BMC Genomics

RESEARCH ARTICLE

High-density genetic map construction and QTLs identification for plant height in white jute (*Corchorus capsularis* L.) using specific locus amplified fragment (SLAF) sequencing

Aifen Tao[1], Long Huang[2], Guifen Wu[3], Reza Keshavarz Afshar[4], Jianmin Qi[1*], Jiantang Xu[1], Pingping Fang[1], Lihui Lin[1], Liwu Zhang[1] and Peiqing Lin[1]

Abstract

Background: Genetic mapping and quantitative trait locus (QTL) detection are powerful methodologies in plant improvement and breeding. White jute (*Corchorus capsularis* L.) is an important industrial raw material fiber crop because of its elite characteristics. However, construction of a high-density genetic map and identification of QTLs has been limited in white jute due to a lack of sufficient molecular markers. The specific locus amplified fragment sequencing (SLAF-seq) strategy combines locus-specific amplification and high-throughput sequencing to carry out *de novo* single nuclear polymorphism (SNP) discovery and large-scale genotyping. In this study, SLAF-seq was employed to obtain sufficient markers to construct a high-density genetic map for white jute. Moreover, with the development of abundant markers, genetic dissection of fiber yield traits such as plant height was also possible. Here, we present QTLs associated with plant height that were identified using our newly constructed genetic linkage groups.

Results: An F_8 population consisting of 100 lines was developed. In total, 69,446 high-quality SLAFs were detected of which 5,074 SLAFs were polymorphic; 913 polymorphic markers were used for the construction of a genetic map. The average coverage for each SLAF marker was 43-fold in the parents, and 9.8-fold in each F_8 individual. A linkage map was constructed that contained 913 SLAFs on 11 linkage groups (LGs) covering 1621.4 cM with an average density of 1.61 cM per locus. Among the 11 LGs, LG1 was the largest with 210 markers, a length of 406.34 cM, and an average distance of 1.93 cM between adjacent markers. LG11 was the smallest with only 25 markers, a length of 29.66 cM, and an average distance of 1.19 cM between adjacent markers. 'SNP_only' markers accounted for 85.54% and were the predominant markers on the map. QTL mapping based on the F_8 phenotypes detected 11 plant height QTLs including one major effect QTL across two cultivation locations, with each QTL accounting for 4.14–15.63% of the phenotypic variance.

(Continued on next page)

---

* Correspondence: qijm863@163.com
[1]Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops; Key Laboratory of Crops by Design, Fujian Agriculture and Forestry University, Fuzhou 350028, People's Republic of China
Full list of author information is available at the end of the article

Tao *et al. BMC Genomics* (2017) 18:355

Page 2 of 12

(Continued from previous page)

Conclusions: To our knowledge, the linkage map constructed here is the densest one available to date for white jute. This analysis also identified the first QTL in white jute. The results will provide an important platform for gene/QTL mapping, sequence assembly, genome comparisons, and marker-assisted selection breeding for white jute.

Keywords: *Corchorus capsularis* L, SLAF, Genetic map, QTL, Plant height

Background

Jute (*Corchorus* sp.) is the second most cultivated fiber crop globally after cotton, and is extensively grown in India, Bangladesh, China, Thailand, Myanmar, and Nepal [1]. White jute (C. *capsularis*) and dark jute (C. *olitorius*) are two most popular cultivated *Corchorus* species. Both species have 2n = 2× = 14 chromosomes [2]. Jute is of importance to textile and paper industries as it is a valuable ingredient for producing paper and fine textiles, as well as being a renewable source for biofuel [3]. Jute fibers exhibit a characteristically high luster, good moisture absorption performance, rapid water loss capacity, and easy degradation [4]. Therefore, the use of jute has received considerable attention in many countries [5]. However, compared with other related bast fiber crops, there has relatively little progress in genetic improvement in jute in recent times, and no new breeding approaches have been developed over the past seven decades [6]. The advances in sequencing technologies have offered an alternative approach to traditional breeding methods and jute breeders and biologists are now giving more attention to the use of molecular tools to increase fiber quality and yield, and to improve agronomic traits [6].

A genetic map, especially a high-density genetic map, provides an important foundation for mapping quantitative trait loci (QTLs) [7–10] and anchoring sequence scaffolds [11–13]. High-density genetic maps have been used to reveal genome composition and for selection of high throughput superior traits in many species [14]. Previous research in jute has shown that plant height is a major component of fiber yield, and is strongly correlated with both fiber content and fiber yield [15]. Thus, construction of a high quality genetic map and identification of QTLs for plant height are necessary for further development of white jute.

Genetic maps have been developed in dark jute and used for identification of QTLs. For example, Kundu et al. developed a linkage map containing 503 restriction site associated DNA (RAD) markers spanning 358.5 cM, and 9 QTLs for histological fiber content were detected [15]. Topdar et al. reported a microsatellite genetic map and identified 26 QTLs for fiber quality, yield, and yield-related traits [16]. Sultana et al. [17], constructed a linkage map with three linkage groups covering 87.3 cM using 10 inter-simple sequence repeats (ISSR) in an F$_2$ population

from a cross between two dark jute genotypes. Haque et al. reported a linkage map including 40 randomly amplified polymorphic DNA (RAPD) markers [18]. Chen et al. developed a genetic linkage map using 122 sequence-related amplified polymorphism (SRAP) loci and three morphological markers, with an average marker interval of 17.86 cM [19]. Das et al. constructed a linkage map covering 784.3 cM using 36 polymorphic simple sequence repeats (SSRs) markers in a recombinant inbred line population from a cross between two dark jute genotypes, and 21 QTLs were identified for eight fiber yield traits and for a fiber quality trait (fiber fineness) [20].

However, genetic map and QTL information for white jute (C. *capsularis*) is much more limited. Chen et al. reported a linkage map comprising 119 markers that covered 2185.7 cM with a mean density of 18.7 cM per locus [2]. Recently, nine linkage groups were identified in a genetic map with an estimated length of 2016 cM and average marker interval of 4.2 cM; this map included a final set of 458 markers (48 SSRs and 410 SNPs) [21].

As described above, genetic maps for white and dark jute exist; however, the total number of markers on the LGs of most of these maps is limited and some of the mapped markers have no sequence information. Only two of the genetic maps relate to white jute and no QTL mapping has been reported [2, 20]. Thus, high-density genetic maps for white jute are lacking; in particular, a map that covers a large number of molecular markers with sufficient sequence information is needed to meet the demand for QTL mapping and other types of research.

In the last decade, many molecular marker technologies have been developed, including RAPD, AFLP, ISSR, SRAP, and SSR [14]. All of these DNA-based molecular markers can be used for linkage map construction. However, these markers are time-consuming and costly to prepare and some have proven to be unstable [22, 23]. Recent developments in sequencing technology have simplified and accelerated the discovery of sequence variants, enabling the development of sequence-based markers including SNPs and insertion/deletion polymorphism (InDel) markers [24]. SNPs are more useful as genetic markers because they are the most abundant and stable form of genetic variation in most genomes [25, 26]. They have been used for genetic linkage mapping in many organisms including soybean, barley, cabbage, oilseed rape, and sugar beet [9, 27–30].

