Vegetos (2022) 35:1036–1046
https://doi.org/10.1007/s42535-022-00385-4

RESEARCH ARTICLES

Evaluation and characterization of EMS induced mutant population of *Gossypium herbaceum*

Umesh Kumar[1,4] · Nilesh S. Chavan[3] · Samir V. Sawant[2,4] · Hemant Kumar Yadav[1,4]

Received: 2 December 2021 / Revised: 30 March 2022 / Accepted: 1 April 2022 / Published online: 3 May 2022

Abstract

Gossypium herbaceum is reported to have innate adaptability to withstand biotic and abiotic stresses. However, the continued selection for desired traits eliminates the rare alleles from advanced cultivars and makes the species vulnerable to several stresses. The induced mutagenesis expands the genetic basis of the existing germplasm. In the present study, 5500 seeds of *G. herbaceum* (cv. WAGAD) were treated by 70 mM ethyl methane sulfonate (EMS), established 2597 fertile M_1 plants, which were used in development of 6600 M_2 plants followed by 5473 M_3 plants. About 5% of the plants showed visible mutant phenotypes like sterility, chlorosis, altered leaf morphology and early flowering/late flowering. The phenotypic evaluation of 4453 plants for 11 agro-morphological traits revealed significant variations. Random amplified polymorphic DNA (RAPD) profiling of a random set of 150 M_3 plants revealed the considerable level of mutation in the genome. The developed mutant population could serve as an important genetic resource for forward and reverse genetic studies. Further, the potential mutant plants could also be used for pyramiding targeted traits and consequently for developing new high yielding varieties.

Keywords Agro-morphological traits · EMS · *Gossypium herbaceum* · Genetic diversity · RAPD

Introduction

Cotton (*Gossypium* spp.) is one of the major sources of natural textile fiber, livestock feed and human foodstuff. The commercial production of cotton comes from two allotetraploid species namely *G. hirsutum* (90%) and *G. barbadense* (8%), rest 2% from two diploid species (*G. herbaceum* and *G. arboreum*) (Kumar et al. 2021). The diploids are known for long maturity time, poor yield and fiber quality, but have a natural ability to withstand biotic and abiotic stresses (Kulkarni et al. 2009; Wang et al. 2012). In spite of better adaptability under harsh conditions, farmers do not prefer to grow diploid cotton due to poor productivity and fiber quality that fetches low price. On the other hand, limited efforts have been made to exploit the genetic potential toward its improvement (Parekh et al. 2018). At present, the demand for diploid cotton mainly *G. herbaceum* is increasing day by day as surgical cotton and technical textiles, which force improvement of its genetic potential towards higher lint production. The existence of satisfactory genetic variability in the gene pool provides a platform to devise suitable breeding strategies for genetic improvement however limited diversity has been reported in *G. herbaceum* (Jena et al. 2012; Parekh et al. 2016). Thus, alternative strategies need to be deployed to increase the genetic variation within the germplasm for further exploitation of genetic potential of this crop. Previously, gene expression studies have been carried out in mannitol-imposed drought stress using cultivar "WAGAD" of *G. herbaceum* and identified several novel genes (Ranjan et al. 2012). Some of the identified genes have characterized and validated their role in drought stress (Trivedi et al. 2012; Gunapati et al. 2016). But, most of them are of unknown functions and no homologs are reported

✉ Samir V. Sawant
samirsawant@nbri.res.in

✉ Hemant Kumar Yadav
h.yadav@nbri.res.in

1 Plant Genetic Resources and Improvement, CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow, Uttar Pradesh, India

2 Genetics and Molecular Biology Division, CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow, Uttar Pradesh, India

3 Agricultural Biotechnology, Mahatma Gandhi Mission College, Aurangabad, Maharashtra, India

4 Academy of Scientific and Innovative Research, Ghaziabad, Uttar Pradesh, India

in the public database. Thus, it is essential to explore the whole genome of WAGAD and characterize newly identified genes to better understand the drought tolerance mechanism at molecular levels. According to Khan et al. (2018), ambient temperature is increasing since the last decades and promoted the drought situation. Thus, WAGAD can be used in the varietal development programs of cotton breeding for drought stress tolerance.

