Construction of A High-Density SNP Genetic Map by Genotyping-By-Sequencing and Location of Seed Quality traits in Flax (*Linum usitatisimum* L.)

Li Zhao (zhyl20020126@163.com)
Crops Research Institute, Gansu Academy of Agricultural Sciences https://orcid.org/0000-0002-5537-2165
Wei Zhao
Crops Research Institute, Gansu Academy of Agricultural Sciences
Limin Wang
Crop Research Institute, Gansu Academy of Agricultural Sciences
Bin Wang
Crops Research Institute, Gansu Academy of Agricultural Sciences
Zhao Dang
Crops Research Institute, Gansu Academy of Agricultural Sciences
Wenjuan Li
Crops Research Institute, Gansu Academy of Agricultural Sciences

Research article

Keywords: Flax (Linum usitatisimum L.), Genotyping-by-sequencing, QTL identification, Seed quality trait, Single nucleotide polymorphism marker

Posted Date: September 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-70683/v1

Abstract

Background

Oil and five main fatty acid content(especially linolenic acid content (LNA))are six main quality traits in flax, and also the important target traits in flax breeding. However, the reports on genetic map construction and QTL mapping of linolenic acid content in flax are limited, which results in low accuracy and efficiency of quality breeding in oil flax. Construction the genetic linkage map and location the QTLs to discover the stable genetic QTLs sites related to oil content and fatty acid content, so as to provide targeted targets for breeders and improve breeding efficiency.

Results

In this study, a final integrated map consisting of 2,239 single nucleotide polymorphism (SNP) markers on 15 linkage groups (LGs) with an average distance of 0.46 cM between adjacent markers was generated using genotyping-by-sequencing (GBS) technique. A total of 21 quantitative trait loci (QTL)s for the six traits, i.e., palmitic acid content (PAL), stearic acid content (STE), oleic acid content (OLE), linoleic acid content (LIN), linolenic acid content (LNA), and oil content (OIL) in the RIL population under three environments. One QTL cluster harbored two QTLs for LIN and LNA trait, respectively was also identified. Especially, QTL qLIN-Group12-2 for LIN and QTL qLNA-Group12-2 for LNA on chromosome 12 were detected in multi-environments.

Conclusions

A high-density SNP genetic map with total 2239 markers was constructed with GBS technique, The total genetic distance of the SNP map was 1032.90 cM, with the average genetic distance of 0.46 cM per marker. A total of 21 quantitative trait loci (QTL)s for the six traits under three environments were determined. One QTL cluster harbored two QTLs for LIN and LNA trait, respectively was also identified. QTL qLIN-Group12-2 for LIN and QTL qLNA-Group12-2 for LNA on chromosome 12 were detected in multi-environments. These results provide more information for determining the seed quality related candidate genes and contribute to the marker-assisted selection in flax breeding.

Background

Flax (*Linum usitatisimum* L. 2n = 30) is one of the earliest used crops for its seed oil and stem fiber. Thus flax varieties are classified into two groups by human beings for their usages of 'fiber type' and 'oil type'. The application history of flax is as early as 30,000 year ago (Zohary et al, 1999; Fu et al, 2011 and 2012; Zhang et al, 2016). For oil-type flax varieties (linseeds), seeds contain highoil content and alpha-linoleic acid (ALA) content, which has been used in food, medicine and feedstuff. ALA is beneficial to body health via reducing cholesterol level, preventing blood clotting, favoring for diabetic, and inhibiting allergic reaction (Gill, 1987; Carter et al, 1993). Breeding new varieties with high oil content and ALA content is one of the key objectives of flax breeding in the world in recent years.

Previous studies indicate that the seed quality traits, such as oil content and ALA content are quantitative traits for flax and could be affected by genotype and environments. To reveal the genetic mechanism of the fatty acid biosynthesis and regulation, Cloutier et al constructed a linkage map that included 24 linkage groups with 113 EST-SSR markers for a DH population, and identified QTLs for seed color, LIN and LAN with SSR markers for the first time (Cloutier et al. 2011). For flax, traditional molecular markers such as random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), restriction fragment length polymorphism (RFLP) have been used to locate the QTLs linked to key agronomic traits since the end of the last century (Spielmeyer et al, 1998; Oh et al, 2000; Gehringer et al, 2006; Cloutier et al, 2011 and 2012; Cloutier et al, 2012; Braulio et al; 2014; Chandrawati and Yadav HK, 2017). In recent a few years, the rapid development of the next-generation sequencing (NGS) and publication of the flax genome sequences impel the molecular genetics in flax (Guo et al, 2015; Gehringer et al, 2006; Zhang et al. 2020). Kumar et al. (2015) constructed the first SNP genetic map for flax based on the selected 329 SNP markers and 362 SSR markers using a RIL population containing 243 individuals. The map is comprised of 691 markers in 15 LGs. A total of 20 quantitative trait loci (QTL) were identified corresponding to 14 traits. Of which 9 QTLs related to seed component content traits were determined (Kumar et al. 2015). Xie et al. (2019) use the genome-wide association analysis (GWAS) method detected 16 SNP loci for seed fatty acid content with a natural population of flax containing 224 samples planted in 3 different environments (Xie et al. 2019). On the whole, the research about identified QTLs for seed quality traits in *Flax* with SNP markers is not many.

There are several methods of NGS such as including restriction site-associated DNA sequencing (RADSeq) (Baird et al, 2008), and specific locus amplified fragment sequencing (SLAF-seq) (Sun et al, 2013) and GBS-Seq (Elshire et al, 2011). Recently, GBS strategy were introduced to construct high diversity SNP-based linkage map especially in large genome species (Elshire et al, 2011; Carrasco et al, 2018; Pereira et al, 2018; Yang et al, 2018).

