



ARTICLE

Construction of EMS-Induced Peanut Mutant Libraries and Identification of Pod-Related Traits Mutant Lines

Hao Chen¹, Faqian Xiong^{2,*}, Rilong Xu¹, Xiangyu Chen¹, Haifeng Zhong¹, Yumei Zhang¹, Xinlong Lan¹, Hong Zhang¹, Yuhua Chen¹, Runfang Hu¹, Guoqiang Lin^{1,*} and Zhaoxiu Tang¹

¹Institute of Crop Sciences, Fujian Academy of Agricultural Sciences, Fujian Research Station of Crop Gene Resource & Germplasm Enhancement, Ministry of Agriculture and Rural Affairs of People's Republic of China, Fujian Engineering Research Center for Characteristic Upland Crops Breeding, Fujian Engineering Laboratory of Crop Molecular Breeding, Fuzhou, 350013, China

²Cash Crops Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, 530007, China

*Corresponding Authors: Faqian Xiong. Email: xfq2002@126.com; Guoqiang Lin. Email: lgq308@163.com

Received: 15 May 2022 Accepted: 22 June 2022

ABSTRACT

Peanut (*Arachis hypogaea* L.) is an oil and economic crop of vital importance, and peanut pod is the key organ influencing the yield and processing quality. Hence, the Pod-related traits (PRTs) are considered as important agronomic traits in peanut breeding. To broaden the variability of PRTs in current peanut germplasms, three elite peanut cultivars were used to construct Ethyl methane sulfonate (EMS)-induced mutant libraries in this study. The optimal EMS treatment conditions for the three peanut varieties were determined. It was found that the median lethal dose (LD50) of EMS treatment varied greatly among different genotypes. Finally, the EMS-induced peanut mutant libraries were constructed and a total of 124 mutant lines for PRTs were identified and evaluated. Furthermore, "M-8070", one of the mutant lines for pod constriction, was re-sequenced via high-throughput sequencing technology. The genome-wide variations between "M-8070" and its wild parent "Fuhua 8" (FH 8) were detected. 2994 EMS-induced single nucleotide polymorphisms (SNPs) and 1188 insertion-deletions (InDels) between "M-8070" and its wild parent were identified. The predominant SNP mutation type was C/G to T/A transitions, while the predominant InDel mutation type was "1-bp". We analyzed the distribution of identified mutations and annotated their functions. Most of the mutations (91.68% of the SNPs and 77.69% of the InDels) were located in the intergenic region. 72 SNPs were identified in the exonic region, leading to 27 synonymous, 43 non-synonymous and 2 stop-gain variation for gene structure. 13 Indels were identified in the exonic region, leading to 4 frame-shift, 8 non-frame-shift and 1 stop-gain variations of genes. These mutations may lead to the phenotypic variation of "M-8070". Our study provided valuable resources for peanut improvement and functional genomic research.

KEYWORDS

Peanut; EMS-induced; mutant line; pod-related traits; re-sequencing; pod constriction

1 Introduction

Peanut (*Arachis hypogaea* L.), which is rich in vegetable fat, proteins, vitamins, mineral elements and other nutrients, supplies excellent raw material for oil processing industry and food productions [1]. It is one



of the most significant oil and economic crops worldwide, with a global production of approximately 47 million metric tons during the year 2020 [2]. Breeding new peanut varieties with high yield production and good quality will be meaningful for consumers, peanut producers, manufacturers and all allied industries.

Pod is the most important organ in peanut from an agronomic perspective. Pod-related traits (PRTs) are also considered as important agronomic traits in peanut breeding. PRTs, such as pod weight, pod length, and pod width have been reported to directly influence peanut yield production [3]. Furthermore, some of the PRTs, such as pod length-width ratio and constriction trait co-determine pod size, shape and appearance of peanut, which are also important in in-shell consumption and processing industry [4,5]. Therefore, it is important for us to broaden the variability of PRTs in the current peanut breeding resource in order to meet various requirements.

Induced mutagenesis, which provides the possibility of generating desired phenotypes that are not present in nature, is a useful tool in plant breeding by broadening the genetic diversity in the existing varieties. It is also widely applied in plant functional genome research [6]. Ethyl methane sulfonate (EMS) is one of the most commonly used mutagens. It induces point base substitutions of guanine-cytosine (G/C) to adenine-thymine (A/T) with high efficiency [7,8]. Compared with other mutant strategies which give rises to chromosomal deletions or aberrations, EMS-induced random point mutations are less detrimental to the organism, and consequently the treated organism has a higher frequency to survive. As a result, it is easier to obtain a saturated mutant population [9]. In peanut, efforts have been made to construct the EMS-induced mutant library, and a lot of mutants with desirable mutant traits, such as high-oleate content [10–12], allergen reduction [13], oil content [14], different testa color [15], salt tolerance [16] and yield-related traits [17,18], have been identified. Some of the new peanut varieties have been developed based on the EMS mutants [19]. However, more mutants with different genetic backgrounds are still urgent needed. It is important for both peanut breeding programs and functional genomic research.

In this study, three elite peanut varieties were treated by EMS to produce new chemical mutants. The optimal EMS concentrations for mutagenesis for different peanut varieties have been investigated. A series of PRTs mutant lines were selected and identified from the mutant libraries. Furthermore, one of the mutants “M-8070”, which showed a deeper constriction phenotype than its wild type, was re-sequenced and analyzed.

2 Materials and Methods

2.1 Plant Materials

Three peanut cultivars, “Fuhua 6” (FH 6), “Fuhua 8” (FH 8) and “Kanghuang 1” (KH 1), were used in this study. These three cultivars were widely cultivated in Fujian Province, South of China. “FH 6” is a peanut cultivar with high-oil; “FH 8” is a high-protein content peanut cultivar, and “KH 1” is one of the earliest released *Aspergillus flavus*-resistant peanut cultivars in China [20]. All the seeds used for EMS treatment were propagated by selfing for over five generations to guarantee the purity and homozygous.

2.2 EMS Treatment and Measurement of the Germination Rate under EMS Treatment

The EMS treatments consisted of 0.3%, 0.6% and 0.9% (v/v) EMS solutions (Sigma-Aldrich (Shanghai) Trading Co., Ltd., Shanghai, China) (0.1 M phosphate buffer (pH 7.0) was used as dissolvent of EMS solutions). In general, peanut seeds were soaked and shaken at low speed in the EMS solution for 8 h. The control seeds were treated by 0.1 M phosphate buffer (pH 7.0) without EMS. Then, all the treated seeds were washed via running water for 4 h. After basically drain surface moisture of seeds, the treated seeds were sowed in the field of Fuzhou (E: 119.28; N: 26.08). In order to calculate the germination rate under the EMS treatments, each treatment had three parallel replications and each replication included

300 seeds. The experiment was carried out in a complete randomized-block design. The germination rates of the treated seeds and their control were evaluated at 17 days after sowing (EMS treatment delayed the emergence of some treated seeds). Germination rate was calculated by dividing the number of germinated seeds by the total number of tested seeds, and then multiplying the number by 100 to get a percentage. The relative germination rate of different treatments in this study was calculated by dividing the germination rate of any treated group by the germination rate of the control group, and then multiplying the number by 100 to get the percentage. The LD₅₀ values for peanut mutagenesis with different EMS concentrations in this study were estimated by linear regression analysis described in a previous study [17], based on the relative germination rate.

2.3 Phenotypic Evaluation

The mutants for PRTs were firstly detected by visual observation during the harvest. The pods from mutants were harvested and dried. The 100-pod weight (HPW), pod length, and pod width were analyzed via a Wanshen seed testing instrument, using at least 10 typical pods. The pod length/width ratio was calculated based on the data of pod length and width. The pod constriction extent was also firstly visual evaluated in the field, and then calculated by a pod constriction index (PCI) described in a previous study [5] (The width of the constricted area was measured by a Vernier caliper, and at least 5 typical pods were measured for a line).

2.4 Identification and Annotation of EMS-Induced Mutations

The young leaves from the mutant line “M-8070” (in the M₄ generation) were collected for DNA extraction. The DNA library was constructed following the protocol of NEB Next Ultra II DNA Library Prep Kit for Illumina. A high-throughput sequencing of “M-8070” DNA library was carried out by Illumina NovaSeq platform with NovaSeq 6000 S4 Reagent Kit. In addition, we got the high-throughput sequencing data of “FH 8” from a previous study [21]. Then, the clean data of “M-8070” and “FH 8” were mapped to a reference genome of cultivated peanut (Tifrunner_V20190521, https://www.peanutbase.org/data/public/Arachis_hypogaea/Tifrunner.gnm2.J5K5/) using the tool of BWA [22], and the comparison results were obtained in a SAM format. “SAMtools” [23] was applied to transform the comparison results from sam to bam format. Picard was used to sort duplicates, add read group information and index the bam file (<http://broadinstitute.github.io/picard/>). Any possible EMS-induced mutation was detected by the “HaplotypeCaller” module of “GATK” tools software [24]. The annotation and effects of mutations on gene function were predicted by “ANNOVAR” [25]. The raw data were deposited on <https://www.ncbi.nlm.nih.gov/sra/PRJNA839216>.

