Development and screening of EMS mutants with altered seed oil content or fatty acid composition in *Brassica napus*

Tang, Shan; Liu, Dong-Xu; Lu, Shaoping; Yu, Liangqian; Li, Yuqing; Lin, Shengli; Li, Long; Du, Zhuolin; Liu, Xiao; Li, Xiao; Ma, Wei; Yang, Qing-Yong; Guo, Liang

2020

Tang, S., Liu, D.-X., Lu, S., Yu, L., Li, Y., Lin, S., ... Guo, L. (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. doi:10.1111/tpj.15003

<https://hdl.handle.net/10356/144905>

<https://doi.org/10.1111/tpj.15003>

This is the accepted version of the following article: Tang, S., Liu, D.-X., Lu, S., Yu, L., Li, Y., Lin, S., ... Guo, L. (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. doi:10.1111/tpj.15003, which has been published in final form at <http://dx.doi.org/10.1111/tpj.15003>. This article may be used for non-commercial purposes in accordance with the Wiley Self-Archiving Policy [<https://authorservices.wiley.com/authorresources/Journal-Authors/licensing/self-archiving.html>].

Development and screening of EMS mutants with altered seed oil content or fatty acid composition in *Brassica napus*

Shan Tang¹, Dong-Xu Liu², Shaoping Lu¹, Liangqian Yu¹, Yuqing Li¹, Shengli Lin¹, Long Li¹, Zhuolin Du¹, Xiao Liu¹, Xiao Li¹, Wei Ma³, Qing-Yong Yang^{2*}, Liang Guo^{1*}

¹National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China

²Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, China

³School of Biological Sciences, Nanyang Technological University, 637551, Singapore

*Corresponding author: yqy@mail.hzau.edu.cn; guoliang@mail.hzau.edu.cn

Running title: Rapeseed EMS mutants with altered oil biosynthesis

SUMMARY

Brassica napus is an important oilseed crop in the world and the mechanism of seed oil biosynthesis in *B. napus* remains unclear. In order to study the mechanism of oil biosynthesis and generate germplasms for breeding, an EMS mutant population with ~100,000 M2 lines was generated using Zhongshuang 11 as the parent line. The EMS-induced genome-wide mutations in M2-M4 plants were assessed. The average number of mutations including single nucleotide polymorphisms and insertion/deletion in M2-M4 was 21,177, 28,675 and 17,915, respectively. The effects of the mutations on gene function were predicted in M2-M4 mutants, respectively. We screened the seeds from 98,113 M2 lines and 9,415 seed oil content and fatty acid mutants were identified. We further confirmed 686 mutants with altered seed oil content and fatty acid in advanced generation (M4 seeds). Five representative M4 mutants with increased oleic acid were re-sequenced and the potential causal variations in *FAD2* and *ROD1* genes were identified. This study generated and screened a large scale of *B. napus* EMS mutant population and the identified mutants could provide useful genetic resources for the study of oil biosynthesis and genetic improvement of seed oil content and fatty acid composition of *B. napus* in the future.

Keywords: *Brassica napus*, EMS, fatty acid, oleic acid, oil content, whole-genome sequencing

Significance Statement: Seed oil content and fatty acid composition are important traits of oil crop *Brassica napus*. In this study, a large number of *Brassica napus* EMS mutants with altered seed oil content or fatty acid composition have been generated and identified. This study provides useful genetic resources for functional genomics study and genetic improvement of *Brassica napus*.

INTRODUCTION

Rapeseed (*Brassica napus* L.; AACC, $2n = 38$) plays an important role in the supply of vegetable oil and rapeseed oil is the third largest source of vegetable oil in the world ranked after soybean oil and palm oil (USDA ERS, 2018). In China, the oil content of rapeseed cultivars is about 40%-45%, 3-5% lower than that of foreign cultivars (Wang and Yang, 2007). It's estimated that an increase of oil content by 1% equals an increase of the seed yield by 2%-3% (Wang *et al.*, 2004). Fatty acid composition is the main factor which determines the nutritional quality of vegetable oils for human health and breeding high oleic acid oil with reasonable level of other fatty acids is one of the major goals of rapeseed breeding (Zhao *et al.*, 2019). Thus, understanding of the mechanism of oil biosynthesis is of great importance for the improvement of seed oil content and fatty acid composition of rapeseed. In *Arabidopsis*, there are more than 700 known genes involved in lipids metabolism and many genes have been shown to regulate the seed oil content or fatty acid composition (Li-Beisson *et al.*, 2013). However, rapeseed is an allotetraploid crop and the oil content is a complex quantitative trait. The mechanism of lipid metabolism remains unclear in rapeseed. In recent years, although seed oil- and fatty acid-related QTLs were mapped by many studies, few valuable genes have been cloned and applied in breeding of high-oil rapeseed varieties (Hua *et al.*, 2016).

Germplasm resources are of great importance in crop functional genomics and breeding application. There are several ways to obtain mutants in large scale, such as T-DNA insertion, physical- or chemical-induced mutagenesis, which offer great convenience for crop research and breeding. Thousands of genes in *Arabidopsis* have been cloned and the mechanisms have been well understood due to the usage of T-DNA mutants (Alonso *et al.*, 2003; Rosso *et al.*, 2003). The most widely used physical mutagens include electromagnetic, ionizing radiations such as gamma rays and X rays (Rédei, 2008). For example, gamma ray ^{60}Co was used to irradiate a popular variety of *Brassica juncea*, resulting in the acquisition of two high yielding mutants which were developed into commercial cultivars subsequently (Malek *et al.*, 2012). Ethyl methanesulfonate (EMS) is a commonly used chemical mutagen which alkylates the guanine base, resulting in primarily C/G to T/A base transitions (Kohalmi and Kunz, 1988). The creation of mutant materials by EMS produces stable genetic variation at the genome level with high density of point mutation and valuable mutants can be

directly applied to breeding without gene-modified (GM) controversy (Steven *et al.*, 2004). EMS-induced mutagenesis has been widely used in different crops such as soybean (Espina *et al.*, 2018; Lakhssassi *et al.*, 2017), maize (Lu *et al.*, 2018b), rice (Serrat *et al.*, 2014; Takagi *et al.*, 2015), wheat (Chen *et al.*, 2012), rapeseed (Amosova *et al.*, 2019; Harloff *et al.*, 2012; Wang *et al.*, 2008) and tomato (Shirasawa *et al.*, 2016). EMS-induced mutagenesis has been shown to be an effective way to obtain germplasm resources in the quality breeding of rapeseed, especially in the development of rapeseed cultivars with high oleic acid and low linolenic acid. The first low linolenic acid mutant M11 was obtained by treating cultivar Oro with the chemical mutagen (Rakow, 1973). The cultivar Stellar producing seed oil with 3% linolenic acid was developed using M11 in Canada (Scarth *et al.*, 1988). Likewise, the first high oleic acid mutant X-82 was produced by EMS from cultivar Cascade (Auld *et al.*, 1992) and it was used to develop the first high oleic acid cultivar Expander in Germany (Rucker and Robbelen, 1996). However, generation and screening of EMS mutants with altered seed oil content are rarely reported or applied in rapeseed breeding.