Tao *et al. BMC Genomics* (2017) 18:355

Page 3 of 12

Specific locus amplified fragment sequencing (SLAF-seq) has been proven to be an efficient method of large-scale *de novo* SNP discovery and genotyping using high-throughput sequencing, and it provides a high-resolution strategy for large-scale genotyping that is applicable to a wide range of species and populations [31]. The efficiency of this approach has been demonstrated in rice and soybean, and it has also been used to create a genetic map for common carp (*Cyprinus carpio* L.) without the benefit of a reference genome sequence [32].

In this study, we used SLAF-seq to develop SNP and InDel markers and then constructed a high-density genetic map for white jute. The characteristics of this map were investigated. Eleven QTLs associated with plant height were identified using this new genetic linkage map.

Methods
Mapping population and phenotyping
To develop an RIL mapping population of *C. capsularis*, a cross was made between elite cultivar '179' (female parent) and local variety 'Aidianyesheng' (male parent, supplied by the National Medium-term Genebank of Jute Germplasm Resources). Cultivar '179' was derived from the cross 'Meifeng No. 2' and 'Minma No. 5'; 'Aidianyesheng' is a long-established local variety. The mapping population consisted of 100 individuals and was developed by single seed descent (F_8) at Fujian Agriculture and Forestry University, Fuzhou and Hainan, China. The F_8 families together with their two founders were grown at two different locations, namely, the experimental fields of Fujian Agriculture and Forestry University, Hongwei in 2011 (26.10°N, 119.01°E, 26 m above m.s.l.) and Yangzhong in 2012 (26.16°N, 118.28°E, 191 m above m.s.l.). At each location, the experiment was conducted from May to October using standard cultural practices; the plants were positioned using a randomized complete block design with three replicates. Each genotype was raised in a double row of 50 plants with a spacing of 50 cm within the rows and 100 cm spacing between the replicates. Five healthy plants were harvested from each replicate at 130 day after sowing. Plant height (m) was recorded as the length of the undivided stem from the base of the plant to the point of bifurcation at the top.
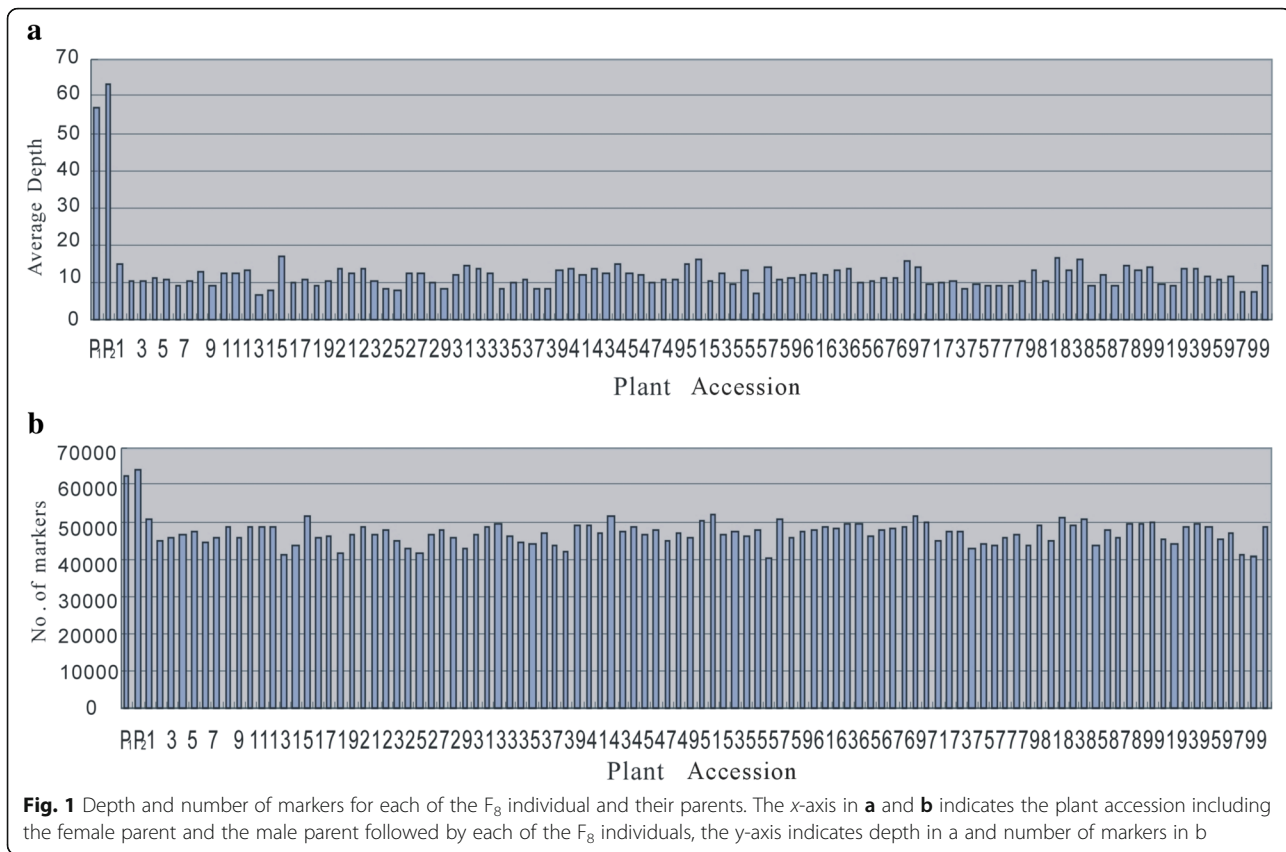
DNA isolation
Leaves from the two parents and the F_8 lines were collected from seedlings and used for DNA extraction. Total genomic DNA was prepared from each plant according to the modified cetyltrimethyl ammonium bromide (CTAB) method [33]. DNA concentration and quality were estimated with a JS-2012 spectrophotometer and by electrophoresis using 1.2% agarose gels.