2.5 Statistical Analysis

The data obtained from different types of mutants were statistically analyzed by one-way ANOVA analysis, using SPSS 16.0 software (SPSS, Chicago, IL, USA). The analyzed data were presented as means (±SD) for at least three replications.

3 Results

3.1 Determination of the Optimal EMS Concentrations for Mutagenesis in Three Peanut Varieties

To determine the optimal EMS concentration of mutagenesis for peanut in this study, three EMS solution concentrations (v/v) of 0.3%, 0.6% and 0.9% were applied to treat seeds from three peanut cultivars, “FH 6”, “FH 8” and “KH 1” for 8 h, respectively. Field germination rates of the treated seeds and their control were evaluated. It was found that the germination rate was negatively correlated with the EMS concentration. However, different genotypes showed different sensitivity to the EMS treatment. Under the treatment of 0.3% EMS concentration for 8 h, about 80% “FH 6” seeds, over 70% “FH 8” seeds and over 40% “KH 1” seeds were germinated. Under the treatment of 0.6% EMS concentration for 8 h, over 60% “FH 6”

seeds, about 50% “FH 8” seeds and about 2% “KH 1” seeds were germinated. When the seeds were treated with 0.9% EMS concentration for 8 h; the germination rate was about 58.33%, 38.77% and 0% for “FH 6”, “FH 8” and “KH 1”, respectively (Fig. 1). Hence, the median lethal dose (LD_{50}) of EMS concentration for “FH 6” was around 0.9% (0.96%), for “FH 8” was around 0.6% (0.68%), and for “KH 1” was around 0.3% (0.34%) under the 8-h treatment.

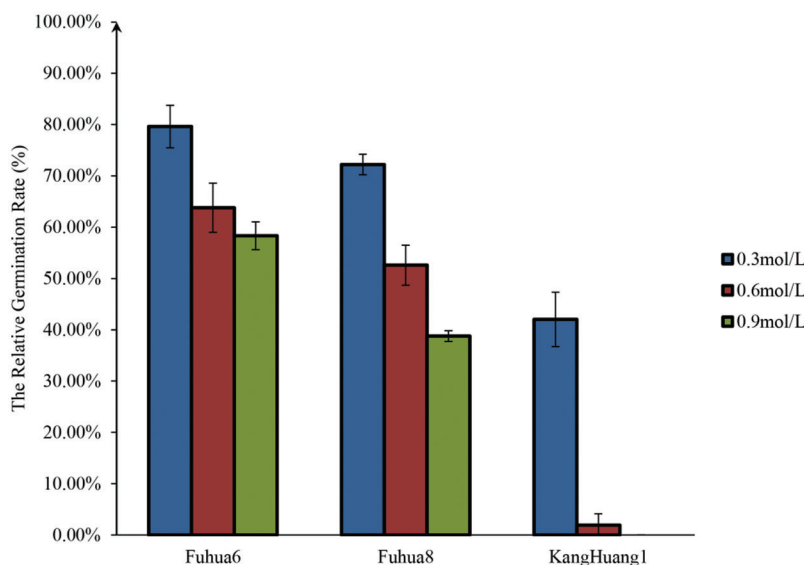


Figure 1: The relative germination rate of M_0 seeds treated by different EMS concentrations. Each histogram represented the means of relative germination rate for different EMS concentration treatments for various peanut genotypes, $n = 3$ replications, each replication included 300 seeds. Vertical bars represented means \pm standard deviations of the results of three independent replications

3.2 Development of the Mutant Peanut Libraries and Identification of Mutants for PRTs

Approximately 5000 treated seeds were germinated as M_0 Plants. 3502 of these plants (1562 plants for “FH 6”, 1711 plants for “FH 8” and for 229 plants for “KH 1”, respectively) were harvested with seeds. The remaining plants did not produce seeds for various reasons, such as sterility, severe dwarfing and all kinds of diseases in the field. The seeds harvested from the M_0 plants which survived were planted as M_1 lines. Preliminary field observations for PRTs variation (including pod length, width, the length-width ratio and pod constriction) were conducted during the harvest period of M_1 lines. The seeds of the M_1 plants which showed variation in PRTs were harvested and planted as M_2 lines. The pods from each of the M_2 lines were evaluated to confirm the results of the M_1 population. As a result, 124 mutant lines (51 lines for “FH 6”, 70 lines for “FH 8” and 3 lines for “KH 1”, respectively) with stable inheritance variation in PRTs were obtained (Fig. 2).

We analyzed the PRTs (including pod weight, pod length, pod width, the length-width ratio and pod constriction) variation in all identified mutants of “FH 6” and “FH 8”. Among accessions of the identified “FH 6” PRTs mutants, the one hundred pod weight ranged from 127.13 to 209.31 g, 5 lines showed heavier pods, and 46 lines showed less pod weights than their wild type. The pod length ranged from 24.30 to 34.64 mm, and 8 lines showed an increase, while the remaining 38 lines showed a decrease in pod length than “FH 6”. The pod width of the “FH 6”-based mutants ranged from 13.82 to 18.47 mm; 2 lines increased in pod width, and the remaining lines decreased in pod length than “FH 6” (Figs. 3A and 3C; Table S1). 14 lines exhibited a visual change in length-width ratio, and 5 lines varied in pod constriction extent (Table S1).

