

# A High-Density SNP Map for Accurate Mapping of Seed Fibre QTL in *Brassica napus* L.

Liezhao Liu<sup>1</sup>, Cunmin Qu<sup>1</sup>, Benjamin Wittkop<sup>2</sup>, Bin Yi<sup>3</sup>, Yang Xiao<sup>1</sup>, Yajun He<sup>1</sup>, Rod J. Snowdon<sup>2\*</sup>, Jiana Li<sup>1\*</sup>

**1** College of Agronomy and Biotechnology, Southwest University, Beibei, Chongqing, China, **2** Department of Plant Breeding, Interdisciplinary Research Centre for Biosystems, Land Use and Nutrition, Justus Liebig University, Giessen, Germany, **3** National Key Laboratory of Crop Genetic Improvement, National Subcenter of Rapeseed Improvement in Wuhan, Huazhong Agricultural University, Wuhan, China

## Abstract

A high density genetic linkage map for the complex allotetraploid crop species *Brassica napus* (oilseed rape) was constructed in a late-generation recombinant inbred line (RIL) population, using genome-wide single nucleotide polymorphism (SNP) markers assayed by the *Brassica* 60 K Infinium BeadChip Array. The linkage map contains 9164 SNP markers covering 1832.9 cM. 1232 bins account for 7648 of the markers. A subset of 2795 SNP markers, with an average distance of 0.66 cM between adjacent markers, was applied for QTL mapping of seed colour and the cell wall fiber components acid detergent lignin (ADL), cellulose and hemicellulose. After phenotypic analyses across four different environments a total of 11 QTL were detected for seed colour and fiber traits. The high-density map considerably improved QTL resolution compared to the previous low-density maps. A previously identified major QTL with very high effects on seed colour and ADL was pinpointed to a narrow genome interval on chromosome A09, while a minor QTL explaining 8.1% to 14.1% of variation for ADL was detected on chromosome C05. Five and three QTL accounting for 4.7% to 21.9% and 7.3% to 16.9% of the phenotypic variation for cellulose and hemicellulose, respectively, were also detected. To our knowledge this is the first description of QTL for seed cellulose and hemicellulose in *B. napus*, representing interesting new targets for improving oil content. The high density SNP genetic map enables navigation from interesting *B. napus* QTL to *Brassica* genome sequences, giving useful new information for understanding the genetics of key seed quality traits in rapeseed.

**Citation:** Liu L, Qu C, Wittkop B, Yi B, Xiao Y, et al. (2013) A High-Density SNP Map for Accurate Mapping of Seed Fibre QTL in *Brassica napus* L. PLoS ONE 8(12): e83052. doi:10.1371/journal.pone.0083052

**Editor:** Tongming Yin, Nanjing Forestry University, China

**Received:** August 17, 2013; **Accepted:** November 7, 2013; **Published:** December 26, 2013

**Copyright:** © 2013 Liu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This research was supported by National Natural Science Foundation of China (number 31171584), Chongqing Natural Science Foundation (number cstc2011jjA8000g), 948 project (number 2011-G23) and the Earmarked Fund for Modern Agro-industry Technology Research System (number CARS-13). This work also funded from Germany with a grant to LL from the "Forschungsfonds Raps" and additional support from Norddeutsche Pflanzenzucht H.G. Lembke KG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The work presented in this paper was funded in part by the commercial breeding company Norddeutsche Pflanzenzucht H.G. Lembke KG, who also provided data for the study from one location in Germany. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

\* E-mail: rod.snowdon@agrar.uni-giessen.de (RJS); ljn1950@swu.edu.cn (JL)

## Introduction

Precise linkage map construction is the first step for mapping of quantitative trait loci (QTL) and comparative genome analysis of interesting QTL regions. In oilseed rape (*Brassica napus* L.) a large number of low-density genetic maps, generated using electrophoretic marker systems like restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR), have been used to map qualitative and quantitative trait loci for a large number of traits [1–9].

A major disadvantage of many of these previous *B. napus* QTL mapping studies was an inability to derive and compare exact chromosomal locations for regions of interest. This situation can be improved using sequence data from tightly-linked markers, particularly as genome sequences become available for *Brassica* crops [10]. It will be extremely useful for closer examination of QTL and potential positional gene cloning to be able to navigate directly from genetic map positions to the genome sequence. In genetic maps with low marker densities, or higher-density maps

based on anonymous markers like AFLP or SRAP, this is possible only by labour-intensive development and addition of sequence-based markers to saturate regions of interest.

The most abundant and simple DNA markers for mapping and other applications are single nucleotide polymorphisms (SNPs). Today SNPs have become the marker of choice in most species for genome-wide association studies (GWAS), phylogenetic analyses, marker-assisted selection, bulked segregant analysis and genomic selection. In 2012 an international *Brassica* SNP consortium produced a 60,000 (60 k) SNP Infinium genotyping array for *B. napus*, in cooperation with Illumina Inc., San Diego, CA, USA [11–12]. This introduced a very low-cost and efficient method for high-density, sequence-based, genome-wide polymorphism screening in *B. napus* populations.

The use of high density genetic maps can greatly improve the precision of QTL localisation and the accuracy of effect estimates for detected QTL, especially for small and medium sized QTL [13]. Especially with the development of automated sequencing and genotyping technologies, many high density linkage maps have been constructed in different crops including oilseed rape/

canola (*B. napus*). For example, Sun et al. [14] constructed an ultradense genetic map consisting of 13351 SRAP markers covering 1604.8 cM in *B. napus*. Raman et al. (2012a) [15] used diversity array technology (DArT) markers to construct a linkage map for a doubled haploid (DH) mapping population of *B. napus*, and further constructed a consensus linkage map for genetic dissection of qualitative and quantitative traits (Raman et al. 2012b) [16]. The polymorphisms in the high throughput DArT marker system derive from restriction enzyme recognition sites and insertion/deletions (InDels). On the other hand, large fixed panels of robust SNP markers on widely-used array platforms are highly reproducible in different genetic backgrounds. A major advantage of genetic maps based on (public) genome-wide SNP screening arrays is the high occurrence of consensus markers for integration and alignment of maps and QTL, both to each other and also to reference genome sequences. Delourme et al. [17] developed an integrated *B. napus* map comprising 5764 SNP and 1603 PCR markers covering a total genetic length of 2250 cM. Chen et al. [18] also constructed a high density *B. napus* bin map, using a modified double-digested restriction-associated DNA sequencing (ddRADseq) approach.

