalt et al. 2008), *Zoysia* (43.7 %, Li et al. 2010) and cowpea (41 %, Muchero et al. 2009), all displayed substantially higher percentages of markers deviating from the expected segregation ratios.

Non-Mendelian segregation ratios arise from chromosomal rearrangements, gametic competition, embryo viability and various physiological causes (Xu et al. 1997; Gonzalo et al. 2005; Portis et al. 2009) and, inadvertently, also from sampling errors (Lorieux et al. 2000). Segregation distortion has been associated more strongly with genetic effects as opposed to population structure or marker type (Anhalt et al. 2008).

Among the three flax populations reported here, the DH SU population had the highest proportion of distorted loci, similar to DH populations of rice (31.8 %, Xu et al. 1997) and *Brassica* (22–49 %, Wang et al. 2011), but lower than alfalfa (68 %, Li et al. 2011). The higher proportion of non-Mendelian markers in DH populations may be attributed to selection for tissue culture responsiveness loci (Xu et al. 1997; Alheit et al. 2011). Even though they were located on all LGs, distorted markers were not randomly distributed but were clustered within LGs (Cloutier et al. 2011; Córdoba et al. 2010; Li et al. 2011) supporting the cause of selection rather than experimental errors (Li et al. 2011), a view further emphasized by higher proportions of distorted markers on specific LGs such as LG2 (36/65), LG8 (52/68) and LG10 (32/46). Chromosome specific uneven distribution of markers has been reported for triticale chromosome 2A and 1R (Alheit et al. 2011); LG1, LG2 and LG3 of *Medicago truncatula* (Studer et al. 2010) and M20 and M32 of *Zoysia* (Li et al. 2010). Clustering of distorted markers was also documented in lettuce (Truco et al. 2007), *Eucalyptus* (Brondani et al. 2006) and peanut (Hong et al. 2010).

Map distance and map order can both be affected by segregation distortion (Lyttle 1991; Zhu et al. 2007) as was observed in the SU population where large gaps were observed between blocks of non-Mendelian markers adjacent to blocks of markers with non-skewed segregation, which accounted in part for the overestimation of the map length for this population. Elimination of non-Mendelian marker loci was suggested to improve mapping accuracies (Zhu et al. 2007; Xu 2008). However, such an approach would decrease the number of markers available and reduce coverage of some genomic regions, hence diminishing the map saturation (Zhu et al. 2007; Xu 2008). Here, we clearly demonstrated that consensus mapping was a

powerful way to correct for mapping inaccuracies caused by non-Mendelian markers because consensus mapping takes into account segregation data from multiple populations including common markers with Mendelian segregation in at least one population.

Anchoring genetic and physical maps

The physical map of flax is comprised of 416 FPC contigs spanning ~368 Mb (Ragupathy et al. 2011). A total of 670 markers were anchored to 204 FPC contigs representing ~274 Mb, i.e. 74 % of the estimated 370 Mb genome of CDC Bethune, comparable to papaya (72.4 %, Yu et al. 2009), apple (60 %, Han et al. 2011) and grapevine (72 %, Scalabrin et al. 2010) genomes and exceeding the extent of anchoring reported in *Medicago truncatula* (32 %, Mun et al. 2006), *Populus trichocarpa* (22 %, Kelleher et al. 2007), *Prunus* (15.5 %, Zhebentyayeva et al. 2008), bean (8 %, Córdoba et al. 2010) and melon (12 %, González et al. 2010). Although sufficient to provide initial ordering of the WGS sequence assembly into bins, the level of anchoring of the physical and genetic maps of flax presented herein falls short of the requirement for high accuracy ordering and orienting of the scaffolds of genomic sequence as was shown in maize (93 %, Wei et al. 2009) and rice (91 %, Chen et al. 2002). Tens of thousands of genome-wide SNPs currently being developed in our lab from the three mapping populations, using the state of the art ‘genotyping by sequencing (GBS)’ approach (Davey et al. 2011) will likely provide the degree of saturation necessary for the task of obtaining an accurate physical map ordering and orientation, a prerequisite for a high quality genome sequence assembly.

