

QTL mapping for pumpkin fruit traits using a GBS-based high-density genetic map

Xiaoxia Han

Hunan Academy of Agricultural Sciences

Ziyang Min

Hunan Academy of Agricultural Sciences

Minghua Wei

Huazhong Agriculture University College of Horticulture and Forestry Sciences

Yongqi Li

Hunan Academy of Agricultural Sciences

Duanhua Wang

Hunan Academy of Agricultural Sciences

Zhuqing Zhang

Hunan Academy of Agricultural Sciences

Xinjun Hu

Hunan Academy of Agricultural Sciences

Qiusheng Kong (✉ qskong@mail.hzau.edu.cn)

Huazhong Agriculture University College of Horticulture and Forestry Sciences

Research Article

Keywords: Pumpkin, Fruit traits, Genotyping by sequencing, High-density genetic map, QTLs

Posted Date: February 21st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1142200/v2>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Pumpkin is a popular vegetable crop and exhibits a broad diversity in fruit shape and size. Fruit-related traits are the decisive factors determining consumer acceptance and market value of pumpkin cultivar. As a result, deciphering the genetic basis of fruit-related traits is of great importance for pumpkin breeding. To address this problem, a F₂ population was generated by two *Cucurbita moschata* inbred lines with contrasting fruit shapes and genotyping by sequencing approach were used to construct a high-density genetic map and localize the QTLs underlying the fruit-related traits in this study. The results showed that a high-quality genetic map was constructed for pumpkin, which comprised of 2413 bins and spanned a total length of 2252.10 cM with an average genetic distance of 0.94 cM. A total of 30 significant QTLs with moderate or small effects were identified for 7 fruit-related traits, including fruit length, fruit diameter, fruit shape index, fruit weight, fruit flesh thickness, seed cavity size, and total soluble solids content. Co-locations were observed between the QTLs underlying different traits, demonstrating that pleiotropic effect plays an important role in genetic control of fruit-related traits. The identified QTLs provide valuable information for further fine mapping of the related genes and pumpkin breeding programs with the aim of improving fruit quality.

Introduction

Pumpkin (*Cucurbita moschata* Duchesne, 2n=2x=40) is an annual dicotyledonous vegetable belonging to the *Cucurbita* genus from the *Cucurbitaceae* family and widely cultivated for food, medicine, and ornamental purposes. The advantages of high yield, good storage life, long period of consumption, high nutritive value, and fitness in transport make pumpkin a very popular vegetable crop (Marie-Magdeleine et al. 2011). Its edible parts include fruits, seeds, flowers, leaves, and young stems. Pumpkin fruits are rich in carotenoids, phenolics, flavonoids, vitamins, carbohydrates, non-cellulosic polysaccharides, and amino acids, which have various effects beneficial to human health, such as anti-diabetic, anti-carcinogenic, antioxidant and anti-inflammatory potential (Yadav et al. 2010; Rolnik and Olas 2020). As a result, pumpkin is recognized as a functional food. Recently, pumpkin was recommended as a complementary food to combat vitamin A deficiency in low and middle income countries (Buzigi et al. 2021). Easy cultivation and high nutritional value make pumpkin constitute an important role in the basic diet of several countries and attractive for production around the world.

Fruit traits, including fruit size, fruit shape, fruit weight (FW), flesh thickness (FTH), seed cavity size (SCS) and sweetness, are the decisive factors for market value of cucurbit crops. Elucidating the genetic architectures of fruit morphology is a popular topic in cucurbit crops, and recently several fruit-related QTLs have been determined in cucumber (Wang et al. 2021), melon (Pereira et al. 2021), bottle gourd (Xu et al. 2021), and bitter melon (Rao et al. 2021). Fruit size and shape are the important external quality not only affecting the yield, but also determining the market classes and consumer acceptance (Weng et al. 2015). Consequently, they are the important targets of selection in cucurbit breeding programs. So far, phytohormone-related genes, microtubule-related genes and cyclin-related genes have been reported to play important roles in controlling fruit size in cucurbit crops (Chomicki et al. 2020). Meanwhile, a total of 150 consensus QTLs underlying fruit size and shape variations were comprehensively summarized in cucurbit crops, especially in cucumber (Pan et al. 2020). Besides the external quality, several QTLs related to fruit internal quality of cucurbit crops were also identified. In cucumber, 6 and 5 QTLs were identified for hollow size and FTH, respectively (Wang et al. 2021). One QTL was identified for FTH in bitter melon (Rao et al. 2021). One QTL determining sugar content was identified in melon (Pereira et al. 2021). These progresses will facilitate the fine-mapping and map-based cloning of the related genes in cucurbit crops.

Pumpkins show a broad diversity in fruit size and shape. Its major fruit shapes include round, oblate, oval, oblong, and pear-shaped. Deciphering the genetic basis of fruit traits is of great importance for pumpkin breeding. In the previous study, 12 pumpkin fruit morphology-related QTLs were identified in a F₂ population derived from the oblate-fruited and round-fruited parents (Zhong et al. 2017). Using a pyriform-shaped maternal parent and a dumbbell-shaped paternal parent, a moderate-effect QTL was detected for pumpkin fruit length (FL) and FW in the F₂ population, respectively (Hashimoto et al. 2020). Using the pumpkin parents with flattened shape and spheroid shape, respectively, a significant QTL underlying SCS was detected in the BC₁ population (Echevarria et al. 2020). These works have provided necessary information to understand the genetic control of fruit traits in pumpkin. However, compared with the other cucurbit crops, such as cucumber, melon and watermelon, knowledge on genetic architecture of fruit traits in pumpkin is very limited (Pan et al. 2020). Long and oblate shapes are the two dominant fruit shapes in commercial

pumpkin cultivars consuming for cuisine purpose. However, genetic controls on the long and oblate fruits of pumpkins are still unknown.