Flax is a self-pollinated annual species with the small genome size of ~ 370 Mb (Soto-Cerdalván et al, 2012). GBS-Seq (Elshire et al, 2011) approach is efficient, low-cost and able to identify thousands SNP markers across the genome (Elshire et al, 2011). Therefore, in this study, we aim to construct a high-density genetic SNP map with a RIL population contained 232 individuals byGBS strategy. Moreover, based on the phenotypes of the RIL population under three environments, we plan to locate the QTLs significantly linked to the 6 seed quality traits including PAL, STE, OLE, LIN, LNA, and OIL and to reveal the genetic characters of the above key traits in flax. We expect the findings could improve the molecular marker-assisted selection breeding in flax.

Methods

Plant Materials

A RIL segregation population containing 232 $F_{6:7}$ individuals was constructed and applied in this study. The two parents of the poulation are STS and DYM. The phenotype data of the six traits in the two parents were shown in Table S1. The cross between DYM and STS was performed in 2011 at Lanzhou experimental field, while the parents and the 232 F_6 individuals of the RIL population were cultured at Dingxi (104°12'E，35°17'N), Zhangjiakou(114°10'E，40°57'N), and Guyuan experimental field (105°48'E,35°21'N) for seed collection and quality analysis. All the lines were cultured with two replications under the cultivation standard: single-row, 0.2m row distance, and 2m row length. Young leaves of each individuals were collected and immersed into liquid nitrogen for DNA extraction.

All the germplasm accessions are available from the flax germplasm reservoir of Gansu Academy of Agricultural Sciences.

Evaluation of the crude fat and fatty acid content

The crude fat content of seed samples was tested using the standard method of GB/T 14772-1993. The fatty acid content was measured by Agilent Gas Chromatograph (7820A). Flame ionization detector (FID) and ATFFAP chromatographic column (30m×320mm×0.33mm) were applied with the following standard: Injector temperature was set at 25°C with the shunting ratio of 60:1 and the current speed set at 2mL·min^{-1}; FID temperature was set at 300°C with the current speed set at H2 40mL·min^{-1}, current air speed at 400mL·min-1, current N_2 speed at 30mL·min^{-1}; Column temperature was set at 210°C for 8 min. Each sample was measured two replications.

Statistical Analysis of Phenotypic Data

Six seed component related traits, i.e., PAL, STE, OLE, LIN, ILNA, and OIL of the RIL population were calculated and analyzed with SPSS 20.0 (https://www.ibm.com/analytics/spss-statistics-software). The mean value, standard deviation, skewness, and kurtosis of the six traits across the population and analyzed. Correlation relationship between the six traits was analyzed.

DNA extraction and GBS Sequencing

Genomic DNA of the two parents and the 232 F_{6-7} individuals was extracted with CTAB method and tested on NanoDrop 2000 (Thermo Fisher Scientific, the USA). For GBS sequencing, two enzymes, *MseI* and *EcoRI* were used to digest the DNA. After each sample was amplified in multiplex, the desired fragments in specific length were collected for library construction. DNA sequencing was performed on the Illumina HiSeqTM sequencing platform (Illumina, the USA). All the reads with paired-end 150bp were quality filtered. the reads containing the retraction site of *MseI* and *EcoRI* enzymes were used for SNP calling (Elshire et al, 2011; Carrasco et al, 2018; Pereira et al, 2018; Yang et al, 2018).

SNP detection and genetic map construction

After cleaned, all the reads were mapped to the flax reference genome (Wang et al. 2012) (http://gigadb.org/dataset/100081) to conduct GBS using BWA (Burrows Wheeler Aligner) (Li and Durbin 2009). SAMtools (Li et al. 2009) was used to transform the file with SAM format to BAM format for the next analysis. GATK (McKenna et al. 2010) was used to filter the BAM files and to detect SNPs. In order to reduce false positive SNPs caused by sequencing errors, the base support for each SNP in the parents was set at ≥4, while the base support for SNPs in the progeny were at ≥2 and ≤1000. Polymorphic SNPs in parents were selected, and the population was genotyped.

The SNP linkage map was constructed using JoinMap 4.0 (van Ooijen et al, 2006) with a logarithm of odds (LOD) threshold from 2 to 20. The maximal distance was set at 50 cM. Recombination frequencies were converted to map distance using the Kosambi map function (Kosambi et al, 1944).

QTL identification

Windows QTL Cartographer 2.5 (Wang et al, 2001) was applied to map the QTLs according to the composite interval mapping method (Zeng et al, 1944). QTLs detected in more than two environments were regarded as stable loci. LOD value for declaring significant QTLs including the QTLs across environments was calculated by a permutation test with the mapping step of 1.0 cM, five control markers, a significance level of $P < 0.05$, and $n = 1000$. The QTLs were nominated as following criteria: QTL name begins with ‘q’, followed by trait abbreviation, LGith “teriaere cleaned, a-7 individuals extraction.n .traits in flax, we ic map. number, and the QTL number. Positive additive effect indicates the favorable allele from the parent DYM, while the negative indicates the favorable allele from the other parent STS.

Results

Phenotypic variation and correlation of the traits relating to fiber quality and yield

A RIL population developed between the two parents STS and DYM was planted in three places (HB, NX and GS). The phenotypic variation ranges of the six traits PAL, STE OLE, LIN, LNA, and OIL for parents and the RIL population are summarized in Table 1 and Fig. 1. All the traits showed approximately

normal distributions with an absolute skewness value of less than one and were characterized by transgressive segregations with respective to their parental performances during the evaluations.