The induced mutagenesis has a significant role in both crop improvement and studies on functional genomics. Among various mutagens, EMS is a widely accepted and utilized chemical mutagen for mutagenesis (Aslam et al. 2016). At the genetic level, it alkylates G nucleotide and leads GC to AC transitions, which results in point mutations throughout the genome. About 1/24 to 1/2500 kb mutation density have been reported in several crops (Kurowska et al. 2011). Among the molecular markers, the RAPD (Randomly amplified Polymorphic DNA) marker is amplified with short arbitrary primer, which amplifies the genomic DNA of targeted species without prior sequence information and produces polymorphic bands that is used in detection of generic diversity (Raj et al. 2014). The binding sites of these primers are changed (either loss or gain) in the mutated lines due to alteration of DNA against the control and detected by gel electrophoresis (Danylchenko and Sorochinsky, 2005). In previous reports, numerous mutant platforms of the model and crop plant species have been developed and identified phenotypic diversity followed by detection of induced mutation through sequencing and RAPD approaches such as wheat (Chen et al. 2012; Slade et al. 2012), tomato (Saito et al. 2011), *Petunia* (Krupa-Małkiewicz and Bienias 2018), alfalfa (Danylchenko and Sorochinsky, 2005), melon (Dahmani-Mardas et al. 2010) and *Jatropha curcas* (Dhakshanamoorthy et al. 2013; Maghuly et al. 2018). In the case of cotton, cultivated allotetraploid species has been used for the development of mutant population and used for identification of genome-wide mutations (Abid et al. 2020; Lian et al. 2020) as well as noble agronomic traits such as herbicide tolerance (Bechere et al. 2010) and good fiber quality (Brown et al. 2015). However, there is a low pace of research on *G. herbaceum* for its exploitation towards functional genomics studies and development of new varieties/lines (Kumar et al. 2021). Considering the importance of the *G. herbaceum*, efforts have been made to create new variability, which may either be used in functional genomics studies or crop improvement programs. Therefore, present work was conducted for phenotypic and molecular characterization of the mutant population and identification of potential variants for further exploitation in genetic improvement.

Materials and methods

Plant material and optimization of EMS concentration

The seed materials of *G. herbaceum* (cv. WAGAD) were obtained from CSIR-NBRI, Lucknow, India (Ranjan et al. 2012). To determine the appropriate optimum lethal concentration (LC_{50}), two rounds of seed treatment tests were performed. A wider range of EMS doses e.g. 50, 100, 150, 200 and 250 mM were tested in the first round. While, lower EMS doses were used i.e. 40, 50, 60, 70, 80 and 90 mM in the second round of test as per the result obtained in the first round. Both rounds of tests were repeated three times. The EMS treatment of seeds were carried in flask kept of shaker for 16 h at 24^0C and 80 rpm. Hundred seeds were used for different concentrations of EMS treatment with control (0.0 mM, untreated). The treated seeds were rinsed under running tap water for 3 h and sown in a cotton growth house (Saveer Biotech Limited, New Delhi, India). Seed germination was recorded on 15 days and expressed in percentage to the nearest whole number.

Development of large mutant population

To develop the M_1 population, 5500 seeds (M_0 seed) were treated with 70 mM EMS (LC_{50}) and harvested plant-wise self-pollinated M_2 seeds from 2597 fertile M_1 plants. Three M_2 seeds from each M_1 plant ($3 \times 2597 = 7791$ M_2 seeds) were sown in the next cropping season to raise M_2 population and collected self-pollinated M_3 seeds from 6600 M_2 plants. Single plant per M_2 derived M_3 generation was used, which is equivalent to 3 plants from each M_1 plant and sown a total of 6,500 M_3 seeds (Table 1). The M_1, M_2 and M_3 generations were grown in research field of CSIR-NBRI at Mahatma Gandhi Mission (MGM), Aurangabad (MS), India (19° 15' N; 75° 23' E). The spacing of plant to plant and line to line was maintained at 75 cm and 90 cm, respectively.