High-density genetic map is an efficient tool to dissect the genetic architecture of interested traits. In the past, traditional molecular makers, such as SSR, were used to construct genetic map (Kong et al. 2020). Due to the limited number of available makers, resolution of the genetic map constructed using the traditional markers is usually very low, which hampers the subsequent cloning of the targeted genes. With the advent of next generation sequencing, the genome-wide makers including SNP and InDel can be called and genotyped simultaneously, which greatly facilitate the construction of high-density genetic map. Genotyping by sequencing (GBS) is a robust, simple, high-throughput, and cost-effective approach for SNP marker discovery and has been widely used for studying genomic diversity (Elshire et al. 2011; Glaubitz et al. 2014). The genome-wide marker discovery capability makes GBS a powerful tool to construct high-density genetic map. A saturated genetic map with an average spacing of 0.4 cM was established for QTL analysis of fruit-related traits in Zucchini (*C. pepo*) using GBS approach. In pumpkin, a GBS-based genetic map with an average of 3.7 cM was constructed to detect the QTLs underlying fruit and agronomic traits (Echevarria et al. 2020). The comparatively low density of genetic map decreased the resolution of QTL mapping in pumpkin.

Pumpkin breeding is a time-consuming and challenging task. The large sizes of the plants also make pumpkin breeding is a resource-intensive job, especially in terms of space (Hernandez et al. 2020). Understanding the genetic basis and cloning of fruit-quality related genes are valuable to improve breeding efficiency. To address this problem, two *C. moschata* inbred lines with the contrasting fruit shapes were selected in this study, to construct a F2 segregating population with the aims of mapping the fruit-related QTLs using a GBS-based high-density genetic map. The results will provide substantial information for understanding the genetic basis of pumpkin fruit traits.

Materials And Methods

Construction and phenotypic evaluation of the F2 population

The *C. moschata* inbred lines N137 and N29, which differ significantly in fruit shape, were used as the materials. N137 has long fruit shape, while N29 has oblate fruit shape. The F1 hybrids were generated using N137 as the female parent and N29 as the male parent. Self-pollinations were performed on the F1 plants to produce an F2 population consisting of 200 individuals. The parental lines, F1 hybrids and F2 individuals were planted in an open field during the spring in Hunan Academy of Agricultural Sciences, China. The fruits were harvested at the stage of commercial mature. The fruit-related traits, including FL, fruit diameter (FD), FTH, SCS, and FW were measured. The fruit shape index (FSI) was calculated as the ratio of FL and FD. The total soluble solids (TSS) were measured using a pocket refractometer (ATAGO). The names and abbreviations of the investigated traits follows the rules proposed in cucumber (Wang et al. 2020). Histogram and density plots were used to display distributions of the traits. Pearson correlation coefficients between the traits were calculated using R (<https://www.r-project.org/>).

DNA isolation and GBS library construction and sequencing

Healthy young leaves from the F2 population and the parental lines were individually collected and stored at -80 °C. Genomic DNA was isolated from each sample using a Plant DNA Extraction Kit (Tiangen, China) following the manufacturer's protocol. Integrity and purity of each DNA sample were measured using the 1% agarose gel electrophoresis and a NanoDrop 1000 spectrophotometer. The high-quality DNA samples were used for sequencing library construction. Sequencings of the parents and F2 individuals were performed according to the GBS protocol (Elshire et al. 2011). The enzyme and sizes of the digested fragments were predicted and evaluated according to the reference genome of *C. moschata* (Sun et al. 2017). Based on the evaluation results, *Haell* was selected as the restrict enzyme. Library preparation and sequencing were carried out with the help of Novogene Co. Ltd on an Illumina HiSeq 4000 platform to produce the paired-end reads in length of 125 bp.

SNP discovery and genotyping

Quality controls on the generated reads, including adapter trimming and low-quality reads removing, were conducted using fastp (Chen et al. 2018). Genome sequence of *C. moschata* was download from the Cucurbit Genomics Database and used as the reference genome (Sun et al. 2017). The bwa-mem2 aligner was used to map the reads to the reference genome (Vasimuddin et al.

2019). Variants were called using GATK4 according to the GATK Best Practices. The mpileup and call modules of BCFtools were also used to identify the variant loci (Danecek et al. 2021). The SNPs identified both by GATK and BCFtools were selected for further filter. Using the SelectVariants module in GATK, the missing and heterozygous loci in each parental line, as well as the monomorphic loci between the parent lines were removed.

Genetic map construction

The high-quality SNPs were used to construct the genetic linkage map using lep-MAP3 (Rastas 2017). Briefly, the parental genotypes were called using the ParentCall2 module. Then the loci with segregation distortion ($p < 0.001$) or missing genotype ratio larger than 0.2 were removed using the Filtering2 module. LOD scores ranging from 20 to 40 were used to separate the SNPs into different linkage groups (LGs). At last, the markers were ordered in each LG and Kosambi genetic distances between the adjacent markers were calculated using the OrderMarkers2 module. The SNPs locating at the interval without recombination were grouped into a bin. The ID of the first SNP in each bin was used to represent the bin.

QTL mapping

QTL analysis was performed using the Composite Interval Mapping (CIM) method provided in R/qtl package (Arends et al. 2010). One thousand permutations were performed to determine the genome-wide significance threshold for QTL detection. QTLs exceeding the threshold value ($p < 0.05$) were considered significant. The QTL effect and the proportion of phenotypic variance explained (PVE) by the QTL were estimated at the highest QTL peaks using *fitqtl* function in R/qtl package. For each detected QTL, the confidence interval was defined using a 1.5-LOD support interval. Naming of QTL for the traits followed the nomenclature recommendations in cucumber (Wang et al. 2020).

Results

Phenotypic variation in the F2 population

To understand the genetic basis of fruit traits, pumpkin inbred lines of N137 with long fruit and N29 with oblate fruit were selected as the parents to generate an F2 segregating population. The long and oblate fruits are the dominant fruit shapes in the commercial pumpkin cultivars consuming for cuisine. The fruits of two parental lines, F1 hybrid, and several F2 individuals are shown in Fig. 1. Phenotypic variations of FL, FD, FSI, FTH, SCS, FW, and TSS were recorded for the parental lines, F1 hybrids and individuals in the F2 population at fruit commercially matured stage. Distributions of these traits are shown in Fig. 2. The FL of N137 was 5 folds higher than that of N29. While, the FD of N29 was 2 folds higher than that of N137. The FSI for N137 and N29 were 3.65 and 0.34, respectively. The FTH and SCS in N29 were higher than that of N137. The FW and TSS were higher in N137 compared with N29. Remarkable differences of these fruit-related traits between N137 and N29 demonstrated that different alleles harbored by the parental lines. Based on the performance of F1 hybrid, FL, FD, FSI, and SCS had high degree of dominance, while FTH, FW, and TSS exhibited transgressive segregation. Both the histograms and density plots displayed the continuous distributions of these fruit-related traits in the F2 population, suggesting their quantitative inheritances.