Table 1
Statistics of the six seed quality traits in the parents and the RIL population

Trait	Environment	Parents				Population						
		DYM	STS	Range	Average	Standard error	standard deviation	variance	kurtosis	skewness	minimum	Maximum
PAL	GS	6.46	5.58	0.88	6.03	0.03	0.45	0.20	-0.71	-0.34	4.90	7.04
	NX	6.87	5.54	1.33	6.26	0.03	0.48	0.23	0.17	0.14	5.13	8.15
	HB	6.48	6.12	0.36	5.85	0.03	0.43	0.19	0.11	0.04	4.76	7.37
STE	GS	4.83	3.78	1.05	4.27	0.05	0.73	0.53	-0.78	0.20	2.89	6.18
	NX	4.71	3.60	1.11	3.90	0.06	0.87	0.75	1.53	-0.37	0.74	6.30
	HB	3.62	3.45	0.17	3.24	0.05	0.71	0.50	-0.16	0.40	1.37	5.19
OLE	GS	21.94	24.91	2.97	23.89	0.17	2.62	6.86	-0.61	-0.08	17.65	31.29
	NX	21.25	24.51	3.26	23.60	0.18	2.70	7.31	0.13	0.31	17.87	31.08
	HB	20.19	20.00	0.19	19.34	0.14	2.14	4.56	0.27	0.20	14.36	27.03
LIN	GS	27.31	13.51	13.80	19.71	0.33	5.01	25.05	-1.45	0.19	11.47	30.87
	NX	28.85	13.39	15.46	19.68	0.32	4.84	23.46	-1.35	0.20	11.70	29.93
	HB	25.41	24.54	0.87	20.10	0.30	4.63	21.45	-1.34	0.16	8.37	29.36
LNA	GS	39.47	52.23	12.76	46.10	0.32	4.81	23.12	-0.89	-0.01	35.56	58.04
	NX	38.32	52.95	14.64	46.56	0.32	4.91	24.08	-0.80	-0.15	33.51	56.39
	HB	44.30	45.90	1.59	51.34	0.30	4.60	21.18	-0.76	-0.27	36.12	61.08
OIL	GS	40.89	42.90	2.01	41.38	0.14	2.09	4.36	0.41	-0.44	34.56	47.50
	NX	42.21	43.99	1.78	42.07	0.13	1.94	3.78	0.71	-0.46	34.61	46.60
	HB	41.29	40.92	0.36	43.20	0.11	1.74	3.01	0.90	-0.62	35.83	46.81

Table 2
The correlation analysis between the six traits

Trait	PAL	STE	OLE	LIN	LNA	OIL
PAL	1	-0.150*	0.034	0.169**	-0.276**	-0.091
STE	-0.150*	1	0.137*	-0.075	-0.127	0.021
OLE	0.033	0.137*	1	-0.355**	-0.183**	-0.224**
LIN	0.169**	-0.075	-0.355**	1	– .831**	-0.039
LNA	-0.276**	-0.127	-0.183**	-0.831**	1	0.160*
OIL	-0.091	0.021	-0.224**	-0.039	0.160*	1
** indicates the significant level of 0.01.						
* indicates the significant level of 0.05.						

For the correlation analysis, the result showed that the PAL showed significant negative correlation to STE and LNA and showed significant positive correlation to LIN; the STE showed significant positive to OLE; the OLE showed significant negative correlation to LIN, LNA and OIL; the LIN showed significant negative correlation to LNA and OIL; the LNA showed significant positive correlation to OIL.

Analysis of Gbs Data and Snp Marker Identification

Genomic DNA from STS, DYM and 232 RIL individuals was double digested by *Mse*I and *EcoR*I to construct libraries and sequencing by Illumina HiSeq2500. Totally,we obtained 2,372,287,104 bp and 2,345,804,352 bp clean data from STS and DYM separately, and 105.6 Gb total raw data from 232

RIL individuals with 451 Mb raw data of each individual. Raw data from both parent and RILs was high quanlity with Q20 ≥ 95%，Q30%≥89% of parents and Q20 ≥ 90%，Q30%≥85% of 232 progenies (Table S2). To further assess data quality, we analysis the average of *Mse*I catch ratio was 98.37% and *Mse*I or *EcoR*I enzyme cut completely ratio was 87.2% which indicated a high effective enzyme production sequenced a depth of approximately 69 × and high quality 25.9 Gb flax genome sequences (Table S3). We used BWA to map the parents and 232 progenies sequencing data to the published flax genome and obtained a high depth and wide coverage mapping results. In STS and DYM, mapping rate was more than 98% and average depth was more than 34 and in 232 RILs, mapping rate was more than 98% and coverage at least 1X was more than 10% (Table S4).

BWA alignment results of parent plants and RILs were further analyzed to identify SNP. As a result, for the parent STS, a total of 45,720 SNPs including 37,925 homozygous SNPs and 7,795 heterozygosis SNPs were identified. For DYM, 51,666 SNPs including 45,231 homozygous SNPs and 6,435 heterozygosis SNPs were identified. Subsequently, we chose 24,641 markers with the genotype of aa × bb and the polymorphism between parents (Table S5 and Table S6) for genetic map construction.

Genetic Map Construction and Structure Analysis

Of the 24,641 markers with aa × bb genotype in progeny, 10,045 with coverage more than 75% were screened and applied to construct the genetic map. The genetic map contains 15 linikage maps ranging from 178.28 cM to 28.14 cM in length. The genetic map contains 2,239 SNP markers, spanned total genetic distance of 1032.90 cM. The largest linkage group (LG) was LG1 that contains 528 SNPs in178.279 cM, while the smallest LG was LG13 that contains 51 SNPs in 28.14 cM. LG4 contains the lowest 43 SNP number. The average density of the SNP marker in the map was 0.46 marker per cM. In this map, there were two gaps, one was in the group 6 (21.48 cM) and the other was in the group 12 (31.927 cM) (Table 3 and Fig. 2).

Table 3
The detail information of the high density genetic map

Linkage Group Number	No. Marker	Genetic Distance (cM)	Average Distance (cM)	Max Gap (cM)
LG 1	528	178.28	0.34	19.1
LG 2	410	105.11	0.26	10.92
LG 3	175	68.02	0.39	10.58
LG 4	43	50.42	1.17	23.18
LG 5	63	45.71	0.73	16.3
LG 6	87	37.06	0.43	21.48
LG 7	93	55.75	0.60	7.7
LG 8	53	65.86	1.24	12.18
LG 9	82	49.52	0.60	7.13
LG 10	79	43.21	0.55	19.16
LG 11	163	41.23	0.25	3.84
LG 12	83	107.25	1.29	31.93
LG 13	51	28.14	0.55	6.16
LG 14	269	122.35	0.45	7.01
LG 15	60	35	0.58	6.2
Total	2239	1032.9	0.46	31.93

QTL location of seed quality related traits in three environments

Based on the phenotypic data of the population and the genetic map, we determined 21 QTLs of six traits two for PAL, four for STE, two for OLE, five for LIN, five for LAN three for OIL. Among them, one for LIN and one for LAN could be identify in multiple environments and the other could be only identifiy in a single environment.