SLAF library generation and sequencing
The SLAF library was generated using the protocol described by Sun et al. [31] with slight modifications. In brief, genomic DNA of each sample was treated with MseI (NEB, Ipswich, MA, USA), T_4 DNA ligase (NEB), ATP (NEB), and MseI adapter at 37 °C. Restriction/ligation reactions were heat-inactivated at 65 °C and digested with BfaI and EcoRI restriction enzymes at 37 °C. Then, polymerase chain reaction (PCR) amplification was carried out in the reaction solutions containing the diluted restriction/ligation samples, dNTP, Taq DNA polymerase (NEB), and MseI-primer containing barcode 1. The PCR products were purified by the E.Z.N.A. Cycle Pure Kit (Omega, London, UK). The purified PCR products were pooled and incubated at 37 °C with MseI, T_4 DNA ligase, ATP, and Solexa adapter. After incubation, the reaction products were purified using a Quick Spin column (Qiagen, Venlo, Netherlands), and electrophoresed on a 2% agarose gel. SLAFs of 350–380 bp and 500–550 bp (including adapter sequence indexes and adaptors) were isolated using Gel Extraction Kits (Qiagen). These SLAFs were subjected to PCR with a Phusion Master Mix (NEB) and Solexa amplification primer mix to add barcode 2. PCR products were gel purified and SLAFs of 280–310 bp and 430–480 bp were selected for paired-end sequencing on an Illumina HiSeq 2500 sequencing platform (Illumina, San Diego, CA, USA). According to the barcode sequences, raw reads were demultiplexed to individual reads. Then, low quality reads (quality score <20) were filtered out and the good quality reads were used for molecular marker discovery.

SLAF-seq data grouping and genotyping
All SLAF pair-end reads with clear index information were clustered using sequence similarity as detected by BLAT [34] (−tilesize = 10 -stepsize = 5). All SLAF markers were filtered four times and quality was assessed by the method described by Sun et al. [31]. Sequences that clustered together were defined as an SLAF locus. All polymorphic SLAF loci were genotyped for consistency in parents and progenies. Jute is a diploid species; therefore, one locus will contain a maximum of four SLAF tags. Groups containing more than four tags were filtered out as repetitive SLAFs. In this study, SLAFs with a sequence depth of less than 10 were defined as low-depth SLAFs and were filtered out. A SLAF which had less than three SNPs and average depth of each sample above three, was used as a high quality SLAF marker. Polymorphic markers were classified into eight segregation patterns (aa × bb, ab × cc, ab × cd, cc × ab, ef × eg, hk × hk, lm × ll and nn × np). An F_8 population is obtained from a cross of two diverse parents with the genotype aa or bb. Therefore, our study only used SLAF markers with the segregation pattern aa × bb for

Tao *et al. BMC Genomics* (2017) 18:355

Page 4 of 12

Fig. 1 Depth and number of markers for each of the F_8 individual and their parents. The *x*-axis in **a** and **b** indicates the plant accession including the female parent and the male parent followed by each of the F_8 individuals, the y-axis indicates depth in a and number of markers in b

genetic map construction. The cut-off value for the proportion of missing data is 30%.

Linkage map construction

In this study, we performed a 2-point linkage analysis using efficient SLAFs after genotyping the 100 RILs. A high-density genetic map including 11 LGs was constructed using the grouping function of Joinmap v4.0 software [35]. Modified logarithm of odds (MLOD) scores between markers were calculated to confirm the robustness of the markers for each LG. Markers with MLOD scores <5 were filtered prior to ordering. A LOD threshold of 3.0 was set as the default. High Map Strategy was used to order SLAF markers and correct genotyping errors within the LGs [36]. The LGs were then analyzed as described by Zhang et al. [37]: primary marker order was established by their location on chromosomes, according to the relationships between ordered

markers; genotyping errors or deletions were corrected using the SMOOTH algorithm; MST map was used to order the map; the new ordered genotypes were corrected with the SMOOTH algorithm; and the Kosambi mapping function was used to estimate map distances [38].

Plant height evaluation and QTL mapping analysis

The frequency distribution of plant heights in the two growing years was analyzed. A QTL analysis was then performed using the mean data and the Ici Mapping package [39]. Plant height QTLs were identified by internal mapping methods with the package R/QTL [40]. LOD significance thresholds for QTL peaks were determined using 1,000 permutations. Results from the interval mapping analysis were used to construct the QTLs, and their positions were used in a default model. Other parameters were set as follows: the step for scanning was 1.0 cM, since many adjacent markers in the SNP linkage map had a map distance of <1 cM; and the largest *P*-value for entering variables in stepwise regression of phenotype on marker variables (PIN) was 0.001. The percentage of variance explained and the additive effect were estimated for each QTL.

Table 1 SLAF markers mining results

Type of markers	Number of SLAF markers	Total depth	Ratio
Polymorphisms	5,074	3,124,714	7.31%
Non-polymorphisms	64,372	42,268,156	92.69%
Total	69,446	45,392,870	100%

Tao *et al. BMC Genomics* (2017) 18:355
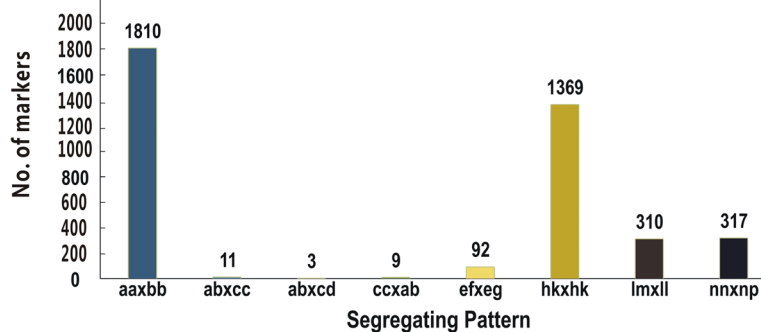
Page 5 of 12

Fig. 2 Number of markers for eight segregation patterns

Results

SLAF sequencing and genotyping

DNA sequencing generated 43.88 Gb of raw data consisting of 135,195,254 pair-end reads, each of about 90 bp after preprocessing. The high-quality base rate was 78.23%, and guanine-cytosine (GC) content was 41.11%, with quality scores of at least Q20 (Q20, indicating a 1% chance of an error, and thus 99% confidence). In total, 69,446 SLAFs were detected, with an average sequencing depth of 56.81 in '179', 62.91 in 'Aidianyesheng', and 11.49 in the progeny (Fig. 1). Of these SLAFs, 5,074 (7.31%) were polymorphic (Table 1). After filtering out SLAFs lacking parental information, 3,921 remained and these were classified into eight segregation patterns (Fig. 2). After filtering out low quality SLAF markers and heterozygous parental markers, 913 markers with high quality were used to construct a linkage map. The

average sequencing depths of these markers were 43.11-fold in '179', 43.33-fold in 'Aidianyesheng', and 9.83-fold in RIL individuals.

High-density linkage map construction

The 913 markers were assigned to 11 LGs, including 7 major LGs and 4 minor LGs. The map spanned a total of 1621.42 cM with an average distance of 1.61 cM between adjacent markers (Table 2, Fig. 3). As shown in Table 2, the largest LG was LG1 with 210 markers, a total length of 406.34 cM, and an average distance of 1.93 cM between adjacent markers. The smallest LG was LG11, with only 25 markers and a length of 29.66 cM. The degree of linkage between markers was reflected by a gap ≤5 ranging from 99.04 to 100% with an average value of 99.91%. The largest gap on the map was 5.18 cM located in LG1.