Whole-genome sequencing serves as a powerful tool in functional genomics research in plants. Applying whole-genome sequencing in EMS mutants can effectively detect EMS-induced mutations and insertion/deletion (InDel) mutations in the genome scale. These mutations existing in different regions contribute to different kinds of mutation types and nonsense and frameshift mutations are most likely to change the gene function (Greene *et al.*, 2003; Li *et al.*, 2017; Lu *et al.*, 2018b). These mutations may be used to explain the phenotypic variations and candidate genes controlling the phenotypic variations can be cloned and characterized (Lu *et al.*, 2018b). In recent years, forward genetic approach such as MutMap allows rapid identification of causal genes in mutant lines. Candidate genes in EMS mutants were successfully identified in rice and maize by MutMap (Abe *et al.*, 2012; Lu *et al.*, 2018b; Takagi *et al.*, 2015).

In this study, a large EMS-mutagenized *B. napus* mutant population was generated and we focused on identification of mutants with altered seed oil content or fatty acid composition. We identified 1,294 mutants with altered morphology during seedling and flowering stage. Thousands of valuable mutants with altered seed oil content or fatty acid composition were identified. This work provides abundant EMS mutants for the study of oil biosynthesis mechanisms in *B. napus* and valuable germplasms for *B. napus* breeding.

RESULTS

Development of EMS-induced mutant population

Zhongshuang 11 was used as the parent line to develop the EMS population because it is an elite semi-winter cultivar in China and its high-quality reference genome is available (Song *et al.*, 2020). The outline of the EMS-induced mutant population was shown in Figure 1. Primarily, seeds were treated with different concentrations of EMS to identify the appropriate condition (Serrat *et al.*, 2014; Wang *et al.*, 2014). Germination rates were recorded at 48 hours. When seeds were treated with 0.4%, 0.8%, 1.2% and 1.6% EMS, the germination rate was 86.5%, 64.0%, 25.5% and 16.0%, respectively (Figure S1). Based on the germination rate, the seeds were mutagenized with 0.8% EMS to produce the M1 population with approximately 10,000 plants which were bagged during flowering and harvested after maturation. M2 seeds were bulked and sowed to produce the M2 population with approximately 100,000 M2 plants. Identification of mutants with phenotypic variations during seedling and flowering stages was performed in M2 population. Meanwhile, ten M2 plants were randomly selected for evaluation of EMS-induced mutations via whole-genome sequencing. M3 seeds from 98,113 individual plants were analyzed by NIRS to screen oil content and fatty acid mutants. We selected 1,886 mutant lines with altered seed oil content or fatty acid composition to confirm the phenotype in next generation. Likewise, analysis of EMS-induced mutations was performed in five representative M3 plants. Stable mutants were confirmed using M4 seeds and five representative high oleic acid mutant lines were used to analyze the EMS-induced mutations and identify genomic variations controlling the level of oleic acid.

Screening of mutants with morphological variations

Morphological mutant survey was performed in the M2 population with ~100,000 plants and 1,294 growth- and development-related mutants were identified. It showed that dwarf, rolled-leaf and male sterility represented the top three types of morphological variations (Figure 2a). During seedling stage, 268 mutants with altered leaf morphology, leaf color and stem were identified. Representative morphological mutants at seedling stage were shown in Figure 2b. During flowering stage, a mass of typical flowering- and plant architecture-related mutants were identified (Figure 2a,c). Flowers with shrunken petal and exposed stigma, flower color with dark yellow and light yellow and double main stems mutants were identified. It is worthwhile to mention

that dwarf mutants have exhibited phenotypic diversities such as dwarf, rod-like without branch; dwarf, rod-like with branch; smaller version with normal plant type (Figure 2c). We focused on identification of the seed oil content and fatty acid composition mutants in this study, and hence M3 seeds harvested from the morphological mutants were kept in the seed storage chamber for future study.

Whole-genome sequencing and characterization of EMS-induced mutations

To analyze the genome-wide EMS-induced mutations, 25 lines were re-sequenced, including five wild-type (WT), ten M2, five M3 and five M4 plants. We obtained a total of 917 Gb paired-end reads and the average sequencing depth was 37.55× for the 20 EMS mutants. The cleaned short read sequences were mapped to the Zhongshuang 11 reference genome (Song *et al.*, 2020) and the average genome coverage was 95.64% (Table S1). We used the genome analysis tool kit (GATK) pipeline (McKenna *et al.* 2010) to detect mutations in the EMS-induced mutants (Table S1). As reported, confirming the genetic identity of research material is necessary to know whether the phenotypic difference was caused by the gene-environment interactions (Bergelson *et al.*, 2016; Michno and Stupar, 2018). To confirm the genetic identity of our WT lines, we re-sequenced 5 WT of ‘Zhongshuang 11’ and analyzed the share and specific mutations in the five WT lines (Figure S2a). The percent of shared mutations by at least two wild-type lines was 72.68% (Figure S2b). The number of mutations in 20 mutant plants was 2,344,536. And 1,409,522 common base changes shared by WT were filtered out as described in the methods. Meanwhile, 490,288 mutations that were common to more than one mutant plants were also removed. Finally, 444,726 loci were considered as induced mutations including SNP and InDel in 20 mutant plants. The average number of SNP and InDel in individual of 20 mutant plants is 21,300 (between 3,162 and 40,651) and 936 (between 419 and 4,204). The average number of homozygous SNP and InDel in 20 mutant plants is 5,167 (between 79 and 17199) and 66 (between 27 and 174). We further characterized the genome-wide mutations in the 20 sequenced mutant lines. The mutation distribution and density on each chromosome of 20 mutants is shown in Figure 3a. The number and ratio of mutations in 20 mutants are shown in Figure 3b. The average density of SNP and InDel in 20 mutant plants is 1 SNP/76.28 kb (between 1 SNP/23.64 kb and 1 SNP/303.86 kb) and 1 InDel/1.29 Mb (between 1 indel/228.55 kb and 1 InDel/2.29 Mb). The induced SNP and InDel density in individual mutant plants is shown in Table S1.