For PAL, there were total two QTLs have been identified. On group two, there were one QTL. the *qPAL-Group2-1* could be detect in the environment HB, located on the position of 27.11 cM, explained 5.69% of the observed PV, and the direction was negative; on group 14, there were one QTL. The *qPAL-Group14-1* could be detected in the environment HB, located on the position of 6.41 cM, explained 5.24% of the observed PV, and the direction was positive.

For STE, there were four QTLs have been identified. On group one, there were three The *qSTE-Group1-1*, *qSTE-Group1-2* and *qSTE-Group1-3* could be detect in the environment GS, GS and NX respectively, located on the position of 23.61 cM, 70.21 cM and 95.71 cM respectively, explained 5.76%, 5.15%

and 5.32% respectively; and the direction of them were negative, negative and positive respectively; on group eight, there were one QTL, *qSTE-Group8-1* could be detected in the environment GS, located on the position of 39.31 cM, explained 7.93% of the observed PV, and the direction of it was positive.

For OLE, there were total two QTLs have been identified. On group one, there were two QTLs. The *qOLE-Group1-1* and *qOLE-Group1-2* could be both detected in the environment GS, located on the position of 15.31 cM and 94.41 cM respectively, explained 7.61% and 5.64% of the observed PV respectively, and their direction were negative and positive respectively.

For LIN, there were five QTLs have been detected. On group one, there were one QTL. The *qLIN-Group1-1* could be detected in the environment HB, located on the position on 16.71 cM, explained 5.43% of the observed PV, and its direction was positive; on group 11, there were one QTL. *The qLIN-Group11-1* could be detected in the environment HB, located on the position on 8.51 cM, explained 4.00% of the observed PV, and the direction were positive; on group 12, there were two QTLs. The *qLIN-Group12-1* could be detected in the environment NX, located on the position of 34.21 cM, explained 6.83% of the observed PV, and the direction was positive; The *qLIN-Group12-2* could be detected in all the three environments, located on the position interval of 46.81–48.81 cM, explained 76.09%-78.05% of the observed PV and the direction was positive; on group 14, there were one QTL. The *qLIN-Group14-1* could be detected in the environment HB, located on the position of 38.41 cM, explained 5.20% of the observed PV, and the direction was negative.

For LNA, there were total five QTLs have been identified. On group one, there were one QTL. The *qLNA-Group1-1* could be detected in the environment HB, located on the position of 144.31 cM, explained 6.62% of the observed PV, and the direction was positive; On group 4, there were one QTL. The *qLNA-Group4-1* could be detected in the environment NX, located on the position of 42.91 cM, explained 5.40% of the observed PV, and the direction was negative; On group 11, there were one QTL. the *qLNA-Group11-1* could be detected in the environment HB, located on the position of 8.51 cM, explained 4.69% of the observed PV, and the direction was negative; On group 12, there were two QTLs, the *qLNA-Group12-1* could be detected in the environment NX, located on the position of 32.21 cM, explained 8.00% of the observed PV and the direction was negative; the *qLNA-Group12-2* could be detected in the environment NX and HB, located on the position interval 46.81–48.81 cM, explained 57.33%-62.78% of the observed PV, and the direction was negative;

For, there were total two QTLs have been identified. On group three, there were two QTLs *qOIL-Group3-1* and *qOIL-Group3-2*, they both could be detected in the environment NX, located on the position of 39.31 cM and 58.71 cM respectively, explained 7.43% and 6.00% of the observed PV respectively and their direction were positive and negative respectively. On group 13, there were one QTL. The *qOIL-Group13-1* could be detected in the environment NX, located on the position of 26.11 cM, explained 4.00% of the observed PV, and its direction was positive (Table 4 and Fig. 3).

Table 4
The detail information of the QTLs for the six traits

QTL Name	Environment	Trait	Group Number	Position (cM)	LOD	Additive	R^2 (%)	Left CI (cM)	Right CI (cM)
qPAL-Group2-1	HB	PAL	2	27.11	2.62	-0.44	5.69	26.70	27.70
qPAL-Group14-1	HB	PAL	14	6.41	2.50	0.13	5.24	2.30	11.90
qSTE-Group1-1	GS	STE	1	23.61	2.85	-0.25	5.76	22.00	24.80
qSTE-Group1-2	GS	STE	1	70.21	2.68	-0.27	5.15	68.80	71.40
qSTE-Group1-3	NX	STE	1	95.71	2.82	0.25	5.32	92.50	97.90
qSTE-Group8-1	GS	STE	8	39.31	3.69	1.00	7.93	38.70	39.60
qOLE-Group1-1	GS	OLE	1	15.31	3.87	-2.06	7.61	14.70	15.60
qOLE-Group1-2	GS	OLE	1	94.41	3.14	1.00	5.64	91.40	97.70
qLIN-Group1-1	HB	LIN	1	16.71	2.67	2.89	5.43	15.90	18.10
qLIN-Group11-1	HB	LIN	11	8.51	2.00	1.45	4.00	7.30	9.70
qLIN-Group12-1	NX	LIN	12	34.21	3.51	1.66	6.83	28.30	34.80
qLIN-Group12-2	NX	LIN	12	46.81	18.85	4.30	76.10	44.00	50.90
	GS			48.81	24.34	4.55	80.28	45.30	52.00
	HB			48.81	16.57	4.15	78.05	45.10	52.00
qLIN-Group14-1	HB	LIN	14	38.41	2.72	-2.64	5.20	38.20	39.80
qLNA-Group1-1	HB	LNA	1	144.31	3.22	1.94	6.62	143.50	144.60
qLNA-Group4-1	NZ	LNA	4	42.91	2.90	-1.33	5.40	42.70	45.90
qLNA-Group11-1	HB	LNA	11	8.51	2.49	-2.15	4.69	7.30	9.30
qLNA-Group12-1	NZ	LNA	12	32.21	4.13	-1.96	8.00	30.70	39.90
qLNA-Group12-2	NZ	LNA	12	46.81	3.45	-3.80	57.33	41.80	51.20
	HB			48.81	5.46	-3.73	62.78	44.90	53.00
qOIL-Group3-1	NZ	OIL	3	39.31	3.31	1.14	7.43	37.80	40.30
qOIL-Group3-2	NZ	OIL	3	58.71	2.93	-1.0149	6.0012	57.5	59.2
qOIL-Group13-1	NZ	OIL	13	26.11	1.81	0.8202	4.0007	28.1	28.1