Table 2 Description on basic characteristics of the 11 LGs

Linkage group ID	Marker number	Total distance (cM)	Average distance (cM)	Largest Gap	Gaps < = 5
LG1	210	406.34	1.93	5.18	99.04%
LG2	161	300.35	1.87	4.16	100%
LG3	127	246.46	1.94	4.16	100%
LG4	108	243.60	2.26	4.42	100%
LG5	64	107.53	1.68	3.77	100%
LG6	62	66.99	1.08	2.76	100%
LG7	55	63.26	1.15	2.8	100%
LG8	39	71.41	1.83	2.82	100%
LG9	34	41.47	1.22	1.57	100%
LG10	28	44.31	1.58	2.84	100%
LG11	25	29.66	1.19	1.24	100%
Max linkage group	210	406.34	1.93	5.18	99.04%
Min linkage group	25	29.66	1.19	1.24	100%
Total	913	1621.42			
Average	83.00	147.40	1.61	3.25	99.91%

'Gap < =5' indicates the percentages of gaps in which the distance between adjacent markers was smaller than 5 cM

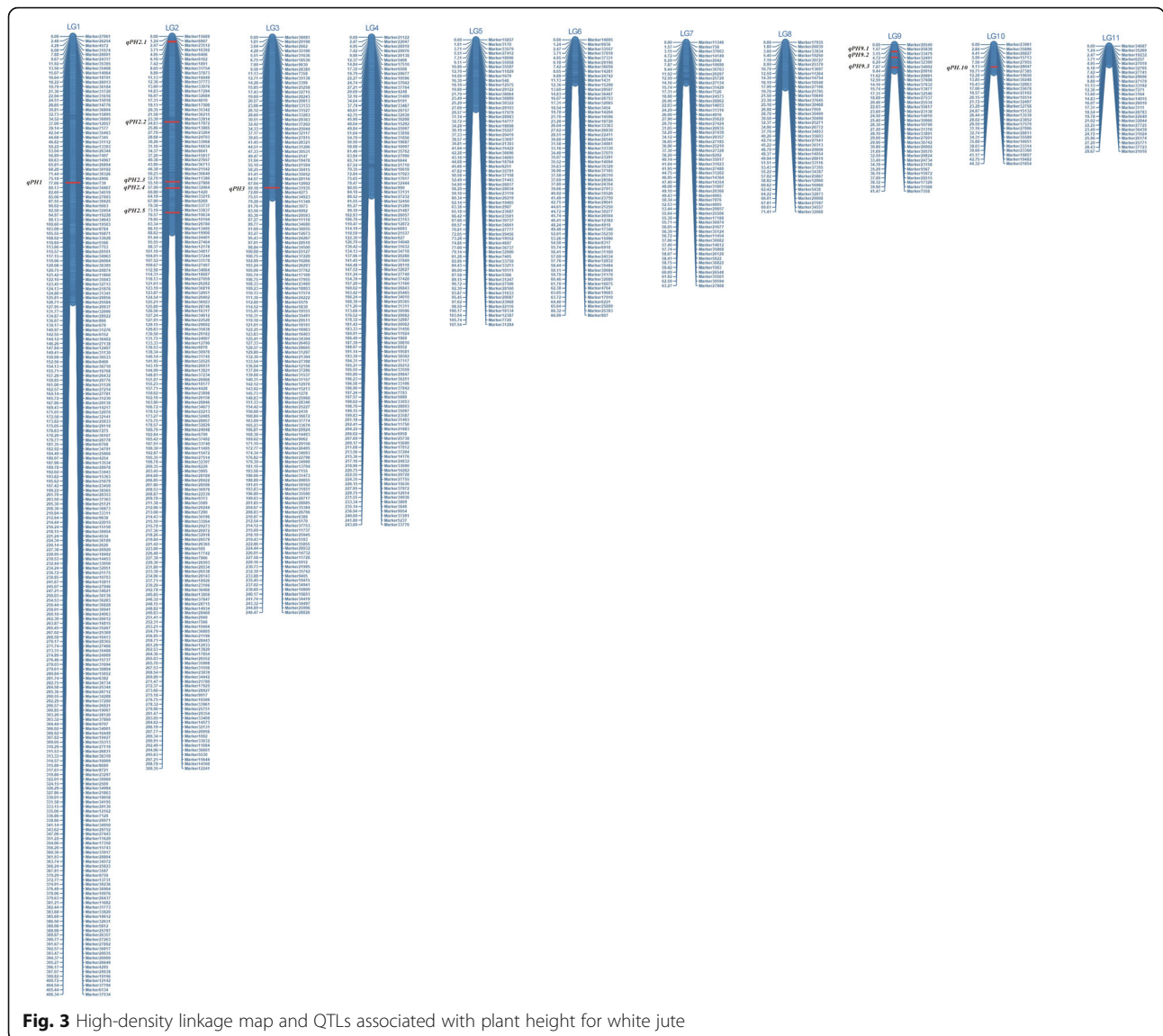Tao *et al. BMC Genomics* (2017) 18:355

Page 6 of 12

Fig. 3 High-density linkage map and QTLs associated with plant height for white jute

Distribution of SLAF markers on the genetic map

The 913 markers included 781 'SNP_only,' 109 'InDel_only,' and 23 'SNP&InDel' markers. The distribution of each of these markers on the genetic map and in each LG was investigated (Table 3, Fig. 4). 'SNP_only' accounted for 85.54% of the markers and was the predominant type. The proportions of the three types of markers on LG1 were 87.14%, 10.95, and 1.90%, respectively, which was similar to the average proportions of marker types for all 11 LGs. LG7 had the lowest proportion of 'SNP_only' markers but had the highest rate of 'InDel_only' markers at 74.55 and 20% respectively. LG4 had the highest proportion of 'SNP_only' markers but had the lowest rate of 'InDel_only' at 88.89 and 9.25%.

The nature of the 781 SNP loci was investigated (Table 4). Most were transition type SNPs of R (G/A) and Y (T/C) types with rates of 34.11% and 34.96%, respectively. Four transversion type SNPs were identified including K (G/T), M (A/C), S (G/C) and W (A/T) with frequencies ranging from 5.95 to 9.25%; these types of SNP accounted for 30.93% of all SNPs. These results are very similar to those reported in sesame [31].

Segregation distortion markers

Analysis of the segregation of the 913 mapped loci showed that 862 (94.4%) deviated significantly ($P \le 0.05$) from the expected 1:1 Mendelian segregation ratio. The result showed that segregation distortion markers were present on every LG (Table 5). LG2 had the highest

Tao *et al. BMC Genomics* (2017) 18:355

Page 7 of 12

Table 3 Description on type of the markers

Linkage group	Number of total markers	SNP_only	InDel_ only	SNP&InDel
1	210	183	23	4
2	161	141	17	3
3	127	106	19	2
4	108	96	10	2
5	64	54	7	3
6	62	53	8	1
7	55	41	11	3
8	39	34	5	0
9	34	29	2	3
10	28	23	4	1
11	25	21	3	1
Total	913	781	109	23
Average	83	71	10	2

proportion of segregation distortion markers at 95.03%, and LG7 had the lowest rate at 83.64%.