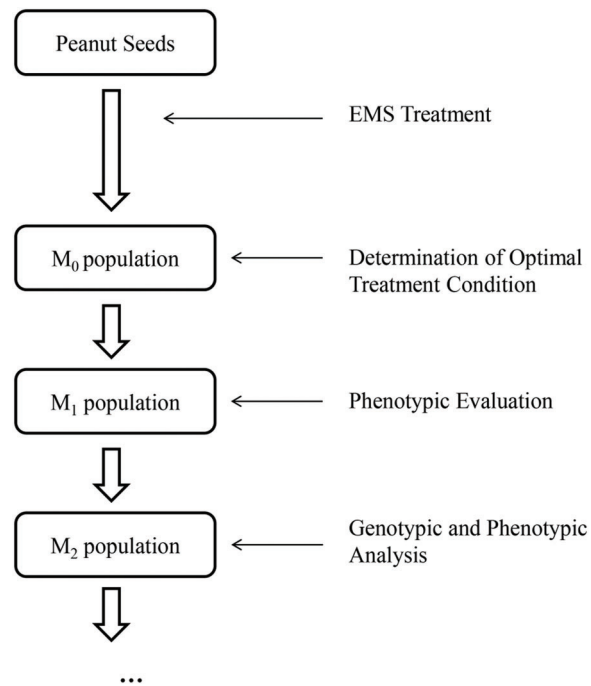


Figure 2: Development of the mutagenized peanut libraries

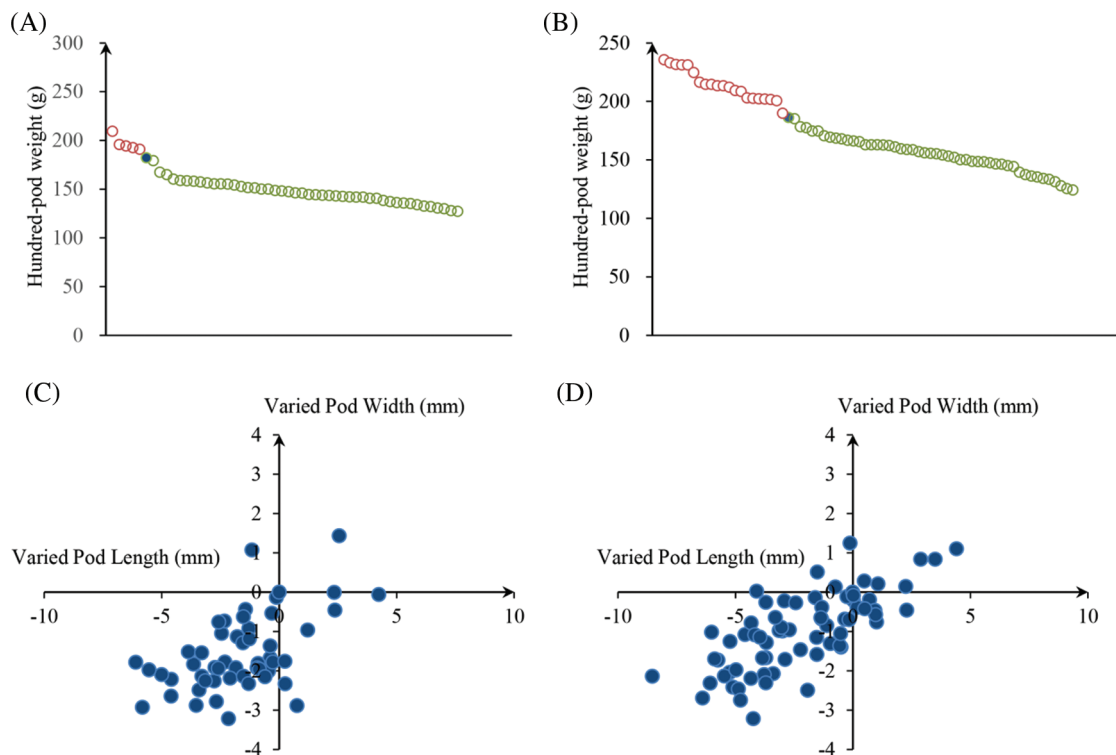


Figure 3: Phenotypic evaluation of PRTs mutants for the “FH 6” and “FH 8” mutagenesis libraries. (A) One hundred-pod weight of PRTs mutant lines from the “FH 6” mutant library. The blue spot represents the wild type parent “FH 6”; (B) One hundred-pod weight of PRTs mutant lines from the “FH 8” mutant library. The blue spot represents “FH 8” (C) Variation in pod length and pod width of PRTs mutant lines from the “FH 6” mutant library. (D) Variation in pod length and pod width of PRTs mutant lines from the “FH 8” mutant library

In the “FH 8”-based mutant, the one hundred pod weight ranged from 124.15 to 239.64 g, and 21 lines increased in pod weight, while 49 lines decreased in it. The pod length of “FH 8”-based mutant lines ranged from 26.47 to 36.37 mm, 16 lines increased in pod length, and 54 lines decreased in pod length. The pod width of the “FH 8”-based mutant lines ranged from 13.96 to 18.42 mm. Among the identified lines, 12 showed an increased pod width, while 58 showed a decrease on it (Figs. 3B and 3D; Table S2). In addition, we also identified 5 lines with a varied length-width ratio mutant, and 1 line with a different pod constriction extent in this “FH 8”-based mutant library (Table S2). The information of PRTs mutants in the “KH 1” background was listed in Table S3.

3.3 Identification of a Pod-Constriction Mutant and Its Whole-Genome Sequencing

Among the identified mutant, “M-8070”, one “FH 8”-based mutant showed deeply constricted pods, while its wild type “FH 8” exhibited slightly or no constriction (Fig. 4). This mutant was re-sequenced, and a total of 33.25 Gb paired-end reads were obtained with $11.69 \times$ average sequencing depth. The obtained clean short-read sequences were mapped to the Tifrunner reference genome, and the average genome coverage was 95.64%. Meanwhile, the “FH 8” re-sequencing data were obtained from our previous study (also used Tifrunner as a reference genome). After filtering out the common base change shared by “M-8070” and “FH 8”, a total of 4,182 loci were considered as EMS-induced sites. Among them, 2,994 were single nucleotide polymorphisms (SNPs), and 1,188 were InDels. As expected, 2,258 SNPs were G/C to A/T transitions, and they were the predominant SNP base changes (75.41%) in our mutant. This is also the typical point base substitutions induced via EMS. The frequency was followed by C/G to A/T transversions (265, 8.85%), A/T to G/C transitions (201, 6.71%), A/T to T/A transversions (165, 5.51%), C/G to G/C transversions (58, 1.94%), and A/T to C/G transversions (57, 1.90%) (Fig. 5A). Among the 1,188 identified InDels, the most frequent InDel length was 1-bp (593), it accounted for 49.92% of the total InDels (Fig. 5B). We further analyzed the distribution and density of EMS-induced sites. It demonstrated that the mutant frequency was 1.65 mutant site/Mb, and those sites were evenly distributed in the whole genome (Figs. 6 and 7, Table 1).

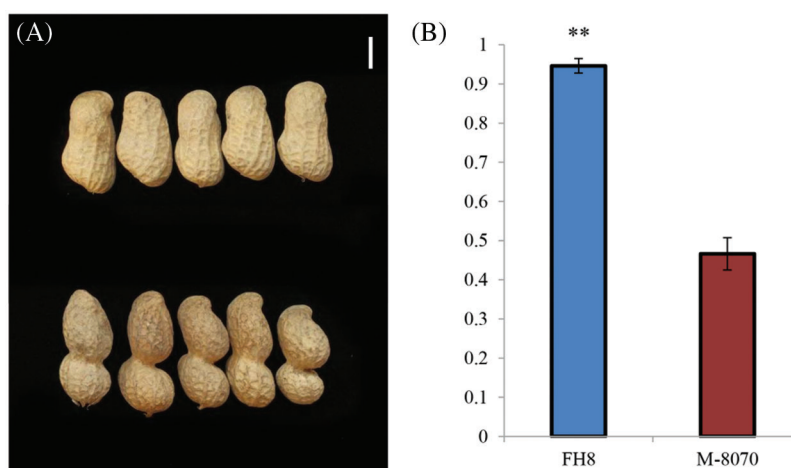


Figure 4: The phenotype of “M-8070” and its wild type parent “FH 8”. (A) Pods of “M-8070” and its wild type parent “FH 8”. The pod of “FH 8” is present in the upper line, while the pod of “M-8070” is present in the lower line. The white bar represents the length of 1 cm. (B) The PCI index of “M-8070” and “FH 8”. The error bars represent the mean \pm SD of five biological replicates. The Student *t*-test was applied to assess the difference of means between the two varieties. “**” indicates a significant difference between “M-8070” and “FH 8” ($P = < 0.01$)

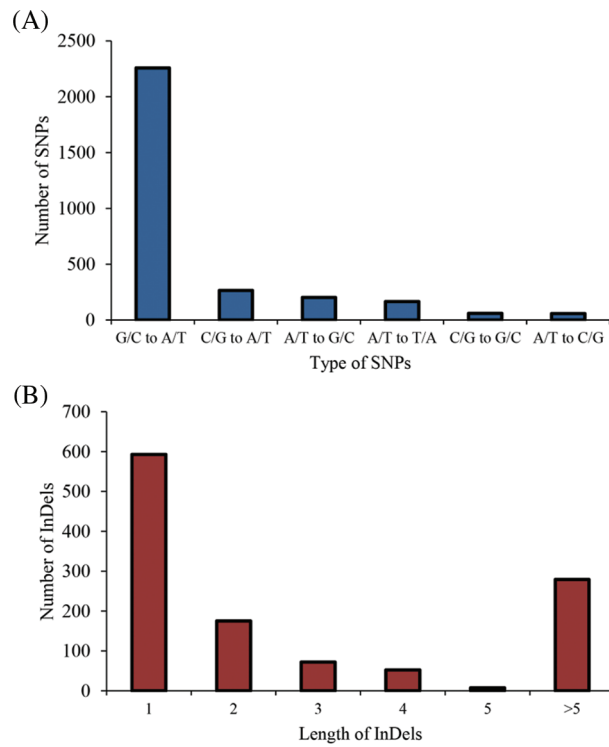


Figure 5: Type of mutations identified in “M-8070”. (A) The type of SNPs identified in “M-8070”. (B) The type of InDels identified in “M-8070”