Discussion

In previous studies, most genetic maps of Flax were constructed with RFLP markers or SSR markers. Due to the characteristics of these two kinds of markes, the number of these markers is not large and the density of these markers is not large and most of these kinds of markers could not be compaired in reference genome with only one clear position (some have no position and some have more than one position), Spielmeyer et al. (1998) used 213 RAPD and RFLP markers constructed a genetic map with 18 LGs and identified two QTLs (Spielmeyer et al. 1998); Oh et al. (2000) also used 99 RFLP markers and RAPD markers contructed a genetic map (Oh et al. 2000); Gehringer et al. (2006) constructed a genetic map with AFLP and SSR markers and identified QTLs for Seed Quality traits (Gehringer et al, 2006) ; Cloutier et al (2011) constructed a genetic map that included 24 LGs with 113 EST-SSR markers and identified QTLs for seed color, linolenic acid content and linoleic acid content (Cloutier et al 2011); Cloutier et al (2012) construct a consensus genetic map by combined three individual linkage maps incorporating 770 markers based on 371 common markers including 114 that were shared by all three populations and 257 shared between any two populations (Cloutier et al 2012). From these results, we could see that the number of LGs and markers and the density of the maps were all not enough. For SNP markers, it could make up the disadvantage of RFLP, RAPD and SSR markers, but the research for constructing genetic map for Flax with SNP markers is not many. Kumar et al. (2015) constructed the first SNP genetic map for flax based on the selected 329 SNP markers and 362 SSR markers using a RIL population (Kumar et al. 2015); Wu et al (2018). constucted a genetic map included 15 LGs with 2,339 SLAF markers, the total length of the genetic map was 1483.25 cM and the average distance between adjacent markers was 0.63 cM (Wu et al. 2018); Zhang et al. (2018) constructed a genetic map which included on 15 LGs with 4497 SNP markers and the average density between adjacent marker was 2.71 cM. In our research, we also constructed a SNP map included 15 LGs thet harbored 2239 SNP markers, spanned total genetic distance of 1032.903 cM and average density between adjacent marker was 0.46 cM. These results could provide informationfor the next step of work such as identifying QTLs, identified candidate genes and function genes, and it could also contribute to MAS.

Untill now, the research about identified QTLs with SNP genetic map even includeded detected SNP loci by GWAS analysis is not many. Kumar et al identified 20 QTLs for fatty acid composition and yield traits based on the RIL population evaluated in eight environments (Kumer et al. 2015); Xie et al use GWAS analysis detected 16 SNP loci significantly associated with seed fatty acid content (Xie et al. 2018). In this research, we identified two stable QTLs (could be detected in more than one environment) and also one QTL clusters, but the number of stable QTLs and QTL clusters was not many. There may be two reasons of that, first was on the genetic map, some group harbored only a few markers with some visible gaps and the length of some group was short. This caused some QTLs and QTL clusters located on the gaps of the map and could not be identified; the second was that most of the fatty acid traits were quantitative trait and influenced by the environments much. In our research, as there were only three environments that were used to identify QTLs, the QTLs that could be detected in at least two environments were only a few. So in the next, we would discover more markers included SNP markers and SSR markers to increase the density of the genetic map and also planted the parents and the RIL population in more environments. For all that, these results could contribute identified candidate genes that related to fatty acid traits and know more about mechanism of fatty acid traits formation.

Declarations

Authors' contributions

ZW, WLM and ML performed the data analysis and drafted the manuscript. WB, DZ and LWJ performed the main experiments. ZL designed the experiments and finished the manuscript for publication. All authors approved the final version of the manuscript for publication.

Competing Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Genepioneer Biotechnologies Company for their technical support.

Funding

This work was financially supported by the earmarked fund for National Natural Science Foundation of China (31760397, 31160291) and China Agriculture Research System (CARS-14).

Availability of data and materials:

The flax reference genome referred in this work were downloaded from http://gigadb.org/dataset/100081. The descriptive statistics of seed quality traits of the parents in this work were presented in Supplementary Table S1. The detail information of the quality control of the sequencing data in this work was listed in Supplementary Table S2. The detail information of enzyme production was list in Table S3. The detail information of mapping ratio was list in Table S4. The detail information of type of SNPs was list in Table S5. The detail information of type of markers was list in Table S6.

“STS” was acquired from Mr. Yong-Bi Fu, Plant Gene Resources of Canada, Saskatoon Research Centre, Agriculture and Agri-Food Canada, 107 Science Place, Saskatoon, SK S7N 0X2, Canada; and “DYM” was acquired from Mr. Xin-Wen LI, Agronomy College of Inner Mongolia Agricultural University Hohhot 010019 China;

Ethics approval and consent to participate

This research was not applicable for ethics approval and consent to participate.

Consent to publish

This research is not applicable to consent for publication.

