BMC Genomics

RESEARCH

Open Access

EMS-induced genomic variation and QTL-seq analysis of safflower spininess through whole genome sequencing (WGS)

Samaneh Karami-Moalem[1], Asadollah Ahmadikhah[1*], Zahra Nemati[2] and Reza Haghi[2*]

Abstract

Phenotypic variation is detrimental for deciphering the genetic control of valuable traits in crop plants. Safflower (*Carthamus tinctorius* L.) is an industrial and medicinal plant with natural stress-resilience. In this study, we used two bulk samples of safflower (spiny and spineless) developed after ethyl methanesulfonate (EMS) induced mutagenesis, aimed to find genomic variation between two phenotypes through whole genome sequencing (WGS) and to find putative associations with spine formation on the plant organs. More than 890,000 EMS-induced SNPs were obtained, and comparison of the G: C > A:T canonical mutation rate in EMS-induced and 93 natural safflower samples obtained via genotyping-by-sequencing (GBS) showed that the canonical mutation rate in two studies was significantly different (38.03% vs. 32.4%). Most of the mutations affected the intergenic regions (46.81%) and only 3.94% were located in exome which could have a high putative impact on gene function. More than 137,000 filtered SNPs were used for marker-trait association (MTA) via QTL-seq analysis which resulted in identifying a genomic region on short arm of chromosome 7 (qSpiny7) to be associated with spininess. The qSpiny7 contains 24 genes carrying 151 SNPs in their exons, 40 of which were functional SNPs. RNA-seq assay indicated that 14 and 8 genes within qSpiny7, respectively, were up- and down-regulated in spineless mutant. The expression patterns of 6 selected genes were confirmed using real-time qRT-PCR. A new candidate gene (*BG1-like E*) involved in cell division and different organ`s enlargement, and a recently identified gene, Lonely Guy (LOG), for controlling spine/prickle formation in higher plants, also locate within qSpiny7 in safflower. The latter gene harbors a SNP in its first intron. Based on our results, the major effect genomic region qSpiny7 harbors important genes for developmental processes and stress responses. These findings can open new insight into the nature of natural and induced genomic variations of safflower and the applicability of them in MTA studies.

Keywords EMS, QTL-seq, BSA, WGS, Safflower, Spiny, Variant calling

*Correspondence:
Asadollah Ahmadikhah
a_ahmadikhah@sbu.ac.ir
Reza Haghi
haghi@ipk-gatersleben.de
[1]Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
[2]Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany

Introduction

Safflower (*Carthamus tinctorius* L.) has emerged as a valuable oilseed crop renowned for its high content of unsaturated fatty acids, particularly linoleic acid, and the accumulation of unique flavonoids such as hydroxysafflor yellow A (HSYA), making it important for both industrial applications and medicinal use [1]. Beyond its nutritional and economic significance, safflower also provides a model for investigating genetic and metabolic adaptations central to crop improvement. However, breeding efforts in safflower have been historically hampered by limited genomic resources and an incomplete understanding of the genetic architecture governing key agronomic traits such as harboring spines, oil composition, yield, and abiotic stress tolerance [2]. Safflower seed oil is odorless, rich in polyunsaturated fatty acids (PUFAs), with the overall qualities believed to be comparable to those of sunflower oil [3].

Recent advances in high-throughput sequencing (HTS) technologies have initiated a transformation in safflower genomics. The first high-quality reference genome and chromosomal assemblies were reported only within the past few years, providing unprecedented resolution for genome-wide characterization of gene families, metabolic pathways, and evolutionary processes [4–6]. These resources have uncovered expansions in crucial gene families such as fatty acid desaturases and chalcone synthases, which underpin the biosynthesis of important oils and flavonoids [1, 6]. Comprehensive re-sequencing and association mapping studies have further revealed substantial genetic diversity among global safflower accessions and identified numerous marker-trait associations (MTAs) that can accelerate marker-assisted selection (MAS) in breeding programs [2, 6].

The integration of whole genome sequencing (WGS) with QTL-seq analysis, a method combining bulked segregant analysis (BSA) with deep sequencing holds particular promise for dissecting the genetic basis of complex quantitative traits in safflower like harboring spines on different plant organs like leaves margins. By contrasting allele frequencies in DNA bulks constructed from individuals showing extreme phenotypes within segregating populations, QTL-seq enables rapid identification of candidate genomic regions associated with target traits, expediting the discovery of linked molecular markers and causal genes. This combined approach has proven effective in other crop species and is poised to drive trait mapping, candidate gene identification, and genomics-assisted breeding in safflower [2, 6]. The sequencing of an entire genome and its ensuing comparison to a wild-type reference genome can potentially directly pinpoint the molecular lesions that result in the mutant phenotype that organism has been selected for [7]. studied on domestication-related traits in safflower, and found only two traits (spininess and flower color) had "major" QTL (i.e., phenotypic variation explained (PVE) > 25%). Phylotranscriptomic and comparative genomic analyses of Asteraceae confirm that morphological traits like spininess are deeply intertwined with the family's evolutionary dynamics and have frequently arisen in tandem with whole-genome duplications and lineage diversification [8–10].

It has been suggested that the improvement of minor crops such as safflower could help to provide food security in the face of climate change and an ever-increasing global population [11]. The spininess trait primarily manifests as sharp, rigid spines along the leaf and in bract margins and has important implications for crop management, mechanization and grower preference. While spiny types may provide protection from herbivory, spineless types are generally favored in breeding due to ease of harvesting and improved worker safety. The manifestation of spininess displays notable variability among safflower accessions, ranging from fully spiny to semi-spiny and non-spiny phenotypes. This trait is highly heritable and was among the earliest targets in safflower improvement programs, motivated by the demands of farmers and processors for more amenable plant types. Early breeding and genetic surveys established that spininess and related traits exhibit clear segregation patterns in biparental crosses, often indicative of simple Mendelian inheritance, though in some backgrounds, more complex or linked inheritance patterns have also been observed [12] [13]. studied on inheritance mode and the number of genes that control flower color and leaf spininess in some Iranian safflower genotypes. In all crosses of spiny × spineless genotypes, the phenotypes of F_1 leaves and bracts were spiny. Therefore, existence of spines on leaves was considered dominant over the spineless type. Development of next-generation sequencing (NGS) based markers, including simple sequence repeats (SSR) and single nucleotide polymorphisms (SNP), has enabled high-resolution mapping and association studies. Genome-wide scans and the construction of dense genetic maps have facilitated trait mapping for a variety of agronomic features, but explicit mapping of the spininess trait remains relatively limited in the published literature.

EMS-based mapping works on a similar yet different principle where instead of using polymorphisms between two different strains to detect a genomic region linked to the causal natural mutation, the nucleotide changes caused by the EMS mutagenesis itself are used [14]. EMS has been successfully applied to a large number of plants to produce morphological diversity and promote the improvement of ideal traits [15]. Based on the mutation of agronomic traits, new materials with excellent quality can be screened from the mutated offspring after

multiple generations of screening and identification. All the mutant varieties are improved for increased yield and yield components, and resistance or tolerance to biotic and abiotic stresses. A large majority of released mutant varieties have traits suitable to face the challenges of climate change and have contributed positively to sustaining crop yields, resulting in positive economic impacts [16].

Here, we present a study employing WGS and QTL-seq to pinpoint genomic regions underpinning the studied trait of interest, spininess, in *C. tinctorius*. Our application of state-of-the-art genomic resources and quantitative mapping frameworks is designed to advance the understanding of safflower's genetic architecture and contribute foundational knowledge and tools for the next generation of molecular breeding efforts in this versatile crop. From an agricultural perspective, the reduction or elimination of spininess via breeding is a prime target in several Asteraceae crops, including safflower and globe artichoke, to facilitate mechanized harvest and improve worker safety [17]. In summary, genes controlling the spininess trait in safflower remains a focal point for both fundamental research and practical breeding, with recent genomic advances providing the foundational resources required to finally unveil its genetic underpinnings, expedite marker-assisted selection, and produce commercially favorable, spineless cultivars for modern agriculture [18]. Nonetheless, identification of the precise genes or regulatory networks directly mediating spine formation remains unresolved, although ongoing genome assemblies and comparative analyses in Asteraceae crops and their wild relatives are expected to close this gap [19].