Person correlation coefficients were calculated between the traits, which are presented in Fig. 2. Highly positive correlations were observed between FSI and FL, FD and SCS, as well as FD and FTH, respectively, suggesting their potentially common genetic basis. While, highly negative correlations was observed between FSI and FD, FSI and SCS, FD and FL, respectively, implying different genetic basis between these traits. TSS exhibited very low or no significant correlation with the other traits, demonstrating its distinct genetic basis. Moderate correlations were observed between the remaining pairs of traits, suggesting different genetic factors may regulate them.

GBS sequencing and SNP calling

To determine the genotype of each plant, a total of 202 GBS libraries, including the 200 F2 individuals and the two parent lines, were sequenced and yielded 158.0 Gb paired-end reads. The clean reads for the parental lines of N29 and N317 were 3.4 Gb and 3.7Gb, representing $12.5 \times$ and $13.7 \times$ sequencing depth, respectively. The clean reads generated by the F2 individuals ranged from 0.45 Gb to 0.90 Gb, with the mean of 0.66 Gb. The corresponding sequencing depths of the F2 individuals varied from $1.7 \times$ to $3.4 \times$ with the mean of $2.5 \times$.

The clean reads were aligned to the *C. moschata* reference genome using bwa-mem2. The successful mapping rates varied from 98.1–99.7% with the mean of 99.3%. Using GATK HaplotypeCaller, a total of 2073737 variants, including 1649104 SNPs and 431304 InDels, were identified. BCFtools identified 2355207 variants, including 2117313 SNPs and 237894 InDels. Intersection analysis was conducted between the variants identified by GATK and BCFtools. The results showed that a total of 1423580 variants, including 1391130 SNPs and 32450 InDels were collectively identified by both GATK and BCFtools. The shared variants were selected and further filtered. The biallelic SNP loci with homozygous genotypes in each parental line and exhibiting polymorphism between the parental lines were kept. The loci with larger than 20% genotype missing ratio were filtered out, resulting in a total of 7082 high-quality SNPs with the Ts/Tv ratio of 2.13.

Genetic map construction

Lep-MAP3 was used to construct the genetic map. The parental genotypes firstly were called using the ParentCall2 module. The loci with significant segregation distortion ($p < 0.001$) were removed. The remaining SNPs were separated into LGs with the LOD scores varying from 20 to 40. At LOD score of 27, a total of 6918 SNPs were classified into 20 LGs and the number of single markers was 13. The other LOD scores resulted in number of LGs different from 20 or larger number of single markers. As a result, LOD score of 27 was adopted to separate the SNPs into LGs. The SNPs were then ordered in each LG and the Kosambi function was used to calculate their genetic distances. The SNPs locating in the interval without recombination were collapsed into a bin. The length and genotype for each bin are show in Fig. S1. All the SNPs were grouped into their corresponding chromosome, demonstrating a chromosome-level genetic map was established (Fig. S2). The number of crossovers in the F2 population varied from 33 to 64 with the mean of 46, which was reasonable in a F2 segregating population with 200 individuals (Fig. S3). The recombination fractions and the corresponding LOD scores for tests of linkage between all pairs of markers are shown in Fig. S4. The largest LOD was observed between the adjacent markers on the LGs, indicating the high quality of the genetic map. Good agreements were observed between the genetic marker positions and physical marker positions on the reference genome, adding confidence in the quality of this map (Fig. 3). All these evidences proved that a high-quality and chromosome-level genetic map was constructed for pumpkin. Summary statistics on the genetic map are list in Table 1. The genetic map comprised of 2413 bins spanning a total length of 2252.10 cM with an average genetic distance of 0.94 cM. Numbers of bins in the LGs varied from 74 (Chr20) to 192 (Chr4). Lengths of the LGs ranged from 78.85 cM (Chr20) to 197.28 cM (Chr4), with the average of 112.61 cM. The averaged spacing in each LG ranged from 0.74 cM to 1.18 cM with the mean of 0.94 cM. The maximum spacing in each LGs varied from 5.27 cM (Chr15) to 11.45 cM (Chr17). The high-density genetic map was integrated with the data of fruit-related traits for QTL analysis.

Table 1
Summary statistics on the genetic map

Chromosome/ linkage group	Number of bins	Length (cM)	Average spacing (cM)	Maximum spacing (cM)
1	171	141.93	0.83	9.36
2	138	124.13	0.91	6.54
3	115	114.89	1.01	6.03
4	192	197.28	1.03	7.30
5	118	99.85	0.85	6.54
6	123	112.60	0.92	5.52
7	101	101.39	1.01	6.79
8	88	102.95	1.18	9.62
9	125	97.32	0.78	6.28
10	125	107.07	0.86	5.52
11	172	140.86	0.82	6.28
12	148	108.61	0.74	6.54
13	82	89.28	1.10	11.45
14	119	125.86	1.07	5.27
15	105	106.85	1.03	5.27
16	98	98.10	1.01	6.28
17	108	103.49	0.97	11.45
18	124	114.85	0.93	6.28
19	87	85.96	1.00	9.88
20	74	78.85	1.08	5.27
Overall	2413	2252.10	0.94	11.45

Identification of QTLs underlying fruit traits

QTL analysis for the fruit traits was performed in R/qtl using the CIM method. Genome-wide views of QTLs detected in the F2 segregating population are illustrated in Fig. 4. Detailed information on the QTLs is reported in Table 2. In total, 30 QTLs were detected for the 7 traits.