Rapeseed is grown worldwide for vegetable oil and biodiesel production, and after oil extraction it also provides a high quality meal used primarily for livestock feeding [19]. Yellow-seeded *B. napus* is considered advantageous for the meal quality due to a thinner seed coat and higher protein content [20] along with reduced quantities of non-energetic fibre (cellulose and hemicellulose) and anti-nutritional polyphenolics (acid detergent lignin: ADL) [21]. Undigestible fiber, a major antinutritional component in rapeseed meal, can be reduced by breeding of light-seeded cultivars, whereas non-energetic cellulose and hemicellulose share photosynthesis products with seed oil and protein and are therefore important pleiotropic contributors to the agronomic value of the seed.

Numerous genetic mapping studies of seed colour loci have been reported in *Brassica* species using different biparental populations and marker technologies [22–35]. Many of these studies revealed QTL with different effects in different genetic backgrounds. A number of studies in *B. napus* suggested that one major locus on chromosome A09 explained most of the phenotype variation for both seed colour and meal quality traits in the most important *Brassica* oilseed crop [36–40]. However, attempts to saturate this QTL with markers [38,41] have revealed possible chromosome rearrangements that make it difficult to find markers and genes with close physical linkage to the QTL. Alignment of QTL from all of these different studies has also been rendered difficult by this complication [41], and because of a lack of consensus markers spanning the QTL in the different studies. Whereas most PCR-based markers are polynary markers, especially in the allopolyploid *B. napus*, the Illumina type II beadtype SNPs utilized in the present study are robust binary markers which can be compared directly among diverse genetic backgrounds.

In this study we constructed a new, high density genetic map of *B. napus*, using a homozygous, F<sub>9</sub> RIL population genotyped with the *Brassica* 60 k Infinium SNP array. To our knowledge this is the first high-density *B. napus* genetic map to be published with genotype data from the new *Brassica* 60 k SNP array. The map was used to precisely map QTL for seed colour, anti-nutritional seed ADL content, non-energetic seed cellulose and hemicellulose content using trait data from four field environments. The results provide tightly-linked and physically adjacent markers for breeding of these important agronomic traits in rapeseed, and a means to navigate directly from interesting map intervals to

*Brassica* genome sequences as a key step in map-based QTL cloning.

## Materials and Methods

### Mapping population and trait analysis

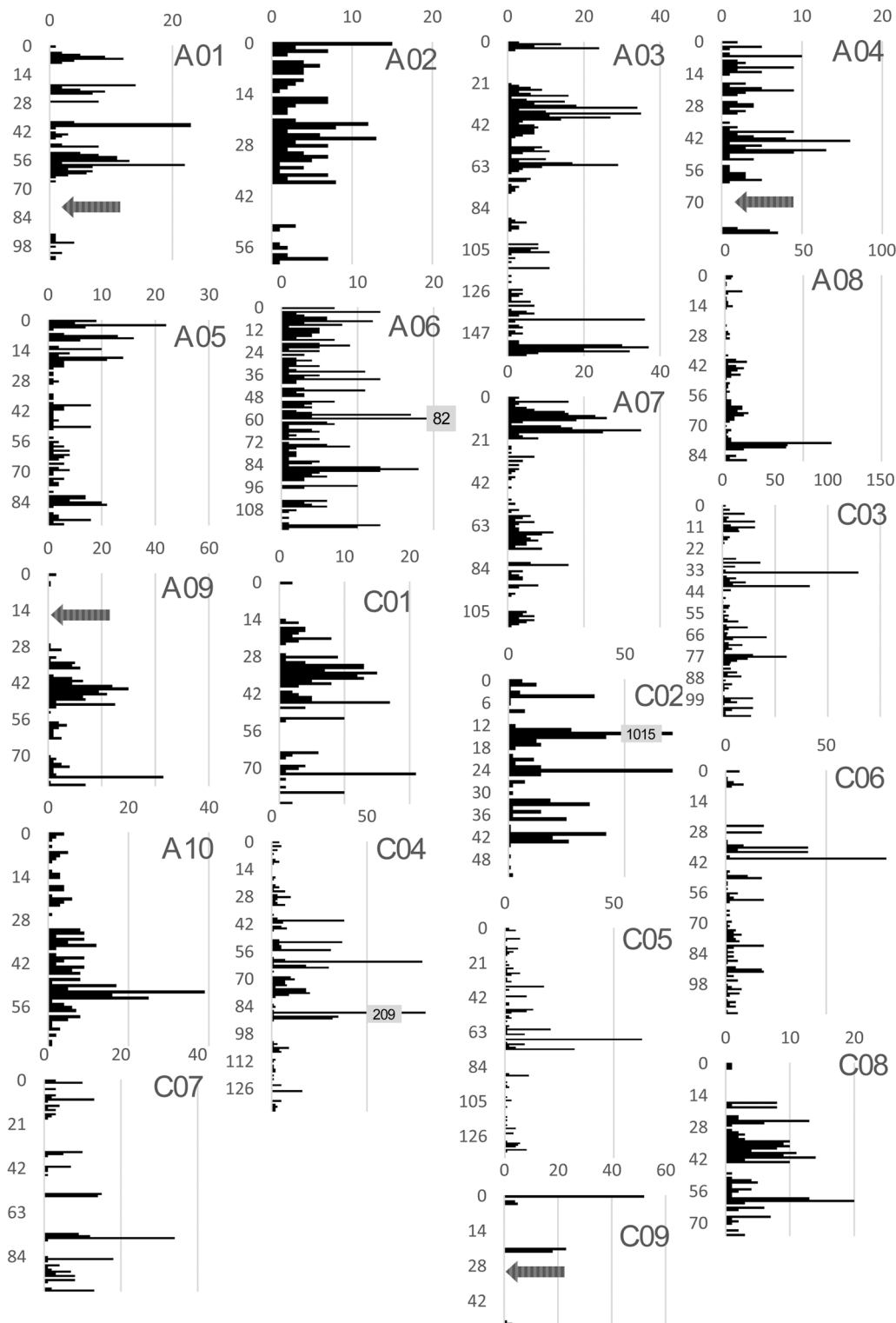
A population of 172 F<sub>9</sub> RILs was derived by single seed descent from F<sub>2</sub> offspring of a cross between the Chinese semi-winter oilseed rape parental lines GH06 (yellow seeds, low seed fibre content) and P174 (black seeds, high seed fibre content). Seed quality traits were measured on self-pollinated seed samples collected from field evaluations of the RIL population in four different environments over two years as follows: 2008 at the agricultural field station of Justus Liebig University Giessen at Weilburger Grenze (central Germany), in the breeding nursery of the German plant breeding company NPZ Lembke KG in Hohenlieth (northern Germany) and in the breeding nursery of Southwest University in Beibei, Chongqing (southwest China), and 2009 again at the field station Weilburger Grenze in Giessen. No specific permissions are required to grow conventional oilseed rape on these agricultural locations. Inflorescences of up to three randomly-chosen plants in each plot were covered in pollen-proof bags at the onset of flowering to prevent cross-pollination. Self-pollinated seeds were collected in the bags at maturity for quality analysis.