QTL mapping of plant height
Frequency distribution for plant height
A comparison of the parental cultivars showed significant differences in plant height (Fig. 5). Average plant height in ' 179' was 3.93 m, which was significantly greater than in 'Aidianyesheng' (2.64 m) in 2011, and a similar difference was also found in 2012. Plant heights in the RIL population showed continuous variation (Fig. 5). A comparison of the results from the two locations showed that plant heights were significantly influenced by the environment.

QTL mapping in white jute
Eleven stable QTLs for plant height were identified using the SLAF linkage map across the two locations and the pooled data (Table 6); these QTLs explained 4.14 –15.63% of phenotypic variance. The QTLs mapped to 5 LGs named LG1, LG2, LG3, LG9, and LG10 (Table 6 and Fig. 3). Five plant height QTLs were detected on LG2, three were mapped on LG9, and one each on LG1, LG3, and LG10. The plant height QTL located at
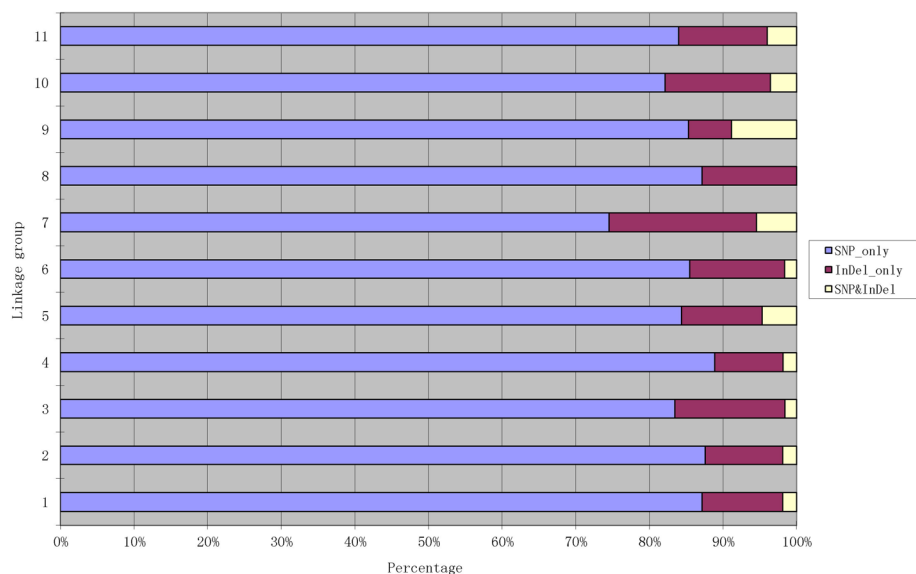

Fig. 4 Percentages of diverse types of markers on each linkage group

Tao *et al. BMC Genomics* (2017) 18:355

Page 8 of 12

Table 4 Statistic of mapped SNP marker types

Type	Number	Ratio
M(A/C)	66	8.40%
R(A/G)	266	34.11%
Y(C/T)	273	34.96%
S(C/G)	46	5.95%
W(A/T)	72	9.25%
K(G/T)	57	7.33%
Total	781	100%

223.0 cM on LG2 (between markers 595 and 17742) was found to have a major effect and accounted 15.63% of the phenotypic variance with a LOD score of 7.20. Most of the additive effects were negative for QTLs of plant height, indicating that increased trait values were conferred by the female ('179') alleles.

Discussion

The need for development of genetic markers for white jute

Previous studies used traditional markers and newly developed markers to construct genetic linkage maps in jute [17, 18, 20]; however, only a few studies were conducted in white jute. Although genetic maps for jute exist, the total number of markers on the LGs is limited [16–19], and do not provide comprehensive coverage of the jute genome. In general, a high number of polymorphic markers is necessary to guarantee the accuracy of a linkage map [2]. The limited number of available markers and their low polymorphism rate made construction of a genetic linkage map with high-density markers for white jute almost impossible. Other markers, such as SSRs and SNPs, clearly need to be

developed to overcome the scarcity of DNA markers for construction of saturated genetic maps in jute [41, 42]. Biswas et al. constructed a linkage map in white jute using SNP markers that were based on expressed sequence tags and not sequencing [21]. Genotyping by sequencing is a high-throughput technique for the efficient development of large numbers of markers in a short time to generate polymorphic markers for high-density genetic map construction. In our study, a high density genetic map of white jute was constructed using SLAF-seq genotyping data. The constructed genetic map had a higher density of the maps made of organism lacking reference genome sequences.

SLAF-seq is an ideal approach for developing markers

The SLAF-seq strategy is better than other sequencing approaches because it combines locus-specific amplification and high-throughput sequencing technology, although reference genome sequences and polymorphism information are not necessary to use this strategy [31]. In contrast to conventional methods, which are inefficient, expensive, and time-consuming, [43, 44], SLAF sequencing can generate large amounts of sequence information and handle whole genome density distributions, which ensures density, uniformity, and efficiency of marker development. Since SLAF-seq methods were first developed, they have been used in map construction for sesame, soybean, oil rape, and other vegetable crops [7, 11]. Our study provides the first development of markers on a large scale for white jute. In total, 69,446 SLAF markers were developed using high throughput sequencing, and 5,074 of these were polymorphic. We obtained 913 markers that were suitable for constructing a linkage map. Marker integrity and accuracy were high and marker quality and quantity met the requirements

Table 5 Distribution of segregation distortion markers

Linkage group	Number of total markers	Percentage	Number of segregation distortion markers	Percentage
1	210	23.00%	184	87.62%
2	161	17.63%	153	95.03%
3	127	13.91%	115	90.55%
4	108	11.83%	102	94.44%
5	64	7.01%	59	92.19%
6	62	6.79%	56	90.32%
7	55	6.02%	46	83.64%
8	39	4.27%	36	92.31%
9	34	3.72%	30	88.24%
10	28	3.07%	26	92.86%
11	25	2.74%	23	92.00%
Total	913			