Materials and methods

Sample Preparation

Safflower cv. Arak 2811 seeds (spiny wild type) were collected from the oilseed section of the Seed and Plant Improvement Institute (SPII) (Alborz, Iran). The spineless mutant line was developed by mutagenesis of Arak 2811 using EMS [20]. Seeds of M_8 mutant of safflower were surface-sterilized using sodium hypochlorite (1%) for 5 min and washed several times with dH_2O. Uniformly sterilized seeds were planted in 800 gr dark pots (with a diameter of 20 cm) with sterilized soil and free of weed seeds in the research greenhouse of the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany at ambient temperature (25 ± 2 °C) and dark/light regime of 16 h/8 h. In each pot, 10 healthy seeds were grown (in total, 100 plants were established in greenhouse). The mutant is spineless, drought-tolerant and with earlier flowering relative to its wild-type counterpart cv. Arak 2811 [21]. Thus, it can be supposed the detectability of candidate genes for trait of interest via integrative WGS and RNA-seq analysis. After normal growth up to flowering and emerging flower heads, the mutant plants were segregated to spineless and spiny types due to residual heterozygosity (Fig. 1). Leaf bulk samples (15 samples in each spiny and spineless bulk) were collected for both types, frozen in liquid nitrogen, powdered and used for DNA extraction. The MagAttract high molecular weight (HMW) DNA Kit was used for

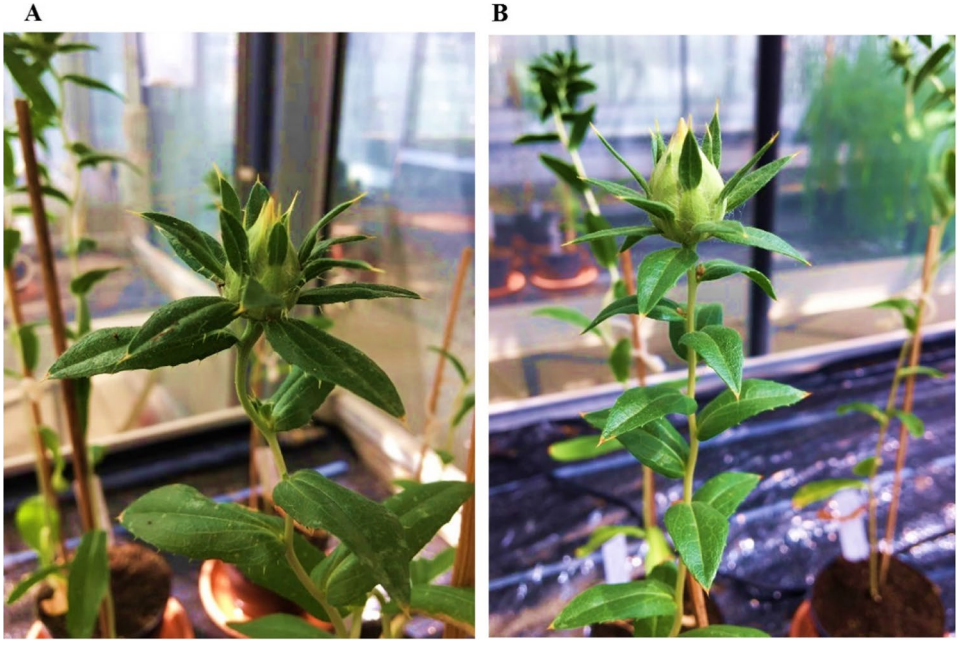


Fig. 1 Spiny wild type (**A**) and spineless mutant (**B**) genotypes cultivated in greenhouse

HMW DNA extraction. Illumina NovaSeq 6000 used for sequencing whole genome of the two bulked samples (https://www.illumina.com/systems/sequencing-platforms/novaseq.html). Short read lengths, standardly 2 × 250 base pairs (paired-end), yield highly accurate sequence data, enabling robust read mapping, variant detection, and precise quantification of expression or allele frequency [22].

Quality control and trimming

The output of sequencing was obtained as short reads in fastq format. In first step, the quality of sequences was assessed using FASTQC software [23] for parameters including per-base sequencing quality, per-sequence quality scores, per-base sequence content, per-sequence guanine and cytosine (GC) content, adaptor contamination, among others. After ensuring the quality of sequences, the low-quality reads (Q30 < 20), reads with adaptor contamination, and duplicated reads were filtered using Trimmomatic software (MIN-LENGH = 35) [24].

Mapping reads to the reference genome

In this study, chromosome-level reference genome "Anhui-1", a spineless mutant with high-LA (linoleic-acid) content, obtained from (http://safflower.scuec.edu.cn) was used as reference genome in alignment step. Then, the cleaned reads were aligned to the reference genome using maximum exact match (MEM) algorithm in Burrows-Wheeler Aligner (BWA) software [25]. Next, to simplify use of the alignment output file, the file size was reduced by converting the format from sequence alignment map (SAM) to binary alignment equivalent (BAM) through SAMtools software [26]. Afterward, duplicated reads were identified and removed by Mark-Duplicates tool in Picard toolkit (http://broadinstitute.github.io/picard).

Variant calling

The mapped reads were used to identify structural variants including single nucleotide polymorphism (SNPs) and InDels (insertion/deletions) via local *de novo* assembly of haplotypes using haplotype-caller algorithm in genome analyzer toolkit (GATK) software [27]. GATK provides a robust framework of command-line tools and best practice workflows that address a spectrum of tasks essential for turning raw sequencing reads into high-confidence variant calls.

GBS data analysis

Ninety-three individuals from GBS data of cultivated safflower (bio-project number PRJNA865057; [28]) were used for comparison to our WGS mutant data. We wanted to compare the mutation type induced by EMS to natural mutation collected in safflower during evolution. Similar to WGS, each GBS paired-end short read accession was aligned against the indexed reference genome using Hisat2 and the sequence variants were called using GATK software. Natural mutation (especially, G > A and C > T) rates in cultivated genotypes were calculated and reported. A chi-square test was conducted for examining the difference between mutation type rates in EMS-WGS and natural-GBS data using "base" package in R.

SNP annotation and filtering

The list of annotated genes (in gff3 format) for safflower were downloaded from safflower database (http://safflower.scuec.edu.cn). After variant calling is performed, SnpEff plays a central downstream role by classifying and interpreting these variants based on their position and predicted impact within annotated genome features [29]. Genomic location, hetero- or homozygosity, and other details of polymorphisms were extracted. The outputs were transferred to excel environment, and SNPs were queued based on their genomic location for haplotype form of both re-sequenced bulks and reference genome. It is notable that the reference genome (Anhui-1) has spineless morphology, so called SNP variants in our spineless mutant sample should be the same with reference genome. On the other hand, SNP variants in our control bulk (wild / spiny) should be different from both spineless mutant and reference genome. So, firstly the heterozygote variants and variants that were different between samples with same phenotype were filtered. Then, the variants that were identical between samples with different phenotype were omitted. Finally, the remained SNPs were considered as true variants (they were different between samples with different phenotype and identical between the samples with the same phenotype). These SNPs are supposed to likely mark the genes involved in loss of spine formation phenotype / drought-tolerance in spineless mutant via QTL-seq analysis.