All trials were performed with a plot size of 4.5 m<sup>2</sup> (1.5 m × 3 m) with 5 or 6 rows per plot depending on the location. Measurements for seed colour and fibre components on the self-pollinated seeds were obtained by near-infrared reflectance spectroscopy (NIRS) using an NIR System 6500 with WinISI II software (FOSS GmbH, Rellingen, Germany). Phenotype values for seed colour (visual light absorbance), acid detergent lignin (ADL; % seed dry weight), acid detergent fibre (ADF; % seed dry weight) and neutral detergent fibre (NDF; % seed dry weight) were extrapolated from NIRS spectra using the NIRS calibrations for these traits as described by Wittkop et al. [20]. NIRS-derived estimates for each trait and genotype were averaged over two technical repetitions. In the RIL population, mean trait values were averaged from up to three self-pollinated plants from each genotype over the four environments. The predicted means (eblups) were used for QTL identification of the seed color and fiber traits. Cellulose concentrations were calculated as the difference between NDF and ADF, hemicellulose as the difference between ADF and ADL. Basic statistical analysis of the phenotype data was performed using Microsoft Excel 2010.

### SNP marker analysis

The *Brassica* 60 K SNP BeadChip Array was used to genotype 172 RIL lines and parental lines. This array, which successfully assays 52,157 Infinium Type II SNP loci in *B. napus*, was developed by an international consortium using preferentially single-locus SNPs contributed from genomic and transcriptomic sequencing in genetically diverse *Brassica* germplasm (Isobel Parkin, Agriculture and AgriFood Canada, unpublished data).

Total genomic DNA was extracted from 150 mg of leaf tissue from 5 young seedlings per RIL and from the two parental lines, using DP321-03 DNA extraction kits (Tiangen, Beijing, China). The DNA samples were diluted to 50 ng/μL. The SNP analysis was done in National Key Laboratory of Crop Genetic Improvement, National Subcenter of Rapeseed Improvement in Wuhan, Huazhong Agricultural University, 430070 Wuhan, China. DNA sample preparation, hybridisation to the BeadChip, washing, primer extension and staining were performed according to the work flow described in the Infinium HD Assay Ultra



**Figure 1. Overview of genome-wide SNP density in the *B. napus* SNP map.** The abscissa indicates the number of SNP markers in each map interval, while the ordinate shows the genetic distance along each of the 19 linkage groups corresponding to the 19 *B. napus* chromosomes. The peak of the QTL was pinpointed to a narrow region at the terminal end of the chromosome flanked by two adjacent SNP markers. doi:10.1371/journal.pone.0083052.g001

manual. Imaging of the arrays was performed using an Illumina HiSCAN scanner after BeadChip washing and coating. Allele calling for each locus was performed using the GenomeStudio genotyping software v2011 (Illumina, Inc.). Cluster definition was

based on genotype data from 216 *B. napus* lines (9 Beadchips) including the 172 RILs and parental lines from this population and 42 additional *B. napus* genotypes. SNP markers were named using SNP plus index numbers assigned by GenomeStudio,

**Table 1.** Summary of locus numbers, marker numbers, multiple-marker bins and linkage group genetic distances of the 19 A-genome (A01–A10) and C-genome (C1–C9) chromosomes in the *Brassica napus* SNP bin map.

Chromosome	Number of multi-marker bins	Number of markers in bins	Total number of markers	Total number of mapped loci (including bins)	Map length (cM)	Average distance between loci (cM)
A01	47	192	285	140	104.2	0.74
A02	37	123	201	116	60.2	0.52
A03	141	579	764	328	158.9	0.49
A04	40	124	191	106	83.9	0.79
A05	58	192	287	152	94.5	0.62
A06	93	392	548	249	119.0	0.48
A07	91	333	462	220	112.7	0.51
A08	81	350	435	166	86.7	0.52
A09	59	239	373	193	81.6	0.42
A10	62	214	301	149	68.5	0.46
C01	47	223	259	82	83.7	1.02
C02	62	1465	1507	103	52.4	0.51
C03	106	947	1033	191	107.7	0.56
C04	86	975	1032	142	137.6	0.97
C05	52	299	340	144	142.2	0.99
C06	80	481	543	142	111.5	0.79
C07	35	219	248	63	100.6	1.59
C08	46	176	222	92	75.1	0.82
C09	9	125	133	17	51.9	3.05
A genome	709	2738	3847	1819	970.2	0.53
C genome	523	4910	5317	976	862.7	0.88
Total (A+C)	1232	7648	9164	2795	1832.9	0.66

Markers showing identical genotype scores across the entire RIL population were grouped into a single bin and a single marker from each bin was used for QTL mapping. The number of mapped loci on each linkage group is the sum of the number of marker bins plus the number of individual markers not assigned to bins. doi:10.1371/journal.pone.0083052.t001

followed by the chromosome number. Positions of A-genome SNPs were provided by the array manufacturer, while C genome SNP source sequences were subjected to a BLAST search against the *B.oleracea* genome database (BRAD, <http://brassicadb.org/>) to locate chromosome positions (E value  $\leq 1e-50$ ).

### Linkage analysis and QTL mapping

Genetic linkage analysis was performed using the software packages MSTmap [42] and JoinMap 4.0 [43]. The algorithm implemented in MSTMap can efficiently handle ultra-dense datasets from 10,000 to 100,000 markers, and independent comparisons of MSTMap with JoinMap have found it to produce the most accurate maps for most experiments with very fast calculation times [44]. According to Wu et al. [42] the software generates extremely accurate map outputs when the data quality is high.

Polymorphic SNP markers were first grouped at LOD 5.0 using MSTmap, and then marker orders were determined by finding the minimum spanning tree of a graph for each linkage group based on pairwise recombination frequencies. The marker order and distance in each linkage group were recalculated and confirmed by Joinmap 4.0, applying the mapping function of Kosambi [45] and a minimum LOD score of 3.0. Marker pairs with zero recombination were assigned to the same genetic bin. A reference genetic map was constructed with SNP bins being designated

according to the index number of the first SNP marker in each bin.

Detection of QTL for seed colour, cellulose, hemicellulose and ADL content were performed in the RIL population by composite interval mapping using the QTL Cartographer software version WinQTLCart2.5 [46]. The LOD threshold for detection of significant QTL was set by permutation analysis with 300 permutations. The linkage map and QTL positions were drawn using the software Mapchart [47].