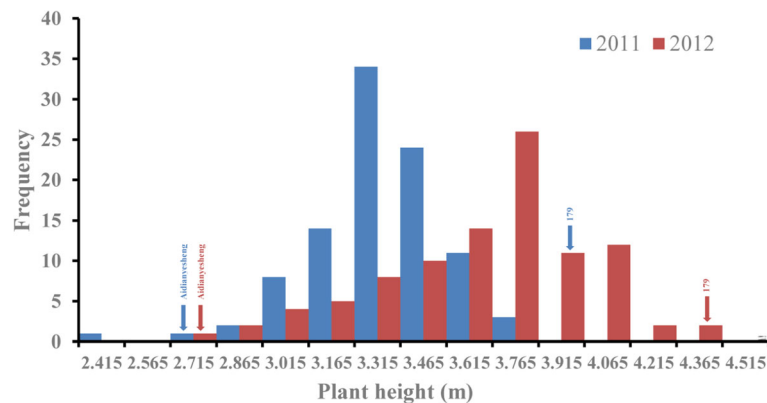Tao *et al. BMC Genomics* (2017) 18:355

Page 9 of 12

Fig. 5 Frequency distribution of plant height in the RIL population of white jute in 2011 and 2012. The x-axis indicates the plant height of the RIL population and the parents, and the y-axis indicates the distributing frequency of plant height. Arrows indicate the distributing frequency of plant height for each parent

for construction of a genetic map. Therefore, the SLAF-seq technology is ideal for developing plant chromosome-specific molecular markers with high success rates, specificity, and stability.

Inconsistency between LGs and chromosomes
To our knowledge, the genetic map presented in this paper is the densest map to date for white jute. However, our map is still not saturated, because it has 11 LGs and not 7 LGs as was expected. Several factors may be responsible, such as the genetic constitution of different mapping populations, mapping strategies, number and type of mapped loci, the choice of mapping software, and ratio between number of markers and population size [20]. In the present study, an unexpectedly high level of heterozygosity in the two parents resulted in some markers that could not be used for linkage analysis, and this may have resulted in larger gaps between adjacent markers. Additionally, if the markers were

unevenly distributed on the map, an increased number of LGs might have been produced. When compared with previous studies in jute [15, 20, 21], the size of the populations used in this study was relatively low, which may have contributed to the inconsistency in numbers of LGs and chromosomes. In addition, the marked markers are clearly not sufficiently abundant. The fairly low genetic diversity at the species level in jute [15] may be a major reason for the low proportion of polymorphic SLAF markers in the present study. We assumed some fragments of chromosomes in two parents have same genetic background, so no polymorphic markers were detected in these sections. The fragments which should be jointed in the same LG may have been separated into different LGs because of this reason. Although we attempted to develop and add more polymorphic markers, the unexpectedly low polymorphism information content in combination with an inherently narrow genetic base makes the reduction in the jute LG number impossible at present.

Table 6 QTLs for plant height in the RIL population in both 2011 and 2012

QTL	Position(cM)	Left Marker	Right Marker	LOD	PVE	ADD
qPH1	222.5–225.5	Marker4534	Marker36189	2.65	4.14	−0.05
qPH2.1	8.5–10.5	Marker31754	Marker37873	4.93	9.37	0.07
qPH2.2	129.5–131.5	Marker29182	Marker24007	4.81	9.51	0.07
qPH2.3	222.5–225.5	Marker595	Marker17742	7.20	15.63	−0.08
qPH2.4	230.5–233.5	Marker28534	Marker26538	3.16	5.41	−0.04
qPH2.5	269.5–270.5	Marker34442	Marker21780	5.27	9.91	−0.08
qPH3	229.5–231.5	Marker21995	Marker35742	3.33	6.33	−0.06
qPH9.1	9.5–12.5	Marker29916	Marker20881	2.84	4.73	−0.06
qPH9.2	17.5–21.5	Marker27537	Marker25534	3.47	5.55	−0.06
qPH9.3	30.5–36.5	Marker5567	Marker15972	3.22	5.35	−0.06
qPH10	33.5–37.5	Marker16051	Marker33314	2.92	4.84	−0.06

Marker segregation distortion
In our study, an unusually high degree of marker segregation distortion was observed. The rate was higher than those reported by Das et al. [20] and Topdar et al. [16]. Segregation distortion is widespread in inter-specific and intra-specific crosses used to construct mapping populations [45]. Distorting factors can be self-incompatibility alleles, deleterious recessive alleles, structural rearrangements, or differences in DNA content [46]. High levels of segregation distortion might also be the result of use of genetically distant parents [47]. Of the two parental lines used in this study, 'Aidianyesheng' is a long-established local line while '179' is a cultivated accession; the differences in their traits indicate they are genotypically divergent and that there is a high degree of genetic variation between them. This may account for the high level of segregation distortion in this

Tao et al. BMC Genomics (2017) 18:355

Page 10 of 12

study. Semagn et al. also reported high segregation distortion in doubled haploid (DH) and RIL populations [47]. In Arabidopsis, a high level of segregation distortion was attributed to a higher frequency of recombination [48]. Nevertheless, segregation distortion has very little effect on marker order and map length [49]. Zhang et al. [50] showed that segregation distortion could result in higher genetic variance than non-distortion and help to improve the detection of linked QTLs, because distortion markers do not have a large effect on the position or effect estimations of QTL analysis.

High-density genetic map and QTL mapping

Our map covers 1,621.42 cM with an average of 83 markers per LG and an average distance of 1.61 cM between adjacent markers. This is the smallest average inter-marker distance reported for white jute. Furthermore, the map constructed in this study has the largest number of markers for jute. More importantly, 85.54% of the markers on this genetic map were SNPs, i.e., sequence tagged markers with co-dominant inheritance, that are suitable for comparative genomic studies [51] and association mapping [52, 53]. The high resolution genetic map developed here will be a useful platform for the assembly of the white jute genome, for the development of sequence-based markers used in breeding programs, and for the identification of genes associated with important agricultural traits.

During the present study, 11 QTLs associated with plant height, including one major effect QTL, were identified on LG1, LG2, LG3, LG9, and LG10. The QTLs detected in this study were stable in the two cultivation locations. The selection of parents can affect the accuracy and utility of QTL mapping. In the present study, plant height in '179' (female parent) was significantly greater than in 'Aidianyesheng' (male parent). We found the female parent contributed the main QTL increasing plant height. Topdar et al. found increased plant height was conferred by the female (taller) allele at three loci and the male (shorter) allele at only one locus [16]. Kundu et al. showed that enhancing alleles of two plant height QTLs were from male and female parent respectively, and they also found plant height was strongly correlated with fiber yield [15].

As we used a different species of jute here compared to previous reports, we could not combine and compare the identified QTLs for plant height among studies [15, 16]. It is unclear whether these QTLs were in the same location or on the same LG as found in previous studies. We expect that further research on QTLs associated with yield traits will improve the breeding efficiency of jute. The molecular markers linked with fiber yield such as plant height may be used in marker assisted selection to accelerate development of cultivars with high fiber yield and quality.

Conclusions

In this study, a high density genetic map of white jute was constructed using SLAF-seq technique. To our knowledge, this map is the densest map to date for white jute. Furthermore, 11 QTLs associated with plant height, including one major effect QTL, were identified. The results will provide an important platform for further development in both molecular biology and breeding of white jute.