Table 2
QTLs for fruit traits detected in pumpkin

Traits	QTL	Chr	Peak marker position (cM)	Peak LOD	P value	1.5 LOD interval left maker	1.5 LOD interval right maker	Additive effects	Dominant effect	PVE
FL	<i>fl2.1</i>	2	1.25	10.08	0.000	Chr02_15483	Chr02_479657	-4.45	-2.18	8.02
	<i>fl6.1</i>	6	97.34	6.95	0.004	Chr06_9951181	Chr06_10445766	3.93	0.68	5.67
	<i>fl14.1</i>	14	40.02	5.73	0.026	Chr14_3471947	Chr14_3832932	-1.92	3.67	3.45
FD	<i>fd1.1</i>	1	65.39	7.09	0.001	Chr01_5335115	Chr01_7205006	0.97	1.02	6.66
	<i>fd5.1</i>	5	18.76	6.49	0.010	Chr05_1998816	Chr05_2265078	0.01	1.61	6.05
	<i>fd12.1</i>	12	24.79	6.95	0.001	Chr12_1829011	Chr12_1931323	-0.38	-1.38	4.55
	<i>fd15.1</i>	15	54.00	6.15	0.014	Chr15_4022423	Chr15_4922896	-0.99	0.38	4.97
	<i>fd19.1</i>	19	28.81	6.33	0.013	Chr19_2048733	Chr19_2518096	0.07	1.19	3.26
FSI	<i>fsi2.1</i>	2	1.25	8.02	0.000	Chr02_15483	Chr02_479657	-0.37	-0.17	8.46
	<i>fsi6.1</i>	6	97.84	11.38	0.000	Chr06_9354982	Chr06_10262644	0.28	0.13	4.95
	<i>fsi15.1</i>	15	80.00	6.55	0.002	Chr15_7914605	Chr15_8691834	0.17	-0.43	6.84
FTH	<i>fth1.1</i>	1	71.40	5.43	0.038	Chr01_8382741	Chr01_9215814	0.15	0.20	4.22
	<i>fth4.1</i>	4	53.83	6.39	0.012	Chr04_4341605	Chr04_5663770	0.04	0.28	3.86
	<i>fyh5.1</i>	5	11.00	5.56	0.031	Chr05_1409933	Chr05_1998816	0.03	0.31	4.46
	<i>fh6.1</i>	6	77.56	7.30	0.003	Chr06_8853925	Chr06_9460491	0.19	0.10	4.14
	<i>fth14.1</i>	14	8.25	7.07	0.004	Chr14_975739	Chr14_1306550	0.10	0.29	4.88
	<i>fth16.1</i>	16	36.53	6.10	0.016	Chr16_2836084	Chr16_3412954	0.10	0.18	2.55
SCS	<i>scs5.1</i>	5	19.01	11.28	0.000	Chr05_2129410	Chr05_2265078	-0.05	1.07	5.23
	<i>scs12.1</i>	12	95.60	8.11	0.000	Chr12_10532303	Chr12_10661046	-0.34	-0.78	3.52
FW	<i>fw1.1</i>	1	13.26	9.19	0.000	Chr01_1149445	Chr01_1255673	0.30	0.25	5.05
	<i>fw2.1</i>	2	14.76	11.2	0.000	Chr02_1355009	Chr02_1712624	-0.08	0.37	4.01
	<i>fw12.1</i>	12	61.33	5.84	0.020	Chr12_7581002	Chr12_11670428	-0.08	-0.38	4.33
	<i>fw16.1</i>	16	37.53	8.14	0.001	Chr16_2513448	Chr16_3198950	0.09	0.41	5.13
TSS	<i>tss2.1</i>	2	56.54	8.96	0.000	Chr02_4883539	Chr02_5529654	-0.29	-0.13	2.60
	<i>tss3.1</i>	3	9.76	5.46	0.043	Chr03_1131739	Chr03_1436846	0.05	-0.54	3.99
	<i>tss4.1</i>	4	137.23	6.28	0.010	Chr04_16131184	Chr04_16757538	0.25	-0.59	6.03
	<i>tss5.1</i>	5	7.25	6.71	0.007	Chr05_1078419	Chr05_1769395	-0.05	-0.68	6.06
	<i>tss7.1</i>	7	82.59	6.02	0.016	Chr07_6470983	Chr07_6998095	0.44	-0.28	6.58
	<i>tss8.1</i>	8	8.00	9.55	0.000	Chr08_1038422	Chr08_1574490	0.20	0.53	4.98
	<i>tss9.1</i>	9	89.00	9.99	0.000	Chr09_8704222	Chr09_10729153	0.19	0.45	3.96

Fruit size was evaluated in terms of FL and FD. Three QTLs, namely *fl2.1*, *fl6.1*, and *fl14.1*, were detected for FL, which could explain 8.02%, 5.67%, and 3.45% of observed phenotypic variance (PVE), respectively. The *fl2.1* had negative additive and dominant effects on FL, whereas, *fl6.1* had positive additive and dominant effects on FL. A total of 5 QTLs were identified for FD,

which located on the chromosomes of 1, 5, 12, 15, and 19, respectively. The total PVE of the 5 QTLs reached up to 25.45%. No co-localized QTLs were found between FL and FD, demonstrating their different genetic basis.

Fruit shape was defined using the ratio of FL to FD, namely FSI. Three QTLs with the total PVE of 20.25% were detected for FSI. The *fsi2.1* and *fsi6.1* were overlapped with *fl2.1* and *fl6.1*, respectively, suggesting that FL plays a major role in determining fruit shape. Meanwhile, *fsi15.1* was closely linked with *fd15.1*.

Fruit internal quality was evaluated in terms of FTH and SCS. Six QTLs with the total PVE of 24.11% and 2 QTLs with the total PVE of 8.75% were detected for FTH and SCS, respectively. The *fth1.1* was closely linked with *fd1.1*. The *fth5.1* and *fth6.1* were co-localized with *fd5.1* and *fsi6.1*, respectively. While, *scs5.1* was overlapped with *fd5.1*. These results demonstrated that the genetic basis of fruit shape and internal quality were closely related.

FW is an important trait that determines the yield. A total of 4 QTLs for FW were identified. These QTLs showed similar effects on FW with the total PVE of 18.52%. The *fw12.1* and *fw16.1* were co-localized with *scs12.1* and *fth16.1*, respectively, indicating that FTH and SCS share the same genetic basis with FW to some extent.

TSS is an important factor determining fruit sensory quality. Seven QTLs were detected for TSS with the total PVE of 34.20%. All the QTLs had small effects on TSS. The *tss5.1* was overlapped with *fth5.1* and closely linked with *fd5.1* and *scs5.1*. These results suggested that this region possibly had pleiotropic effect on fruit size and quality.