In conclusion, we reported on the construction of the first consensus genetic map of flax using 411 individuals from three populations and grouping and ordering 770 markers in 15 LGs spanning 1,551 cM. The vast majority of the markers are SSRs, a highly reproducible marker system which should prove its usefulness as an important resource for the flax research community, especially flax breeders. The overall map density averaged one marker every 2.0 cM. The consensus genetic map has been anchored to the flax physical map, a first step in the ordering of the scaffolds that currently make up the WGS sequence assembly of the flax genome. This integrated map will enable structural and functional genomic studies including fine mapping of genes of interest, marker-assisted flax breeding, map-based gene cloning, comparative/synteny mapping, QTL analysis and association mapping in flax and other related species.

Acknowledgments We sincerely thank Rae-Ann Trudeau for technical assistance, Joanne Schiavoni for manuscript editing and

Michael Shillinglaw for figure preparation. This research is part of the Total Utilization Flax GENomics (TUFGEN) project funded by Genome Canada and co-funded in part by the Agriculture Development Fund of Saskatchewan, the Governments of Manitoba and Saskatchewan, the Flax Council of Canada and the Manitoba Flax Growers Association. Project management and support by Genome Prairie are also acknowledged.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Adugna W, Labuschagne MT, Viljoen CD (2006) The use of morphological and AFLP markers in diversity analysis of linseed. *Biodivers Conserv* 15:3193–3205
- Alheit KV, Reif JC, Maurer HP, Hahn V, Weissmann EA, Miedaner T, Würschum T (2011) Detection of segregation distortion loci in triticale (x *Triticosecale* Wittmack) based on a high-density DArT marker consensus genetic linkage map. *BMC Genomics* 12:380
- Anhalt UCM, Heslop-Harrison PJS, Byrne S, Guillard A, Barth S (2008) Segregation distortion in *Lolium*: evidence for genetic effects. *Theor Appl Genet* 117:297–306
- Antonio BA, Inoue T, Kajiya H, Nagamura Y, Kurata N, Minobe Y, Yano M, Nakagahra M, Sasaki T (1996) Comparison of genetic distance and order of DNA markers in five populations of rice. *Genome* 39:946–956
- Ball AD, Stapley J, Dawson DA, Birkhead TR, Burke T, Slate J (2010) A comparison of SNPs and microsatellites as linkage mapping markers: lessons from the zebra finch (*Taeniopygia guttata*). *BMC Genomics* 11:218
- Banik M, Duguid S, Cloutier S (2011) Transcript profiling and gene characterization of three fatty acid desaturase genes in high, moderate, and low linolenic acid genotypes of flax (*Linum usitatissimum* L.) and their role in linolenic acid accumulation. *Genome* 54:471–483
- Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B (2010) PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 327:836–840
- Bickel CL, Gadani S, Lukacs M, Cullis CA (2011) SSR markers developed for genetic mapping in flax (*Linum usitatissimum* L.). *Res Rep Biol* 2011:23–29
- Blanc G, Wolfe KH (2004) Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* 16:1667–1678
- Brondani RPV, Williams ER, Brondani C, Grattapaglia D (2006) A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biol* 6:20