Table 1 Seed germination, survival and fertility details of the M_3 population of *G. herbaceum*

Treatment	Number of sown seeds	Number of germinated seeds	Number of surviving seedlings	Number of plants harvested seed
Control	400	377	377	377
M_3 population	6,500	6,073	5,527	5,473

Phenotyping of M_3 population and RAPD analysis

The visible observations were made at regular time intervals from germination to harvesting stage and noticed mutant phenotypes. In addition, quantitative data for 11 agronomic traits were recorded for induced diversity analysis against the control. The name of traits and their methods for data collection are described in supplementary table 1. To perform RAPD analysis, the genomic DNA was isolated from a random set of 150 M_3 plants using plant DNeasy mini kit (Qiagen, Valencia, California) as per manufacturer's instructions. The quality of DNA was assessed on 0.8% agarose gel and quantified using Qubit 2.0 fluorometer (Life Technologies, USA). A set of 32 RAPD primers were used for preliminary screening. Finally, 14 RAPD primers with clear and reproducible amplification were selected for genotyping (Table 2). The PCR amplification was carried in 15 µL reaction volume with 50 ng genomic DNA, 10 pmol primer and 1X TopTaq PCR master mix (Qiagen Inc, USA). The PCR condition included pre-denaturation at 94 °C for 3 min, followed by 43 cycles of denaturation at 94 °C for 1 min, annealing at 35 °C for 1 min, extension at 72 °C for 1.30 min and a final extension at 72 °C for 10 min. The PCR product was separated on 1.2% agarose gel in 0.5X TBE buffer (Tris–Borate-EDTA). After electrophoresis, the gel was stained in ethidium bromide and visualized on Gel Doc XR (BioRad Laboratories, Inc). The mutation had considered from the change of RAPD amplified band position with respect to the control. RAPD bands were scored as 0–1 matrix, where '0' indicates absence of band while '1' indicates presence of band for each primer. The unclear (faint) bands were not considered for data scoring.

Statistical analysis

Microsoft Excel 2016 was used for the statistical analysis of the treatment viz., range of dispersion, mean, standard error and coefficient of variance (CV%). The mean values of M_3 population and the control were analyzed by student's t-test ($p \leq 0.05$). The frequency distribution graph was plotted with the help of PAST-3 software (Hammer et al. 2001). The 0–1 binary data of RAPD was used to compute pair-wise genetic dissimilarities among the selected plants using Jaccard's similarity coefficient. The neighbor-joining (NJ) tree was constructed along with principal component analysis (PCA) using DARwin 5.0.157 software (Perrier et al. 2003).

Results and discussion

Optimization of lethal concentration of EMS

The rate of induced mutation depends on the effectiveness and efficiency of a mutagen followed by the duration of treatment as well as plant materials being treated (e.g., seed, pollen and vegetative part). The EMS is absorbed by treated samples which shows a genotoxic effect, can induce either DNA damage or inhibit the physiological processes that block/reduce enzyme activity and RNA synthesis (Kumar et al. 2013). The higher doses increase frequency of mutation but dramatically decrease the recovery of viable

Table 2 RAPD analysis among selected M_3 plants and related statistical parameters

S.N	RAPD primer	Nucleotide sequences (5'–3')	Number of amplified bands	Number of polymorphic bands	Percent polymorphic bands
1	OPE 06	AAGACCCCTC	5	4	80.0
2	OPE -07	AGATGCAGCC	6	5	83.3
3	OPE-20	AACGGTGACC	8	7	87.5
4	OPA-13	CAGCACCCAC	8	7	87.5
5	OPF-10	GGAAGCTTGG	8	6	75.0
6	OPZ-13	GACTAAGCCC	4	3	75.0
7	OPW-02	ACCCCGCCAA	6	5	83.3
8	OPW-05	GGCGGATAAG	5	4	80.0
9	OPW-09	GTGACCGAGT	5	4	80.0
10	OPT-08	AACGGCGACA	8	6	75.0
11	OPT-15	GGATGCCACT	5	4	80.0
12	OPG-01	CTACGGAGGA	6	2	33.3
12	OPG-02	GGCACTGAGG	11	5	45.5
14	OPG-15	ACTGGGACTC	7	6	85.7
Sum			**92**	**68**	
Average			**6.57**	**4.8**	**75.1**