Discussion

Pumpkin is an important vegetable with high nutritional value and wide adaptability. Fruit is its major edible part, and shows a broad diversity in size, shape and weight. As a result, fruit-related traits have been a subject of research for a long time. However, compared with the other cucurbit crops, especially cucumber and melon, knowledge on the genetic basis of fruit-related traits in pumpkins is very limited (Pan et al. 2020). To fill this gap, a F2 population and GBS sequencing approach were used to elucidate the genetic architecture of pumpkin fruit-related traits in this study.

Genetic map is a powerful tool to localize the QTLs underlying the targeted traits. Several genetic maps were constructed in pumpkins using different approaches. A genetic map spanning a total genetic distance of 3087.03 cM on 20 LGs with an average genetic distance of 0.89 cM was constructed to map QTLs for fruit-related traits in pumpkin (Zhong et al. 2017). Moreover, using the GBS approach, a genetic map consisting of 24 LGs with an overall genetic length of 2665.0 cM and an average spacing of 3.7 cM was constructed for a pumpkin BC1 population (Echevarria et al. 2020). These genetic maps were constructed without using the information of reference genome, which hampered their comparisons with the other genetic maps. Recently, based on a F2 population, a genetic map with total length of 2026.7 cM and an average marker interval of 4.6 cM was constructed in pumpkin (Hashimoto et al. 2020). The resolution of this map was relatively low, which decreased the accuracy of QTL mapping. In the present study, a genetic map comprising of 2413 bins and spanning 2252.10 cM with an average genetic distance of 0.94 cM was constructed (Table 1). Tests on crossover number, relationships between recombination fraction and LOD scores, as well as synteny between the genetic and physical maps proved that it was a high-quality and chromosome-level genetic map, which greatly facilitated the downstream QTL mapping and comparison with the other genetic maps.

Fruit size is an important composition of fruit external quality, and determines product classes and consumer acceptance (Weng et al. 2015). FL and FD are usually used to describe the fruit size and represent fruit elongation and radial growth, respectively (Pan et al. 2020). FL was positively correlated with FD in *C. maxima* (Kazminska et al. 2020). However, negative correlation of phenotypic variations between FL and FD was observed in the present study, probably due to their different genetic backgrounds. Two QTLs for FD were previously detected in pumpkin (Zhong et al. 2017). Due to lack of the physical position information in pumpkin reference genome, it was difficult to compare the above-mentioned QTLs with the other studies. Moreover, 3 and 2 QTLs were identified for FL and FD in *C. maxima*, respectively (Kazminska et al. 2018). In the present study, 3 QTLs for FL and 5 QTLs for FD were detected in pumpkin. Of them, a QTL for FL locating on chromosome 14 was overlapped with the one detected in *C. maxima*, while no consensus QTLs for FD were observed between pumpkin and *C. maxima* (Kazminska et al. 2018). In addition, no co-location of QTLs for FL and FD were found in the present study, suggesting that elongation of fruit and increase of diameter might be under different genetic mechanisms in pumpkin.

Fruit shape is usually represented by FSI, which is the ratio of FL and FD. Three QTLs with the total PVE of 20.25% were detected for FSI in this study, and two of them overlapped with the FL related QTLs, indicating that fruit elongation played an important role in determination of final fruit shape. Similar results were also observed in cucumber (Weng et al. 2015; Pan et al. 2020). However, in *C. maxima*, the FSI was reported not to reflect the variation of fruit shape appearing in RILs (Kazminska et al. 2020). Since FSI is a composite trait and calculated from FL and FD, its genetic basis is unknown. As a result, caution should be taken in using the FSI related QTLs in marker-assisted selection, and more sophisticated approach for fruit shape phenotyping is needed to accurately map fruit shape QTLs (Xu et al. 2021).

FW is an important yield-related trait. In the present study, FW exhibited transgressive segregation and was positively correlated with FL, FD, SCS, and FTH, respectively. Similar results were also observed in *C. maxima* (Kazminska et al. 2020). In the previous reports, 1 and 6 QTL related to FW was detected in pumpkin and *C. maxima* (Kazminska et al. 2020; Hashimoto et al. 2020), respectively. In the present study, 4 QTLs were detected for FW, and two of them were co-localized with the SCS and FTH related QTLs, respectively. Similar result was also observed in *C. maxima* (Kazminska et al. 2020). These results demonstrated that FTH and SCS have the same genetic basis with FW to some extent.

Fruit internal quality includes the traits of FTH, SCS, TSS, etc. A fruit with small SCS, thick FTH, and high TSS is preferred for both the producers and consumers. In the present study, the parental line N137 had thicker FTH, smaller SCS, and higher TSS than N25. FTH was positively correlated with SCS in the F2 population. Similar results were also observed in cucumber (Wang et al. 2021). Three and 4 QTLs for FTH were previously detected in pumpkin and *C. maxima* (Zhong et al. 2017; Kazminska et al. 2020), respectively. In the present study, 6 and 2 QTLs were detected for FTH and SCS, respectively. Co-locations of QTLs between FTH and TSS, FD, FSI, as well as FW were observed, respectively. Close locations of QTLs between FTH and FD, SCS were also observed. Two QTLs for SCS were overlapped with the QTLs of FW and FD, respectively. These results demonstrated that the genetic basis of fruit external and internal qualities were closely related.

TSS is an important component affecting fruit quality and taste, which has been studied in several crops. No QTLs related to Brix were detected in pumpkin in the previous report (Zhong et al. 2017). However, 1 QTL related to Brix was detected in pumpkin in a recent report (Hashimoto et al. 2020). Similarly, 1 QTL for Brix was also identified in melon (Pereira et al. 2021). In the present study, 7 QTLs for TSS with the total PVE of 34.20% were identified. The large number of QTLs for TSS identified in the present study is probably due to the significant difference of TSS contents between the parental lines. Moreover, co-location and close location were observed between TSS and FTH and FD on chromosome 5, respectively, indicating that pleiotropic effect plays an important role in genetic control of fruit traits. Taken together, the QTLs identified in the present study provide valuable information for further fine mapping of the related genes and pumpkin breeding programs with the aim of improving fruit quality.

Conclusion

In the present study, a chromosome-level high-quality linkage map was constructed for pumpkin, which comprised of 2413 bins and spanned a total length of 2252.10 cM with an average genetic distance of 0.94 cM. Using the genetic map, a total of 30 significant QTLs with moderate or small effects were identified for 7 traits related to the external or internal quality of pumpkin fruit. Co-locations and close locations were observed between the QTLs underlying different traits. The identified QTLs will provide substantial information for further fine mapping of the related genes and pumpkin breeding programs aiming at improving fruit quality.