## Results

### SNP map construction

A total of 16613 SNP markers from the Brassica 60 K array showed polymorphisms between the mapping parents GH06 and P174. Of these, 9804 homologous SNP markers showing the expected segregation 1:1 ratio in the RIL population were used for genetic linkage analysis AND linkage map construction using MSTmap and Joinmap 4.0. SNP markers expected to be specific to the *B. napus* A and C genome were cleanly separated into different linkage groups according to their chromosomal origin by using MSTmap at LOD 5, and some groups remained intact even at LOD 10. Finally, a set of 9164 SNP markers were successfully assigned to linkage groups representing the 19 *B. napus* chromosomes of the A genome (A01–A10) and C genome (C01–C09), respectively. As expected from the variable size of the *B. napus*

**Table 2.** Significant QTL associated with seed colour and fiber traits in the GH06 x P174 F<sub>9</sub> RIL population (n = 172) after phenotyping in 4 different field environments.

Trait	Environment	QTL	Chromosome	Marker interval	LOD score	Additive effect	R <sup>2</sup> (%)
ADL content (% DW)	2008 Cq	<b>qADLA09</b>	A09	SNP21195A09-SNP21172A09	14.31	-0.66	32.8
		<b>qADLC05</b>	C05	SNP37372-SNP37409	4.52	-0.32	9.1
	2008 Gi	<b>qADLA09</b>	A09	SNP21195A09-SNP21172A09	15.39	-0.45	33.3
		<b>qADLC05</b>	C05	SNP37372-SNP37409	9.54	-0.28	14.1
	2008 Ho	<b>qADLA09</b>	A09	SNP21195A09-SNP21172A09	20.72	-0.75	42.8
		<b>qADLC05</b>	C05	SNP37372-SNP37409	5.33	-0.31	8.1
2009 Gi	<b>qADLA09</b>	A09	SNP21195A09-SNP21172A09	15.2	-0.53	31.6	
	<b>qADLC05</b>	C05	SNP37372-SNP37409	6.95	-0.28	13.1	
Seed colour	2008 Gi	<b>qCoIA09</b>	A09	SNP21195A09-SNP21172A09	17.8	-0.44	36.9
	2008 Ho	<b>qCoIA09</b>	A09	SNP21195A09-SNP21172A09	23.6	-0.62	47.2
	2009 Gi	<b>qCoIA09</b>	A09	SNP21195A09-SNP21172A09	18.56	-0.53	39.1
Cellulose content (% DW)	2008 Cq	<b>qCelA08</b>	A08	SNP17683A08-SNP17656A08	6.51	-0.18	13.5
		<b>qCelA09</b>	A09	SNP20823A09-SNP20748A09	11.33	0.23	21.9
		<b>qCelA08</b>	A08	SNP17683A08-SNP17656A08	6.27	-0.32	12.1
	2008 Gi	<b>qCelA09</b>	A09	SNP19925A09-SNP20492A09	5.38	0.55	4.7
		<b>qCelA08</b>	A08	SNP17683A08-SNP17656A08	8.86	-0.31	15.1
		<b>qCelC02</b>	C02	SNP37983C02-SNP35881C02	4.41	-0.20	6.4
	2008 Ho	<b>qCelA08</b>	A08	SNP17683A08-SNP17656A08	8.86	-0.31	15.1
		<b>qCelC02</b>	C02	SNP37983C02-SNP35881C02	4.41	-0.20	6.4
		<b>qCelC03</b>	C03	SNP34516C03-SNP34515C03	5.85	-0.25	7.9
2009 Gi	<b>qCelA08</b>	A08	SNP17683A08-SNP17656A08	6.91	-0.31	14.0	
	<b>qCelC03</b>	C03	SNP34516C03-SNP34515C03	5.95	-0.25	7.5	
	<b>qHemA03</b>	A03	SNP7833A03-SNP5891A03	6.29	-0.22	8.3	
Hemi-cellulose (% DW)	2008 Gi	<b>qHemA03</b>	A03	SNP7833A03-SNP5891A03	6.29	-0.22	8.3
	2008 Ho	<b>qHemA03</b>	A03	SNP7833A03-SNP5891A03	4.23	-0.15	7.3
	2009 Gi	<b>qHemA03</b>	A03	SNP7833A03-SNP5891A03	5.18	-0.18	8.3
		<b>qHemA06</b>	A06	SNP12895A06-SNP13359A06	5.51	0.19	10.2
<b>qHemA09</b>	A09	SNP20527A09-SNP20457A09	4.68	-0.24	16.9		

QTL were localized on high-density SNP map to allow navigation directly from QTL confidence intervals to genomic sequence regions of interest. Negative additive effect values indicate a decrease in the trait value caused by QTL alleles from parent GH06, while positive additive values indicate increasing trait values with QTL alleles from parent GH06. R<sup>2</sup> is the percentage of variation explained by each QTL. Cq: Chongqing, China; Ho, Hohenlieth, Germany; Gi: Giessen Germany. QTL consistent over different environments are shown in bold font.

doi:10.1371/journal.pone.0083052.t002

chromosomes [48] and from previous genetic mapping studies, the marker number and density varied considerably across the different chromosomes (Figure 1). In the A genome three chromosomes displayed gaps of more than 20 cM, while 1 chromosome in the C genome also showed a gap over 20 cM. The highest marker density was found on chromosome C02, with 1507 markers distributed over a genetic map distance of 52.4 cM.