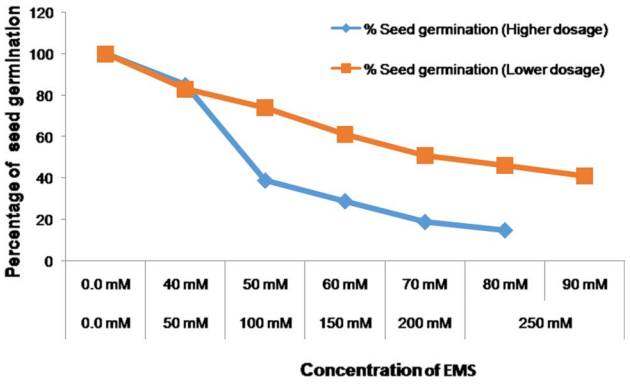


Fig. 1 Kill-curve plot of *G. herbaceum* (cv. WAGAD) during the optimization of lethal concentration (LC_{50})

seeds, while too low concentration recovered less mutation (Chawade et al. 2010). In the present investigation the seed germination was decreased as the EMS concentrations increased (Fig. 1). Similar pattern was also reported in previous studies which have shown the effectiveness of mutagenesis in sunflower (Sabetta et al. 2011) and wheat (Rawat et al. 2018). According to Abid et al. (2020), the optimum lethal concentration of a mutagen provides maximum mutation density without paying off plant viability and fertility. For this, different concentrations of EMS were tested and tried to find LC_{50} before treating the bulk materials (Fig. 1). In the first round of treatment, about 30–70% of seed germination was assumed from the control between 50 to 100 mM (Fig. 1). Too higher doses of EMS (150 to 250 mM) were very effective but dramatically killed the seed and were discarded (Fig. 1). In the second round of lower doses test about 51% of seed germination was noticed at 70 mM (Fig. 1). While, 83, 74, 61, 46 and 41% of seed germination were counted at 40, 50, 60, 80 and 90 mM of EMS, respectively (Fig. 1). Based on these results, we concluded that 70 mM was the best as LC_{50} in the subsequent experiments. In previous studies, higher concentrations of EMS ranging from 1.5 to 2% have been tested for allotetraploid cotton (Abid et al. 2020; Lian et al. 2020). This may be due to genome buffering as reported for wheat mutagenesis (Rawat et al. 2018).

Phenotyping of M_3 population

Out of 6500 M_3 seeds sown, 5527 of them had survived till the maturity. Finally, M_4 seeds were harvested from 5473 fertile M_3 plants (Table 1). The control was well grown and 100% fertile. The higher mortality of germinated seedlings might be due to damage of cotyledon, root and shoot. Such mortality in the initial phase of mutagenesis may have occurred due to the increased homozygosity of the lethal mutation (Schreiber et al. 2019). Other studies by Espina et al. (2018) and Chantreau et al. (2013) have reported similar results for soybean and linseed, respectively. According to Saito et al. (2011), the population size is necessary to develop a saturated mutant resource, where every locus of the genome is mutated and needed more than thousands of

Table 3 Classification of visible phenotype and number of mutant plants in the M_3 population

S.N	Group	Subgroup	Numbers of plant	Mutation frequency in population
1	Leaf color	Chimeric leaf (white)	5	0.09%
		Chimeric leaf (yellow)	49	0.89%
		Green patches on leaf	3	0.05%
		White leaf	1	0.02%
		Chimeric bract (yellow)	1	0.02%
2	Leaf architecture	Highly serrated leaf	4	0.07%
		Curly leaf	46	0.85%
		Hairless leaf (smooth)	84	1.52%
3	Flowering	Late flowering (120 days)	14	0.25%
		Early flowering (30 days)	3	0.05%
4	Plant growth	Spreading habit	2	0.04%
		Bushy habit	3	0.05%
5	Sterility	No seed formation	1	0.02%
		No flowering (Sterile)	47	0.85%
		No boll setting (Partial sterile)	6	0.11%
6	Boll shape	Round boll shape	1	0.02%
		Conical boll shape	6	0.11%
Subtotal			**276**	**~5%**

* Observed mutation % = Percentage of observed visible (noticeable) mutation among 5,527 M_3 plants