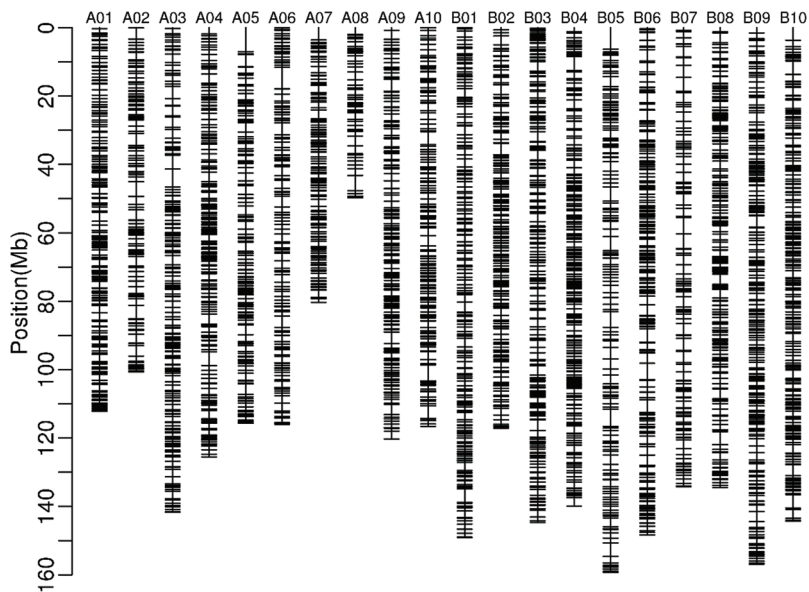


Figure 6: The identified mutation of “M-8070” in chromosomes view

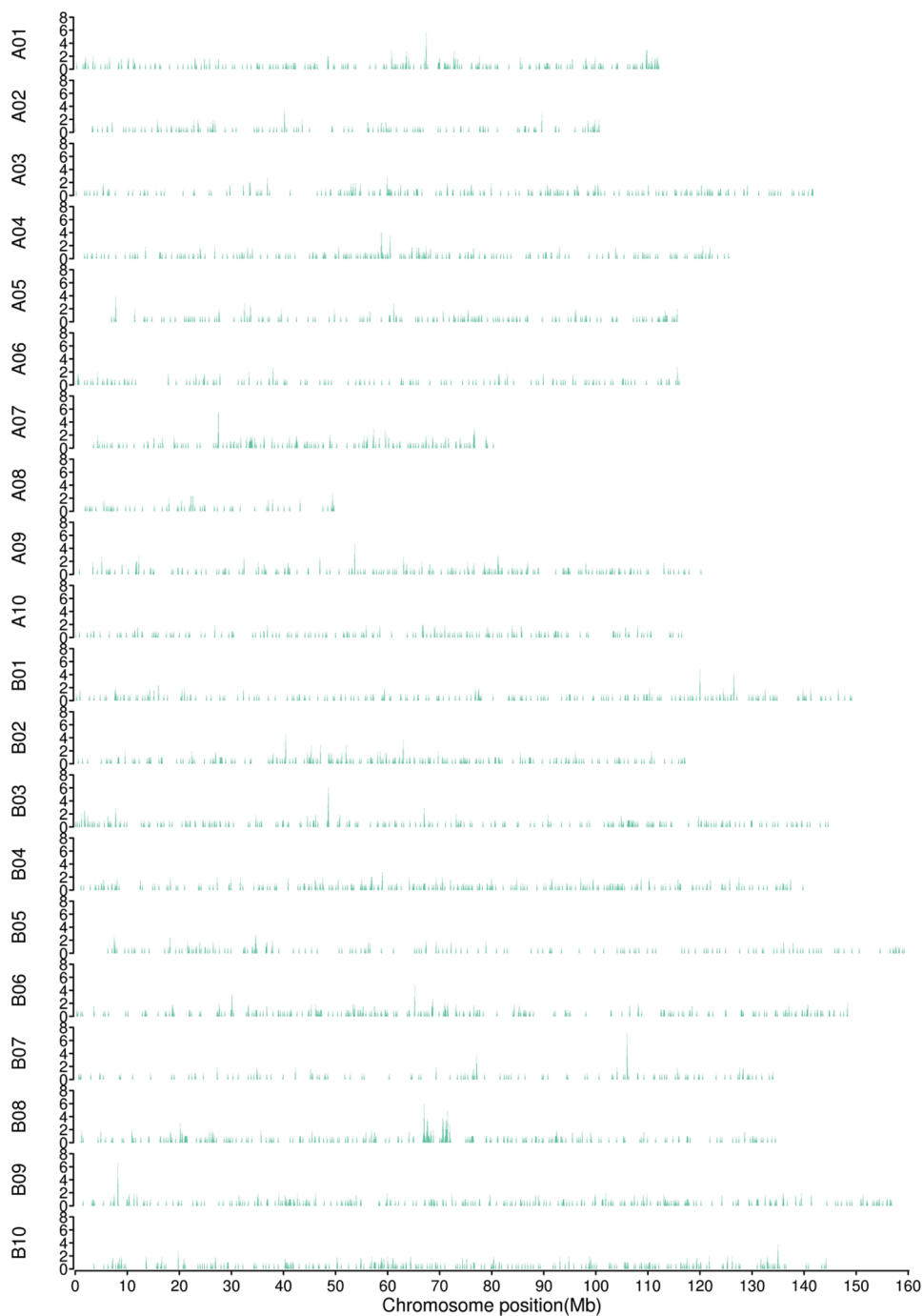


Figure 7: The distribution of identified mutation in the genome of “M-8070”

3.4 Determination of the Functional Effects of EMS-Induced Mutation

We used the software “ANNOVAR” to annotate these mutations, and predicted the effect of mutations on gene function. Most SNPs (2,745 of 2,994, 91.68%) mutations were located in the intergenic region. 70 of 2,994 SNPs (2.37%) were identified in the upstream/downstream region. There were 178 of 2,994 SNPs (5.95%) identified in the genic region. 91 of them were in the intronic region, 72 were in the exonic

region, and the rest were identified in the UTR region or splicing region (Table 2). The 72 SNPs identified in the exonic region led to 27 synonymous, 43 non-synonymous and 2 stop-gain variations of genes. For InDels, most of them (923 of 1,188, 77.69%) also occurred in the intergenic region. Nevertheless, 77 InDels (6.48%) occurred in the upstream/downstream region, and 188 InDels (15.82%) were identified in the genic region (Table 2). 13 InDels occurred in the exonic region led to 4 frame-shift, 8 non-frame-shift and 1 stop gain variations of genes.

Table 1: The distribution and density of SNPs and InDels in “M-8070”

Chr	Length	No. SNPs	SNP density	No. InDels	InDel density
A01	112,420,854	180	1.60	60	0.53
A02	103,302,290	97	0.94	42	0.41
A03	143,109,472	158	1.10	63	0.44
A04	128,801,742	167	1.30	46	0.36
A05	116,542,366	119	1.02	51	0.44
A06	118,975,115	109	0.92	40	0.34
A07	81,752,458	144	1.76	30	0.37
A08	51,529,986	33	0.64	34	0.66
A09	120,499,698	177	1.47	38	0.32
A10	117,076,737	134	1.14	41	0.35
B01	149,287,806	178	1.19	55	0.37
B02	120,530,088	167	1.39	68	0.56
B03	146,301,462	168	1.15	78	0.53
B04	143,237,272	194	1.35	80	0.56
B05	160,028,458	105	0.66	74	0.46
B06	151,242,074	194	1.28	69	0.46
B07	134,191,082	62	0.46	78	0.58
B08	135,027,066	161	1.19	117	0.87
B09	159,361,216	238	1.49	56	0.35
B10	145,034,356	209	1.44	68	0.47
Whole	2,538,251,598	2,994	1.18	1,188	0.47

Some of the EMS-induced SNPs were located on the exonic region of genes; we collected the information of these genes and annotated them (Table 3). In addition, we further annotated genes which EMS-induced Indels mutations led to the “stop-gain”, “nonframeshift insertion”, “frameshift insertion”, or “frameshift-deletion” effect on gene structure (Table 4). Among these genes, two “stop-gain” type genes induced by SNPs, which encode “zinc finger MYM-type protein 1-like” and “probable dolichyl pyrophosphate Man⁹GlcNAc₂ alpha-1, 3-glucosyltransferase-like”, respectively, showed the typical point base substitutions of guanine-cytosine (G/C) to adenine-thymine (A/T) induced via EMS treatment. We speculated that they were the possible candidate genes for the deep constriction phenotype in our mutant.