Declarations

Funding

This work was supported by the National Natural Science Foundation of China (grant number 31601755), China Agriculture Research System of MOF and MARA (grant number CARS-23-G-29) and Natural Science Foundation of Hunan Province-Youth Science Fund (grant number 2021JJ40300).

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

All authors contributed to the study conception and design. Material preparation and data collection were performed by Xiaoxia Han, Ziyang Min, Yongqi Li, Duanhua Wang, Zhuqing Zhang and Xinjun Hu. Data analysis was performed by Minghua Wei and Qiusheng Kong. The first draft of the manuscript was written by Qiusheng Kong and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Arends D, Prins P, Jansen RC, Broman KW (2010) R/qtl: high-throughput multiple QTL mapping. *Bioinformatics* 26 (23):2990-2992. doi:10.1093/bioinformatics/btq565
2. Buzigi E, Pillay K, Siwela M (2021) Potential of pumpkin to combat vitamin A deficiency during complementary feeding in low and middle income countries: variety, provitamin A carotenoid content and retention, and dietary reference intakes. *Critical Reviews in Food Science and Nutrition*. doi:10.1080/10408398.2021.1896472
3. Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34 (17):884-890. doi:10.1093/bioinformatics/bty560
4. Chomicki G, Schaefer H, Renner SS (2020) Origin and domestication of Cucurbitaceae crops: insights from phylogenies, genomics and archaeology. *New Phytologist* 226 (5):1240-1255. doi:https://doi.org/10.1111/nph.16015
5. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. *Gigascience* 10 (2). doi:10.1093/gigascience/giab008
6. Echevarria ARDV, Campbell A, Radovich TJK, Silvasy T, Moore S, Kantar MB (2020) Quantitative Trait Loci (QTL) Analysis of Fruit and Agronomic Traits of Tropical Pumpkin (*Cucurbita moschata*) in an Organic Production System. *Horticulturae* 6 (1). doi:10.3390/horticulturae6010014
7. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *Plos One* 6 (5). doi:10.1371/journal.pone.0019379
8. Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *Plos One* 9 (2). doi:10.1371/journal.pone.0090346
9. Hashimoto T, Kubo N, Nishimura K, Nagano AJ, Sasaki A, Nakamura Y, Mimura Y (2020) Quantitative Trait Locus Analysis in Squash (*Cucurbita moschata*) Based on Simple Sequence Repeat Markers and Restriction Site-Associated DNA Sequencing Analysis. *Horticulturae* 6 (4). doi:10.3390/horticulturae6040071
10. Hernandez CO, Wyatt LE, Mazourek MR (2020) Genomic Prediction and Selection for Fruit Traits in Winter Squash. *G3-Genes Genomes Genetics* 10 (10):3601-3610. doi:10.1534/g3.120.401215
11. Kazminska K, Hallmann E, Korzeniewska A, Niemirowicz-Szczytt K, Bartoszewski G (2020) Identification of Fruit-Associated QTLs in Winter Squash (*Cucurbita maxima* Duchesne) Using Recombinant Inbred Lines. *Genes* 11 (4). doi:10.3390/genes11040419
12. Kazminska K, Hallmann E, Rusaczek A, Korzeniewska A, Sobczak M, Filipczak J, Kuczerski KS, Steciuk J, Sitarek-Andrzejczyk M, Gajewski M, Niemirowicz-Szczytt K, Bartoszewski G (2018) Genetic mapping of ovary colour and quantitative trait loci for carotenoid content in the fruit of *Cucurbita maxima* Duchesne. *Molecular Breeding* 38 (9). doi:10.1007/s11032-018-0869-z
13. Kong Q, Liu Y, Xie J, Bie Z (2020) Development of Simple Sequence Repeat Markers from De Novo Assembled Transcriptomes of Pumpkins. *Plant Molecular Biology Reporter* 38 (1):130-136. doi:10.1007/s11105-019-01189-8

14. Marie-Magdeleine C, Mahieu M, Archimède H (2011) Chapter 110 - Pumpkin (*Cucurbita moschata* Duchesne ex Poir.) Seeds as an Anthelmintic Agent? In: Preedy VR, Watson RR, Patel VB (eds) Nuts and Seeds in Health and Disease Prevention. Academic Press, San Diego, pp 933-939. doi:<https://doi.org/10.1016/B978-0-12-375688-6.10110-0>
15. Pan Y, Wang Y, McGregor C, Liu S, Luan F, Gao M, Weng Y (2020) Genetic architecture of fruit size and shape variation in cucurbits: a comparative perspective. *Theoretical and Applied Genetics* 133 (1):1-21. doi:10.1007/s00122-019-03481-3
16. Pereira L, Santo Domingo M, Argyris J, Mayobre C, Valverde L, Montserrat Martin-Hernandez A, Pujol M, Garcia-Mas J (2021) A novel introgression line collection to unravel the genetics of climacteric ripening and fruit quality in melon. *Scientific Reports* 11 (1). doi:10.1038/s41598-021-90783-6
17. Rao PG, Behera TK, Gaikwad AB, Munshi AD, Srivastava A, Boopalakrishnan G, Vinod (2021) Genetic analysis and QTL mapping of yield and fruit traits in bitter melon (*Momordica charantia* L.). *Scientific Reports* 11 (1). doi:10.1038/s41598-021-83548-8
18. Rastas P (2017) Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics* 33 (23):3726-3732. doi:10.1093/bioinformatics/btx494
19. Rolnik A, Olas B (2020) Vegetables from the Cucurbitaceae family and their products: Positive effect on human health. *Nutrition* 78. doi:10.1016/j.nut.2020.110788
20. Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, Zhang J, Zhang H, Gong G, Jia Z, Zhang F, Tian J, Lucas WJ, Doyle JJ, Li H, Fei Z, Xu Y (2017) Karyotype Stability and Unbiased Fractionation in the Paleo-Allotetraploid *Cucurbita* Genomes. *Molecular Plant* 10 (10):1293-1306. doi:10.1016/j.molp.2017.09.003
21. Vasimuddin M, Misra S, Li H, Aluru S, lee (2019) Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In: 33rd IEEE International Parallel and Distributed Processing Symposium (IPDPS), Rio de Janeiro, BRAZIL, 2019. International Parallel and Distributed Processing Symposium IPDPS. pp 314-324. doi:10.1109/ipdps.2019.00041
22. Wang Y, Bo K, Gu X, Pan J, Li Y, Chen J, Wen C, Ren Z, Ren H, Chen X, Grumet R, Weng Y (2020) Molecularly tagged genes and quantitative trait loci in cucumber with recommendations for QTL nomenclature. *Horticulture Research* 7 (1). doi:10.1038/s41438-019-0226-3
23. Wang Y, Jiang B, Dymerski R, Xu X, Weng Y (2021) Quantitative trait loci for horticulturally important traits defining the Sikkim cucumber, *Cucumis sativus* var. *sikkimensis*. *Theoretical and Applied Genetics* 134 (1):229-247. doi:10.1007/s00122-020-03693-y
24. Weng Y, Colle M, Wang Y, Yang L, Rubinstein M, Sherman A, Ophir R, Grumet R (2015) QTL mapping in multiple populations and development stages reveals dynamic quantitative trait loci for fruit size in cucumbers of different market classes. *Theoretical and Applied Genetics* 128 (9):1747-1763. doi:10.1007/s00122-015-2544-7
25. Xu P, Wang Y, Sun F, Wu R, Du H, Wang Y, Jiang L, Wu X, Wu X, Yang L, Xing N, Hu Y, Wang B, Huang Y, Tao Y, Gao Q, Liang C, Li Y, Lu Z, Li G (2021) Long-read genome assembly and genetic architecture of fruit shape in the bottle gourd. *Plant Journal*. doi:10.1111/tpj.15358
26. Yadav M, Jain S, Tomar R, Prasad GBKS, Yadav H (2010) Medicinal and biological potential of pumpkin: an updated review. *Nutrition Research Reviews* 23 (2):184-190. doi:10.1017/s0954422410000107
27. Zhong Y-J, Zhou Y-Y, Li J-X, Yu T, Wu T-Q, Luo J-N, Luo S-B, Huang H-X (2017) A high-density linkage map and QTL mapping of fruit-related traits in pumpkin (*Cucurbita moschata* Duch.). *Scientific Reports* 7. doi:10.1038/s41598-017-13216-3