Determination of functional effects of EMS-induced mutations

The type of base changes of mutants in M2, M3 and M4 generations were shown in Figure 3b. In M2 mutants, C/G to T/A transitions were predominant (86.52%) followed by A/T to G/C transitions (4.13%), A/T to T/A transversions (3.21%), C/G to A/T transversions (3.01%), A/T to C/G transversions (2.00%) and C/G to G/C transversions (1.12%). In M3 mutants, C/G to T/A transitions were predominant (87.90%) followed by A/T to G/C transitions (3.79%), C/G to A/T transversions (2.78%), A/T to T/A transversions (2.66%), A/T to C/G transversions (1.75%) and C/G to G/C transversions (1.11%). In M4 mutants, C/G to T/A transitions were predominant (86.11%) followed by A/T to G/C transitions (4.44%), A/T to T/A transversions (3.21%), C/G to A/T transversions (3.11%), A/T to C/G transversions (1.88%) and C/G to G/C transversions (1.25%). The transition to transversion ratios of M2, M3 and M4 mutants were 9.71 (184,777/19,038), 11.03 (124,867/11,316) and 9.59 (77,876/8,124), respectively.

The size of the InDels is mostly between 1 bp and 12 bp. The most frequent InDel observed was 1 bp deletion (2,629) and 1 bp insertion (2,420) in M3 mutants (Figure 4a). The mutations in 20 mutants occur more often in intergenic regions (38.39%) and depleted in 5'UTR (0.28%) and 3'UTR (0.29%) (Figure 4b). The functional effects of the mutations were predicted utilizing the SnpEff program (Cingolani *et al.* 2012) and those mutations were classified into four categories based on mutation types: high-impact (nonsense mutations and frameshift mutations), moderate-impact (missense mutations), modifier-impact (intron and intergenic mutations) and low-impact (synonymous mutations). The ratio of the mutations with various functional effects in 20 mutants was shown in Figure 4c. The predominant effect was intron and intergenic mutations (92.91%). There was only 0.52% high-impact (nonsense mutations and frameshift mutations). The number and ratio of homozygous and heterozygous mutations in 20 mutants is shown in Figure S3a and Figure S3b. As shown in Figure 4d, the homozygous percent of mutations in the M2, M3 and M4 mutants were 1.81%, 34.97% and 57.03%.

Screening of mutants with altered seed oil content or fatty acid composition

M3 seeds harvested from 98,113 M2 open-pollinated plants were analyzed by NIRS. As shown in Table S2, the seed oil content of WT was 51.94±1.11%, and the compositions of oleic acid (C18:1), linoleic acid (C18:2) and linolenic acid (C18:3) of WT were 75.45±1.66%, 14.06±0.99% and 6.46±0.45%, respectively. Samples with

abnormal NIRS values were not included in further round of mutant screening. We set two screening criteria for selection of the mutants based on the NIRS data of 98,113 seed samples. Under screening criteria 1, 7,280 mutants with altered oil content were obtained. There were 1,177 mutants with increased oil content and 6,103 mutants with decreased oil content. In addition, 3,540 C18:1, 1,190 C18:2 and 658 C18:3 mutants were achieved. In total, 9,415 seed oil content- and fatty acid-related mutants were identified. Under screening criteria 2, 1,430 mutants with altered oil content were obtained. There were 120 mutants with increased oil content and 1,310 mutants with decreased oil content. In addition, 1,213 C18:1, 355 C18:2 and 204 C18:3 mutants were obtained. In total, 1,886 seed oil content- and fatty acid-related mutants were identified under screening criteria 2 (Table S2).

Identification of stable mutants

To further identify stable mutants in advanced generations, the above 1,886 seed oil content- and fatty acid-related mutants were chosen, and each mutant line was planted in the field (ten plants per line). M4 open-pollinated seeds were harvested from six individual plants for each mutant line and analyzed by NIRS to determine the content of oil and composition of fatty acids. The average oil content and fatty acid composition for each mutant line was calculated based on six individual plants of each mutant line. Figure 5 depicted the distribution of the oil content and fatty acids level in M4 seeds of the 1,886 mutant lines. Specifically, 63 mutant lines displayed an increase by 2% oil content and 151 mutant lines had a decrease by 5% oil content; C18:1 increased by 2% in 164 mutant lines and decreased by 5% in 417 mutant lines; For C18:2, 150 mutant lines showed an increase by 2% and 14 mutant lines had a decrease by 5%; For C18:3, 63 mutant lines displayed an increase by 2% and 38 mutant lines had a decrease by 2% (Table S3). In total, 686 mutant lines were identified as stable mutants based on analysis of the M4 seeds using 6 individual plants of each line. These identified mutants represented stable mutants with high reliability and some mutants could be chosen for conducting gene mapping and breeding. Furthermore, we will continue to plant the mutants to further confirm the phenotypes and other vital agronomic traits.

Characterization of mutants with extreme oil content and fatty acid

We selected representative M4 seeds with altered oil content or fatty acid composition from stable mutants. As shown in Table 1, these representative mutants were selected and re-analyzed by NIRS. The highest oil content was $55.09 \pm 0.77\%$ and the lowest was

33.91±1.24%. The highest C18:1 was 83.42±0.92% and the lowest was 52.22%. The highest C18:2 was 21.07±0.76% and the lowest was 7.99±0.22%. The highest C18:3 was 12.51% and the lowest was 2.87±0.26%. Oil content and fatty acid composition of above representative mutants were analyzed and confirmed by GC-FID (Figure 6a-d). The results indicated that oil content in WT was 48.70±1.16%, M4-1027 and M4-1451 reached 51.64±1.01% and 51.41±1.03%, respectively. M4-42 and M4-236 reached 37.43±0.49% and 30.92±0.52%, respectively. Meantime, fatty acid composition in these oil content mutants was significantly modified compared to WT (Figure 6a). The composition of C18:1 in WT was 65.75±1.03% while M4-1039, M4-42, M4-371 and M4-1446 were 61.42±1.11, 56.89±4.69, 76.70±1.66% and 69.57±0.02%, respectively. In addition to C18:1, other fatty acid composition was significantly altered in M4-1039, M4-42, M4-371 and M4-1446 (Figure 6b). The level of C18:2 in WT was 18.38±0.01% while M4-1876, M4-1875, M4-236 and M4-42 were 7.58±0.59%, 8.55±0.93%, 20.96±0.10% and 22.83±1.61%, respectively. Meantime, other fatty acids, especially C18:1 and C18:3, possessed a significant difference in the C18:2 mutants (Figure 6c). The level of C18:3 in WT was 8.24±0.04%, while C18:3 levels in M4-947, M4-926, M4-42 and M4-1039 were 7.18±0.58%, 7.06±0.73%, 11.20±0.81% and 10.67±0.08%, respectively (Figure 6d). The results analyzed by GC-FID indicated that the oil content and fatty acids data collected by NIRS were reliable, which exhibited a strong consistent trend.