Table 2: Distribution and percentage of mutations in the genome

SNPs			InDels		
Position	Number	Ratio (%)	Position	Number	Ratio (%)
Intergenic	2745	91.68%	Intergenic	923	77.69%
Upstream	42	1.40%	Upstream	38	3.20%
Downstream	26	0.87%	Downstream	38	3.20%
Up/down	2	0.07%	Up/down	1	0.08%
Genic	178	5.95%	Genic	188	15.82%
Intronic	91	3.04%	Intronic	132	11.11%
Exonic	72	2.40%	Exonic	13	1.09%
UTR5	6	0.20%	UTR5	21	1.77%
UTR3	6	0.20%	UTR3	20	1.68%
Splicing	3	0.10%	Splicing	2	0.17%

Table 3: EMS-induced SNPs located on the exonic region and their effect on gene function

Chr	Position	Alt	Ref	Annotated genes ID	Function effect
A01	6651746	G	A	<i>arahy.36HDXC</i>	Nonsynonymous SNV
A01	95115500	G	A	<i>arahy.Y01GTZ</i>	Synonymous SNV
A01	106168680	G	A	<i>arahy.ZP1L4J</i>	Synonymous SNV
A01	11206955	C	T	<i>arahy.2D2JRR</i>	Synonymous SNV
A02	58822935	G	A	<i>arahy.74EFZ5</i>	Nonsynonymous SNV
A02	99506837	C	T	<i>arahy.JXQ8LX</i>	Nonsynonymous SNV
A02	100347946	C	T	<i>arahy.4NCV9J</i>	Nonsynonymous SNV
A03	87014831	G	A	<i>arahy.FKNV1Q</i>	Nonsynonymous SNV
A03	120306286	G	A	<i>arahy.4PZU7W</i>	Synonymous SNV
A03	47333363	A	G	<i>arahy.LZ43MP</i>	Nonsynonymous SNV
A03	2819807	C	T	<i>arahy.LS9MUN</i>	Nonsynonymous SNV
A03	140850761	C	T	<i>arahy.DBKV77</i>	Synonymous SNV
A04	3325232	C	T	<i>arahy.N8S0IJ</i>	Nonsynonymous SNV
A04	20822870	C	T	<i>arahy.XDA15Z</i>	Synonymous SNV
A04	114373666	C	T	<i>arahy.RCQ4HC</i>	Nonsynonymous SNV
A05	27745830	G	A	<i>arahy.4M9DN5</i>	Synonymous SNV
A05	78001184	C	T	<i>arahy.LW0R14</i>	Nonsynonymous SNV
A05	86949125	C	T	<i>arahy.VH9XRT</i>	Synonymous SNV
A05	112320302	C	T	<i>arahy.UM4T3G</i>	Nonsynonymous SNV
A06	114168565	C	A	<i>arahy.X97WXI</i>	Nonsynonymous SNV

(Continued)

Table 3 (continued)					
Chr	Position	Alt	Ref	Annotated genes ID	Function effect
A06	115785640	A	C	<i>arahy.2TEC20</i>	Nonsynonymous SNV
A06	115785650	A	G	<i>arahy.2TEC20</i>	Synonymous SNV
A06	115785646	A	T	<i>arahy.2TEC20</i>	Nonsynonymous SNV
A06	76905	C	T	<i>arahy.XSF16F</i>	Nonsynonymous SNV
A06	3416849	C	T	<i>arahy.3M3KZ3</i>	Synonymous SNV
A06	93326873	C	T	<i>arahy.RZ4PVC</i>	Synonymous SNV
A07	6116346	G	A	<i>arahy.3MHP6W</i>	Nonsynonymous SNV
A07	8379780	G	A	<i>arahy.2X25WT</i>	Stopgain
A07	15129083	G	A	<i>arahy.0PI3PU</i>	Nonsynonymous SNV
A07	73920297	G	A	<i>arahy.4JR4GS</i>	Nonsynonymous SNV
A07	7168648	C	T	<i>arahy.M9Q6QA</i>	Nonsynonymous SNV
A08	6722519	G	C	<i>arahy.V55W7Y</i>	Nonsynonymous SNV
A09	115938063	T	C	<i>arahy.Q6727A</i>	Nonsynonymous SNV
A09	5107771	C	T	<i>arahy.6GXK25</i>	Nonsynonymous SNV
A09	6067020	C	T	<i>arahy.EFT9AB</i>	Synonymous SNV
A09	31550399	C	T	<i>arahy.H0NCP1</i>	Nonsynonymous SNV
A09	108695001	C	T	<i>arahy.ZJJ54E</i>	Nonsynonymous SNV
A10	3583664	G	A	<i>arahy.I1TJ3H</i>	Nonsynonymous SNV
A10	103846494	G	A	<i>arahy.K7FNKV</i>	Nonsynonymous SNV
A10	110548042	G	A	<i>arahy.T9QV62</i>	Nonsynonymous SNV
B01	143406854	T	C	<i>arahy.Q5N9NX</i>	Synonymous SNV
B01	8273553	C	T	<i>arahy.57ZBEA</i>	Synonymous SNV
B01	34464634	C	T	<i>arahy.TC99UD</i>	Nonsynonymous SNV
B02	115955722	G	A	<i>arahy.RL25CR</i>	Nonsynonymous SNV
B02	110414469	T	G	<i>arahy.D4T6QG</i>	Nonsynonymous SNV
B03	131958401	G	A	<i>arahy.T94LMP</i>	Synonymous SNV
B03	131574658	T	C	<i>arahy.C5DY5V</i>	Synonymous SNV
B03	1854275	C	T	<i>arahy.356BXD</i>	Synonymous SNV
B03	8783675	C	T	<i>arahy.V9TZWK</i>	Synonymous SNV
B03	61402045	C	T	<i>arahy.Z58MGP</i>	Synonymous SNV
B04	107646158	C	T	<i>arahy.Y46IFH</i>	Stopgain
B04	112697161	C	T	<i>arahy.N5W5LK</i>	Nonsynonymous SNV
B04	136079673	C	T	<i>arahy.Z0VDVK</i>	Nonsynonymous SNV
B05	19377867	G	A	<i>arahy.UE7NK8</i>	Nonsynonymous SNV
B05	24202403	G	A	<i>arahy.ML4V41</i>	Nonsynonymous SNV

(Continued)

Table 3 (continued)

Chr	Position	Alt	Ref	Annotated genes ID	Function effect
B06	1465071	C	T	<i>arahy.Y86I8S</i>	Synonymous SNV
B06	140526535	C	T	<i>arahy.08UKT4</i>	Synonymous SNV
B06	144902137	C	T	<i>arahy.798C6B</i>	Synonymous SNV
B07	127642840	G	A	<i>arahy.UITY6Q</i>	Nonsynonymous SNV
B07	73475588	T	C	<i>arahy.3I18X0</i>	Synonymous SNV
B08	20193704	G	A	<i>arahy.RBK0JN</i>	Nonsynonymous SNV
B08	7217142	C	T	<i>arahy.X1RCBJ</i>	Nonsynonymous SNV
B09	8239656	G	A	<i>arahy.9NZ7QY</i>	Nonsynonymous SNV
B09	8239674	G	A	<i>arahy.9NZ7QY</i>	Nonsynonymous SNV
B09	153235201	G	A	<i>arahy.KV6MGK</i>	Nonsynonymous SNV
B09	153565374	G	A	<i>arahy.19YJEB</i>	Nonsynonymous SNV
B09	153945439	G	A	<i>arahy.R2BN40</i>	Synonymous SNV
B09	68506213	T	A	<i>arahy.BAL03A</i>	Nonsynonymous SNV
B09	8239712	T	G	<i>arahy.9NZ7QY</i>	Synonymous SNV
B09	8268721	C	T	<i>arahy.84IRST</i>	Synonymous SNV
B10	24082209	T	A	<i>arahy.1A1CKU</i>	Nonsynonymous SNV
B10	131725482	C	T	<i>arahy.FBF79B</i>	Synonymous SNV

Note: "Alt" is the sequence from "M-8070", while "Ref" is the sequence from "FH 8".

Table 4: EMS-induced InDels with large effect on gene function

Chr	Pos	Ref	Alt	Gene	Annotation-effect
A01	93319428	C	CCTTCTT	<i>arahy.VAS4HN</i>	nonframeshift insertion
A05	112910960	T	TGTTGGGAATAAGACAC-AATTCCCCCTTGAGAAA-ACACCTTTGA	<i>arahy.GW217K</i>	stopgain
A06	633363	T	TAACAG	<i>arahy.V481EC</i>	frameshift insertion
A08	27342401	AGAGAGAGAGC	A	<i>arahy.S7J6Q0</i>	frameshift deletion
A09	61462085	C	CCTTCTTCTTCTTCTTCT-TCTTCTTCTT	<i>arahy.L0C43K</i>	nonframeshift insertion
A09	6511203	T	TTCTGGTTTTGGAATTC-AGGGACTTTAGACAAC	<i>arahy.F66B8V</i>	nonframeshift insertion
B04	119970377	G	GGAA	<i>arahy.HQ8SVE</i>	nonframeshift insertion
B04	132222186	GCTGGAACGGC-GAGTTCGACTG-GGCGAGGGCGA-CGGCA	G	<i>arahy.EX98MN</i>	frameshift deletion

(Continued)

Table 4 (continued)

Chr	Pos	Ref	Alt	Gene	Annotation-effect
B05	158240870	TAACCCTAGAA	T	<i>arahy.63SMXT</i>	frameshift deletion
B06	140799753	A	AATAAGACACAATCCCC	<i>arahy.DD5I6P</i>	frameshift insertion
B06	124958265	C	CAGAGAGAGAGAGAG	<i>arahy.U0LKD0</i>	frameshift insertion
B06	140799750	T	TGGTGTTG	<i>arahy.DD5I6P</i>	frameshift insertion
B10	118607079	G	GTGTTAGGATTTGGTTGAATT	<i>arahy.DAV6SS</i>	frameshift insertion

Note: "Alt" is the sequence from "M-8070", while "Ref" is the sequence from "FH 8".