Figures

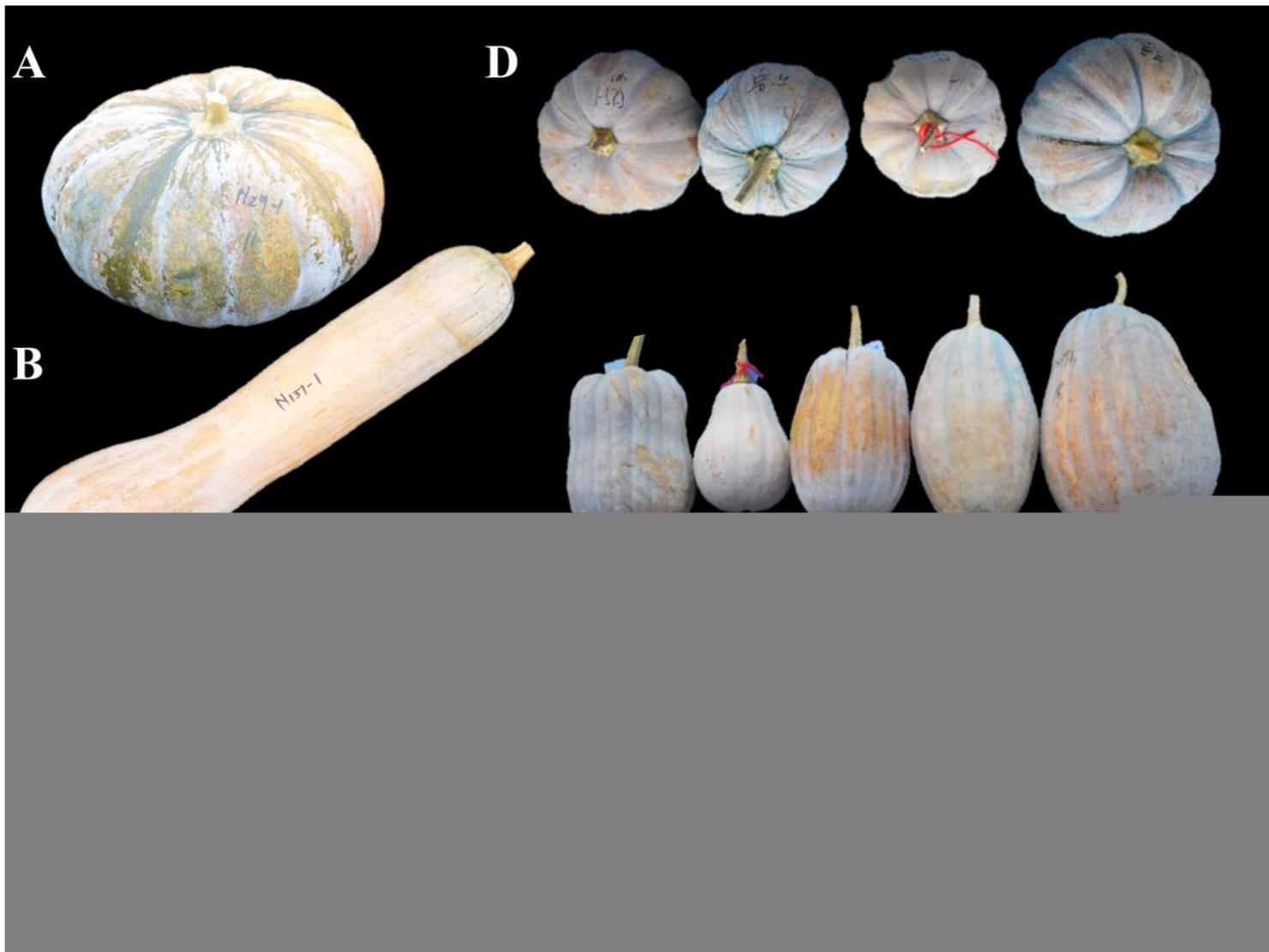


Figure 1

Representative fruits of the parental lines N29 (A), N317 (B), F1 hybrid (C), and some F2 individuals (D).