QTL-seq analysis

QTL-seq is a method developed in recent years that uses high-throughput sequencing technology for cluster segregation analysis (BSA) to locate the main genes of quantitative traits or genes of qualitative traits [30]. QTL-seqr is an R package that performs BSA using NGS data [31]. In this method for each SNP, the total reference allele frequency, per bulk SNP-index, and Δ(SNP-index) are calculated. SNP index (ΔSNP) refers to the subtraction of the alternate allele frequency value of the low bulk from the high bulk [30].

Local mapping of chr7 QTL (qSpiny7)

Some features of qSpiny7 were assessed using IRanges package in R including physical position, number of

SNPs, type of SNPs, number of genes which had overlap with SNP positions. The physical maps were drawn using rMVP package in R.

Validation of identified candidate genes using RNAseq and real-time qRT-PCR

The RNASeq mega data of two contrasting genotypes viz. spiny wild-type Arak 2811 and its EMS-induced drought tolerant spineless mutant was obtained from short read archive (SRA) at NCBI under project PRJNA833048. All plants were raised to flowering stage under no-stress condition, and flower leaves (bracts) were collected for RNA extraction. The results of mapping short reads of two genotypes to the reference genome and differentially expressed genes (DEGs) between two genotypes were previously reported by us [21]. The co-location of the DEGs with the identified QTLs on different chromosomes is reported in "Results" section.

To confirm the RNAseq results, we selected 2 up-regulated and 4 down-regulated genes and examined their expression patterns in real-time qRT-PCR using specific primer pairs designed by Primer 3.0 Plus online tool (https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/) (supplementary Table S1). The relative expression values were calculated using REST software.

Results

WGS Raw data statistics

WGS of two contrasting bulks generated 106.78 Gb raw data for spiny (control) and spineless (mutant) bulks (Table 1). The megadata were deposited in NCBI short read archive (SRA) under project PRJNA872295 (accession numbers of SRX17239309-312).

SNP annotation report

After alignment of the re-sequenced genomes with the reference genome by BWA, more than 3 million SNPs between two opposite bulks (spiny and spineless) were detected. After filtering heterozygote SNPs in both samples, 893,975 homozygote SNPs retained. SNP distribution across safflower chromosomes is demonstrated in Fig. 2A. In the genome with size of 1.05 Gbp, variation rate was 1,182 bp/SNP. It is well known that EMS causes G: C to A: T transitions induced by guanine alkylation [32]. Several studies [6, 33, 34] showed that EMS-induced canonical transitions (G: C > A: T) has a high frequency (38.03%) among other types of substitutions (Fig. 3A). Transition/ transversion (Ts/ Tv) ratio > 1.9 in our results approves this phenomenon (Fig. 3B). Moreover, SNP distribution over safflower chromosomes and genome is shown in Table 2; Fig. 3C-D, respectively. As seen, SNP frequency rate over safflower chromosomes was 1,182 bp/snp. Intergenic regions had highest mutation frequency and 3'UTR and 5'UTR had lowest mutation frequency. Missense (47.08%), nonsense (1.03%) and silent (51.88%) mutation rate showed the effectiveness of EMS mutagenesis. EMS typically induces G: C to A: T transition mutations, resulting in a spectrum of all three mutation types and a corresponding range of functional consequences depending on the gene context and specific molecular change.

Table 1 Whole genome sequencing (WGS) Raw data statistics in safflower

Bulks	Number of reads	Number of bases	GC (%)	Sequence length (bp)	Sequencing depth (mean)
Spiny (wild type)	109,384,028	51.7 G	39	35–251	19.88 X
Spineless (mutant)	115,598,848	55.0 G	39	35–251	21.15 X
Total	224,982,876	106.7 G	39	35–251	20.51 X

Comparison of EMS-induced mutation types to natural mutation detected by GBS assay

In GBS study, 20,755 high-quality SNPs were detected in cultivated genotypes of safflower. The density map of GBS-derived SNPs is shown in Fig. 2B. The rate of canonical SNPs was 32.41% in cultivated safflowers (Table 3). Meanwhile, the rate of canonical SNPs in EMS-induced WGS data reached 38.03%. A chi-square test showed that the EMS-induced mutation types were significantly higher than natural mutation types inherited during evolution of safflower.

QTL-seq analysis

893,975 SNPs that could be tentatively assumed as being correlated with spine formation in safflower were used as entry to the QTL-seq analysis. After filtering for reference allele frequency ($0.05 \le REF\_FRQ \le 0.95$), and total sample read depth ($minTotalDepth \ge 65$), 225,819 and 137,387 SNPs, respectively, retained at $GQ \ge 95\%$ and $GQ \ge 99\%$ (Table 4). QTL-seq analysis for all SNPs at the two different genotype qualities ($GQ \ge 95\%$ and $GQ \ge 99\%$) was run through QTL-seqr package in R. In $GQ \ge 95\%$, G' method showed signals for 3 QTLs on chromosomes 3, 7 and 12 at q-value ≤ 0.01 (Fig. 4A-B). Detailed information of these QTLs (hereafter, GQ95 QTLs) is shown in Tables 5 and supplementary TableS2. Moreover, for $GQ \ge 99\%$ at q-value ≤ 0.01, with G' method chromosomes 1, 3, 6, 7 and 12 showed 7 association signals for spininess (hereafter, GQ99 QTLs). In this case, four QTLs on chromosomes 1 and 12 (each 2), and three on chromosomes 3, 6 and 7 were detected (Table 5; Fig. 4C-D). However, on the basis of Δ(SNP-index), only the QTL on Chr7 (hereafter, qSpiny7) was significant at both GQ levels. Details on every feature of these 7 QTLs were brought in supplementary Table S3.

Fig. 2 Distribution pattern of SNPs on 12 safflower chromosomes. (**A**) Density plot of > 137,000 SNPs in 1 Mb window size after filtering which were used for marker-trait association (MTA) mapping. (**B**) Density plot of GBS-derived SNPs (> 20,000 SNPs) in 1 Mb window size
(**A**)

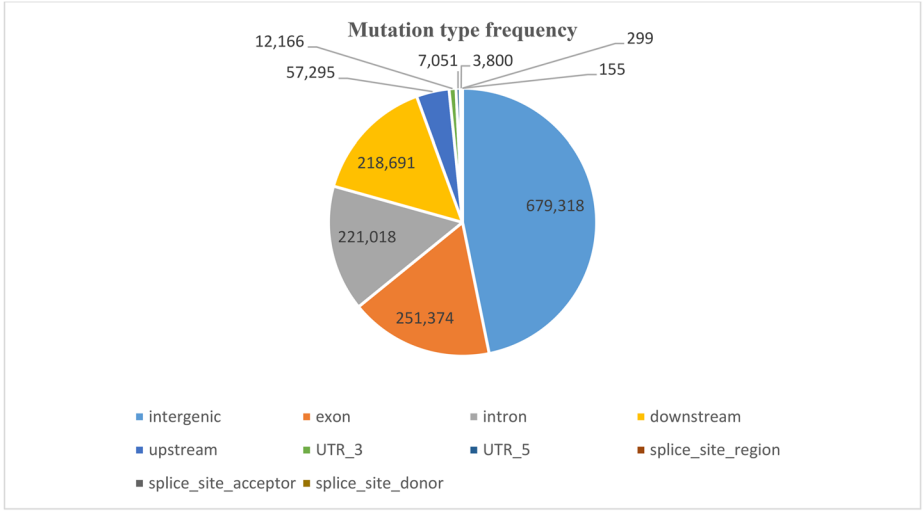
(A)

	>A	>C	>G	>T
>A	0	32,714	126,283	46,793
>C	41,949	0	29,482	170,061
>G	169,915	29,255	0	41,561
>T	46,723	126,358	32,881	0

(B)