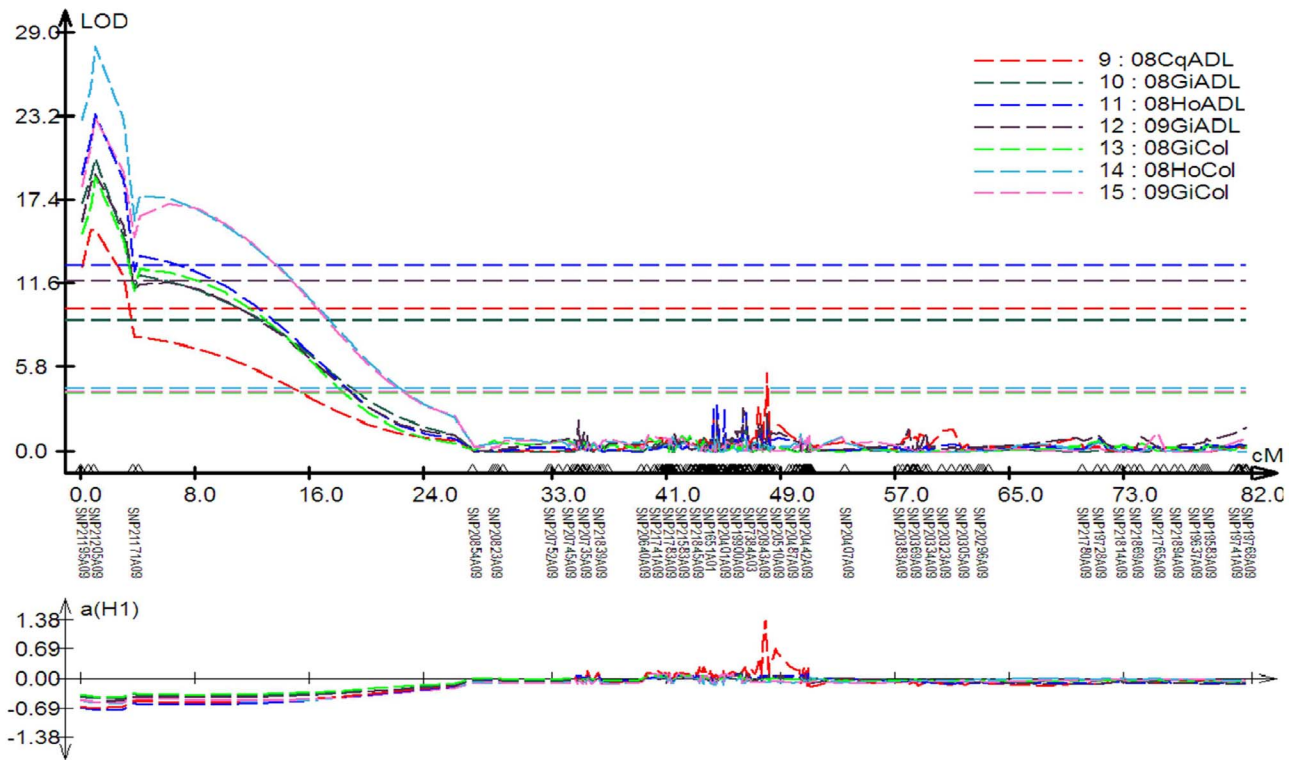
A SNP bin map comprising 1232 bins was constructed (File S1). The number of SNP markers per bin varied from 2 to 950. The largest bin on C02 containing 950 markers, and only 504 markers from this bin gave successful BLAST hits in the *B.oleracea* database to chromosome C02.

The final high density *B. napus* linkage map for the GH06 x P174 RIL population contains a total of 2795 distinct loci, including 976 loci on C genome chromosomes and 1819 loci in the A genome. The map covers 1832.9 cM, with total chromosome lengths of 970.2 cM and 862.7 cM for the A and C genomes, respectively (File S1). The marker density in A genome was thus considerably higher than in the C genome, with an average

distance between markers of 0.53 cM in the A genome and 0.93 cM in the C genome, respectively (Table 1).

### QTL analysis using the high-density SNP map

A single major QTL explaining 36.9% to 47.2% of the phenotypic variation for seed colour across the three different environments was detected at the end of linkage group A09 (Table 2, Figure 2, Figure 3). Two QTL for seed ADL content were localised on chromosomes A09 and C05, accounting for 31.6–42.8% and 8.1–14.1% of the phenotypic variation in the four different environments, respectively (Table 2, Figure 2 and Figure 3). The negative additive effects of these three QTL indicated that the yellow-seeded parent GH06 contributed to a strong decrease in seed colour and seed ADL content. The QTL for seed colour and ADL on chromosome A09 showed perfect colocalisation- This is presumably indicative of a strong pleiotropic effect caused by reduced seed coat thickness [36], which also explains their high correlation of R = 0.67 in the RIL population (File S2). The confidence interval of these two



**Figure 2. Graph showing a major QTL with large additive effects acid detergent lignin (ADL, four environments) and seed colour (Col, three environments) plotted according to logarithmic odds (LOD) score across the length of linkage group A09.** Locations: Cq, Chongqing China; Ho, Hohenlieth Germany; Gi, Giessen Germany. Years: 08, 2008; 09, 2009. doi:10.1371/journal.pone.0083052.g002

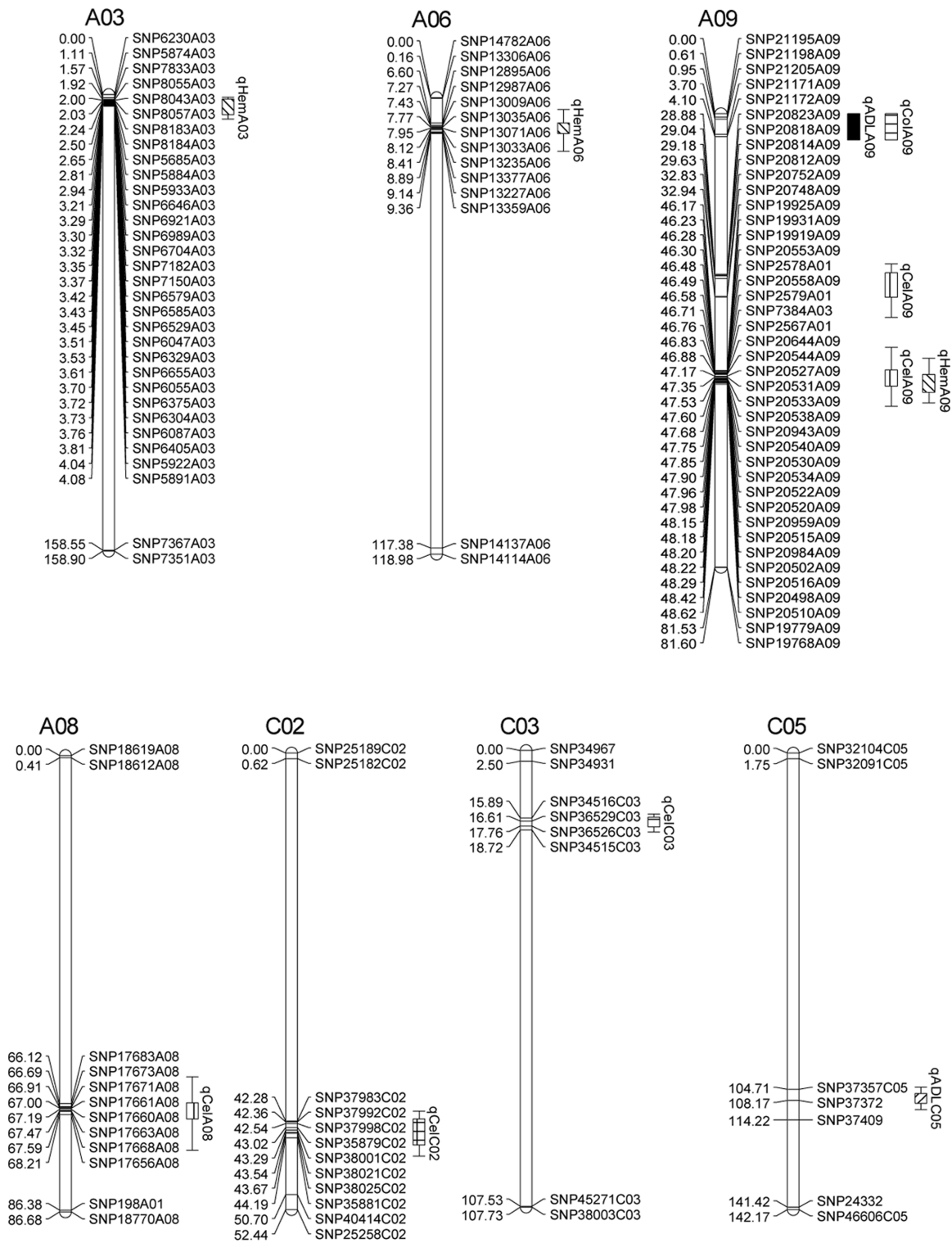
QTL spans a physical region from 32.22 Mbp to 32.84 Mbp in the *B. rapa* A genome reference sequence (Figure 2). We scanned the 620 kbp physical interval between the SNP markers flanking the QTL confidence interval for potential candidate genes involved in seed phenylpropanoid (lignin biosynthesis) or proanthocyanin pigmentation (flavanoid biosynthesis). The flavonoid biosynthesis gene Bra007813, homologous to *FLAVONOL 3-HYDROXYLASE (F3H)*, also known as *TRANSPARENT TESTA 6, TT6* was located 32.52 Mbp on A09. In the flavonoid biosynthetic pathway F3H catalyzes naringenin to dihydrokaempferol, a key precursor of proanthocyanins that accumulate as dark seed pigments (condensed tannins) in the testa of *Brassica* seeds. Other lignin biosynthesis and flavonoid genes which previous studies have suggested as potential candidates for the major QTL for seed colour and fibre content on chromosome A09 [36,38,41], were shown to be outside the narrow QTL confidence interval calculated using the high-density SNP map.