Acquisition of mutants with increased oil content and C18:1

Breeding cultivars with high oil and high C18:1 is one of the main goals in breeding program of *B. napus* (Hua *et al.*, 2016; Wu *et al.*, 2018). In order to obtain mutants with increased oil content and C18:1, we screened and identified four mutant lines (M4 seeds) with both oil content and C18:1 increased by more than 2% (Table S4). Among these mutant lines, M4-1121 possessed 3.13% increase of oil content and 4.78% increase of C18:1. GC-FID analysis was conducted to verify the oil content and C18:1 level. The results showed that the oil content of M4-1120, M4-1121, M4-1346 and M4-1444 was 55.23±1.09%, 55.78±1.07%, 52.75±1.00% and 52.82±1.04%, respectively. These four mutants displayed a significant increase in oil content compared to WT (Figure S4a). In addition, C18:1 level in M4-1120, M4-1121, M4-1346 and M4-1444 was 70.45±2.02%, 69.54±1.50%, 69.01±1.40% and 70.38±1.70%, respectively. These four mutants had an increase of C18:1 by 3.26 to 4.7% compare to WT (Figure S4b). Taken

together, these results indicated that we indeed acquired four reliable mutants which possessed elevated oil content and C18:1.

Identification of mutations in *FAD2* and *ROD1* genes responsible for elevated C18:1 by whole-genome sequencing

FAD2 and *ROD1* play direct roles in the regulation of C18:1 level (Lu *et al.*, 2009). We have identified 164 mutant lines (M4 seeds) with elevated C18:1 (Table S3). In order to identify the potential causal mutations in genes of *FAD2* and *ROD1*, we re-sequenced five mutants whose C18:1 was elevated by 7.50-9.95% (Table 2). The genomic variations of *FAD2* and *ROD1* of five high C18:1 mutants were identified by analyzing the mapped reads in four *FAD2* and eight *ROD1* gene regions. The results indicated that C/G to T/A transitions were detected in exon and/or intergenic regions of *FAD2* and/or *ROD1*. One missense variant in *FAD2* and one synonymous variant in *ROD1* were detected in M4-371, which might result in the increase of C18:1 by 9.95% compared to WT. One missense variant was existed in *FAD2* and one upstream gene variant was existed in *ROD1* in M4-1875. One upstream gene variant existed in *FAD2* and one missense variant existed in *ROD1* in M4-1449. One missense variant and one upstream variant in *FAD2* were detected in M4-1876. In M4-1429, one intron variant was detected in *ROD1* (Table 2). Our results obtained here suggested that these mutations in *FAD2* and *ROD1* might potentially alter the function of *FAD2* and *ROD1*, which led to increase of C18:1 by 7.50-9.95%. In addition, we found that the major agronomic traits such as plant height and yield per plant of these four mutants were comparable to the WT, indicating that these mutants might be promising mutants in rapeseed breeding (Table S5). These new alleles of *FAD2* and *ROD1* which had been generated through EMS mutations in this study may have important values in breeding new *B. napus* varieties with high oleic acid oil.

DISCUSSION

The success of obtaining desired mutants depends on the plant material characteristics, EMS concentration, population size, etc. In this study, Zhongshuang 11 has been chosen as the parent line because it is an elite semi-winter cultivar with released high-quality reference genome (Song *et al.*, 2020). The concentration of EMS plays a vital role in determination of the mutation frequencies. Based on previous EMS mutagenesis studies in *B. napus* (Amosova *et al.*, 2019; Harloff *et al.*, 2012; Wang *et al.*, 2008) and our examinations of various EMS concentrations on the germination rate, 0.8% EMS was

applied in this study. Previous EMS mutagenesis studies in crops usually screened a few thousands of mutagenized plants and numerous mutants have been successfully identified in peanut (Fang *et al.*, 2012), cucumber (Wang *et al.*, 2014), eggplant (Xi-ou *et al.*, 2017), soybean (Espina *et al.*, 2018), rice (Takagi *et al.*, 2015) and *B. napus* (Amosova *et al.*, 2019; Harloff *et al.*, 2012; Wang *et al.*, 2008). In order to identify adequate mutants for gene cloning and breeding, we generated a M2 population with approximately 100,000 plants in this study. With this large M2 population, we identified 1,294 mutants with morphological variations during seedling and flower developmental stages, among which, dwarf phenotype accounted for the highest ratio, which was similar to previous studies in soybean and eggplant (Tsuda *et al.*, 2015; Xi-ou *et al.*, 2017). Most importantly, we successfully identified 9,415 mutant lines (M3 seeds) with altered oil content or fatty acid composition. Thus, our study confirmed that EMS-induced mutation was an effective way to create new germplasm resources to broaden the genetic pool in *B. napus*. *B. napus* is a predominant self-pollination crop and the outcrossing rate could be over 12% depending on the environment and genotype (Becker *et al.*, 1992). To harvest healthy seeds for screening of mutants, we did not bag the M2 and M3 plants to avoid cross-pollination because seed development and oil biosynthesis may be affected. Thus, small chance of cross-pollination may occur in the population.

Whole-genome sequencing has been applied in the survey of mutations induced by mutagens in several crop species and mutations including SNP, InDel and CNV could be identified and characterized (Sevanthi *et al.*, 2018; Shirasawa *et al.*, 2016; Tsuda *et al.*, 2015; Xiao *et al.*, 2019). In our study, whole-genome sequencing was applied to survey the mutations in 20 mutant plants in different generations. The type, number and region of mutations were characterized and compared among individual lines and among different generations. As a mutagen, EMS alkylates the guanine base, mainly resulting in C/G to T/A transitions (Greene *et al.*, 2003). In our research, mutation analysis in the M2, M3 and M4 mutants showed that C/G to T/A transitions accounted for the predominant type (Figure 3b), which was consistent with the previous observations in soybean, tomato and maize (Lu *et al.*, 2018b; Shirasawa *et al.*, 2016; Tsuda *et al.*, 2015). However, we also detected a variety of non-C/G to-T/A-transitions in the EMS mutants as reported in previous studies (Lu *et al.*, 2018b; Tsuda *et al.*, 2015). Mutants in different generations showed similar pattern of mutagenesis except M2-2, M2-4 and M2-7. The number of mutations is similar between individual M2 and M3

mutants except M2-2, M2-4 and M2-7 (Figure 3b). We suspect this is because some seeds may not completely absorb EMS solution and the mutagenesis is efficiently induced in M2-2, M2-4 and M2-7. The number of mutations in M4 mutants is smaller which is possible that some of the mutations lost during reproduction. Point mutations occurring in the exon region is most likely to disrupt the gene function. In our study, 6.69% of whole variations existed in exon (Figure 4b), which is similar to the results in maize with 5.4% (Lu *et al.*, 2018b). Besides, high-impact mutations are more likely to affect gene function disruption and approximately 0.52% of the variations had high-impact effects on gene function in our study (Figure 4c). The ratio of high-impact mutations was low and a large M2 EMS population guaranteed successfully obtaining of adequate mutants in our study.