Authors' contributions
AT, JQ and PF designed and organized the entire research. GW, LL, PL and JX performed the experiments and acquired the data. LH, RA, GW and LZ analyzed the data. AT, LH, LZ and JX drafted the manuscript. JQ, PF, RA, LL and PL revised the manuscript critically. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests (financial or non-financial).

Consent to publish
Not applicable

Ethics approval and consent to participate
Not applicable

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops; Key Laboratory of Crops by Design, Fujian Agriculture and Forestry University, Fuzhou 350028, People's Republic of China. [2]Biomarker Technologies Corporation, 101300 Beijing, China. [3]Guangxi University, 530000 Nanning, China. [4]Eastern Agricultural Research Center, Montana State University 59270, Sidney, Montana, USA.

References
1. Islam AS TM, Lee CT, Ingram C, Montalvo RJ, Ende G, Alam S, Siddiqui J, Sathasivan K: Preliminary progress in jute (Corchorus species) genome analysis. Plant Tissue Cult Biotechnol 2005, 15:145–156

Tao *et al. BMC Genomics* (2017) 18:355

Page 11 of 12

2. Chen Y, Zhang L, Qi J, Chen H, Tao A, Xu J, Lin L, Fang P. Genetic linkage map construction for white jute (Corchorus capsularis L.) using SRAP, ISSR and RAPD markers. Plant Breed. 2014;133:777–81.
3. Wazni MW, Islam AS, Taliaferro JM, Anwar N, Sathasivan K. Novel ESTs from a Jute (Corchorus olitorius L.) cDNA Library. Plant Tissue Cult Biotech. 2007;17(2):173–82.
4. Zhang G, Qi J, Xu J, Niu X, Zhang Y, Tao A, Zhang L, Fang P, Lin L. Overexpression of UDP-glucose pyrophosphorylase gene could increase cellulose content in Jute (Corchorus capsularis L.). Biochem Biophys Res Commun. 2013;442(3–4):153–8.
5. HP X. Breeding sciences of bast and leaf fiber crops, 1st edn. Beijing: Agricultural Science and Technology Press of China; 2008.
6. A K: Recent Agricultural Developments in Jute, Kenaf and Mesta Through Traditional and Biotechnological Approaches. a seminar talk held in Myanmar in February 2007 under the auspices of International Jute Study Group 2007, 1(1):1–5
7. Li B, Tian L, Zhang J, Huang L, Han F, Yan S, Wang L, Zheng H, Sun J. Construction of a high-density genetic map based on large-scale markers developed by specific length amplified fragment sequencing (SLAF-seq) and its application for QTL analysis for isoflavone content in Glycine max. BMC Genomics. 2014;15(1):1086.
8. Petroli CD, Sansaloni CP, Carling J, Steane DA, Vaillancourt RE, Myburg AA, da Silva Jr OB, Pappas Jr GJ, Kilian A, Grattapaglia D. Genomic characterization of DArT markers based on high-density linkage analysis and physical mapping to the Eucalyptus genome. Plos One. 2012;7(9):e44684.
9. Esteras C, Gomes CP, Monforte AJ, Blanca J, Vicente-Dolera N, Roig C, Nuez F, Pico B. High-throughput SNP genotyping in Cucurbita pepo for map construction and quantitative trait loci mapping. BMC Genomics. 2012;13:80.
10. Song W, Pang R, Niu Y, Gao F, Zhao Y, Zhang J, Sun J, Shao C, Liao X, Wang L, Tian Y, Chen S. Construction of high-density genetic linkage maps and mapping of growth-related quantitative trail loci in the Japanese flounder (Paralichthys olivaceus). Plos One. 2012;7(11):e50404.
11. Wei Q, Wang Y, Qin X, Zhang Z, Wang J, Li J, Lou Q, Chen J. An SNP-based saturated genetic map and QTL analysis of fruit-related traits in cucumber using specific-length amplified fragment (SLAF) sequencing. BMC Genomics. 2014;15(1):1158.
12. Wang Y, Sun S, Liu B, Wang H, Deng J, Liao Y, Wang Q, Cheng F, Wang X, Wu J. A sequence-based genetic linkage map as a reference for Brassica rapa pseudochromosome assembly. BMC Genomics. 2011;12:239.
13. Ren Y, Zhao H, Kou QH, Jiang J, Guo SG, Zhang HY, Hou WJ, Zou XH, Sun HH, Gong GY, Levi A, Xu Y. A high resolution genetic Map anchoring scaffolds of the sequenced watermelon genome. Plos One. 2012;7(1):e29453.
14. Wang N, Fang L, Xin H, Wang L, Li S. Construction of a high-density genetic map for grape using next generation restriction-site associated DNA sequencing. BMC Plant Biol. 2012;12:148.
15. Kundu A, Chakraborty A, Mandal NA, Das D, Karmakar PG, Singh NK, Sarkar D. A restriction-site-associated DNA (RAD) linkage map, comparative genomics and identification of QTL for histological fibre content coincident with those for retted bast fibre yield and its major components in jute (Corchorus olitorius L., Malvaceae s. l). Molecular Breeding. 2015;35(1):19.
16. Topdar N, Kundu A, Sinha MK, Sarkar D, Das M, Banerjee S, Kar CS, Satya P, Balyan HS, Mahapatra BS, et al. A complete genetic linkage map and QTL analyses for bast fibre quality traits, yield and yield components in jute (Corchorus olitorius L.). Cytol Genet. 2013;47(3):129–37.
17. Sultana N, Khan H, Ashraf N, Sharkar MTK. Construction of an intraspecific linkage map of jute. Asian J Plant Sci. 2006;5(5):758–62.
18. Haque S, Ashraf N, Begum S, Sarkar RH, Khan H. Construction of genetic map of Jute (Corchorus olitorius L.) based on RAPD markers. Plant Tissue Cult Biotech. 2008;18:165–72.
19. Chen H, Chen M, Tao A, Zhang G, Xu J, Qi J, Fang P. Construction of a molecular linkage map using SRAP and mapping of three qualitative traits in Corchorus olitorius L. Scien Agric Sinica. 2011;44(12):2422–30.
20. Das M, Banerjee S, Dhariwal R, Vyas S, Mir RR, Topdar N, Kundu A, Khurana JP, Tyagi AK, Sarkar D, Sinha MK, Balyan HS, Gupta PK. Development of SSR markers and construction of a linkage map in jute. J Genet. 2012;91(1):21–31.
21. Biswas C, Dey P, Karmakar PG, Satpathy S. Discovery of large-scale SNP markers and construction of linkage map in a RIL population of jute (Corchorus capsularis). Mol Breeding. 2015;35(5):119.
22. Gerber S, Mariette S, Streiff R, Bodenes C, Kremer A. Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. Mol Ecol. 2000;9(8):1037–48.