Based on the high density linkage map, five QTL with phenotypic effects on seed cellulose content ranging from 4.7% to 21.9% were detected across the four different environments. For four of these QTL the alleles from GH06 contributed to a decrease in seed cellulose content, whereas the decrease in cellulose caused by QTL *qCelA09* on chromosome A09 was contributed by P174. The QTL with the greatest effect on cellulose content, *qCelA08*, was detected across all four environments on chromosome A08 with an effect of 12.1% to 15.1% on the phenotypic variation. A single QTL for hemicellulose content was localised on chromosome A03, with a phenotypic effect of around 8% on this trait in the three German environments.

## Discussion

The identification of genome-wide SNP markers for *B. napus* based on growing quantities of next-generation sequencing data has opened the possibility for development of cost-effective, high-density SNP screening arrays for high-throughput genotyping of large oilseed rape mapping and breeding populations [12]. Whereas most QTL studies in *B. napus* in the past were based on lower-density genetic maps, often without sequence annotation of the marker loci, high-density SNP maps open the possibility for accurate alignment of *B. napus* genetic maps to physical chromosome segments in assembled *Brassica* genomes. This greatly increases the utility of genetic linkage maps for characterisation and comparison of QTL controlling important agronomic traits.

The marker density of a genetic map cannot be increased solely by increasing the number of markers used for genotyping, since physically linked markers will co-segregate on chromosome segments that show no recombination in a given population. By use of a F9 RILs mapping population we were able to capture a large amount of recombination in our mapping study, greatly increasing the number of individual loci that were able to be mapped to discrete genetic map positions. On this regard, one great advantage of robust SNP array screening platforms like the Brassica 60 k SNP BeadArray is the low proportion of missing data. Other high-density marker types, like Diversity Array Technology (DArT) or sequencing-based restriction site-associated DNA (RAD) markers, can be problematic for accurate genetic mapping because of the large amount of missing data normally associated with such methods: Missing data can be very difficult to distinguish from real polymorphisms and can therefore introduce significant errors into a high-density genetic map. To reduce



**Figure 3. Locations of significant QTL for acid detergent lignin (ADL), cellulose (Cel), hemicellulose (Hem) and seed colour (Col) in four different environments on the high-density SNP map.** For simplicity only the markers in the QTL confidence intervals, along with the terminal two markers at each end of the QTL-containing chromosomes, are shown. Full map data is provided in File S1. doi:10.1371/journal.pone.0083052.g003

mapping errors and avoid artificial exaggeration of map distances, missing data form must be dealt with using complex imputation analyses [49]. Furthermore, the use of high-throughput screening

arrays enables automated assaying of very large numbers of markers in a very short time (in our case 52157 SNP loci in 174 genotypes in around 3 days).

At 1832.4 cM the length of the SNP linkage map was somewhat longer than a previous 1589 cM map from a younger generation of the same population which was constructed using SSR, SRAP and AFLP markers [50]. This minor increase in map length is expected due to the huge increase in the number of mapped loci. The length is similar to that of the map reported by Chen et al. [16] for 8827 sequence-derived SNP loci mapped over 1860 cM. About 80 cM of our map was accounted for by 4 gaps of around 20 cM. These may result from distorted segregation caused by rearrangements among homoeologous chromosomes in one or both of the mapping parents [51]. Interestingly, the marker density and recombination frequency was considerably higher in A genome than the C genome in this population. This may reflect a frequent implementation of diverse *B. rapa* (A genome) germplasm in breeding of Chinese rapeseed, which may have reduced the extent of linkage disequilibrium and higher polymorphism in the A genome of the mapping parents used in this study. The largest bin, containing 950 SNP markers on chromosome C02, provides an example for a particularly recombination-poor region of the C genome in this cross. Interestingly, a relatively large number of SNP markers from the C genome of *B. napus* gave no BLAST hit in the *B. oleracea* database. Such results suggest significant genome-scale differences between *B. napus* and its C-genome progenitor *B. oleracea*.