This study aimed to generate a large EMS-induced mutant population and focus on identification of mutants with altered seed oil content or fatty acid composition. Seed oil content is a complex trait which is controlled by hundreds of genes in *B. napus* (Lu *et al.*, 2018a). Although we set criteria 1 to screen the seeds of mutants, 9,415 seed oil content- and fatty acid-related mutants were identified from 98,113 M2 mutant lines. It is possible that the growth environment and the measurement by NIRS might cause variations in the quantification of oil content and fatty acid composition. In order to select reliable mutants and reduce the workload, we selected 1,886 mutant lines based on criteria 2 demonstrated in Table S2. We continued to grow 1,886 mutant lines to confirm the traits and 686 seed oil content- and fatty acid-related mutants were further confirmed by NIRS analysis of the M4 seeds. M4 seeds collected from six M3 plants for each mutant line were analyzed to confirm the phenotype of the mutant lines. It is possible that multiple mutant lines out of confirmed 686 mutant lines may come from shared M1 parent plants and show same phenotype because bulked M2 seeds were sowed for screening of mutants. Thus, these 686 mutant lines will be further grown and characterized in future studies.

The identified mutants with altered seed oil content or fatty acid composition were valuable resources for studying the mechanism of plant oil biosynthesis. Stable mutants can be used to clone the causal genes responsible for the phenotypes by forward genetics approach such as map-based cloning and MutMap. MutMap has been shown to be a rapid and effective approach for identifying the mutated gene of a mutant by whole-genome sequencing of bulked DNA from segregating progenies (Abe *et al.*, 2012; Takagi *et al.*, 2015). The high-quality reference genome was released for the

parent line Zhongshuang 11, the mutants identified in this study could be used for rapid cloning of the genes controlling lipid metabolism in *B. napus*. For example, M4-42 possessed the lowest oil content ($33.91 \pm 1.24\%$), which had a reduction of 17.41% of oil content compared to WT ($51.32 \pm 1.52\%$). M4-42 mutant line could be crossed with the parent line to produce F2 progeny to be used in MutMap for identifying the genes which resulted in decreased oil content. There are some promising mutants with elevated C18:1 and these types of mutants have been reported in previous studies of oilseed crops such as peanut and soybean (Fang *et al.*, 2012; Lakhssassi *et al.*, 2017; Lu *et al.*, 2009). In our study, whole-genome sequencing was performed on five mutants with elevated C18:1 and sequence variations of *FAD2* and *ROD1* were identified (Table 2). We speculated the elevated C18:1 phenotype was due to either the altered expression of *FAD2* or *ROD1* or the changes of the enzyme activity of FAD2 or ROD1. Of course, we could not exclude a possibility that alternative genes may cause the elevated C18:1 level in these mutants. Hence, candidate genes with mutations could be cloned and verified by experimental analysis in the future. In addition, there are multiple functional *FAD2* and *ROD1* genes in the genome of *B. napus*, stacking of the mutated *FAD2* and *ROD1* genes may be further increase the level of oleic acid for breeding utilization.

In this study, we have built a large EMS mutant population of *B. napus* and identified a large scale of oil content- or fatty acid composition-related mutants. These mutants can not only be exploited to study the lipid metabolism in *B. napus*, but also have the translational values to be applied in the genetic improvement of *B. napus*.

EXPERIMENTAL PROCEDURES

Plant material, EMS mutagenesis, growth condition and mutant survey in the field

Mutagenesis was performed as described by Zhang *et al.* with modifications (Zhang *et al.*, 2010). An elite cultivar Zhongshuang11 released in 2008 by Oil Crop Research Institute (China) was chosen as parent line. Zhongshuang 11 was propagated by selfing for 6 generations after release and bulked seeds collected from multiple selfed plants were used for EMS treatment. The overall scheme of EMS-induced mutant population is shown in Figure 1. Approximately 37,500 seeds were soaked in a 0.8% (v/v) EMS (Sigma-Aldrich (Shanghai) Trading Co. Ltd) for 8h. EMS-treated seeds were shaken at low speed and washed with running tapped water for 4h. The treated seeds were sowed in Lanzhou, Gansu province in May 2017. The seeds of approximately 10,000 bagged individual M1 plants were harvested in September 2017. Bulk M2 seeds were planted

in Ezhou, Hubei province in October 2017. Seeds were harvested from approximately 100,000 M2 plants in May 2018. 1,886 selected M3 mutant lines were sowed in Ezhou, Hubei province in October 2018. Open-pollinated seeds were harvested in May 2019. Mutants with altered leaf morphology, leaf color, stem, flower, plant architecture, siliques and flowering time were surveyed during the seedling stage (50 days after sowing) and flowering stage (160 days after sowing) in the field. The identified mutants with growth phenotype were recorded and then bagged during flowering stage. The field management including weeding, fertilization, irrigation and pest control was carried out in a conventional way.

DNA extraction and whole-genome sequencing

Young leaves at seedling stage were randomly sampled from ten M2, five M3, and five high oleic acid M4 mutants and frozen in liquid nitrogen immediately. Total DNA extraction was performed using DNasecure Plant Kit (TIANGEN BIOTECH (BEIJING) CO., LTD). Qualified 50 ng genomic DNA were randomly interrupted by Covaris crusher with a length of 350 bp. TruePrep DNA Library Prep Kit V2 for Illumina was used to build the library. The whole library was prepared through the steps of terminal repair, ployA tail, sequencing adapter, purification, and PCR amplification. The qualified library was sequenced by Illumina HiSeq 2000.

Identification and functional annotation of EMS-induced mutations

Clean reads were mapped to the *B. napus* reference genome (Song *et al.*, 2020) using BWA-MEM (Li and Durbin, 2009) with default parameters. SAMtools (Etherington *et al.* 2015) was used to transform the alignment outputs from sam to bam format. SNPs and InDels were identified using Genome Analysis Tool Kit (McKenna *et al.* 2010). Picard (version 1.115) was used to sort, mark duplicates, add read group information, and index the bam file, following which local realignment was performed using the GATK InDelRealigner tool. Based on the realigned bam results, SNPs and InDels were called using GATK UnifiedGenotyper and SAMtools mpileup, respectively. Raw mutations in M2, M3 and M4 mutants were identified after mapping the clean reads to the *B. napus* reference genome of Zhongshuang 11. In raw mutations, the common SNP/InDel polymorphisms shared by five WT plants was filtered and the mutations that were common to more than one mutants were also removed using in-house python scripts. The annotation and effects of mutations on gene function were predicted using SnpEff (Cingolani *et al.* 2012).