References

1. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. Rapid 286 SNP discovery and genetic mapping using sequenced RAD markers. PLoS One. 2008;3(10):e3376.
2. Carrasco B, González M, Gebauer M, García-González R, Maldonado J, Silva H. Construction of a highly saturated linkage map in Japanese plum (Prunus salicina L.) using GBS for SNP marker calling. PlosOne. 2018;13(12):e0208032.
3. Carter J. Flax seed as a source of alpha linolenic acid. J Am Coll Nutr. 1993;12(5):551.
4. Chandrawati, Yaday HK. Development of linkage map and mapping of QTLs for oil content and yield attributes in linseed (*Linum usitatissimum* L.). Euphytica. 2017;213:258.
5. Cloutier S, Ragupathy R, Niu ZX, Duguid S. SSR-based linkage map of flax (*Linum usitatissimum* L.) and mapping of QTLs underlying fatty acid composition traits. Mol Breeding. 2011;28:437–51.

6. Cloutier S, Ragupathy R, Miranda E, Radovanovic N, Rermer E, Walichnowski A, Ward K, Rowland G, Duguid S, Banik M. Integrated consensus genetic and physical maps of flax (*Linum usitatissimum* L.). Theor Appl Genet. 2012;125:1783–95.
7. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. A 297 robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One. 2011;6(5):e19379.
8. Fu YB. Genetic evidence for early flax domestication with capsular dehiscence. Genet Resour Crop Evol. 2011;58:1119–28.
9. Fu YB, Diederichsen A, Allaby RG. Locus-specific view of flax domestication history. Ecology Evolution. 2012;2(1):139–52.
10. Gehringer A, Friedt w, Lühs W, Snowdon RJ. Genetic mapping of agronomic traits in false flax (*Camelina sativa subsp. sativa*). Genome. 2006;46:1555–63.
11. Gill KS. Linseed. Publications and Information Division. New Delhi: Indian Council of Agricultural Research; 1987. p. p386.
12. Guo SY, Shi GL, Liu ZD, Zhao YH, Yang XX, Zhu JC, Li K, Guo XW. Using specific length amplified fragment sequencing to construct the high-density genetic map for Vitis (*Vitis vinifera* L. × *Vitis amurensis* Rupr.). Front Plant Sci. 2015. doi:10.3389.
13. Kosambi DD. The estimation of map distance from recombination values. Ann Eugen. 1994;12:172–5.
14. Kumar S, You FM, Cloutier S. Genome wide SNP discovery in flax through next generation sequencing of reduced representation libraries. BMC Genomics. 2012;13.
15. Kumar S, You FM, Duguid S, Booker H, Rowland G, Cloutier S. QTL for fatty acid composition and yield in linseed (*Linum usitatissimum* L.). Theor Appl Genet. 2015;128(5):965–84.
16. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60.
17. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.
18. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297–303.
19. Oh TJ, Gorman M, Cullis CA. RFLP and RAPD mapping in flax (Linum usitatissimum). Theor Appl Genet. 2000;101(4):590–3.
20. Perira L, Ruggieri V, Pérez S, Alexiou KG, Fernández M, Jahrmann T, Pujol M, Garcia-Mas J. QTL mapping of melon fruit quality traits using a high-density GBS-based genetic map. BMC Plant Biol. 2018;18:324.
21. Soto-Cerdalván BJ, Maureira-Butler I, Muñoz G, Rupayan A, Cloutier S. SSR-based population structure, molecular diversity and linkage disequilibrium analysis of a collection of flax (*Linum usitatissimum* L.) varying for mucilage seed-coat content. Mol Breeding. 2012;30:875–88.
22. Soto–Cerdalván BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S. Association mapping of seed quality traits using the Canadian flax (*Linum usitatissimum* L.) core collection. Theor Appl Genet. 2014;127:881–89.
23. Spielmeyer W, Green AG, Bittisnich D, Mendham N, Lagudah ES. Identification of quantitative trait loci contributing to fusarium wilt resistance on an AFLP linkage map of flax (*Linum usitatissimum*). Theor Appl Genet. 1998;97(4):633–41.
24. Sun XW, Liu DY, Zhang XF, Li WB, Liu H, Hong WG, Jiang acb, Guan N, Ma CX, Zeng HP, Xu CH, Song J, Huang L, Wang CM, Shi JJ, Wang R, Zheng XH, Lu CY, Wang XW, Zheng HK. SLAF–seq: An Efficient Method of Large–Scale De Novo SNP Discovery and Genotyping Using High–Throughput Sequencing. PLoS One. 2013;8(3):e58700.
25. Wang S, Basten CJ, Zeng ZB, Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. 2001. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm. Accessed 2013 Feb 1.
26. Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, Yang L, Hawkins S, Neutelings G, Datla R, Lambert G, Galbraith DW, Grassa CJ, Geraldes A, Cronk QC, Cullis C, Dash PK, Kumar PA, Cloutier S, Sharpe AG, Wong GK, Wang J, Deyholos MK. Genomic data of Flax (*Linum usitatissimum*). GigaScience Database. 2012. http://dx.doi.org/10.5524/100081.
27. Wu JZ, Zhao Q, Zhang LY, Li SY, Ma YH, Pan LY, Lin H, Wu GW, Yuan HM, Yu Y, Wang X, Yang X, Li ZG, Jiang TB, Sun DQ. QTL Mapping of Fiber-Related Traits Based on a High-Density Genetic Map in Flax (*Linum usitatissimum L.*). Frontier in Plant Science. 2018;9:885.
28. van Ooijen JW. JoinMap 4.0: Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Wageningen: Kyazma B.V; 2006.
29. Yang Q, Yang ZJ, Tang HF, Yu Y, Chen ZY, Wei SH, Sun QX, Peng ZS. High-density genetic map construction and mapping of the homologous transformation sterility gene (hts) in wheat using GBS markers. BMC Plant Biol. 2018;18:301.
30. Xie DW, Dai ZG, Yang ZM, Tang Q, Deng CH, Xu Y, Wang J, Chen J, Zhao DB, Zhang SL, Zhang SQ, Su JG. Combined genome-wide association analysis and transcriptome sequencing to identify candidate genes for flax seed fatty acid metabolism. Plant Sci. 2019;286:98–107.
31. Zeng ZB. Precision mapping of quantitative trait loci. Genetics. 1994;136:1457–68.
32. Zhang JP, Xie YP, Dang Z, Wang LM, Li WJ, Zhao W, Zhao L, Dang ZH. Oil content and fatty acid components of oilseed flax under different environments in China. Agron J. 2016;108(1):365–72.
33. Zhang JP, Long Y, Wang LM, Dang Z, Zhang TB, Song XX, Dang ZH, Pei XW. Consensus genetic linkage map construction and QTL mapping for plant height-related traits in linseed flax.
34. (*Linum usitatissimum* L). BMC Plant Biol. 2018;18:160.
35. Zohary D. Monophyletic vs. polyphyletic origin of the crops on which agriculture was founded in the near east. Gen Resour Crop Evol. 1999;46(2):133–42.