4 Discussion

In this study, we investigated the optimal EMS treatment condition for different peanut genotypes, and constructed an EMS-induced mutant library. The mutants underlying PRTs were identified. One of the PRTs mutants which exhibited deeply constricted pods was selected for re-sequencing.

According to our study, there were great differences in EMS sensitivity among peanut varieties. Although previous studies have reported that the sensitivity to EMS mutagens varied with genotype in peanut [17], such great variation in EMS sensitivity was not addressed. It is reported that a higher frequency of nucleotide variation would be induced by a higher EMS concentration treatment [26,27]. Since the multiple mutations may mask the possible varied phenotype induced by EMS treatment, and the extraordinary high frequency mutations may lead to a higher possibility for lethal phenotypes in the mutant library [28], an optimal EMS treatment concentration should be determined [29]. The LD₅₀ value is regarded as the evaluating indicator for the EMS treatment. Hence, it is worthwhile to evaluate the LD₅₀ value for different peanut genotypes before large scale application of EMS inducement for the new genotypes.

Pod-related traits (PRTs) are important agronomic traits not only influencing peanut yield, but also affecting pod appearance and processing quality. Although many QTLs underlying PRTs were identified through natural populations [30–35], more information was still needed for elucidating the genetic mechanism underlying these traits. EMS-induced mutagenesis is a suitable choice to supplement the study in the natural population. On the one hand, mutants may create haplotypes not existing in the natural environment. On the other hand, fewer variations in the genome and typical point mutations in the EMS-induced mutant make it easier to find the candidate gene than using the natural materials. In this study, we also produced a variety of peanut PRTs mutants via this method. They exhibited a varying degree of variations in PRTs, and they will be good materials for further peanut breeding programs and functional genomics research.

Whole-genome sequencing is a powerful tool for surveying the mutations induced by EMS mutagens [36]. It has been successfully applied in several crop species. However, few reports were listed about the whole-genome sequencing for EMS induced mutants in peanut. In our study, whole-genome sequencing was applied to survey the mutations in one of our identified mutants. A total of 2,994 SNPs and 1,188 InDels have been identified and characterized. The mutation frequency in our study was 1.65 mutant site/Mb, and this was similar to that of in other crop species, such as *Arabidopsis* (11.24 mutant site/Mb) [37], foxtail millet (4.69 mutant site/Mb) [28], rice (3.77 mutant site/Mb) [38], maize (2.06 mutant site/Mb) [39], sorghum (1.90 mutant site/Mb) [40], and tomato (1.36 mutant site/Mb) [41]. It indicates that our EMS mutagens may induce proper density of mutation in the genome. On one hand, it can induce enough mutations to create variation in phenotypes. On the other hand, a moderate frequency of mutations can minimize the possible masked phenotypes resulted by multiple mutations.

Based on our result, the majority of mutations were C/G to T/A transitions in our mutant. It is consistent with the previous studies, and fit the working mechanism of EMS alkylating agent. However, a series of non-C/G to T/A transitions have also been detected in our mutant. This condition has also been reported in other studies [27,36,42]. What is certain is that some of the non-C/G to T/A transitions may be the false positive result during re-sequencing and data analysis. However, the mechanism on how other non-C/G to T/A transitions are induced is still unclear to us. Nevertheless, in the traditional reverse-genetic study, non-C/G to T/A type mutations were filtered out, and only C/G to T/A type mutations were considered to be the reason for variation [36,43]. In this study, after filtering out the non-C/G to T/A type mutations, 2,248 C/G to T/A-type mutations remained. 2,057 of the 2,248 SNPs were identified in the intergenic region, 54 were in the upstream/downstream of gene, and 137 were identified in the genic region. Among the 137 SNPs identified in the genic region, 68 are identified in the intronic region, 11 were identified in the UTR region or splicing region and the remaining 58 were identified in the exonic region, leading to 27 synonymous, 29 non-synonymous and 2 stop-gain variations of genes, respectively.

The degree of pod constriction influences seed development in peanut. It is also an important agronomic trait affecting peanut processing, especially for peanut in-shell processing industry. The genetic mechanism for peanut constriction is known to be very complex [5,44,45]. Although some QTLs underlying the peanut pod constriction degree have been identified via linkage analysis [5], more genes controlling constriction should be addressed to dissect its genetic mechanism. In this study, we identified the deep constriction mutant lines, and re-sequenced one of them. We propose that those mutations leading to the stop-gain variations are the most likely cause for the phenotype variation. In our study, two genes exhibited the “stop-gain” variation in gene structure resulting from the typical G/C to A/T transitions point base substitutions induced via EMS treatment. We speculated that they were the possible candidate genes for deep constriction phenotypes in our mutant. These two genes encode “zinc finger MYM-type protein 1-like” and “probable dolichyl pyrophosphate Man9GlcNAc2 alpha-1, 3-glucosyltransferase-like”, respectively. Their specific roles in peanut pod development are still not known based on the current knowledge. According to the study in other species, the zinc finger MYM-type proteins are involved in a protein-protein interaction module, a function in regulating the transcriptional ability [46], telomeric fusions [47], and DNA repair [48]. The dolichyl pyrophosphate Man9GlcNAc2 alpha-1, 3-glucosyltransferase catalyzes the addition of the second glucose residue to the lipid-linked oligosaccharide precursor for N-linked glycosylation of proteins [49]. Of course, the candidate gene for phenotypic variation should be confirmed via further study, such as mut-map, and so on.

Authorship: The authors confirm contribution to the paper as follows: study conception and design: H.C, G.L, Z.T; data collection: H.C, F.X, R.X, X.C, H.Z, Y.Z, X.L, Y.C, R.H; analysis and interpretation of results: H.C, F.X; draft manuscript preparation: H.C, F.X. All authors reviewed the results and approved the final version of the manuscript.