- Chen M, Presting G, Barbazuk WB, Goicoechea JL, Blackmon B, Fang G, Kim H, Frisch D, Yu Y, Sun S, Higingbottom S, Phimphilai J, Phimphilai D, Thurmond S, Gaudette B, Li P, Liu J, Hatfield J, Main D, Farrar K, Henderson C, Barnett L, Costa R, Williams B, Walser S, Atkins M, Hall C, Budiman MA, Tomkins JP, Luo M, Bancroft I, Salse J, Regad F, Mohapatra T, Singh NK, Tyagi AK, Soderlund C, Dean RA, Wing RA (2002) An integrated physical and genetic map of the rice genome. *Plant Cell* 14:537–545
- Cheng X, Xu J, Xia S, Gu J, Yang YF, Fu J, Qian X, Zhang S, Wu J, Liu K (2009) Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*. *Theor Appl Genet* 118:1121–1131
- Cloutier S, Niu Z, Datla R, Duguid S (2009) Development and analysis of EST-SSRs for flax (*Linum usitatissimum* L.). *Theor Appl Genet* 119:53–63
- Cloutier S, Ragupathy R, Niu Z, Duguid S (2011) SSR-based linkage map of flax (*Linum usitatissimum* L.) and mapping of QTLs underlying fatty acid composition traits. *Mol Breeding* 28:437–451
- Cloutier S, Miranda E, Ward K, Radovanovic N, Reimer E, Walichnowski A, Datla R, Rowland G, Duguid S, Ragupathy R (2012) Simple sequence repeat marker development from bacterial artificial chromosome end sequences and expressed sequence tags of flax (*Linum usitatissimum* L.). *Theor Appl Genet* 125:685–694
- Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142:169–196
- Cone KC, McMullen MD, Bi IV, Davis GL, Yim YS, Gardiner JM, Polacco ML, Sanchez-Villeda H, Fang Z, Schroeder SG, Havermann SA, Bowers JE, Paterson AH, Soderlund CA, Engler FW, Wing RA, Coe EH Jr (2002) Genetic, physical, and informatics resources for maize. On the road to an integrated map. *Plant Physiol* 130:1598–1605
- Córdoba JM, Chavarro C, Schlueter JA, Jackson SA, Blair MW (2010) Integration of physical and genetic maps of common bean through BAC-derived microsatellite markers. *BMC Genomics* 11:436
- Danan S, Veyrieras J-B, Lefebvre V (2011) Construction of a potato consensus map and QTL meta-analysis offer new insights into the genetic architecture of late blight resistance and plant maturity traits. *BMC Plant Biol* 11:16
- Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* 12:499–510
- de Campos T, Oblessuc PR, Sforça DA, Cardoso JMK, Baroni RM, de Sousa ACB, Carbonell SAM, Chioratto AF, Garcia AAF, Rubiano LB, de Souza AP (2011) Inheritance of growth habit detected by genetic linkage analysis using microsatellites in the common bean (*Phaseolus vulgaris* L.). *Mol Breeding* 27:549–560
- Deng X, Long SH, He DF, Li X, Wang YF, Liu J, Chen XB (2010) Development and characterization of polymorphic microsatellite markers in *Linum usitatissimum*. *J Plant Res* 123:119–123
- Deng X, Long SH, He DF, Li X, Wang YF, Hao DM, Qiu CS, Chen XB (2011) Isolation and characterization of polymorphic microsatellite markers from flax (*Linum usitatissimum* L.). *Afr J Biotechnol* 10:734–739
- Diaz A, Fergany M, Formisano G, Ziarsolo P, Blanca J, Fei Z, Staub JE, Zalapa JE, Cuevas HE, Dace G, Oliver M, Boissot N, Dogimont C, Pitrat M, Hofstede R, van Koert P, Harel-Beja R, Tzuri G, Portnoy V, Cohen S, Schaffer A, Katzir N, Xu Y, Zhang H, Fukino N, Matsumoto S, Garcia-Mas J, Monforte AJ (2011) A consensus linkage map for molecular markers and quantitative trait loci associated with economically important traits in melon (*Cucumis melo* L.). *BMC Plant Biol* 11:111
- Diederichsen A (2001) Comparison of genetic diversity of flax (*Linum usitatissimum* L.) between Canadian cultivars and a world collection. *Plant Breed* 120:360–362