36. Zhang JP, Qi YN, Wang LM, Wang LL, Yan XC, Dang Z, Li WJ, Zhao W, Pei XW, Li XM, Liu M, Tan ML, Wang L, Long Y, Wang J, Zhang XW, Dang ZH, Zheng HK. Genomic comparison and population diversity analysis provide insights into the domestication and improvement of Flax iScience. 2020;6:23(4):100967.

Figures

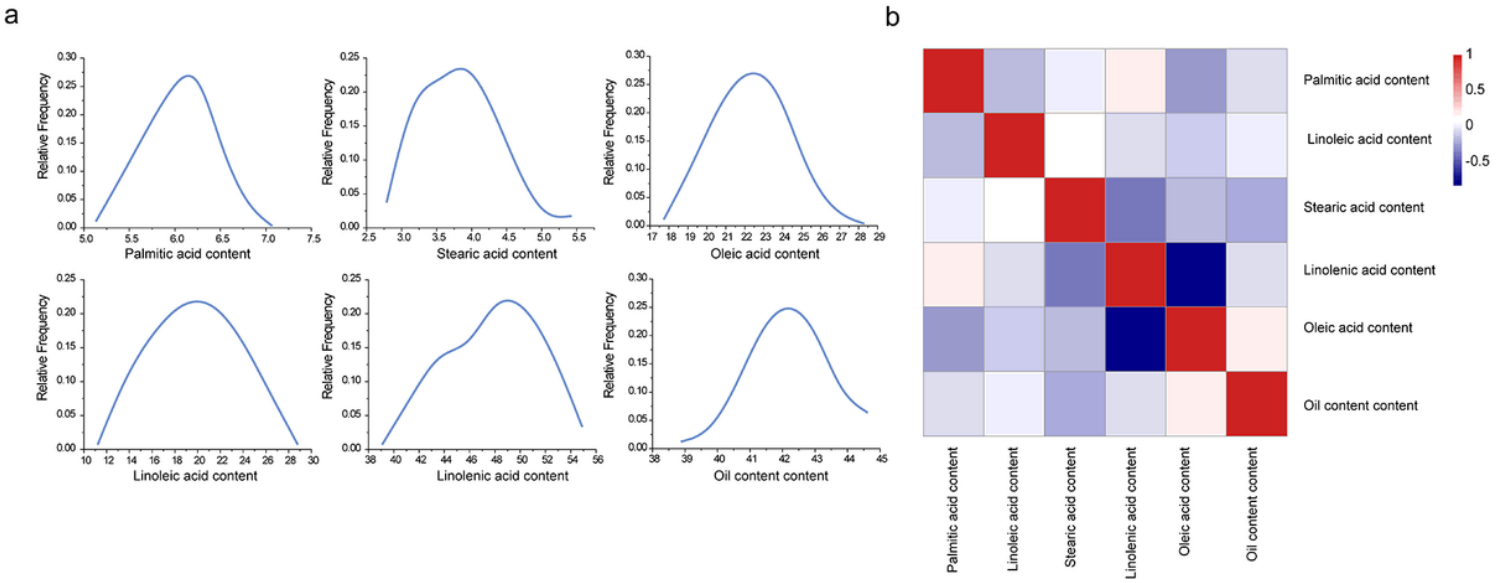

Figure 1

The distribution of the six traits in the RIL population and the correlation analysis between the six traits. a. The distribution of the six traits in the RIL population. b. The correlation analysis between the six traits.