Mutation	Count	%
Transition	592,617	66.30
Transversion	301,358	33.70

(C)



(D)

Mutation	Count	%
Missense	27,185	47.077
Nonsense	599	1.037
Silent	29,962	51.886

Fig. 3 Base substitution rate (**A**) and Transition/ Transversion (Tr/ Tv) ratio among SNPs (**B**), SNP distribution over genome (**C**) and mutation type frequency among SNPs (**D**)

Genomic annotation of qSpiny7

As mentioned earlier, the only highly significant QTL in QTL-seq analysis was qSpiny7. The physical location of this QTL is 8,283 – 15,207,861 bp on short arm of chromosome 7. 3,218 SNPs were identified in this region, 1,060 of them are canonical SNPs (G: C > A:T). Among 1,008 annotated genes in this QTL region, 92 genes had overlap with 569 SNPs, from which 418 were intronic and 151 were exonic SNPs. The exonic SNPs were evaluated in view point of the mutation type (including missense, nonsense, stop-codon, start-codon, intron-exon junction), and it showed that these base substitutions occurred in 24 genes (Table 6), in which 3 start codon gain (introducing Met codon) and 37 missense mutations (amino acid substitution) were found. A gene involved in cell division and different organ`s enlargement (*Ct07T0035400*; *BG1-like E*) at position 4,446,965 bp along with a LONELY GUY gene (*CtLOG: Ct07G0075400*) at position 10,636,511 – 10,640,422 bp, both locate within qSpiny7 on chromosome 7. Based on SnpEff annotation results, a SNP (T > C) was located in the first intron of *CtLOG* gene.

Mapping DEGs to the detected QTLs

RNA-seq data of the drought tolerant spineless mutant and the spiny wild type (cv. Arak 2811) obtained from bracts at flowering stage (all plants cultivated under no-stress condition) was used for investigating the relative

Table 2 SNP distribution in safflower genome based on WGS

Chromosome	Length (bp)	Initial SNPs	Filtered SNPs	SNP rate (bp/1 SNP)
Ct01	88,212,955	57,801	9,746	1,526
Ct02	88,221,803	69,803	11,481	1,263
Ct03	106,782,788	79,752	16,328	1,338
Ct04	92,479,977	71,368	12,602	1,295
Ct05	97,127,944	166,313	16,423	584
Ct06	93,909,653	75,082	11,372	1,250
Ct07	73,178,102	54,336	10,900	1,346
Ct08	87,883,749	73,789	10,255	1,191
Ct09	87,195,570	63,148	9,993	1,380
Ct10	84,812,270	72,935	8,965	1,162
Ct11	74,724,136	39,448	7,824	1,894
Ct12	82,586,919	70,200	11,498	1,176
Total	**1,057,115,866**	**893,975**	**137,387**	**1,182**

Table 3 The rate of canonical SNPs in the WGS EMS-induced mutant and natural GBS panel of safflower

Mutation type	WGS-based induced mutations		GBS-based natural mutations	
	count	%	count	%
G > A	169,915	19.01	3,338	16.08
C > T	170,061	19.02	3,389	16.33
Sum of canonical mutations	339,976	38.03	6,727	32.41
Other types	553,999	61.97	14,028	67.59
Difference between observed (EMS-induced) and expected (natural) mutations	χ-squared = 12,888**, p-value < 2.2e-16			

**: significant difference at 0.01 level of probability

Table 4 SNP annotation report obtained using SnpEff for different levels of filtering (including GQ95 and GQ99)

#	hetero in both	hetero in mutant	total SNPs (GQ95)	total SNPs (GQ99)
Original SNPs	582,738	700,693	893,968	893,968
Filtered SNPs	435,118	529,251	668,149	756,581
Remained SNPs	147,620	171,442	225,819	137,387

expression of the genes locating within QTLs. Thirty DEGs were mapped to the GQ95 QTLs. 11 and 19 out of them showed up- and down-regulation in spineless mutant, respectively (supplementary Table S4). However, most of DEGs ($n = 19$) were mapped to qSpiny7, that 7 and 12 of them were up- and down-regulated in spineless mutant, respectively. In the case of GQ99 QTLs, 52 DEGs were mapped to different QTLs, that 22 and 30 of them were up- and down-regulated in spineless mutant, respectively. From these 52 DEGs, 22 DEGs were mapped to qSpiny7, that 8 and 14 of them showed up- and down-regulation in spineless mutant, respectively.

Among 22 DEGs mapped to qSpiny7 QTL, eight of them which were up-regulated in spineless mutant include *Ct07T0009500* (protein phosphatase 2 A subunit A2), *Ct07T0020100* (nitrate transporter), *Ct07T0024800* (ribonuclease), *Ct07T0026100* (NAC-like activated by AP3/PI), *Ct07T0057200* (6-phosphogluconate dehydrogenase family protein), *Ct07T0058100* (phosphatidylinositol 3- and 4-kinase family protein), *Ct07T0073400* (transcription factor bHLH75-like) and *Ct07T0091400* (uncharacterized). However, it must be noticed that spine formation may be not in the result of altered expression of genes within the detected QTLs; it may be in the result of altered protein structure/conformation due to a mutation event introduced via EMS mutagenesis.

Validation of relative expression pattern of selected genes

Based on RNAseq assay, in the qSpiny7 genomic region 2 and 4 genes which showed up- and down-regulation in spineless mutant were selected for confirming the relative expression pattern (supplementary Table S1). Real-time qRT-PCR tests showed that the 2 up-regulated genes (*bHLH75-like* and *NAC-like*) in spineless mutant were expressed 3.4 to 4.4 times higher than spiny wild type; similarly, the 4 down-regulated genes (*MLP-like 28*, *BG1-like E*, *UGT73C1* and *SAM methyltransferase*) in spineless mutant were expressed 1.95 to 4.95 times lower than spiny wild type (Fig. 5). This examination well confirmed the results of RNAseq assay.

Discussion

QTL mapping using dense SNPs potentially guided by expression or association data was increasingly applied to dissect the genetic basis of spininess and similar epidermal traits in both wild and cultivated species [4, 40]. For instance, mapping of trichome density and length QTLs in wheat has allowed researchers to filter for candidate genes with annotated roles in cell proliferation, hormone signaling, and cell wall modification within QTL intervals [40]. Such paradigms are readily applicable to identify spine-related candidate genes in Asteraceae family (including safflower) once relevant segregating populations and high-resolution maps are available.

Over 3 million EMS-induced SNPs were discovered between two opposite bulks (spiny and spineless) in this research. After filtering, 893,975 SNPs remained between two bulks that were used in downstream analysis. The efficiency of chemical mutagenesis is variable, depending not only on the nature and dosage of the mutagen, but also on the considered species and the background genotype [41]. Currently, EMS mutagenesis can achieve a large number of mutants in a short period of time, which can facilitate the study of quantitative traits in plants [6]. According to SnpEff report, prevailing EMS-induced mutations were C > T and G > A (38.03%) that were found