23. Woodhead M, Russell J, Squirrell J, Hollingsworth PM, Mackenzie K, Gibby M, Powell W. Comparative analysis of population genetic structure in Athyrium distentifolium (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. Mol Ecol. 2005;14(6):1681–95.
24. Hyten DL, Cannon SB, Song Q, Weeks N, Fickus EW, Shoemaker RC, Specht JE, Farmer AD, May GD, Cregan PB. High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. BMC Genomics. 2010;11(1):38.
25. Troggio M, Malacarne G, Coppola G, Segala C, Cartwright DA, Pindo M, Stefanini M, Mank R, Moroldo M, Morgante M, Grando MS, Velasco R. A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (Vitis vinifera L.) anchoring Pinot Noir bacterial artificial chromosome contigs. Genetics. 2007;176(4):2637–50.
26. Liu J, Huang S, Sun M, Liu S, Liu Y, Wang W, Zhang X, Wang H, Hua W. An improved allele-specific PCR primer design method for SNP marker analysis and its application. Plant Methods. 2012;8(1):34.
27. Qi Z, Huang L, Zhu R, Xin D, Liu C, Han X, Jiang H, Hong W, Hu G, Zheng H, Chen Q. A high-density genetic map for soybean based on specific length amplified fragment sequencing. Plos One. 2014;9(8):e104871.
28. Wang W, Huang S, Liu Y, Fang Z, Yang L, Hua W, Yuan S, Liu S, Sun J, Zhuang M. Construction and analysis of a high-density genetic linkage map in cabbage (Brassica oleracea L. var. capitata). BMC Genomics. 2012;13:523.
29. Chutimanitsakun Y, Nipper RW, Cuesta-Marcos A, Cistue L, Corey A, Filichkina T, Johnson EA, Hayes PM. Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. BMC Genomics. 2011;12:4.
30. Holtgrawe D, Sörensen TR, Viehover P, Schneider J, Schulz B, Borchardt D, Kraft T, Himmelbauer H, Weisshaar B. Reliable in silico identification of sequence polymorphisms and their application for extending the genetic map of sugar beet (Beta vulgaris). Plos One. 2014;9(10):e110113.
31. Sun X, Liu D, Zhang X, Li W, Liu H, et al. SLAF-seq: An efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. PLoS ONE. 2013;8(3):e58700.
32. Zhang Y, Wang L, Xin H, Li D, Ma C, Ding X, Hong W, Zhang X. Construction of a high-density genetic map for sesame based on large scale marker development by specific length amplified fragment (SLAF) sequencing. BMC Plant Biol. 2013;13:141.
33. Zhou D, Ji Q, Wu W, Li W. Studies on DNA extraction and establishment of RAPD reaction system in jute. J Fujian Agric Univ. 2001;30(3):334–9.
34. Ken WJ. BLAT–the BLAST-like aligment tool. Genome Res. 2002;12(4):656–64.
35. Stam P. Construction of integrated genetic linkage maps by means of a new computer package: Join Map. Plant J. 1993;3(5):6.
36. Liu D, Ma C, Hong W, Huang L, Liu M, et al. Construction and analysis of high-density linkage map using high-throughput sequencing data. Plos One. 2014;9(6):e98855.
37. Zhang J, Zhang Q, Cheng T, Yang W, Pan H, Zhong J, Huang L, Liu E. High-density genetic map construction and identification of a locus controlling weeping trait in an ornamental woody plant (Prunus mume Sieb. et Zucc). DNA Res. 2015;22(3):183–91.
38. Ooijen JW. JoinMap 4, Software for the calculation of genetic linkage maps in experimental populations. Wageningen: Kyazma B.V.; 2006
39. Wang J, Li H, Zhang L, Meng L: Users' Manual of QTL IciMapping Version 3. 2(2012). The Quantitative Genetics Group, Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100081, China, and Genetic Resources Program, International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6–641, 06600 Mexico, D.F., Mexico.
40. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. Bioinformatics. 2003;19(7):889–90.
41. Trick M, Adamski NM, Mugford SG, Jiang CC, Febrer M, Uauy C. Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. BMC Plant Biol. 2012;12:14–23.
42. Raman H, Raman R, Nelson MN, Aslam M, Rajasekaran R, Wratten N, Cowling WA, Kilian A, Sharpe AG, Schondelmaier J. Diversity array technology markers: genetic diversity analyses and linkage map construction in rapeseed Brassica napus L. DNA Res. 2012;19:51–65.
43. Kennedy GC, Matsuzaki H, Dong S, Liu WM, Huang J, et al. Large scale genotyping of complex DNA. Nat Biotechnol. 2003;21:1233–7.
44. Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D, Weng Q, Huang T, Dong G, Sang T, Han B. High-throughput genotyping by whole-genome resequencing. Genome Res. 2009;19(6):1068–76.

Tao *et al. BMC Genomics* (2017) 18:355

Page 12 of 12

45. Hall MC, Willis JH. Transmission ratio distortion in intraspecific hybrids of Mimulus guttatus: implications for genomic divergence. Genetics. 2005;170:12.

46. Jenczewski E, Gherardi M, Bonnin I, Prosperi JM, Olivieri I, Huguet T. Insight on segregation distortions in two intraspecific crosses between annual species of *Medicago* (Leguminosae). Theor Appl Genet. 1997;94:10.

47. Semagn K, Bjørnstad Å, Ndjiondjop MN. Principles, requirements and prospects of genetic mapping in plants. Afr J Biotechnol. 2006;5(25):19.

48. Liu SC, Kowalski SP, Lan TH, Feldmann KA, Paterson AH. Genome-wide high-resolution mapping by recurrent intermating using Arabidopsis thaliana as a model. Genetics. 1996;142:12.

49. Hackett CA, Broadfoot LB. Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity. 2003;90:6.

50. Zhang L, Deng Q, Li P, Wang S, Zheng A, Li Z, Li H, Li S, Wang J. Effects of missing marker and segregation distortion on QTL mapping in F_2 populations. Theor Appl Genet. 2010;121:12.

51. Luo MC, Deal KR, Akhunov ED, Akhunova AR, Anderson OD, Anderson JA, Blake N, Clegg MT, Coleman-Derr D, Conley EJ, et al. Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in Triticeae. Proc Natl Acad Sci. 2009;106(37):15780–5.

52. Chan EK, Rowe HC, Kliebenstein DJ. Understanding the evolution of defense metabolites in Arabidopsis thaliana using genome-wide association mapping. Genet Mol Res. 2010;185(3):991–1007.

53. Cogan NO, Ponting RC, Vecchies AC, Drayton MC, George J, Dracatos PM, Dobrowolski MP, Sawbridge TI, Smith KF, Spangenberg GC, et al. Gene-associated single nucleotide polymorphism discovery in perennial ryegrass (Lolium perenne L.). Mol Genet Genomics. 2006;276(2):101–12.