Almost two decades after the first genetic mapping of seed colour loci in *B. napus* [22] the responsible genes are still unclear. The considerably higher accuracy for QTL detection enabled by a higher map density [13] now gives us the opportunity to more accurately localise candidate genes and compare different mapping studies via the genome sequence. In a previous study we fine-mapped the major locus for seed colour and ADL to one side of SSR marker KBrH092O19.5\_400, however no polymorphic reference markers could be identified in the nearby chromosome sequence on the other side of the QTL. In the present study the SNP marker SNP21195A09 was identified to the other side of the QTL peak, enabling the QTL to be clearly located on the *B. rapa* genome sequence. Xiao et al. [39] mapped a recessive gene for seed colour on chromosome A09 in *B. rapa*, however the physical region containing this gene (between 19.3 and 22.1 Mbp) does not correspond to the locus identified in our study. Furthermore, Kebede et al [40] mapped a major seed colour QTL in *B. rapa* on the middle of A09. These contrasting results suggest that different genes on chromosome A09 are involved in seed colour variation in different genetic backgrounds in *B. rapa* and *B. napus*. A major locus influencing seed colour and ADL content in a genetically diverse winter oilseed rape background was also mapped by Stein et al. [41] on A09. The most closely and consistently linked marker in that study, Ni4D09,

is located at 32.89 Mbp in the *B. rapa* sequence, only 4 kbp from the plasma membrane proton-ATPase Bra039228. This gene is an orthologue of another Arabidopsis *TRANSPARENT TESTA* gene, *AHA10*, which is a seed-expressed H<sup>+</sup>-ATPase pathway, Bra007813, was found within the QTL confidence interval, Bra007813 is a homolog of *FLAVANONE 3-HYDROXYLASE* (*F3H*, also known as *TRANSPARENT TESTA 6*) and thus another highly interesting candidate for seed variation. These results indicate that independent mutations in different phenylpropanoid genes may cause similar seed coat phenotypes in different genetic backgrounds.

To our knowledge this study is the first to report QTL for seed cellulose and hemicellulose content in *B. napus*. The environmentally consistent QTL on A08, which can reduce the seed cellulose content more than 10%, may be an interesting target for breeding to improve oil content. Both cellulose and hemicellulose showed negative correlations with seed oil content in the GH06 × P174 RIL population, demonstrating that reduction of cellulose and hemicellulose can improve the seed oil content due to a redirection of photosynthetic assimilates from sugar biosynthesis into seed oil biosynthesis. The two QTL qCelA09 and qHemA09 were found overlapped on A09 with opposite additive effects for cellulose and hemicellulose content, respectively. These two QTL may represent the same QTL with negative pleiotropic effects. Accurate mapping of this QTL represents a first step towards marker development and gene identification for potential improvement of oil content.

## Supporting Information

**File S1 Map data with bins.** Full list of SNP markers and bin map positions for the 19 *B. napus* chromosomes along with a summary of map distances and bin numbers. (XLSX)

**File S2 Fibre trait correlations.** Scatter plots and correlation coefficients among fiber traits across the different environments. (XLSX)

## Acknowledgments

The authors thank Martin Frauen and Felix Dreyer, NPZ Lembke KG, for providing seed samples for quality analysis from one location in Germany.

## Author Contributions

Conceived and designed the experiments: RJS JL LL. Performed the experiments: LL CQ BY BW. Analyzed the data: YX YH. Contributed reagents/materials/analysis tools: BW. Wrote the paper: LL RJS.

## References

- Landry BS, Hubert N, Etoh T, Harada JJ, Lincoln SE (1991) A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. *Genome* 34: 543–552.
- Dion Y, Gugel RK, Rakow GFW, SeguinSwartz G, Landry BS (1995) RFLP mapping of resistance to the blackleg disease [causal agent, *Leptosphaeria maculans* (Desm.) Ces, et de Not] in canola (*Brassica napus* L). *Theor Appl Genet* 91: 1190–1194.
- Ecke W, Uzunova M, Weissleder K (1995) Mapping the genome of rapeseed (*Brassica napus* L). II. Localization of genes controlling erucic acid synthesis and seed oil content. *Theor Appl Genet* 91: 972–977.
- Ferreira ME, Williams PH, Osborn TC (1995) Mapping of a locus controlling resistance to *Albugo candida* in *Brassica napus* using molecular markers. *Phytopathology* 85: 218–220.
- Ferreira ME, Satagopan J, Yandell BS, Williams PH, Osborn TC (1995) Mapping loci controlling vernalization requirement and flowering time in *Brassica napus*. *Theor Appl Genet* 90: 727–732.
- Ferreira ME, Rimmer SR, Williams PH, Osborn TC (1995) Mapping loci controlling *Brassica napus* resistance to *Leptosphaeria maculans* under different screening conditions. *Phytopathology* 85: 213–217.
- Foisset N, Delourme R, Barret P, Renard M (1995) Molecular tagging of the dwarf breizh (Bzh) gene in *Brassica napus*. *Theor Appl Genet* 91: 756–761.
- Uzunova M, Ecke W, Weissleder K, Robbelen G (1995) Mapping the genome of rapeseed (*Brassica napus* L). I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. *Theor Appl Genet* 90: 194–204.
- Lombard V, Delourme R (2001) A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. *Theor Appl Genet* 103: 491–507.
- Wang XW, Wang HZ, Wang J, Sun RF, Wu J, et al. (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat Genet* 43: 1035–1157.
- Snowdon RJ, Iniguez Luy FL (2012) Potential to improve oilseed rape and canola breeding in the genomics era. *Plant Breeding* 131: 351–360.
- Edwards D, Batley J, Snowdon RJ (2013) Accessing complex crop genomes with next-generation sequencing. *Theor Appl Genet* 126: 1–11.