- Diederichsen A (2007) *Ex-situ* collections of cultivated flax (*Linum usitatissimum* L.) and other species of the genus *Linum* L. *Genet Resour Crop Evol* 54:661–678
- Diederichsen A, Hammer K (1995) Variation of cultivated flax (*Linum usitatissimum* L. subsp. *usitatissimum*) and its wild

- progenitor pale flax (subsp. *angustifolium* (Huds.) Thell.). Genet Resour Crop Evol 42:263–272
- Diederichsen A, Raney JP (2006) Seed colour, seed weight and seed oil content in *Linum usitatissimum* accessions held by plant gene resources of Canada. Plant Breed 125:372–377
- Doligez A, Adam-Blondon AF, Cipriani G, Di Gasparo G, Laucou V, Merdinoglu D, Meredith CP, Riaz S, Roux C, This P (2006) An integrated SSR map of grapevine based on five mapping populations. Theor Appl Genet 113:369–382
- Fu YB (2006) Redundancy and distinctness in flax germplasm as revealed by RAPD dissimilarity. Plant Genet Resour 4:117–124
- Fu YB, Allaby RG (2010) Phylogenetic network of *Linum* species as revealed by non-coding chloroplast DNA sequences. Genet Resour Crop Evol 57:667–677
- Fu YB, Peterson GW (2010) Characterization of expressed sequence tag-derived simple sequence repeat markers for 17 *Linum* species. Bot 88:537–543
- Fu YB, Rowland GG, Duguid SD, Richards KW (2003) RAPD analysis of 54 North American flax cultivars. Crop Sci 43:1510–1515
- García-Dorado A, Gallego A (1992) On the use of classical tests for detecting linkage. J Heredity 83:143–146
- Gill KS (1987) Linseed. Publications and Information Division. Indian Council of Agricultural Research, New Delhi, p 386
- Gill KS, Yermanos DM (1967) Cytogenetic studies on the genus *Linum* I. Hybrids among taxa with 15 as the haploid chromosome number. Crop Sci 7:623–627
- Gong L, Stift G, Kofler R, Pachner M, Lelley T (2008) Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. Theor Appl Genet 117:37–48
- González VM, García-Mas J, Arús P, Puigdomènech P (2010) Generation of a BAC-based physical map of the melon genome. BMC Genomics 11:339
- Gonzalo MJ, Oliver M, García-Mas J, Monforte AJ, Dolcet-Sanjuan R, Katzir N, Arus P, Monforte A (2005) Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.). Theor Appl Genet 110:802–811
- Gu YQ, Ma Y, Huo N, Vogel JP, You FM, Lazo GR, Nelson WM, Soderlund C, Dvorak J, Anderson OD, Luo M-C (2009) A BAC-based physical map of *Brachypodium distachyon* and its comparative analysis with rice and wheat. BMC Genomics 10:496
- Gustafson JP, Ma X-F, Korzun V, Snape JW (2009) A consensus map of rye integrating mapping data from five mapping populations. Theor Appl Genet 118:793–800
- Han Y, Zheng D, Vimolmangkang S, Khan MA, Beever JE, Korban SS (2011) Integration of physical and genetic maps in apple confirms whole-genome and segmental duplications in the apple genome. J Exp Bot 62:5117–5130
- Hauge BM, Hanley SM, Cartinhour S, Cherry JM, Goodman HM (1993) An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant J 3:745–754
- Hong Y, Chen X, Liang X, Liu H, Zhou G, Li S, Wen S, Holbrook CC, Guo B (2010) A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. BMC Plant Biol 10:17
- Hwang T-Y, Sayama T, Takahashi M, Takada Y, Nakamoto Y, Funatsuki H, Hisano H, Sasamoto S, Sato S, Tabata S, Kono I, Hoshi M, Hanawa M, Yano C, Xia Z, Harada K, Kitamura K, Ishimoto M (2009) High-density integrated linkage map based on SSR markers in soybean. DNA Res 16:213–225