Measurement of seed oil content by near infrared spectrometer (NIRS)

B. napus seeds were harvested from approximately 100,000 M1, 1,886 M3 and representative M4 plants after maturation and then naturally air dried before NIRS analysis. About 5 g desiccated seeds were used to measure the oil content and fatty acid composition by a Foss NIRSystems 5000 NIRS. In order to be consistent, we used one instrument for analysis of the seed samples. We included one WT seed sample for every 200 mutant samples to make sure the instrument was stable and the data was reliable. The relevant parameter settings were described by Gan et al (Gan *et al.*, 2003).

Gas chromatography-flame ionization detector (GC-FID) analysis of fatty acid and oil content

Fatty acid and oil content of M4 mutant lines and WT were analyzed by GC-FID. About 10 mg seeds of six individual plant for each mutant line were crushed and used for GC-FID analysis. Fatty acids were methylated and extracted with methanol containing 5% (v/v) H₂SO₄ and 0.01% butylated hydroxyl toluene according to the method described by Lu et al (Lu *et al.*, 2016). Oil content and fatty acids were quantified using gas chromatography supplied with a hydrogen flame ionization detector and a capillary RESTEK Rtx®-Wax column (0.25 mm × 30 m) with helium carrier at 20 ml/min. The GC-FID parameters are described as following: oven temperature was maintained at 170°C for 1 min and then increased to 210°C with an increase of 3°C/min. Finally, fatty acids were identified according to their retention times and peak areas were analyzed. Fatty acids were calculated as nmol%. Heptadecanoic acid (17:0) was used as the internal standard to quantify the fatty acid composition and oil content of the seeds and oil content was calculated as % of the dry weight.

DATA AVAILABILITY STATEMENT

All relevant data can be found within the manuscript and its supporting materials.

ACKNOWLEDGEMENTS

This research was supported by The National Key Research and Development Program of China (2016YFD0101000) and the National Natural Science Foundation of China (31871658).

AUTHOR CONTRIBUTIONS

L.G. and Q.Y.Y. conceived and supervised the study. S.T., L.Y., S.L., L.L., Z.D., X.L., X.L. performed the experiments. S.T. and S.L. analyzed the data. Q.Y.Y. and L.D.X.

performed bioinformatics analysis. S.T. and L.D.X. wrote the manuscript. L.G., W.M. and Q.Y.Y. discussed and revised the manuscript. All authors read and approved the manuscript.

CONFLICT OF INTERESTS

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Effects of EMS concentration on seed germination.

Figure S2. Number of mutations in five WT plants.

Figure S3. Number and ratio of EMS-induced homozygous and heterozygous mutations.

Figure S4. Identification of mutants with high oil content and oleic acid by GC-FID.

Table S1. Whole-genome sequencing of the five wild-type and twenty mutant lines.

Table S2. Summary of screening oil content and fatty acid mutants of M3 seeds from 98,113 mutant lines.

Table S3. Summary of screening stable oil content and fatty acid mutants of M4 seeds from 1,886 mutant lines.

Table S4. Mutants with increased both oil content and oleic acid.

Table S5. Agronomical traits of four high oleic acid mutants.

REFERENCES

Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, Matsumura H, Yoshida K, Mitsuoka C, Tamiru M, Innan H, Cano L, Kamoun S, Terauchi R. 2012. Genome sequencing reveals agronomically important loci in rice using MutMap. *Nature Biotechnology* **30**, 174-178.

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653-657.

Amosova AV, Zoshchuk SA, Volovik VT, Shirokova AV, Horuzhiy NE, Mozgova GV, Yurkevich OY, Artyukhova MA, Lemesh VA, Samatadze TE, et al. 2019. Phenotypic, biochemical and genomic variability in generations of the rapeseed (*Brassica napus* L.) mutant lines obtained via chemical mutagenesis. *PLoS One* **14**, e0221699.

Auld DL, Heikkinen MK, Erickson DA, Sernyk JL, Romero JE. 1992. Rapeseed Mutants with Reduced Levels of Polyunsaturated Fatty-Acids and Increased Levels of Oleic-Acid. *Crop Science* **32**, 657-662.

Becker HC, Damgaard C, Karlsson B. 1992. Environmental variation for outcrossing rate in rapeseed (*Brassica napus*). *Theoretical and Applied Genetics* **84**, 303-306.

Bergelson J, Buckler ES, Ecker JR, Nordborg M, Weigel D. 2016. A Proposal Regarding Best

Practices for Validating the Identity of Genetic Stocks and the Effects of Genetic Variants. *Plant Cell* **28**, 606-609.

Chen L, Huang L, Min D, Phillips A, Wang S, Madgwick PJ, Parry MA, Hu YG. 2012. Development and characterization of a new TILLING population of common bread wheat (*Triticum aestivum* L.). *PLoS One* **7**, e41570.

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. 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 (Austin)* **6**, 80-92.

Espina MJ, Ahmed CMS, Bernardini A, Adeleke E, Yadegari Z, Arelli P, Pantalone V, Taheri A. 2018. Development and Phenotypic Screening of an Ethyl Methane Sulfonate Mutant Population in Soybean. *Frontiers in Plant Science* **9**, 394.

Etherington GJ, Ramirez-Gonzalez RH, MacLean D. 2015. bio-samtools 2: a package for analysis and visualization of sequence and alignment data with SAMtools in Ruby. *Bioinformatics* **31**, 2565-2567.

Fang CQ, Wang CT, Wang PW, Tang YY, Wang XZ, Cui FG, Yu SL. 2012. Identification of a Novel Mutation in *FAD2B* from a Peanut EMS Mutant with Elevated Oleate Content. *Journal of Oleo Science* **61**, 143-148.

Gan L SX, Jin L, Wang G, Xiu J, Wei Z, Fu T. 2003. Establishment of Math Models of NIRS Analysis for Oil and Protein Contents in Seed of *Brassica napus*. *Scientia Agricultura Sinica* **36**, 1609-1613.

Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, Enns LC, Chris B, Johnson JE, Odden AR. 2003. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* **164**, 731-740.

Harloff HJ, Lemcke S, Mittasch J, Frolov A, Wu JG, Dreyer F, Leckband G, Jung C. 2012. A mutation screening platform for rapeseed (*Brassica napus* L.) and the detection of sinapine biosynthesis mutants. *Theoretical and Applied Genetics* **124**, 957-969.

Hua W, Liu J, Wang H. 2016. Molecular regulation and genetic improvement of seed oil content in *Brassica napus* L. *Frontiers of Agricultural Science and Engineering* **3**.

Kohalmi SE, Kunz BA. 1988. Role of neighbouring bases and assessment of strand specificity in ethylmethanesulphonate and N-methyl-N'-nitro-N-nitrosoguanidine mutagenesis in the SUP4-o gene of *Saccharomyces cerevisiae*. *Journal of Molecular Biology* **204**, 561-568.

Lakhssassi N, Zhou Z, Liu S, Colantonio V, AbuGhazaleh A, Meksem K. 2017. Characterization of the *FAD2* Gene Family in Soybean Reveals the Limitations of Gel-Based TILLING in Genes with High Copy Number. *Frontiers in Plant Science* **8**, 324.

Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, Baud S, Bird D, Debono A, Durrett TP, et al. 2013. Acyl-lipid metabolism. *Arabidopsis Book* **11**, e0161.

Li G, Jain R, Chern M, Pham NT, Martin JA, Wei T, Schackwitz WS, Lipzen AM, Duong PQ, Jones KC, et al. 2017. The Sequences of 1504 Mutants in the Model Rice Variety Kitaake Facilitate Rapid Functional Genomic Studies. *Plant Cell* **29**, 1218-1231.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.

Lu C, Xin Z, Ren Z, Miquel M, Browse J. 2009. An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of *Arabidopsis*. *Proceeding of the National Academy of Science* **106**, 18837-18842.

Lu S, Sturtevant D, Aziz M, Jin C, Li Q, Chapman KD, Guo L. 2018. Spatial analysis of lipid metabolites and expressed genes reveals tissue-specific heterogeneity of lipid metabolism in high-

and low-oil *Brassica napus* L. seeds. Plant Journal **94**, 915-932.

Lu S, Yao S, Wang G, Guo L, Zhou Y, Hong Y, Wang X. 2016. Phospholipase D ϵ enhances *Braasca napus* growth and seed production in response to nitrogen availability. Plant Biotechnology Journal **14**, 926-937.

Lu X, Liu J, Ren W, Yang Q, Chai Z, Chen R, Wang L, Zhao J, Lang Z, Wang H, et al. 2018. Gene-Indexed Mutations in Maize. Molecular Plant **11**, 496-504.

Malek MA, Begum HA, Begum M, Sattar MA, Ismail MR, Rafii MY. 2012. Development of two high yielding mutant varieties of mustard [*Brassica juncea* (L.) Czern.] through gamma rays irradiation. Australian Journal of Crop Science **6**, 922-927.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research **20**, 1297-1303.

Michno JM, Stupar RM. 2018. The importance of genotype identity, genetic heterogeneity, and bioinformatic handling for properly assessing genomic variation in transgenic plants. BMC Biotechnology **18**, 38.

Rakow G. 1973. Selektion auf Linol-und Linolensäuregehalt in Rapssamen nach mutagener Behandlung. Z. Pflanzenzüchtg **69**, 62-82.

Rédei GP. 2008. Encyclopedia of genetics, genomics, proteomics, and informatics: Springer Science & Business Media.

Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B. 2003. An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Molecular Biology **53**, 247-259.

Rucker B, Robbelen G. 1996. Impact of low linolenic acid content on seed yield of winter oilseed rape (*Brassica napus* L). Plant Breeding **115**, 226-230.

Scarth R, McVetty PBE, Rimmer SR, Stefansson BR. 1988. STELLAR LOW LINOLENIC-HIGH LINOLEIC ACID SUMMER RAPE. Canadian Journal of Plant Science **68**, 509-511.

Serrat X, Esteban R, Guibourt N, Moysset L, Nogues S, Lalanne E. 2014. EMS mutagenesis in mature seed-derived rice calli as a new method for rapidly obtaining TILLING mutant populations. Plant Methods **10**.

Sevanthi AMV, Kandwal P, Kale PB, Prakash C, Ramkumar MK, Yadav N, Mahato AK, Sureshkumar V, Behera M, Deshmukh RK, et al. 2018. Whole Genome Characterization of a Few EMS-Induced Mutants of Upland Rice Variety Nagina 22 Reveals a Staggeringly High Frequency of SNPs Which Show High Phenotypic Plasticity Towards the Wild-Type. Frontiers in Plant Science **9**, 1179.

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**, 51-60.

Song JM, Guan Z, Hu J, Guo C, Yang Z, Wang S, Liu D, Wang B, Lu S, Zhou R, et al. 2020. Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. Nature Plants **6**, 34-45.

Steven H, Till BJ, Luca C. 2004. TILLING. Traditional mutagenesis meets functional genomics. Plant Physiology **135**, 630-636.

Takagi H, Tamiru M, Abe A, Yoshida K, Uemura A, Yaegashi H, Obara T, Oikawa K, Utsushi H, Kanzaki E, et al. 2015. MutMap accelerates breeding of a salt-tolerant rice cultivar. Nature Biotechnology **33**, 445-449.

Tsuda M, Kaga A, Anai T, Shimizu T, Sayama T, Takagi K, Machita K, Watanabe S,

Nishimura M, Yamada N, et al. 2015. Construction of a high-density mutant library in soybean and development of a mutant retrieval method using amplicon sequencing. *BMC Genomics* **16**, 1014.

USDA ERS (2018) <https://www.ers.usda.gov/data-products/oil-crops-yearbook/>

Wang HZ, Wang XF, Liu GH, Zheng YB, Yang Q. 2004. Studies on microspore culture of hybrid parents in *Brassica napus*. *Chinese Journal of Oil Crop Scieives* **26**, 1-4.

Wang L, Zhang B, Li J, Yang X, Ren Z. 2014. Ethyl methanesulfonate (EMS)-mediated mutagenesis of cucumber (*Cucumis sativus* L.). *Agricultural Sciences* **5**, 716-721.

Wang N, Wang Y, Tian F, King GJ, Zhang C, Long Y, Shi L, Meng J. 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**, 751-765.

Wu X. J., Hu B, C., and Chen F. X. 2018. Research progress of high oleic acid in *Brassica napus* L. *Crops*, **28**, 5-9. (In Chinese)

Xi-ou X, Wenqiu L, Wei L, Xiaoming G, Lingling L, Feiyue M, Yuge L. 2017. The Analysis of Physiological Variations in M2 Generation of *Solanum melongena* L. Mutagenized by Ethyl Methane Sulfonate. *Frontier in Plant Science* **8**, 17.

Xiao XO, Lin W, Li K, Feng X, Jin H, Zou H. 2019. Genome-Wide Analysis of Artificial Mutations Induced by Ethyl Methanesulfonate in the Eggplant (*Solanum melongena* L.). *Genes (Basel)* **10**.

Zhao Q, Wu J, Cai G, Yang Q, Shahid M, Fan C, Zhang C, Zhou Y. 2019. A novel quantitative trait locus on chromosome A9 controlling oleic acid content in *Brassica napus*. *Plant Biotechnology Journal* **17**, 2313-2324.

FIGURE LEGENDS

Figure 1. Outline of EMS-induced mutant population. Seeds were mutagenized with EMS to produce the M1 population, bulked M2 seeds were sowed to produce the M2 population. Identification of mutants with phenotypic variations during seedling and flowering stages was performed in the M2 population. Then, M3 seeds from 98,113 mutant lines was analyzed by NIRS to screen mutants with altered seed oil content or fatty acid composition. 1,886 M3 mutant lines were selected to confirm stable mutants. In the M4 population, seeds were analyzed by NIRS to identify stable mutants. Meantime, whole-genome sequencing was performed for M2, M3 and M4 plants to analyze EMS-induced mutations.