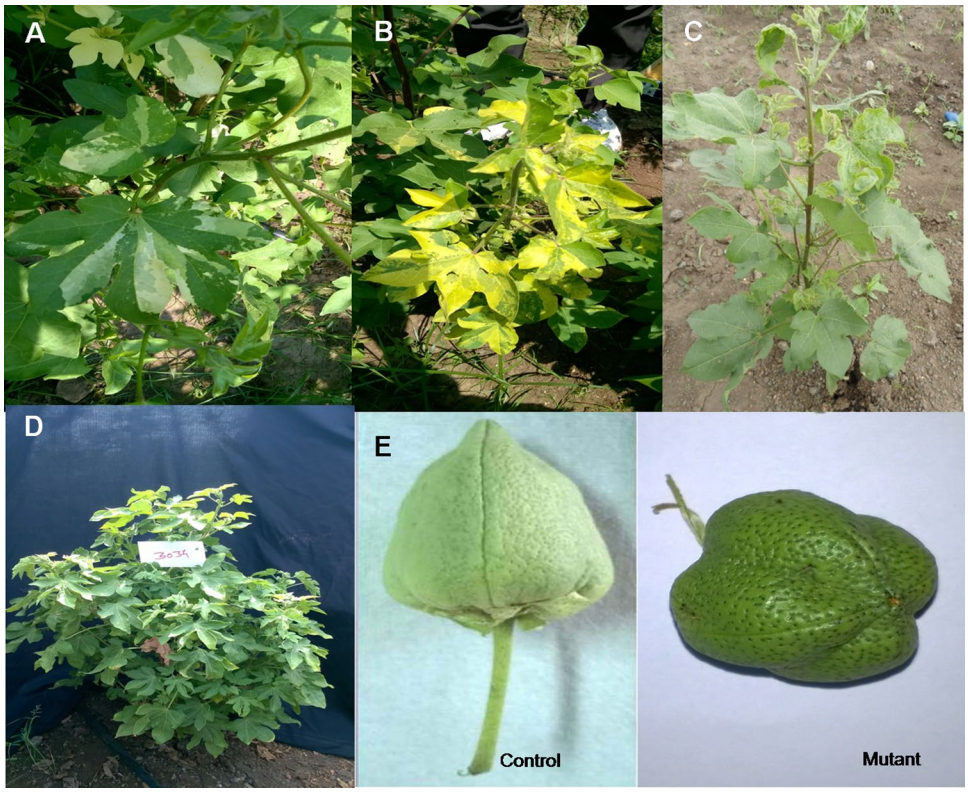


Fig. 2 Visual screening of mutations in M_3 population. **(a)** Chimeric white leaf **(b)** Chimeric yellow leaf **(c)** Curly leaf **(d)** Bushy plant **(e)** Round shape boll

individuals in the case of the EMS mutagenesis platform, suggesting that the current population was almost saturated. Furthermore, perceptible phenotypic variation was noticed in 276 plants (~5%), which were categorized into 6 groups and 17 subgroups (Table 3). According to Mardas et al. (2010), M_1 mutation becomes homozygous in the subsequent generations, exploring the fact that all phenotypes might be due to homozygous condition of mutated gene/s that control a defined pathway/s. The major mutant phenotypes showed changes in leaf color, curly leaf, bushy plant growth (shorter internode and more branches) and round shape of boll (Fig. 2a–e), others identified mutations are listed in Table 3. The leaf color and plant growth-related mutations are reported as primary signal of induced mutagenesis. These have also been observed in other experiments that elucidated sufficient mutagenesis and reported a

Table 4 Performance of 4,453 M_3 plants for 11 different agro-morphological traits with the control