- Isobe S, Kölliker R, Hisano H, Sasamoto S, Wada T, Klimenko I, Okumura K, Tabata S (2009) Construction of a consensus linkage map for red clover (*Trifolium pratense* L.). BMC Plant Biol 9:57
- Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack L, dePamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473:97–100
- Kale SM, Pardeshi VC, Kadoo NY, Ghorpade PB, Jana MM, Gupta VS (2012) Development of genomic simple sequence repeat markers for linseed using next-generation sequencing technology. Mol Breeding 30:597–606
- Kelleher CT, Chiu R, Shin H, Bosdet IE, Krzywinski MI, Fjell CD, Wilkin J, Yin TM, DiFazio SP, Ali J, Asano JK, Chan S, Cloutier A, Grin N, Leach S, Lee D, Mathewson CA, Olson T, O'Connor K, Prabhu AL, Smailus DE, Stott JM, Tsai M, Wye NH, Yang GS, Zhuang J, Holt RA, Putnam NH, Vrebalov J, Giovannoni JJ, Grinwood J, Schmutz J, Rokhsar D, Jones SJM, Marra MA, Tuskan GA, Bohlmann J, Ellis BE, Ritland K, Douglas CJ, Schein JE (2007) A physical map of the highly heterozygous *Populus* genome: integration with the genome sequence and genetic map and analysis of haplotype variation. Plant J 50:1063–1078
- Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175
- Lacape J-M, Jacobs J, Arioli T, Derijcker R, Forestier-Chiron N, Llewellyn D, Jean J, Thomas E, Viot C (2009) A new interspecific, *Gossypium hirsutum* X *G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. Theor Appl Genet 119:281–292
- Langridge P, Karakousis A, Collins N, Kretschmer J, Manning S (1995) A consensus linkage map of barley. Mol Breeding 1:389–395
- Li M, Yuyama N, Hirata M, Wang Y, Han J, Cai H (2010) An integrated SSR based linkage map for *Zoysia matrella* L. and *Z. japonica* Steud. Mol Breeding 26:467–476
- Li X, Wang X, Wei Y, Brummer EC (2011) Prevalence of segregation distortion in diploid alfalfa and its implications for genetics and breeding applications. Theor Appl Genet 123:667–679
- Lin L, Paterson AH (2011) Size variation in homologous segments across divergent plant genomes. Mob Genet Elements 1:92–96
- Llaca V, Campbell MA, Deschamps S (2011) Genome diversity in maize. J Bot. doi:10.1155/2011/104172
- Lorieux M, Ndjioudjop M-N, Ghesquière A (2000) A first interspecific *Oryza sativa* X *O. glaberrima* microsatellite-based genetic linkage map. Theor Appl Genet 100:593–601
- Lu H, Romero-Severson J, Bernardo R (2002) Chromosomal regions associated with segregation distortion in maize. Theor Appl Genet 105:622–628
- Lyttle TW (1991) Segregation distorters. Annu Rev Genet 25:511–555
- Mace ES, Rami J-F, Bouchet S, Klein PE, Klein RR, Kilian A, Wenzl P, Xia L, Halloran K, Jordan DR (2009) A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. BMC Plant Biol 9:13
- Marchenkov A, Rozhmina T, Uschapovsky I, Muir AD (2003) Cultivation of flax. In: Muir AD, Westcott ND (eds) Flax: the genus *Linum*. CRC, New York, pp 74–91
- McDill J, Repplinger M, Simpson BB, Kadereit JW (2009) The phylogeny of *Linum* and Linaceae subfamily Linoideae, with implications for their systematics, biogeography, and evolution of heterostyly. Syst Bot 34:386–405
- Messing J, Bharti AK, Karlowski WM, Gundlach H, Kim HR, Yu Y, Wei F, Fuks G, Soderlund CA, Mayer KFX, Wing RA (2004) Sequence composition and genome organization of maize. Proc Natl Acad Sci USA 101:14349–14354
- Millan T, Winter P, Jüngling R, Gil J, Rubio J, Cho S, Cobos MJ, Iruela M, Rajesh PN, Tekeoglu M, Kahl G, Muehlbauer FJ