Funding Statement: This research was funded by the Basic Scientific Research Special Project for Fujian Provincial Public Research Institutes (2020R10310011; 2019R1031-12); the National Natural Science Foundation of China (32001577); the Foreign Cooperation Projects for FAAS (DWHZ2021-20); the Opening Foundation of Fujian Key Laboratory for Vegetable Breeding (FJVRC2020-02); the Free Explore Program for FAAS (ZYTS2019007); the Guangxi Natural Science Foundation Program (2018GXNSFDA281027), and the Science and Technology Innovative Team in Fujian Academy of Agricultural Sciences (CXTD2021008-3).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Ojiewo, C. O., Janila, P., Bhatnagar-Mathur, P., Pandey, M. K., Desmae, H. et al. (2020). Advances in crop improvement and delivery research for nutritional quality and health benefits of groundnut (*Arachis hypogaea* L.). *Frontiers in Plant Science*, 11, 29.
2. FAO (2021). *FAOSTAT statistical database*. Rome, Italy: FAO.
3. Gomes, R. L. F., de Almeida Lopes, Â. C. (2005). Correlations and path analysis in peanut. *Crop Breeding and Applied Biotechnology*, 5, 105–112.
4. Zhang, S., Hu, X., Miao, H., Chu, Y., Cui, F. et al. (2019). QTL identification for seed weight and size based on a high-density SLAF-seq genetic map in peanut (*Arachis hypogaea* L.). *BMC Plant Biology*, 19(1), 1–15.
5. Patil, A. S., Popovsky, S., Levy, Y., Chu, Y., Clevenger, J. et al. (2018). Genetic insight and mapping of the pod constriction trait in Virginia-type peanut. *BMC Genetics*, 19(1), 1–9.
6. Oladosu, Y., Rafii, M. Y., Abdullah, N., Hussin, G., Ramli, A. et al. (2016). Principle and application of plant mutagenesis in crop improvement: A review. *Biotechnology Biotechnological Equipment*, 30(1), 1–16.
7. Krieg, D. R. (1963). Ethyl methanesulfonate-induced reversion of bacteriophage T4rII mutants. *Genetics*, 48(4), 561–580.
8. Taheri, S., Abdullah, T. L., Jain, S. M., Sahebi, M., Azizi, P. (2017). TILLING, high-resolution melting (HRM), and next-generation sequencing (NGS) techniques in plant mutation breeding. *Molecular Breeding*, 37(3), 1–23.
9. Gilchrist, E., Haughn, G. (2010). Reverse genetics techniques: Engineering loss and gain of gene function in plants. *Briefings in Functional Genomics*, 9(2), 103–110.
10. Karaman, K., Kizil, S., Başak, M., Uzun, B., Yol, E. (2021). Development of EMS-induced mutagenized groundnut population and discovery of point mutations in the *ahFAD2* and *Ara h 1* genes by TILLING. *Journal of Oleo Science*, 70(11), 1631–1640.
11. Fang, C. Q., Wang, C. T., Wang, P. W., Tang, Y. Y., Wang, X. Z. et al. (2012). Identification of a novel mutation in *FAD2B* from a peanut EMS mutant with elevated oleate content. *Journal of Oleo Science*, 61(3), 143–148.
12. Mondal, S., Badigannavar, A. (2013). A narrow leaf groundnut mutant, TMV2-NLM has a G to A mutation in *AhFAD2A* gene for high oleate trait. *Indian Journal of Genetics and Plant Breeding*, 73(1), 105–109.
13. Knoll, J. E., Ramos, M. L., Zeng, Y., Holbrook, C. C., Chow, M. et al. (2011). TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.). *BMC Plant Biology*, 11(1), 1–13.
14. Tang, Y., Wang, X., Wu, Q., Fang, C., Guan, S. et al. (2013). Identification of differentially expressed genes from developing seeds of a normal oil peanut cultivar and its high oil EMS mutant. *Research Crops*, 14(2), 511–516.
15. Wan, L., Li, B., Pandey, M. K., Wu, Y., Lei, Y. et al. (2016). Transcriptome analysis of a new peanut seed coat mutant for the physiological regulatory mechanism involved in seed coat cracking and pigmentation. *Frontiers in Plant Science*, 7, 1491.
16. Zhao, M., Sun, H., Ji, R., Hu, X., Sui, J. et al. (2013). *In vitro* mutagenesis and directed screening for salt-tolerant mutants in peanut. *Euphytica*, 193(1), 89–99.
17. Chen, T., Huang, L., Wang, M., Huang, Y., Zeng, R. et al. (2020). Ethyl methyl sulfonate-induced mutagenesis and its effects on peanut agronomic, yield and quality traits. *Agronomy*, 10(5), 655.
18. Wang, C., Wang, X., Tang, Y., Chen, D., Zhang, J. et al. (2010). High yielding mutants achieved by injecting EMS into peanut flower organs. *Journal of Nuclear Agricultural Sciences*, 24(2), 239–242.
19. Wu, L., Li, Z., Qiu, Q., Miao, H. (2002). Application of mutation techniques in peanut breeding in China. *Journal of Nuclear Agricultural Sciences*, 16(5), 334–336.
20. Liao, B., Zhuang, W., Tang, R., Zhang, X., Shan, S. et al. (2009). Peanut aflatoxin and genomics research in China: Progress and perspectives. *Peanut Science*, 36(1), 21–28.
21. Chen, H., Chen, X., Xu, R., Liu, W., Liu, N. et al. (2021). Fine-mapping and gene candidate analysis for *AhRt1*, a major dominant locus responsible for testa color in cultivated peanut. *Theoretical and Applied Genetics*, 134(11), 3721–3730.
22. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997.

23. Etherington, G. J., Ramirez-Gonzalez, R. H., MacLean, D. (2015). bio-samtools 2: A package for analysis and visualization of sequence and alignment data with SAMtools in Ruby. *Bioinformatics*, 31(15), 2565–2567.
24. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K. et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303.
25. Wang, K., Li, M., Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16), e164.
26. Wang, N., Wang, Y., Tian, F., King, G. J., Zhang, C. et al. (2008). A functional genomics resource for *Brassica napus*: Development of an EMS mutagenized population and discovery of *FAE1* point mutations by TILLING. *New Phytologist*, 180(4), 751–765.
27. Tang, S., Liu, D., Lu, S., Yu, L., Li, Y. et al. (2020). Development and screening of EMS mutants with altered seed oil content or fatty acid composition in *Brassica napus*. *Plant Journal*, 104(5), 1410–1422.
28. Sun, J., Luu, N. S., Chen, Z., Chen, B., Cui, X. et al. (2019). Generation and characterization of a foxtail millet (*Setaria italica*) mutant library. *Frontiers in Plant Science*, 10, 369.
29. Yadav, P., Meena, H., Meena, P., Kumar, A., Gupta, R. et al. (2016). Determination of LD₅₀ of ethyl methanesulfonate (EMS) for induction of mutations in rapeseed-mustard. *Journal of Oilseed Brassica*, 1(1), 77–82.
30. Chen, W., Jiao, Y., Cheng, L., Huang, L., Liao, B. et al. (2016). Quantitative trait locus analysis for pod-and kernel-related traits in the cultivated peanut (*Arachis hypogaea* L.). *BMC Genetics*, 17(1), 1–9.
31. Chavarro, C., Chu, Y., Holbrook, C., Isleib, T., Bertioli, D. et al. (2020). Pod and seed trait QTL identification to assist breeding for peanut market preferences. *G3: Genes, Genomes, Genetics*, 10(7), 2297–2315.
32. Gangurde, S. S., Wang, H., Yaduru, S., Pandey, M. K., Fountain, J. C. et al. (2020). Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). *Plant Biotechnology Journal*, 18(6), 1457–1471.
33. Luo, H., Ren, X., Li, Z., Xu, Z., Li, X. et al. (2017). Co-localization of major quantitative trait loci for pod size and weight to a 3.7 cM interval on chromosome A05 in cultivated peanut (*Arachis hypogaea* L.). *BMC Genomics*, 18(1), 1–12.
34. Luo, H., Guo, J., Ren, X., Chen, W., Huang, L. et al. (2018). Chromosomes A07 and A05 associated with stable and major QTLs for pod weight and size in cultivated peanut (*Arachis hypogaea* L.). *Theoretical and Applied Genetics*, 131(2), 267–282.
35. Gomez Selvaraj, M., Narayana, M., Schubert, A. M., Ayers, J. L., Baring, M. R. et al. (2009). Identification of QTLs for pod and kernel traits in cultivated peanut by bulked segregant analysis. *Electronic Journal of Biotechnology*, 12(2), 3–4.
36. Shirasawa, K., Hirakawa, H., Nunome, T., Tabata, S., Isobe, S. (2016). Genome-wide survey of artificial mutations induced by ethyl methanesulfonate and gamma rays in tomato. *Plant Biotechnology Journal*, 14(1), 51–60.
37. Martín, B., Ramiro, M., Martínez-Zapater, J. M., Alonso-Blanco, C. (2009). A high-density collection of EMS-induced mutations for TILLING in Landsberg erecta genetic background of Arabidopsis. *BMC Plant Biology*, 9(1), 1–9.
38. Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H. et al. (2008). MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Molecular Genetics Genomics*, 279(3), 213–223. DOI 10.1007/s00438-007-0293-2.
39. Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C. et al. (2004). Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biology*, 4(1), 1–8. DOI 10.1186/1471-2229-4-12.
40. Xin, Z., Li, W., Barkley, N. A., Burow, G., Franks, C. et al. (2008). Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. *BMC Plant Biology*, 8(1), 1–14. DOI 10.1186/1471-2229-8-103.
41. Okabe, Y., Asamizu, E., Saito, T., Matsukura, C., Ariizumi, T. et al. (2011). Tomato TILLING technology: Development of a reverse genetics tool for the efficient isolation of mutants from Micro-Tom mutant libraries. *Plant Cell Physiology*, 52(11), 1994–2005. DOI 10.1093/pcp/pcr134.

42. Lu, X., Liu, J., Ren, W., Yang, Q., Chai, Z. et al. (2018). Gene-indexed mutations in maize. *Molecular Plant*, 11(3), 496–504. DOI 10.1016/j.molp.2017.11.013.
43. Uchida, N., Sakamoto, T., Kurata, T., Tasaka, M. (2011). Identification of EMS-induced causal mutations in a non-reference *Arabidopsis thaliana* accession by whole genome sequencing. *Plant Cell Physiology*, 52(4), 716–722.
44. Cofflet, T. A., Hammons, R. O. (1974). Inheritance of pod constriction in peanuts. *Journal of Heredity*, 65(2), 94–96.
45. Hake, A. A., Shirasawa, K., Yadawad, A., Sukruth, M., Patil, M. et al. (2017). Mapping of important taxonomic and productivity traits using genic and non-genic transposable element markers in peanut (*Arachis hypogaea* L.). *PLoS One*, 12(10), e0186113.
46. Guzzo, C. M., Ringel, A., Cox, E., Uzoma, I., Zhu, H. et al. (2014). Characterization of the SUMO-binding activity of the myeloproliferative and mental retardation (MYM)-type zinc fingers in *ZNF261* and *ZNF198*. *PLoS One*, 9(8), e105271.
47. Xhemalce, B., Riising, E. M., Baumann, P., Dejean, A., Arcangioli, B. et al. (2007). Role of SUMO in the dynamics of telomere maintenance in fission yeast. *PNAS*, 104(3), 893–898.
48. Kunapuli, P., Somerville, R., Still, I. H., Cowell, J. K. (2003). ZNF198 protein, involved in rearrangement in myeloproliferative disease, forms complexes with the DNA repair-associated HHR6A/6B and RAD18 proteins. *Oncogene*, 22, 3417–3423. DOI 10.1038/sj.onc.1206408.
49. Burda, P., Aebi, M. (1999). The dolichol pathway of *N*-linked glycosylation. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1426(2), 239–257. DOI 10.1016/S0304-4165(98)00127-5.