S.N		Range		Mean ± SE		CV	
	Trait	Control	M_3 population	Control	M_3 population	Control	M_3 population
1	Number of branch/plant	15.0–22.0	5.0–47.0	18.63 ± 0.59	21.90 ± 0.09*	14	29.27
2	Plant height (cm)	102.0–135.0	33.3–240.0	115.05 ± 2.66	155.87 ± 0.34*	10.08	16.04
3	Stem circumference (cm)	2.2–4.5.0	2.0–12.6	3.13 ± 0.14	5.65 ± 0.02*	12.13	27.18
4	Leaf size(cm^2)	54.9–94.8	16.0–153.6	77.11 ± 2.8	67.05 ± 0.29*	15.97	29.29
5	Number of node/plant	11.0–18.0	7.0–49.0	13.78 ± 0.5	25.37 ± 0.08*	16.31	23
6	Internodal length (cm)	10.2–15.2	5.9–25.4	11.37 ± 0.28	11.66 ± 0.03 ns	10.94	17.38
7	Number of boll/plant	21.0–28.0	5.0–63.0	24.94 ± 0.56	28.06 ± 0.11*	9.86	26.54
8	Test weight (gm)	55.0–70.0	35.0–79.0	61.47 ± 1.3	60.86 ± 0.10 ns	9.28	11.42
9	Lint weight/plant (gm)	18.0–28.0	3.7–56.0	22.88 ± 0.67	26.98 ± 0.11*	12.88	28.17
10	Seed weight/plant (gm)	25.0–38.0	10.0–111.6	29.57 ± 0.83	55.33 ± 0.22*	12.3	27.57
11	Biological yield/plant (gm)	44.0–61.0	13.7–165.9	52.45 ± 1.13	82.31 ± 0.33*	9.43	27.25

*Significant differences between the mean performance of M_3 population and the control plants ($p \leq 0.05$)

ns = Non-significant differences between the mean performance of M_3 population and the control plants ($p \leq 0.05$)