- (2010) A consensus genetic map of chickpea (*Cicer arietinum* L.) based on 10 mapping populations. *Euphytica* 175:175–189
- Mozo T, Dewar K, Dunn P, Ecker JR, Fischer S, Kloska S, Lehrach H, Marra M, Martienssen R, Meier-Ewert S, Altmann T (1999) A complete BAC-based physical map of the *Arabidopsis thaliana* genome. *Nat Genet* 22:271–275
- Muchero W, Diop NN, Bhat PR, Fenton RD, Wanamaker S, Pottorff M, Hearne S, Cisse N, Fatokun C, Ehlers JD, Roberts PA, Close TJ (2009) A consensus genetic map of cowpea [*Vigna unguiculata* (L) Walp.] and synteny based on EST-derived SNPs. *Proc Natl Acad Sci USA* 106:18159–18164
- Mun J-H, Kim D-J, Choi H-K, Gish J, Debelle F, Mudge J, Denny R, Endré G, Saurat O, Duzde A-M, Kiss GB, Roe B, Young ND, Cook DR (2006) Distribution of microsatellites in the genome of *Medicago truncatula*: a resource of genetic markers that integrate genetic and physical maps. *Genetics* 172:2541–2555
- Mun J-H, Kwon S-J, Yang T-J, Kim H-S, Choi B-S, Baek S, Kim JS, Jin M, Kim JA, Lim M-H, Lee SI, Kim H-I, Kim H, Lim YP, Park B-S (2008) The first generation of a BAC-based physical map of *Brassica rapa*. *BMC Genomics* 9:280
- Muravenko OV, Lemesh VA, Samatadze TE, Amosova AV, Grushetskaya ZE, Popov KV, Semenova OYu, Khotuyleya LV, Zelenin AV (2003) Genome comparisons with chromosomal and molecular markers for three closely related flax species and their hybrids. *Russ J Genet* 39:414–422
- Myers S, Freeman C, Auton A, Donnelly P, McVean G (2008) A common sequence motif associated with recombination hot spots and genome instability in humans. *Nat Genet* 40:1124–1129
- Oh TJ, Gorman M, Cullis CA (2000) RFLP and RAPD mapping in flax (*Linum usitatissimum*). *Theor Appl Genet* 101:590–593
- Paigen K, Petkov P (2010) Mammalian recombination hot spots: properties, control and evolution. *Nat Rev Genet* 11:221–233
- Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335:721–726
- Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc Natl Acad Sci USA* 101:9903–9908
- Pfeil BE, Schlueter JA, Shoemaker RC, Doyle JJ (2005) Placing paleopolyploidy in relation to taxon divergence: a phylogenetic analysis in legumes using 39 gene families. *Syst Biol* 54:441–454
- Portis E, Mauromicale G, Mauro R, Acquadro A, Scaglione D, Lanteri S (2009) Construction of a reference molecular linkage map of globe artichoke (*Cynara cardunculus* var. *scolymus*). *Theor Appl Genet* 120:59–70
- Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. *Trends Plant Sci* 1:215–222
- Ptak SE, Hinds DA, Koehler K, Nickel B, Patil N, Ballinger DG, Przeworski M, Frazer KA, Pääbo S (2005) Fine-scale recombination patterns differ between chimpanzees and humans. *Nat Genet* 37:429–434
- Rachinskaya OA, Lemesh VA, Muravenko OV, Yurkevich OYu, Guzenko EV, Bol'sheva NL, Bogdanova MV, Samatadze TE, Popov KV, Malyshev SV, Shostak NG, Heller K, Hotyleva LV, Zelenin AV (2011) Genetic polymorphism of flax *Linum usitatissimum* based on the use of molecular cytogenetic markers. *Russ J Genet* 47:56–65
- Ragupathy R, Rathinavelu R, Cloutier S (2011) Physical mapping and BAC-end sequence analysis provide initial insights into the flax (*Linum usitatissimum* L.) genome. *BMC Genomics* 12:217
- Reiter RS, Williams JG, Feldmann KA, Rafalski JA, Tingey SV, Scolnik PA (1992) Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc Natl Acad Sci USA* 89:1477–1481