Figure 2. Morphological mutants. (a) Summary of typical mutant phenotypes detected in the M2 population. All mutants were bagged for future research. (b) Representative morphological mutants during seedling stage. (c) Representative morphological mutants during flowering stage.

Figure 3. Genomic features of EMS-induced mutations of M2, M3 and M4 plants. (a) Mutation distribution and density on chromosomes. (b) Number and ratio of mutations in 20 mutants. Both heterozygous and homozygous mutations are included in the analysis.

Figure 4. Characterization of the EMS-induced mutations in the M2, M3 and M4 mutants. (a) Genome-wide analyses of the size of InDel mutations. (b) Distribution of the mutations in the genome. (c) Percentage of mutation effect on gene functions. (d) The homozygous percent of mutations in the M2, M3 and M4 mutants.

Figure 5. Phenotype distribution of 1,886 mutant lines (M4 seeds). Variations in oil content (a), oleic acid (b), linoleic acid (c) and linolenic acid (d) among the offspring of 1,886 M3 mutant lines were displayed. The average oil content and fatty acid composition in WT plants were showed. M4 seeds were collected and analyzed from six individual plants for each mutant line.

Figure 6. Analysis of mutants with extreme oil content and fatty acid by GC-FID method. Analysis of oil content (a), oleic acid (b), linoleic acid (c) and linolenic acid (d) mutants by GC-FID. Values are means $\pm$ SD ($n = 6$). Asterisk denotes significant difference at $P < 0.05$ and $P < 0.01$ using t-test.

Table 1. Representative mutants with altered seed oil content or fatty acid composition traits

	Oil content	Oleic acid	Linoleic acid	Linolenic acid	Protein	Stearic acid	Palmitic acid	Eicosenoic acid	Mutant type
WT	51.32±1.52	73.47±2.77	13.87±0.65	6.44±0.90	18.58±1.75	0.20±0.13	4.26±0.13	8.08±4.31	
M4-42	33.91±1.24	57.72±2.46	21.07±0.76	12.28±1.86	26.79±2.05	0.68±0.19	5.25±0.49	12.59±1.51	Oil content mutants
M4-236	36.45±1.75	65.68±0.4	20.15±0.77	8.92±0.72	26.4±1.42	0.27±0.07	5.02±0.14	1.9±0.63	
M4-1027	54.96±1.34	58.13±3.09	11.71±0.87	7.67±0.47	17.92±0.45	0.49±0.19	3.83±0.14	10.67±1.83	
M4-1451	55.09±0.77	63.62±1.96	11.7±0.51	6.06±0.97	16.7±0.99	0.44±0.1	4.08±0.21	8.19±2.89	
M4-1039*	41.86	52.22	17.08	12.51	24.44	1.37	5.06	9.64	Oleic acid mutants
M4-42	33.91±1.24	57.72±2.46	21.07±0.76	12.28±1.86	26.79±2.05	0.68±0.1	5.25±0.49	12.59±1.51	
M4-1446*	52.98	82.94	13.37	4.5	16.26	0.04	4.52	0.77	
M4-371	50.36±0.71	83.42±0.92	8.86±0.71	4.49±0.54	18.53±0.38	0.56±0.06	4.15±0.08	10.75±2.57	
M4-1876	51.26±0.39	80.97±0.97	7.99±0.22	4.12±0.14	18.27±0.54	0.30±0.08	3.80±0.07	6.78±1.04	Linoleic acid mutants
M4-1875	52.41±0.73	82.23±2.71	8.54±0.81	4.26±0.48	18.17±0.97	0.35±0.07	3.87±0.10	9.34±1.04	
M4-236	36.45±1.75	65.68±0.4	20.15±0.77	8.92±0.72	26.4±1.42	0.27±0.07	5.02±0.14	1.9±0.63	
M4-42	33.91±1.24	57.72±2.46	21.07±0.76	12.28±1.86	26.79±2.05	0.68±0.1	5.25±0.49	12.59±1.51	
M4-947	51.6±1.31	80.7±0.74	10.45±0.16	2.87±0.26	18.75±0.42	0.13±0.09	4.1±0.04	2.73±0.42	Linolenic acid mutants
M4-926	50.2±0.66	80.37±1.35	10.67±0.41	3.18±0.28	18.75±0.35	0.08±0.07	4.12±0.05	2.71±0.33	
M4-42	33.91±1.24	57.72±2.46	21.07±0.76	12.28±1.86	26.79±2.05	0.68±0.1	5.25±0.49	12.59±1.51	
M4-1039*	41.86	52.22	17.08	12.51	24.44	1.37	5.06	9.64	

These are the representative phenotypic data of M4 seeds. Seeds were analyzed by NIRS and the data are the average value of four to six individual plants of each mutant line. Asterisk denotes significant difference at $P < 0.05$ and $P < 0.01$ using t-test. Green bold showed the extreme low traits and red bold showed the extreme high traits. *Only one line was analyzed.

Table 2. Mutation detections of *FAD2* and *ROD1* in five high oleic acid mutants

	Accession	Gene	Position (bp) and Base Change	Region	Mutation Type	C18:1 content	C18:2 content
	Wild-type	—	—	—	—	73.47±2.77	13.87±0.65
M4-1	1429	<i>BnaC03.ROD1</i>	g.514-2634C>T ^a	intron	intron variant	81.32±1.52	10.52±0.80
M4-2	1449	<i>BnaC05.FAD2</i>	g.-592C>T ^b	upstream	upstream gene variant	82.12±4.61	11.65±1.57
		<i>BnaA03.ROD1</i>	g.247G>A ^c	exon	missense variant		
M4-3	1875	<i>BnaC01.FAD2</i>	g.762G>A	exon	missense variant	82.23±2.71	8.54±0.8
		<i>BnaC02.ROD1</i>	g.-518C>T	upstream	upstream gene variant		
M4-4	1876	<i>BnaA05.FAD2</i>	g.817G>A	exon	missense variant	80.97±0.97	7.99±0.22
		<i>BnaC01.FAD2</i>	g.-47C>T	upstream	upstream gene variant		
M4-5	371	<i>BnaA05.FAD2</i>	g.995C>T	exon	missense variant	83.42±0.92	8.86±0.87
		<i>BnaA03.ROD1</i>	g.267G>A	exon	synonymous variant		

a "g." for genomic sequence, the number represent the mutation nucleotide in the exon.

b "g." for genomic sequence, the number of the first nucleotide of the following exon, a minus sign, and the position upstream in the intron.

c "g." for genomic sequence, the nucleotide 5' of the ATG-translation initiation codon is -1.