13. Stange M, Utz HF, Schrag TA, Melchinger AE, Würschum T (2013) High-density genotyping: an overkill for QTL mapping? Lessons learned from a case study in maize and simulations. *Theor Appl Genet* 126: 2563–2574.
14. Sun ZD, Wang ZN, Tu JX, Zhang JF, Yu FQ, et al. (2007) An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. *Theor Appl Genet* 114: 1305–1317.
15. Raman H, Raman R, Nelson MN, Aslam MN, Rajasakaran R, et al. (2012) Diversity array technology markers: Genetic diversity analyses and linkage map construction in rapeseed (*Brassica napus* L.). *DNA Res* 19: 51–65.
16. Raman H, Raman R, Kilian A, Detering F, Long Y, et al. (2013) A consensus map of rapeseed (*Brassica napus* L.) based on diversity array technology markers: Applications in genetic dissection of qualitative and quantitative traits. *BMC Genomics* 14: 277.
17. Delourme R, Falentin C, Fomeju BF, Boillot M, Lassalle G, et al. (2013) High-density SNP-based genetic map development and linkage disequilibrium assessment in *Brassica napus* L. *BMC Genomics* 14: 120.
18. Chen X, Li X, Zhang B, Xu J, Wu Z, et al. (2013) Detection and genotyping of restriction fragment associated polymorphisms in polyploid crops with a pseudo-reference sequence: a case study in allotetraploid *Brassica napus*. *BMC Genomics* 14: 346.
19. Downey RK, Rakow G (1987) Rapeseed and mustard; Fehr WR, editor. New York: Macmillan Publishing Co. pp. 437–486.
20. Wittkop B, Snowdon RJ, Friedt W (2009) Status and perspectives of breeding for enhanced yield and quality of oilseed crops for Europe. *Euphytica* 170: 131–140.
21. Simbaya J, Slominski BA, Rakow G, Campbell LD, Downey RK, et al. (1995) Quality Characteristics of yellow-seeded *Brassica* seed meals: Protein, carbohydrates, and dietary fiber components. *J Agr Food Chem* 43: 2062–2066.
22. Van Deynze AE, Landry BS, Pauls KP (1995) The identification of restriction-fragment-length-polymorphisms linked to seed color genes in *Brassica napus*. *Genome* 38: 534–542.
23. Negi MS, Devic M, Delseny M, Lakshmikumaran M (2000) Identification of AFLP fragments linked to seed coat colour in *Brassica juncea* and conversion to a SCAR marker for rapid selection. *Theor Appl Genet* 101: 146–152.
24. Somers DJ, Rakow G, Prabhu VK, Friesen KRD (2001) Identification of a major gene and RAPD markers for yellow seed coat colour in *Brassica napus*. *Genome* 44: 1077–1082.
25. Liu ZW, Fu TD, Tu JX, Chen BY (2005) Inheritance of seed colour and identification of RAPD and AFLP markers linked to the seed colour gene in rapeseed (*Brassica napus* L.). *Theor Appl Genet* 110: 303–310.
26. Mahmood T, Rahman MH, Stringam GR, Raney JP, Good AG (2005) Molecular markers for seed colour in *Brassica juncea*. *Genome* 48: 755–760.
27. Padmaja KL, Arumugam N, Gupta V, Mukhopadhyay A, Sodhi YS, et al. (2005) Mapping and tagging of seed coat colour and the identification of microsatellite markers for marker-assisted manipulation of the trait in *Brassica juncea*. *Theor Appl Genet* 111: 8–14.
28. Badani AG, Snowdon RJ, Wittkop B, Lipsa FD, Baetzel R, et al. (2006) Colocalization of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*). *Genome* 49: 1499–1509.
29. Fu FY, Liu LZ, Chai YR, Chen L, Yang T, et al. (2007) Localization of QTLs for seed color using recombinant inbred lines of *Brassica napus* in different environments. *Genome* 50: 840–854.
30. Rahman M, McVetty PBE, Li G (2007) Development of SRAP, SNP and Multiplexed SCAR molecular markers for the major seed coat color gene in *Brassica rapa* L. *Theor Appl Genet* 115: 1101–1107.
31. Rahman M, Li GY, Schroeder D, McVetty PBE (2010) Inheritance of seed coat color genes in *Brassica napus* (L.) and tagging the genes using SRAP, SCAR and SNP molecular markers. *Mol Breeding* 26: 439–453.
32. Xu AX, Huang Z, Ma CZ, Xiao ES, Tian GW, et al. (2010) Inheritance of seed color and molecular markers linked to the seed color gene in *Brassica juncea*. *Mol Breeding* 25: 57–65.
33. Zhang Y, Li X, Chen W, Yi B, Wen J, et al. (2011) Identification of two major QTL for yellow seed color in two crosses of resynthesized *Brassica napus* line No. 2127–17. *Mol Breeding* 28: 335–342.
34. Guo SM, Zou J, Li RY, Long Y, Chen S, et al. (2012) A genetic linkage map of *Brassica carinata* constructed with a doubled haploid population. *Theor Appl Genet* 125: 1113–1124.
35. Snowdon RJ, Wittkop B, Rezaiddad A, Hasan M, Lipsa F, et al. (2010) Regional association analysis delineates a sequenced chromosome region influencing antinutritive seed meal compounds in oilseed rape. *Genome* 53: 917–928.
36. Rahman M, McVetty PBE (2011) A review of *Brassica* seed color. *Can J Plant Sci* 91: 437–446.
37. Liu LZ, Stein A, Wittkop B, Sarvari P, Li JN, et al. (2012) A knockout mutation in the lignin biosynthesis gene *CCR1* explains a major QTL for acid detergent lignin content in *Brassica napus* seeds. *Theor Appl Genet* 124: 1573–1586.
38. Xiao L, Zhao Z, Du DZ, Yao YM, Xu L, et al. (2012) Genetic characterization and fine mapping of a yellow-seeded gene in Dahuang (a *Brassica rapa* landrace). *Theor Appl Genet* 124: 903–909.
39. Kebede B, Cheema K, Greenshields DL, Li CX, Selvaraj G, et al. (2012) Construction of genetic linkage map and mapping of QTL for seed color in *Brassica rapa*. *Genome* 55: 813–823.
40. Huang Z, Ban YY, Yang L, Zhang Y, Li HQ, et al. (2012) Fine mapping of the yellow seed locus in *Brassica juncea* L. *Genome* 55: 8–14.
41. Stein A, Wittkop B, Liu L, Obermeier C, Friedt W, et al. (2013) Dissection of a major QTL for seed colour and fibre content in *Brassica napus* reveals colocalization with candidate genes for phenylpropanoid biosynthesis and flavonoid deposition. *Plant Breeding* 132: 382–389.
42. Wu Y, Bhat PR, Close TJ, Lonardi S (2008) Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS genetics* 4: e1000212.
43. Van Ooijen J, Voorrips R (2006) JoinMap 4.0. Software for the calculation of genetic linkage maps in experimental populations.
44. Cheema J, Dicks J (2009) Computational approaches and software tools for genetic linkage map estimation in plants. *Brief Bioinform* 10: 595–608.
45. Kosambi DD (1943) The estimation of map distances from recombination values. *Ann Eugenetic* 12: 172–175.
46. Wang S, Basten CJ, Zeng ZB (2006) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC, USA.
47. Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. *J Hered* 93: 77–78.
48. Snowdon RJ, Friedrich T, Friedt W, Köhler W (2002) Identifying the chromosomes of the A and C genome diploid *Brassica* species *B. rapa* and *B. oleracea* in their amphidiploid *B. napus*. *Theor Appl Genet* 104: 533–538.
49. Spindel J, Wright M, Chen C, Cobb J, Gage J, et al. (2013) Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theor Appl Genet* 126: 2699–2716.
50. Yan XY, Li JN, Fu FY, Jin MY, Chen L, et al. (2009) Co-location of seed oil content, seed hull content and seed coat color QTL in three different environments in *Brassica napus* L. *Euphytica* 170: 355–364.
51. Lu W, Liu J, Xin Q, Wan LL, Hong DF, et al. (2013) A triallelic genetic male sterility locus in *Brassica napus*: an integrative strategy for its physical mapping and possible local chromosome evolution around it. *Ann Bot-London* 111: 305–315.