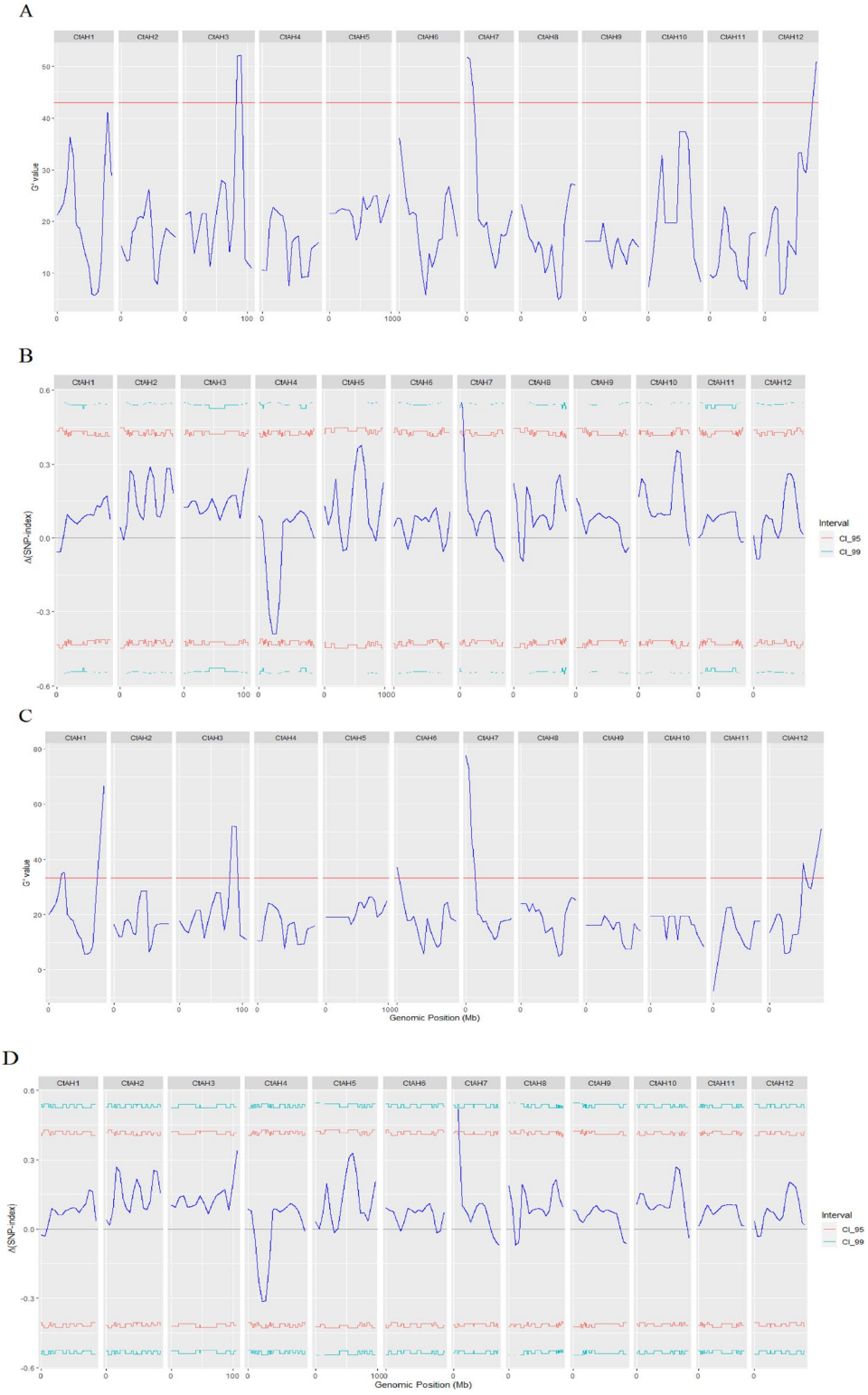


Fig. 4 Genomic regions with significant association (at q value ≤ 0.01) to spininess detected using QTL-seq analysis. (**A**) GQ ≥ 95% using G′ method. (**B**) GQ ≥ 95% using Δ(SNP-index) method. (**C**) GQ ≥ 99% using G′ method. (**D**) GQ ≥ 99% using Δ(SNP-index) method. The setup filterations were refAlleleFreq = 0.05, minTotalDepth = 50, maxTotalDepth = 180, minSampleDepth = 20

Table 5 Results of QTLseq analysis using G' and Δ(SNP-index) approaches at different levels of filtering on SNPs (GQ95 and GQ99) at q-value < 0.01

Filter level	QTL count by G' method	Initial QTLs	QTLs approved by Δ(SNP-index)
GQ95	3	qSpiny3, qSpiny7, qSpiny12	qSpiny7
GQ99	7	qSpiny1.1, qSpiny1.2, qSpiny3, qSpiny6, qSpiny7, qSpiny12.1, qSpiny12.2	qSpiny7

to be randomly distributed along 12 chromosomes of safflower (Fig. 2A). Most of the mutations affected the intergenic regions (46.81%) but 3.94% were located in exome which could have a high putative impact on gene function. The occurrence and proportion of canonical mutations is consistent with findings from other studies in *Hordeum vulgare* [42], *Drosophila melanogaster* [43, 44], *Solanum lycopersicum* [45, 46], and *Oryza sativa* [34], suggesting that although the precise mechanisms remain unclear, these mutations are also attributable to EMS. This indicates that EMS-induced mutagenesis generates reproducible results across species, which enhances confidence in using this method for functional genomics and breeding. This strengthens EMS's value as a tool in forward genetic screens and crop improvement studies by generating predictable and traceable mutations. A comparison of the G: C > A:T mutation rate between EMS-induced mutant and natural safflower samples (analyzed via WGS and GBS, respectively) revealed a significant difference (38.03% vs. 32.41%; χ^2 = 12,909, p < 2.2e-16; Table 3). This difference indicates that EMS not only increases mutation load but also skews the mutation spectrum, favoring transitions that are characteristic of its mode of action. Nonsense, missense and silent mutations caused by EMS mutagen can have distinct biological implications. Nonsense mutations introduce premature stop codons that truncate protein synthesis, often leading to nonfunctional proteins or triggering nonsense-mediated mRNA decay, which reduces the abundance of the corresponding transcripts and can result in loss-of-function phenotypes [35, 36]. Missense mutations alter the encoded amino acid, potentially impairing protein function or structure, and depending on the context, may have variable phenotypic effects; certain missense mutations, such as dominant negative mutations in crucial genes, can drastically affect biological processes and disease susceptibility [37, 38]. Silent mutations do not change the encoded amino acid but can nevertheless impact gene expression or function by altering mRNA splicing, stability, or translation efficiency, particularly if they disrupt exonic splicing enhancers or other regulatory motifs [39].

Spines and prickles are often regarded as modified trichomes or resulted from similar developmental pathways. Thus, genes governing trichome initiation, patterning and differentiation are logical candidates for the control of spininess [47]. Early QTL-based studies suggested that domestication traits were predominantly controlled by a small number of large-effect QTLs [48, 49]. Genetic analysis of domestication traits in safflower [7] found two large-effect QTLs corresponded to flower color and leaf spininess. These researchers constructed a genetic map of safflower genome along with the corresponding QTLs on 12 linkage groups (A to L). The spine size and the number of days to flowering were placed on the linkage groups E, H and L, and D, H and I, respectively, where L and D had the highest amount of PVE for spine size. At that time (2014), it wasn't clear these linkage groups were which chromosomes. Hence, we conducted a homology analysis between the probes used in abovementioned study and the newly released genome sequence of the safflower, and we found the two linkage groups L and D are equivalent to chromosomes number 7 and 4, respectively. In the present research, one QTL on short arm of chromosome 7 of safflower (qSpiny7) passed the significance thresholds in both G' and Δ(SNP-index) methods and can be a good candidate for phenotypic and molecular differences between two contrasting genotypes.

Exonic SNPs were evaluated in terms of nucleotide substitution (nt > nt) and amino acid change (aa > aa). 24 annotated genes carrying 151 exonic SNPs were identified which within their exons, 3 start codon gain mutations and 37 missense mutations (amino acid substitutions) were discovered (Table 6). The genes harboring SNPs with start-codon-gain effect include *Ct07G0008200* (MAK16 protein-like protein), *Ct07G0011600* (Cysteine-rich RLK (receptor-like protein kinase 2)) and *Ct07G0012200* (protein kinase 1B). Protein kinases constitute one of the largest gene families in plant genomes. These enzymes catalyze the reversible phosphorylation of specific amino acids (serine, threonine, and tyrosine) to regulate the activity of their target proteins [50]. In the vast majority of cases, the mutant phenotype is due to premature stop codons, splice site disruption, deleterious amino acid substitutions, and loss of gene function [51]. Therefore, it is critical to ensure that a large quantity of mutant alleles is present in a small mutant population and that the requirements for efficient homogenization are adequately met [6]. The effect of each point mutation on protein structure of each protein should be assessed through complementary functional analyses such as gene transfer, gene knockout or gene editing. It is also possible to trace a series of mutations and extract the related phenotypic changes.