Table S1: Evaluation of PRTs for PRTs-mutant lines in “FH 6” mutant library

No.	HPW (g)	APL (mm)	APL (mm)	APL/APW	PCI
60211	209.31	32.96	18.47	1.79	0.83
60142	195.80	32.74	17.03	1.94	0.71
60521	194.23	28.98	16.59	1.76	0.89
60631	192.55	30.29	16.90	1.79	0.94
60201	190.77	34.64	16.98	2.05	0.49
FH6 (Wild Type)	182.08	30.40	17.04	1.79	0.89
60731	179.29	29.25	18.11	1.62	0.76
60361	167.27	31.62	16.08	1.97	0.91
60713	164.77	29.11	16.11	1.82	0.84
60061	160.23	26.87	14.16	1.91	0.99
60543	158.75	30.65	14.71	2.09	0.85
60143	158.67	32.77	16.58	2.00	0.32
60182	158.00	30.02	15.36	1.96	0.69
60171	157.44	30.02	15.67	1.93	0.42
60342	156.33	28.62	15.91	1.81	0.88
60111	155.36	28.89	16.41	1.76	0.95
60443	155.33	29.51	15.22	1.95	0.89
60832	155.20	27.96	15.99	1.76	0.90
60471	154.19	28.10	15.26	1.84	0.98

(Continued)

Table S1 (continued)					
No.	HPW (g)	APL (mm)	APL (mm)	APL/APW	PCI
60701	152.67	28.07	16.30	1.73	0.84
60743	151.50	30.09	16.50	1.83	0.80
60861	151.33	27.64	14.78	1.88	0.91
60041	150.00	31.16	14.15	2.21	0.86
60072	150.00	26.76	15.20	1.76	0.97
60292	148.56	26.99	14.55	1.86	0.93
60391	148.00	25.82	14.82	1.75	0.89
60891	147.33	27.67	15.12	1.85	0.97
60801	146.25	27.83	16.28	1.74	0.94
60671	145.86	26.54	15.52	1.71	0.91
60831	144.46	27.81	15.10	1.86	0.85
60333	144.31	28.86	15.75	1.84	0.99
60522	143.83	28.56	15.11	1.90	0.90
60233	143.44	27.11	15.49	1.75	0.46
60352	143.07	28.32	14.85	1.91	0.97
60371	142.40	29.95	15.03	2.00	0.84
60633	141.90	29.14	15.85	1.85	0.87
60838	141.86	29.47	15.11	1.97	0.66
60263	141.67	27.73	14.26	1.95	0.94
60202	140.65	29.56	15.08	1.97	0.77
60052	140.47	29.80	14.89	2.01	0.72
60133	138.47	30.13	15.26	1.99	0.89
60601	137.23	27.12	14.90	1.83	0.96
60883	136.06	25.80	14.39	1.80	0.96
60421	135.79	27.26	14.78	1.86	0.89
60621	135.31	28.88	14.90	1.94	0.89
60312	133.94	25.40	14.94	1.72	0.97
60092	132.50	24.86	15.06	1.66	0.97
60151	132.00	29.10	14.71	1.99	0.76
60031	130.50	30.65	15.28	2.01	0.37
60835	129.88	28.25	13.82	2.05	0.70
60872	128.00	24.59	14.11	1.75	0.93
60841	127.13	24.30	15.25	1.57	0.86

Table S2: Evaluation of PRTs for PRTs-mutant lines in “FH 8” mutant library

No.	HPW (g)	APL (mm)	APL (mm)	APL/APW	PCI
80631	239.64	33.50	17.55	1.91	0.85
80606	235.64	33.57	16.99	1.98	0.88
80105	232.93	35.75	18.01	1.99	0.84
80796	231.50	32.12	17.31	1.86	0.88
80684	231.27	33.35	17.46	1.91	0.86
80782	231.23	37.27	18.28	2.04	0.85
80751	224.80	31.35	17.69	1.78	0.91
80681	216.40	35.15	16.72	2.11	0.86
80253	214.61	32.89	16.97	1.94	0.86
80286	214.54	35.11	17.32	2.04	0.80
80692	213.33	36.37	18.02	2.02	0.76
80625	213.31	33.80	16.71	2.03	0.98
80302	212.13	32.74	18.42	1.78	0.73
80493	209.25	32.36	15.78	2.05	0.99
80291	208.45	33.87	16.42	2.07	0.80
80296	203.00	31.91	15.87	2.01	0.94
80812	202.60	31.26	17.04	1.84	0.92
80491	202.25	33.05	16.79	1.98	0.78
80463	202.08	29.95	16.95	1.78	0.77
80621	201.57	31.32	16.03	1.96	0.95
80273	200.57	32.62	17.06	1.92	0.98
FH8 (Wild Type)	189.87	32.86	17.18	1.91	0.95
80451	186.21	28.77	17.20	1.68	0.80
80122	185.25	33.84	16.61	2.04	0.77
80794	178.36	31.54	16.50	1.92	0.97
80756	177.54	33.93	17.39	1.96	0.94
80262	174.64	31.54	16.79	1.88	0.75
80121	174.58	32.52	16.49	1.97	0.78
80644	170.55	27.64	15.92	1.74	0.97
80591	169.36	29.18	15.89	1.84	0.84
80263	168.58	31.75	16.33	1.96	0.81
80502	167.92	29.87	16.19	1.85	0.83
80256	166.82	31.50	16.53	1.91	0.83
80623	166.07	31.33	15.59	2.02	0.93
80851	165.54	30.63	15.72	1.96	0.93

(Continued)

Table S2 (continued)					
No.	HPW (g)	APL (mm)	APL (mm)	APL/APW	PCI
80271	163.00	29.47	15.10	1.97	0.75
80222	163.00	28.28	16.10	1.76	0.93
80203	162.88	28.54	16.39	1.74	0.93
80243	162.75	29.72	16.20	1.83	0.87
80671	162.29	29.16	15.51	1.89	0.91
8070	161.15	33.37	16.76	2.00	0.47
80287	159.38	30.94	14.68	2.11	0.75
80652	158.77	30.16	16.22	1.87	0.91
80061	158.70	32.30	15.82	2.06	0.92
80633	157.10	27.62	15.13	1.83	0.92
80301	156.08	28.69	16.10	1.79	0.81
80421	155.57	29.16	16.92	1.73	0.84
80813	155.17	32.89	17.09	1.93	0.96
80073	154.27	29.98	15.47	1.95	0.87
80341	153.17	27.12	15.45	1.76	0.97
80181	152.27	29.09	15.08	1.94	0.87
80276	150.22	27.76	14.76	1.88	0.96
80391	150.11	29.16	14.87	1.97	0.94
80353	148.77	26.47	14.49	1.83	0.97
80465	148.50	26.79	14.87	1.81	0.86
80281	148.36	32.33	16.12	2.02	0.74
80022	147.43	29.83	16.28	1.83	0.72
80581	146.46	27.99	14.71	1.90	0.86
80204	146.40	26.98	15.48	1.75	0.90
80695	145.07	27.39	15.05	1.83	0.93
80124	144.31	32.70	16.48	1.99	0.84
80223	139.50	29.56	16.54	1.79	0.90
80081	137.31	27.89	15.20	1.84	0.89
80321	136.08	28.53	14.99	1.91	0.97
80842	135.40	30.43	16.91	1.81	0.90
80231	134.25	28.09	14.43	1.96	0.86
80013	133.40	28.91	16.03	1.82	0.91
80761	131.27	29.02	15.50	1.88	0.87
80542	127.77	28.63	13.96	2.05	0.76
80342	125.43	24.33	15.04	1.63	0.73
80464	124.15	26.85	16.16	1.67	0.84

Table S3: Evaluation of PRTs for PRTs-mutant lines in “KH 1” mutant library

No.	HPW (g)	APL (mm)	APL (mm)	APL/APW	PCI
1025	172.35	31.27	16.92	1.85	0.91
10188	160.65	30.21	16.72	1.81	0.90
11237	171.22	31.71	17.26	1.84	0.89
KH 1	192.25	33.26	18.29	1.82	0.95