- Roose-Amsaleg C, Cariou-Pham E, Vautrin D, Tavernier R, Solignac M (2006) Polymorphic microsatellite loci in *Linum usitatissimum*. *Mol Ecol Notes* 6:796–799
- Saeidi G (2008) Genetic variation and heritability for germination, seed vigour and field emergence in brown and yellow-seeded genotypes of flax. *Int J Plant Prod* 2:15–22
- Scalabrin S, Troglio M, Moroldo M, Pindo M, Felice N, Coppola G, Prete G, Malacarne G, Marconi R, Faes G, Jurman I, Grando S, Jesse T, Segala C, Valle G, Policriti A, Fontana P, Morgante M, Velasco R (2010) Physical mapping in highly heterozygous genomes: a physical contig map of the Pinot Noir grapevine cultivar. *BMC Genomics* 11:204
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X-C, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the paleopolyploid soybean. *Nature* 463:178–183
- Sewell MM, Sherman BK, Neale DB (1999) A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. *Genetics* 151:321–330
- Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS (2009) Polyploidy and angiosperm diversification. *Am J Bot* 96:336–348
- Somers DJ, Isaac P, Edwards K (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor Appl Genet* 109:1105–1114
- Song KM, Suzuki JY, Slocum MK, Williams PH, Osborn TC (1991) A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci. *Theor Appl Genet* 82:296–304
- Soto-Cerda BJ, Cloutier S (2012) Association mapping in plant genomes. In: Caliskan M (ed) Genetic diversity in plants. InTech, Rijeka, pp 29–54
- Soto-Cerda BJ, Carrasco RA, Aravena GA, Urbina HA, Navarro CS (2011a) Identifying novel polymorphic microsatellites from cultivated flax (*Linum usitatissimum* L.) following data mining. *Plant Mol Biol Rep* 29:753–759
- Soto-Cerda BJ, Urbina Saavedra H, Navarro Navarro C, Mora Ortega P (2011b) Characterization of novel genic SSR markers in *Linum usitatissimum* (L.) and their transferability across eleven *Linum* species. *Electronic J Biotechnol* 14:6
- Spielmeier W, Green AG, Bittisnich D, Mendham N, Lagudah ES (1998) Identification of quantitative trait loci contributing to Fusarium wilt resistance on an AFLP linkage map of flax (*Linum usitatissimum*). *Theor Appl Genet* 97:633–641
- Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. *Plant J* 3:739–744
- Sterck L, Rombauts S, Jansson S, Sterky F, Rouzé P, Van de Peer Y (2005) EST data suggest that poplar is an ancient polyploid. *New Phytol* 167:165–170
- Studer B, Kölliker R, Muylle H, Asp T, Frei U, Roldán-Ruiz I, Barre P, Tomaszewski C, Meally H, Barth S, Skøt L, Armstead IP, Dolstra O, Lübberstedt T (2010) EST-derived SSR markers used as anchor loci for the construction of a consensus linkage map in ryegrass (*Lolium* spp.). *BMC Plant Biol* 10:177
- Thoquet P, Ghérardi M, Journet E-P, Kereszt A, Ané J-M, Prospero J-M, Huguet T (2002) The molecular genetic linkage map of the

- model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. *BMC Plant Biol* 2:1
- Truco MJ, Antonise R, Lavelle D, Ochoa O, Kozik A, Witsenboer H, Fort SB, Jeuken MJW, Kesseli RV, Lindhout P, Michelmore RW, Peleman J (2007) A high-density, integrated genetic linkage map of lettuce (*Lactuca* spp.). *Theor Appl Genet* 115:735–746
- Uysal H, Fu Y-B, Kurt O, Peterson GW, Diederichsen A, Kusters P (2010) Genetic diversity of cultivated flax (*Linum usitatissimum* L.) and its wild progenitor pale flax (*Linum bienne* Mill.) as revealed by ISSR markers. *Genet Resour Crop Evol* 57:1109–1119