Twenty four overlapped genes on qSpiny7 are brought in Table 6, eight of them had SNPs both in exonic and promoter regions including *Ct07G0009500* (protein

Table 6 Annotation of exonic SNPs in terms of mutation type within qSpiny7. The genes showing significant up-regulation in spineless mutant are marked with asterisk (*)

Gene	Gene description	SNP position (bp)	nt > nt	Mutation type	aa > aa
Ct07G0008200	MAK16 protein-like protein	895,612	AcG > ATG	Start codon gain	. > Met
		898,042	A > G	Missense	Ile93Val
		903,751	A > G	Missense	Asn242Asp
Ct07G0011600	Cysteine-rich RLK (RECEPTOR-like protein kinase) 2	1,382,393	AcG > ATG	Start codon gain	. > Met
Ct07G0012200	protein kinase 1B	1,481,788	ATa > ATG	Start codon gain	. > Met
Ct07G0003900	AMP-dependent synthetase and ligase family protein	469,144	A > G	Missense	Val272Ala
		469,279	C > T	Missense	Gly227Asp
Ct07G0006600	COP1-interacting protein-like protein	756,027	G > A	Missense	Pro618Ser
		756,828	A > G	Missense	Lys885Glu
		757,056	A > G	Missense	Thr961Ala
		757,234	C > T	Missense	Thr1020Met
Ct07G0007200	phosphoribosyl pyrophosphate (PRPP) synthase 3	806,626	G > A	Missense	Gly37Ser
Ct07G0007900	Nucleotide-diphospho-sugar transferase family protein	865,472	C > T	Missense	Ser240Phe
		866,714	C > T	Missense	Thr462Ile
Ct07G0008500	DORNROSCHEN-like protein; ethylene-responsive transcription factor ESR2-like	940,101	C > T	Missense	Ser163Leu
Ct07G0008800	hypothetical protein E3N88_37334	971,484	C > T	Missense	His203Tyr
Ct07G0008900	hypothetical protein LSAT_8 × 153,060	987,527	A > G	Missense	Lys98Arg
Ct07G0009000	hypothetical protein LSAT_8 × 153,060	995,300	C > T	Missense	His135Tyr
		995,363	A > G	Missense	Ile156Val
Ct07G0009200	lysine ketoglutarate reductase trans-splicing protein (DUF707)	998,596	T > C	Missense	.Met372Thr
		1,000,276	C > T	Missense	Val112Ile
Ct07G0009300	Protein phosphatase 2 C family protein	1,017,346	C > T	Missense	Met340Ile
*Ct07G0009500**	protein phosphatase 2 A subunit A2	1,040,305	A > G	Missense	Val1129Ala
		1,050,337	A > G	Missense	Leu492Pro
		1,057,572	C > T	Missense	Val608Ile
Ct07G0009600	GDSL-like Lipase/Acylhydrolase superfamily protein	1,058,405	C > T	Missense	Asp470Asn
		1,068,106	C > T	Missense	Gly272Ser
Ct07G0009700	eukaryotic translation initiation factor 4B1	1,070,824	A > G	Missense	Leu163Ser
Ct07G0009800	Melibiase family protein	1,079,311	A > G	Missense	Met138Thr
		1,082,961	A > G	Missense	.Phe469Leu
Ct07G0012000	homeobox protein knotted-1-like 6	1,444,846	A > G	Missense	Phe196Leu
		1,445,673	C > T	Missense	Gly90Asp
Ct07G0012100	pantothenate kinase 2	1,463,861	C > T	Missense	Gly137Ser
Ct07G0031400	NA	3,861,775	A > G	Missense	Phe9Ser
Ct07G0060800	NB-ARC domain-containing disease resistance protein	8,270,124	T > C	Missense	Met763Thr
		8,270,132	A > G	Missense	Ile766Val
Ct07G0086500	calponin-like domain protein	12,653,264	G > A	Missense	Ala5Thr
		12,655,664	G > A	Missense	Val805Ile
*Ct07G0091400**	NA	13,387,876	G > A	Missense	Ala133Thr
Ct07G0092000	NA	13,537,835	G > A	Missense	Cys82Tyr

phosphatase 2 A subunit A2), *Ct07G0009600* (GDSL-like Lipase / Acylhydrolase superfamily protein), *Ct07G0012000* (homeobox protein knotted-1-like 6), *Ct07G0006600* (COP1-interacting protein-like protein), *Ct07G0008500* (DORNROSCHEN-like protein; ethylene-responsive transcription factor ESR2-like), *Ct07G0008800* (hypothetical protein E3N88_37334), *Ct07G0008900* (hypothetical protein LSAT_8 × 153060) and *Ct07G0009300* (protein phosphatase 2 C family protein). These genes can be a putative candidate for spininess.

We examined relative expression of four down-regulated genes (*MLP-like 28*, *BG 1-like E*, *UGT73C1* and *SAM-dependent methyltransferase*) in spineless mutant. Real-time results showed that the relative expression of all 4 genes was decreased in the spineless mutant compared to spiny wild type (Fig. 5). Among these genes, *BIG GRAIN 1* (*BG1*), found in rice and other crops, acts as

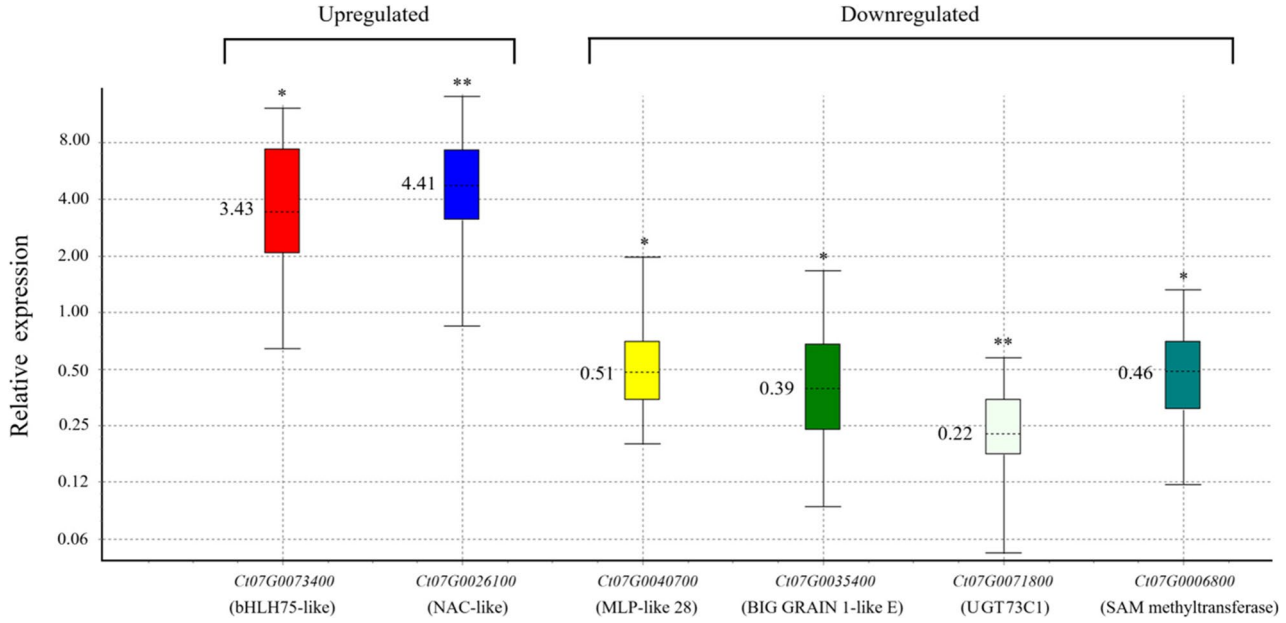


Fig. 5 Confirming the relative expression pattern of six RNAseq-based selected genes using real-time qRT-PCR. The 2 up-regulated genes were expressed in spineless mutant 3.4–4.4 times higher than spiny wild type, and the 4 dow-regulated genes were expressed in spineless mutant 1.95 to 4.95 times lower than spiny wild type. * and ** indicate significant differences at $P < 0.05$ and $P < 0.01$, respectively