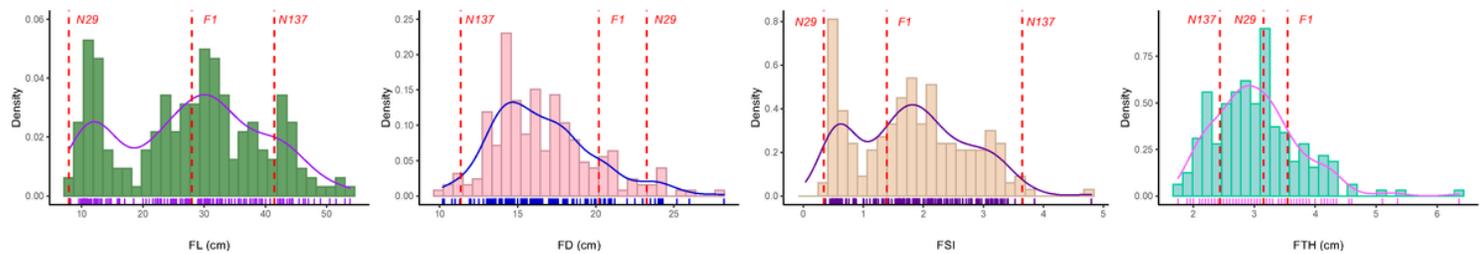


Figure 2

Histogram and density plots of pumpkin fruit traits and their correlation matrix

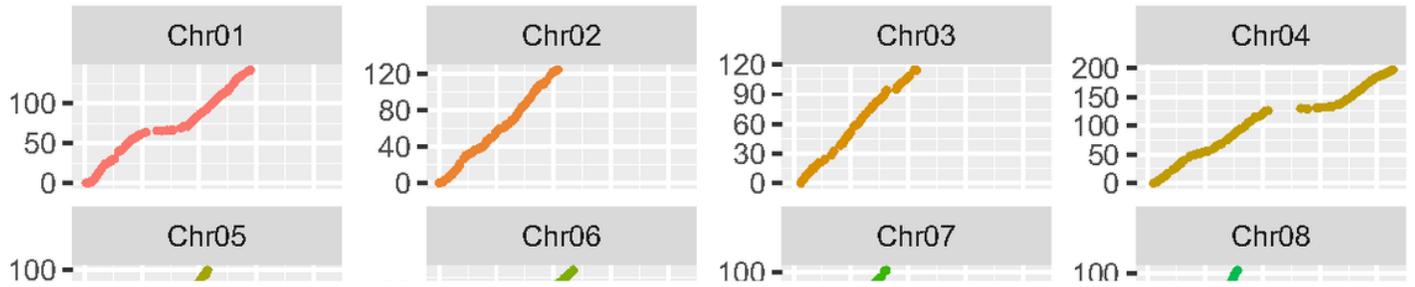


Figure 3

The synteny between linkage map and physical map

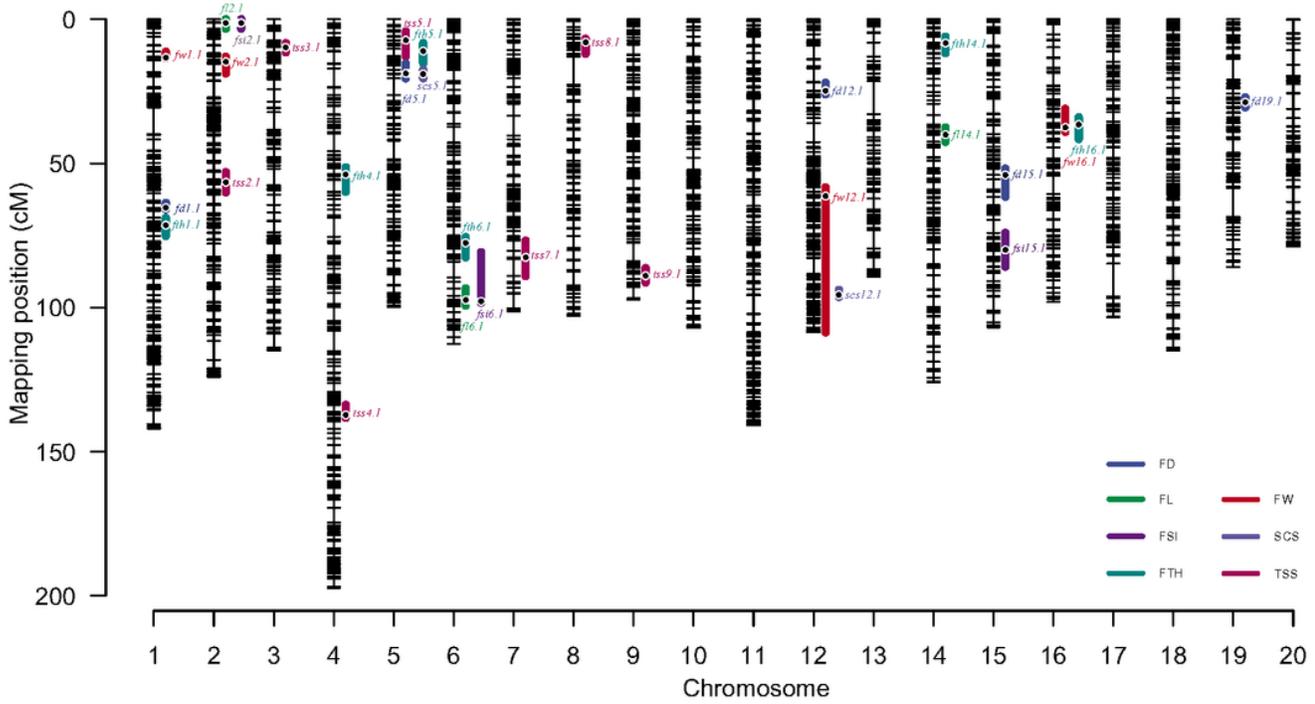


Figure 4

Locations of the identified fruit traits related QTLs on pumpkin chromosomes

Supplementary Files

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

- [Fig.S1Thelengthandgenotypeofbins.pdf](#)
- [Fig.S2Pumpkingeneticmap.pdf](#)
- [Fig.S3Thedistributionofcrossovernumber.pdf](#)
- [Fig.S4TherecombinationsandLODscores.pdf](#)