- Van Ooijen JW (2006) JoinMap[®] 4.0, Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, Netherlands
- Varshney RK, Sigmund R, Borner A, Korzun V, Stein N, Sorrells ME, Langridge P, Graner A (2005) Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Sci* 168:195–202
- Varshney RK, Marcel TC, Ramsay L, Russell J, Röder MS, Stein N, Waugh R, Langridge P, Niks RE, Graner A (2007) A high density barley microsatellite consensus map with 775 SSR loci. *Theor Appl Genet* 114:1091–1103
- Venglat P, Xiang D, Qiu S, Stone SL, Tibiche C, Cram D, Alting-Mees M, Nowak J, Cloutier S, Deyholos M, Bekkaoui F, Sharpe A, Wang E, Rowland G, Selvaraj G, Datla R (2011) Gene expression analysis of flax seed development. *BMC Plant Biol* 11:74
- Vezzulli S, Troggio M, Coppola G, Jermakow A, Cartwright D, Zharkikh A, Stefanini M, Grando MS, Viola R, Adam-Blondon AF, Thomas M, This P, Velasco R (2008) A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers. *Theor Appl Genet* 117:499–511
- Wang J, Lydiate DJ, Parkin IAP, Falentin C, Delourme R, Carion PWC, King GJ (2011) Integration of linkage maps for the Amphidiploid *Brassica napus* and comparative mapping with *Arabidopsis* and *Brassica rapa*. *BMC Genomics* 12:101
- Wei F, Zhang J, Zhou S, He R, Schaeffer M, Collura K, Kudrna D, Faga BP, Wissotski M, Golser W, Rock SM, Graves TA, Fulton RS, Coe E, Schnable PS, Schwartz DC, Ware D, Clifton SW, Wilson RK, Wing RA (2009) The physical and genetic framework of the maize B73 genome. *PLoS Genet* 5(11):e1000715
- Wiesner I, Wiesnerova D, Tejklova E (2001) Effect of anchor and core sequence in microsatellite primers on flax fingerprinting patterns. *J Agric Sci* 137:37–44
- Wu C, Sun S, Nimmakayala P, Santos FA, Meksem K, Springman R, Ding K, Lightfoot DA, Zhang H-B (2004) A BAC- and BIBAC-based physical map of the soybean genome. *Genome Res* 14:319–326
- Xu S (2008) Quantitative trait locus mapping can benefit from segregation distortion. *Genetics* 180:2201–2208
- Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F₂, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Mol Gen Genet* 253:535–545
- Xu Z, Kohel RJ, Song G, Cho J, Yu J, Yu S, Tomkins J, Yu JZ (2008) An integrated genetic and physical map of homoeologous chromosomes 12 and 26 in Upland cotton (*G. hirsutum* L.). *BMC Genomics* 9:108
- Xu J, Qian X, Wang X, Li R, Cheng X, Yang Y, Fu J, Zhang S, King GJ, Wu J, Liu K (2010) Construction of an integrated genetic linkage map for the A genome of *Brassica napus* using SSR markers derived from sequenced BACs in *B. rapa*. *BMC Genomics* 11:594
- Yu Q, Tong E, Skelton RL, Skelton RL, Bowers JE, Jones MR, Murray JE, Hou S, Guan P, Acob RA, Luo M-C, Moore PH, Alam M, Paterson AH, Ming R (2009) A physical map of the papaya genome with integrated genetic map and genome sequence. *BMC Genomics* 10:371
- Zhebentyayeva TN, Swire-Clark G, Georgi LL, Garay L, Jung S, Forrest S, Blenda AV, Blackmon B, Mook J, Horn R, Howad W, Arús P, Main D, Tomkins JP, Sosinski B, Baird WV, Reighard GL, Abbott AG (2008) A framework physical map for peach, a model Rosaceae species. *Tree Genet Genomes* 4:745–756
- Zhu C, Wang C, Zhang Y-M (2007) Modeling segregation distortion for viability selection I. Reconstruction of linkage maps with distorted markers. *Theor Appl Genet* 114:295–305
- Zraidi A, Stift G, Pachner M, Shojaeiyan A, Gong L, Lelley T (2007) A consensus map for *Cucurbita pepo*. *Mol Breeding* 20:375–388