a positive regulator of plant growth and grain development via auxin transport and signaling [52]. It promotes cell division not only in grains but also in roots, shoots and panicles, and its suppression or knockdown leads to reduced grain size and plant stature [53].

In a recent study on QTL mapping of trichome traits and analysis of candidate genes in leaves of wheat [40], 150 and 285 genes, respectively, were found in a 15.1 cM and a 18.8 cM intervals on chromosomes 7D and 3 A, and based on WheatOmics five genes were highly expressed in leaves, suggesting that these genes may have the function of regulating leaf trichome traits. Hormonal regulation via auxin and gibberellin pathways is also intimately involved in specifying spine and trichome fate, as supported by QTL and functional studies in several taxa [54]. Genes involved in localized auxin biosynthesis (e.g., YUCCA family), polar auxin transport (e.g., PIN1), and GA-mediated growth regulation frequently emerge as candidate loci in QTL intervals controlling spiny structures. More recently [55], reported the functionality of LONELY GUY (LOG) in spine/prickle formation in some crop plants and their wild relatives such as eggplant and foraged berry. The enzyme encoded by the LOG gene catalyzes the final step of cytokinin hormone biosynthesis which is involved in plant cell proliferation and differentiation [56]. In present research, a LOG gene (*CtLOG: Ct07G0075400*) is located within qSpiny7 at position 10.64 Mbp. We locate a SNP (T > C) within the first intron of *CtLOG* gene. Understanding the structure and function of these candidate genes is crucial for unraveling the complex regulatory networks that govern gene expression and protein function.

Conclusion

In Asteraceae, and particularly in safflower (*Carthamus tinctorius*), recent advances in whole-genome sequencing, high-density SNP mapping, and GWAS / QTL-seq techniques have set the stage for systematic candidate gene discovery for spininess and related morphological innovations. This study unveils thousands of SNPs introduced by chemical mutagen EMS in safflower genome as revealed by WGS. The trait-associated SNPs were detected by QTL-seq analysis which further were characterized by subsequent analyses. A major effect genomic region on short arm of chromosome 7 (qSpiny7) was detected which harbors important genes for developmental processes and stress responses. Subsequent functional assays (CRISPR-Cas9 mutagenesis, gene silencing, or overexpression in model species) will be essential for confirming candidate gene function. In conclusion, while direct functional validation of spininess-related genes in Asteraceae crops like safflower is ongoing, current evidence strongly highlights trichome regulatory genes as prime candidates. Their involvement in other plant families, coupled with the new genome-wide association, QTL, and transcriptome resources which are now available for key members of Asteraceae, will facilitate their rapid identification and exploitation for marker-assisted breeding of spine formation.

Abbreviations

BSA	Bulked segregant analysis
EMS	Ethyl methanesulfonate
GWAS	Genome-wide association study
QTL-seq	Quantitative trait loci sequencing
WGS	Whole genome sequencing

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-12488-8.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Acknowledgements
We acknowledge the help of Dr. Frank Blattner at IPK, Germany for providing instrument for whole genome sequencing and RNA-seq.

Authors' contributions
Samaneh Karami-Moalem performed experiments, conducted formal analyses and prepared the draft of the manuscript. Asadollah Ahmadikhah supervised the research, provided all samples used in the project and contributed in all analyses, Zahra Nemati contributed in formal analyses, Reza Haghi provided resources and contributed in formal analyses. All authors read and approved the final manuscript.

Funding
Not applicable.

Data availability
All relevant data have been provided as Tables, Figures with in the text and in the following supplementary data. The megadata were deposited in NCBI short read archive (SRA) under project PRJNA872295.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that the research was conducted in the absence of any commercial or financial relationships and that they have no conflict of interest.

Received: 10 June 2025 / Accepted: 23 December 2025
Published online: 30 December 2025

References

1. Wu Z, Liu H, Zhan W, Yu Z, Qin E, Liu S, et al. The chromosome-scale reference genome of safflower (*Carthamus tinctorius*) provides insights into Linoleic acid and flavonoid biosynthesis. Plant Biotechnol J. 2021;19(9):1725–42.
2. Ambreen H, Kumar S, Kumar A, Agarwal M, Jagannath A, Goel S. Association mapping for important agronomic traits in safflower (*Carthamus tinctorius* L.) core collection using microsatellite markers. Front Plant Sci. 2018;9:402.
3. Bhat R, Reddy KR. Challenges and issues concerning Mycotoxins contamination in oil seeds and their edible oils: updates from last decade. Food Chem. 2017;215:425–37.
4. Bowers JE, Pearl SA, Burke JM. Genetic mapping of millions of SNPs in safflower (Carthamus tinctorius L.) via whole-genome resequencing. G3: Genes, Genomes, Genet. 2016;6(7):2203-11.
5. Wu ZH, Liao R, Dong X, Qin R, Liu H. Complete Chloroplast genome sequence of Carthamus tinctorius L. from PacBio sequel platform. Mitochondrial DNA Part B. 2019;4(2):2635–6.
6. Chen L, Duan L, Sun M, Yang Z, Li H, Hu K, et al. Current trends and insights on EMS mutagenesis application to studies on plant abiotic stress tolerance and development. Front Plant Sci. 2023;13:1052569.
7. Pearl SA, Bowers JE, Reyes-Chin-Wo S, Michelmore RW, Burke JM. Genetic analysis of safflower domestication. BMC Plant Biol. 2014;14(1):43.
8. Karis PO. Morphological data indicates two major clades of the subtribe gorteriinae (Asteraceae-Arctotideae). Cladistics. 2006;22(3):199–221.
9. Shen J, Zhang X, Landis JB, Zhang H, Deng T, Sun H, et al. Plastome evolution in dolomiaea (Asteraceae, Cardueae) using phylogenomic and comparative analyses. Front Plant Sci. 2020;11:376.
10. Zhang C, Huang CH, Liu M, Hu Y, Panero JL, Luebert F, et al. Phylotranscriptomic insights into Asteraceae diversity, polyploidy, and morphological innovation. J Integr Plant Biol. 2021;63(7):1273–93.
11. Amini H, Arzani A, Karami M. Effect of water deficiency on seed quality and physiological traits of different safflower genotypes. Turk J Biol. 2014;38(2):271–82.
12. Golkar P, Arzani A, Rezaei AM. Inheritance of flower colour and spinelessness in safflower (*Carthamus tinctorius* L). J Genet. 2010;89(2):259–62.
13. Pahlavani M, Mirlohi A, Saeidi G. Inheritance of flower color and spininess in safflower (*Carthamus tinctorius* L). J Hered. 2004;95(3):265–7.
14. Zuryn S, Jarriault S. Deep sequencing strategies for mapping and identifying mutations from genetic screens. Worm. 2013;2(3):e25081.
15. Lethin J, Shakil SS, Hassan S, Sirijovski N, Töpel M, Olsson O, et al. Development and characterization of an EMS-mutagenized wheat population and identification of salt-tolerant wheat lines. BMC Plant Biol. 2020;20:15.
16. Manjaya JG, Gupta SK. Mutation breeding for adaptation to climate change in seed-propagated crops. Advanced crop Improvement, volume 2: case studies of economically important crops. Cham: Springer International Publishing; 2023. pp. 197–229.
17. Kammili A. Genetic linkage between male sterility and non-spiny trait in safflower (*Carthamus tinctorius* L). Plant Breed. 2013;132:180–4.
18. Lee G, Sung J, Lee S, Chung J, Yi J, Kim Y, et al. Genetic assessment of safflower (Carthamus tinctorius L.) collection with microsatellite markers acquired via pyrosequencing method. Mol Ecol Resour. 2014;14:36–45.
19. El-Gazzar A, El-Husseini N, Khafagi A, Mostafa N. Computer-generated keys to the flora of Egypt. 9. The spiny taxa of Asteraceae. Egypt J Bot. 2018;58:23–35.
20. Barakati M, Ahmadikhah A. Assessment of pleiotropic effects of spinelessness mutation on other morphological traits of cultivated safflower. In: Proceedings of the 1st international and 13th Iranian genetics congress. Tehran, Iran; 2014. https://biotecher.ir/.
21. Karami-Moalem S, Ahmadikhah A, Nemati Z, Haghi R. Transcriptome differential display of a drought-tolerant early flowering spineless mutant of safflower (*Carthamus tinctorius*) and identification of candidate genes. Crop Sci. 2023;63(4):2329–46.
22. Jeon M, Jeong D, Doh H, Kang H, Jung H, Eyun S. A practical comparison of the next-generation sequencing platform and assemblers using yeast genome. Life Sci Alliance. 2023;6:e202201234.
23. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. Available from. https://www.bioinformatics.babraham.ac.uk/projects/fastqc.
24. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
25. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv Preprint ArXiv:1303.3997. 2013. https://doi.org/10.48550/arXiv.1303.3997.
26. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and samtools. Bioinformatics. 2009;25(16):2078–9.
27. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–303.
28. Sardouei-Nasab S, Nemati Z, Mohammadi-Nejad G, Haghi R, Blattner FR. Phylogenomic investigation of safflower (*Carthamus tinctorius*) and related species using genotyping-by-sequencing (GBS). Sci Rep. 2023;13(1):6212.
29. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly. 2012;6(2):80–92.
30. Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, et al. QTL-seq: rapid mapping of quantitative trait loci in rice by whole