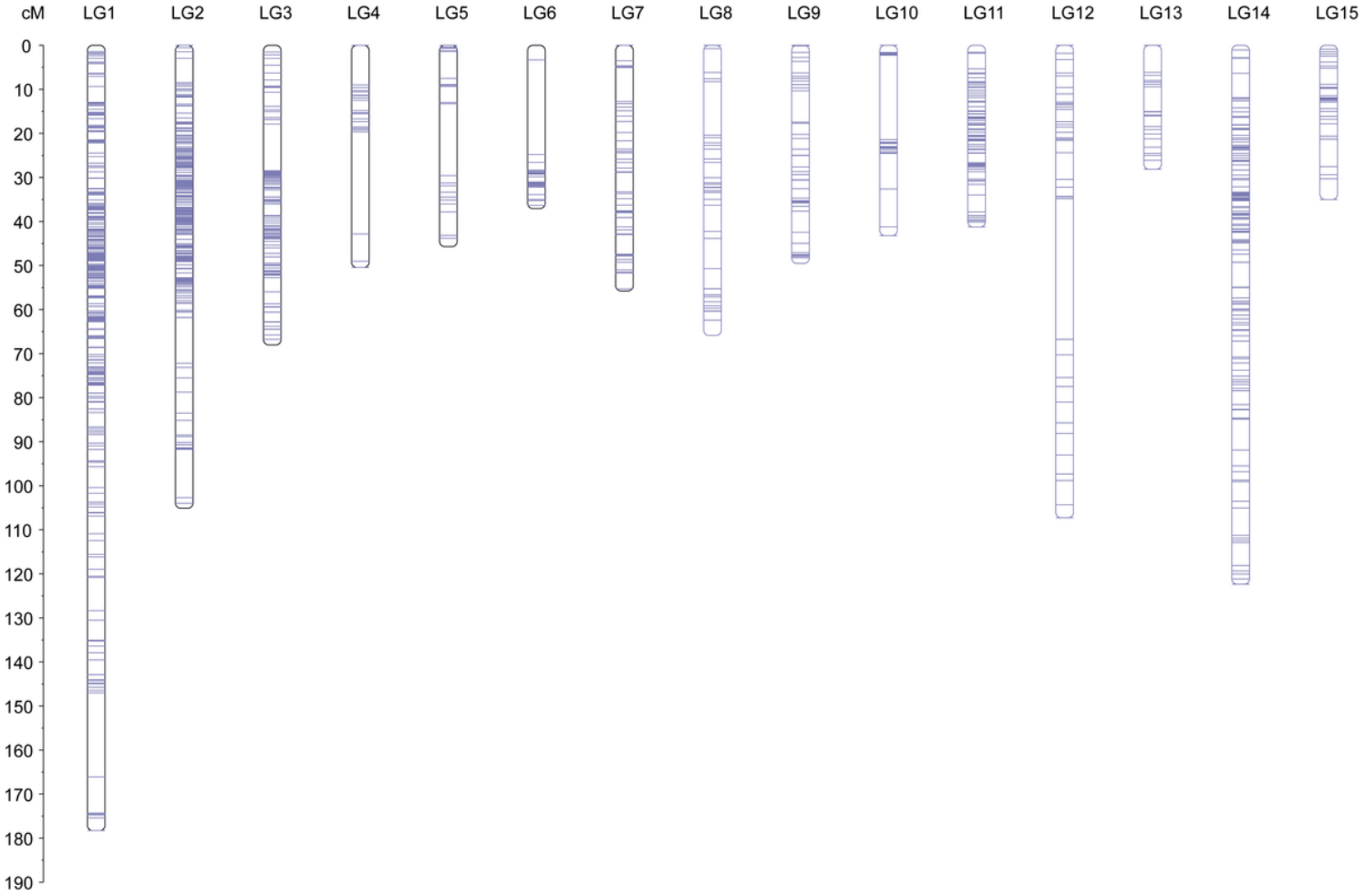

Figure 2

The high density genetic map.

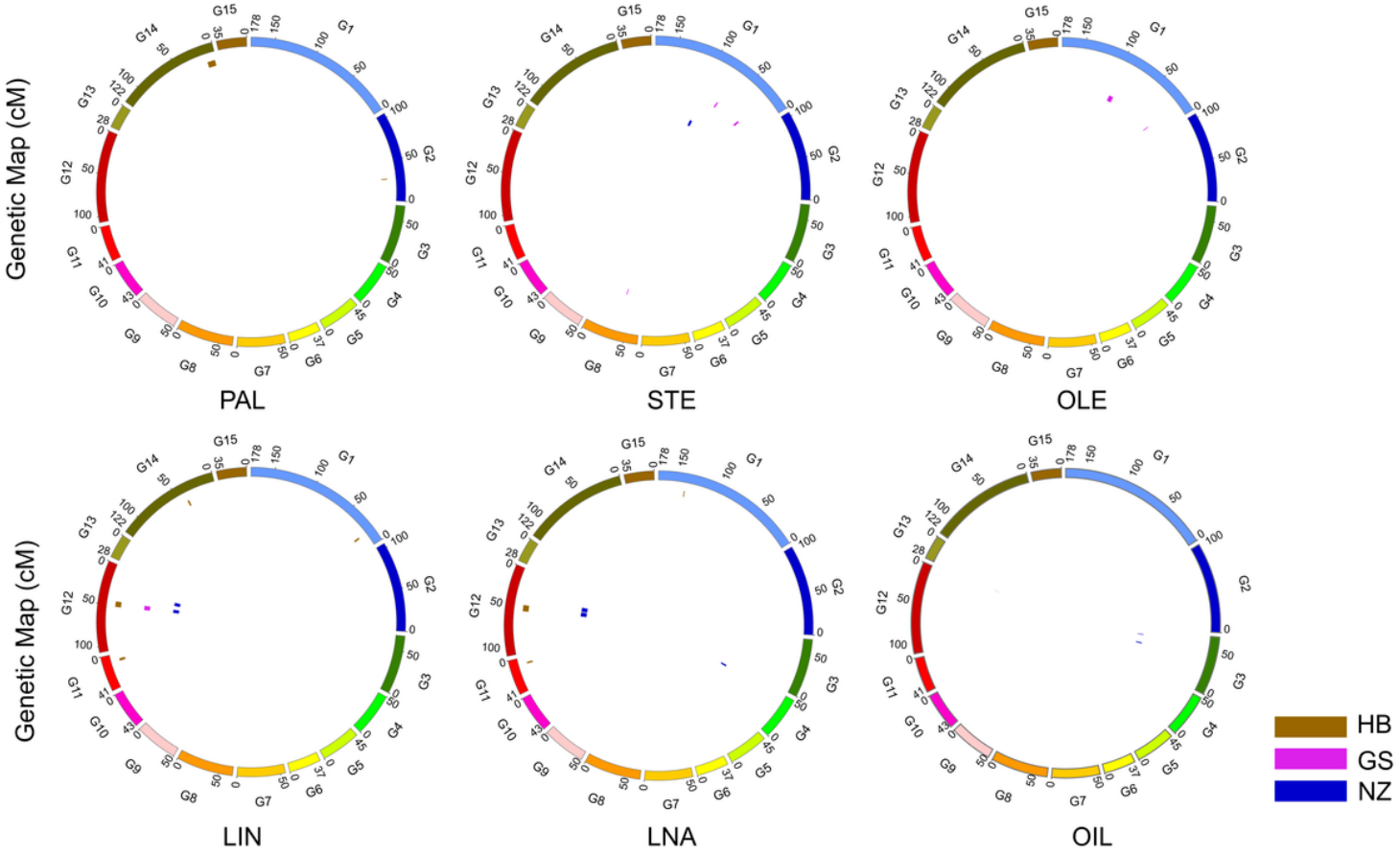

Figure 3

The QTLs of the six traits.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS6.xlsx
- TableS5.xlsx
- TableS4.xlsx
- TableS3.xlsx
- TableS2.xlsx
- TableS1.xlsx