genome resequencing of DNA from two bulked populations. Plant J. 2013;74(1):174–83.
31. Mansfeld BN, Grumet R. QTL-seqr: an R package for bulk Segregant analysis with next-generation sequencing. Plant Genome. 2018;11(2):180006.
32. Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, et al. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics. 2003;164(2):731–40.
33. García A, Aguado E, Cebrián G, Garrido D, Jamilena M. Phenomic and genomic characterization of a mutant platform in *Cucurbita Pepo*. Front Plant Sci. 2018;9:376580.
34. Yan W, Deng XW, Yang C, Tang X. The genome-wide EMS mutagenesis bias correlates with sequence context and chromatin structure in rice. Front Plant Sci. 2021;12:579675.
35. Urlaub G, Mitchell P, Ciudad C, Chasin L. Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. Mol Cell Biol. 1989;9:2868–80.
36. Mort M, Ivanov D, Cooper D, Chuzhanova N. A meta-analysis of nonsense mutations causing human genetic disease. Hum Mutat. 2008;29:103–13.
37. Greene E, Codomo CA, Taylor NE, Henikoff J, Till B, Reynolds S, et al. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics. 2003;164(2):731–40.
38. Zerdoumi Y, Aury-Landas J, Bonaïti-Pellié C, Derambure C, Sesboüé R, Renaux-Petel M, et al. Drastic effect of germline TP53 missense mutations in Li-Fraumeni patients. Hum Mutat. 2013;34:453–65.
39. Auclair J, Busine MP, Navarro C, Ruano E, Montmain G, Desseigne F, et al. Systematic mRNA analysis for the effect of MLH1 and MSH2 missense and silent mutations on aberrant splicing. Hum Mutat. 2006;27:145–54.
40. Fan H, Xu J, Ao D, Jia T, Shi Y, Li N, et al. QTL mapping of trichome traits and analysis of candidate genes in leaves of wheat (*Triticum aestivum* L). Genes. 2023;15:45–57.
41. Chen L, Hao L, Parry MA, Phillips AL, Hu YG. Progress in TILLING as a tool for functional genomics and improvement of crops. J Integr Plant Biol. 2014;56(5):425–43.
42. Caldwell DG, McCallum N, Shaw P, Muehlbauer GJ, Marshall DF, Waugh R. A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L). Plant J. 2004;40(1):143–50.
43. Cooper JL, Greene EA, Till BJ, Codomo CA, Wakimoto BT, Henikoff S. Retention of induced mutations in a drosophila reverse-genetic resource. Genetics. 2008;180(1):661–7.
44. Blumenstiel JP, Noll AC, Griffiths JA, Perera AG, Walton KN, Gilliland WD, et al. Identification of EMS-induced mutations in *Drosophila melanogaster* by whole-genome sequencing. Genetics. 2009;182(1):25–32.
45. Minoia S, Petrozza A, D'Onofrio O, Piron F, Mosca G, Sozio G, et al. A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC Res Notes. 2010;3:69.
46. Shirasawa K, Hirakawa H, Nunome T, Tabata S, Isobe S. Genome-wide survey of artificial mutations induced by Ethyl methanesulfonate and gamma rays in tomato. Plant Biotechnol J. 2016;14(1):51–60.
47. Swarnkar MK, Kumar P, Dogra V, Kumar S. Prickle morphogenesis in Rose is coupled with secondary metabolite accumulation and governed by canonical MBW transcriptional complex. Plant Direct. 2021;5(6):e00325.
48. Doebley J. Molecular evidence and the evolution of maize. Econ Bot. 1990;44(Suppl 3):6–27.
49. Koinange EM, Singh SP, Gepts P. Genetic control of the domestication syndrome in common bean. Crop Sci. 1996;36(4):1037–45.
50. Romero-Hernandez G, Martinez M. Plant kinases in the perception and signaling networks associated with arthropod herbivory. Front Plant Sci. 2022;13:824422.
51. Sabetta W, Alba V, Blanco A, Montemurro C. SunTILL: a TILLING resource for gene function analysis in sunflower. Plant Methods. 2011;7:20.
52. Liu L, Tong H, Xiao Y, Che R, Xu F, Hu B, et al. Activation of big Grain1 significantly improves grain size by regulating auxin transport in rice. Proc Natl Acad Sci U S A. 2015;112(35):11102–7.
53. Lo SF, Cheng ML, Hsing YIC, Chen YS, Lee KW, Hong YF, et al. Rice big grain 1 promotes cell division to enhance organ development, stress tolerance and grain yield. Plant Biotechnol J. 2020;18(9):1969–83.
54. Remington D, Purugganan M. Candidate genes, quantitative trait loci, and functional trait evolution in plants. Int J Plant Sci. 2003;164(S3):S7–20.
55. Satterlee JW, Alonso D, Gramazio P, Jenike KM, He J, Arrones A, et al. Convergent evolution of plant prickles by repeated gene co-option over deep time. Science. 2024;385(6708):eado1663.
56. Kurakawa T, Ueda N, Maekawa M, Kobayashi K, Kojima M, Nagato Y, et al. Direct control of shoot meristem activity by a cytokinin-activating enzyme. Nature. 2007;445(7128):652–5.

Publisher's Note