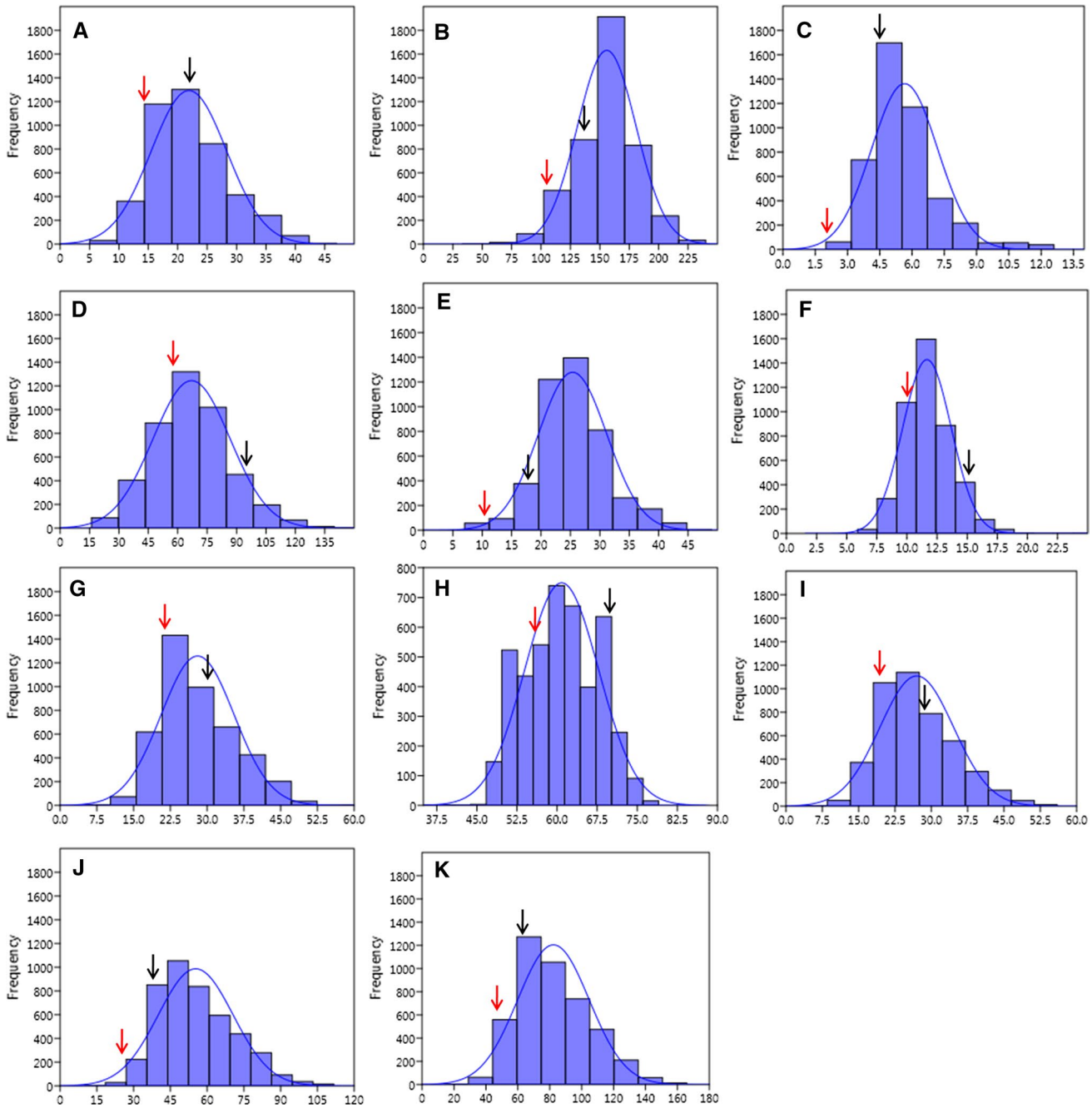


Fig. 3 Frequency distribution of 11 agro-morphological traits for 4,453 M_3 plants of *G. herbaceum* (cv. WAGAD). (**a**) Number of branches/plant, (**b**) Plant height, (**c**) Stem circumference, (**d**) Leaf size, (**e**) Number of nodes/plant, (**f**) Internodal length, (**g**) Number of bolls/plant, (**h**) Test weight, (**i**) Lint weight/plant, (**j**) Seed weight/plant, (**k**) Biological yield/plant. The red and blue arrow indicates the lower and higher range of dispersion, respectively of the controls

high rate of point mutations (Chantreau et al. 2013; Kumar et al. 2013). The validation of these mutant phenotypes can be useful in identifying the genes responsible for photosynthetic efficiency, plant growth and development (Abid et al. 2020). The most striking phenotypic variations observed in the field were 47 complete sterile and 7 partial sterile plants (Table 3).

The quantitative data were recorded from 4453 M_3 plants. The lower and higher range of measured traits were highly scattered than control (Table 4) and formed a bell-shaped curve (Fig. 3a–k). This result suggested that the EMS mutagenesis has introduced additional variation in which intermediate phenotypes are more common than the extremes. Moreover, assessed CV had varied from 11.42 to

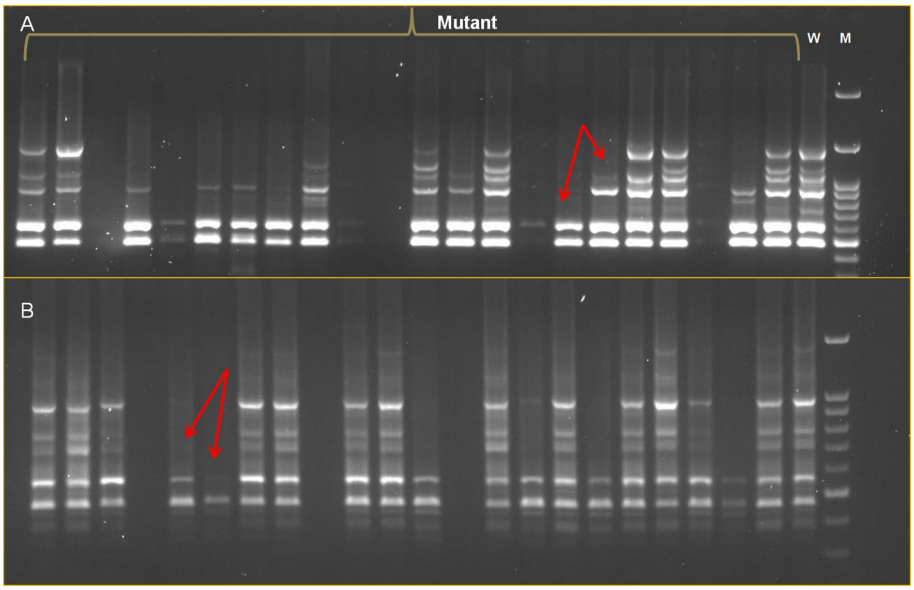
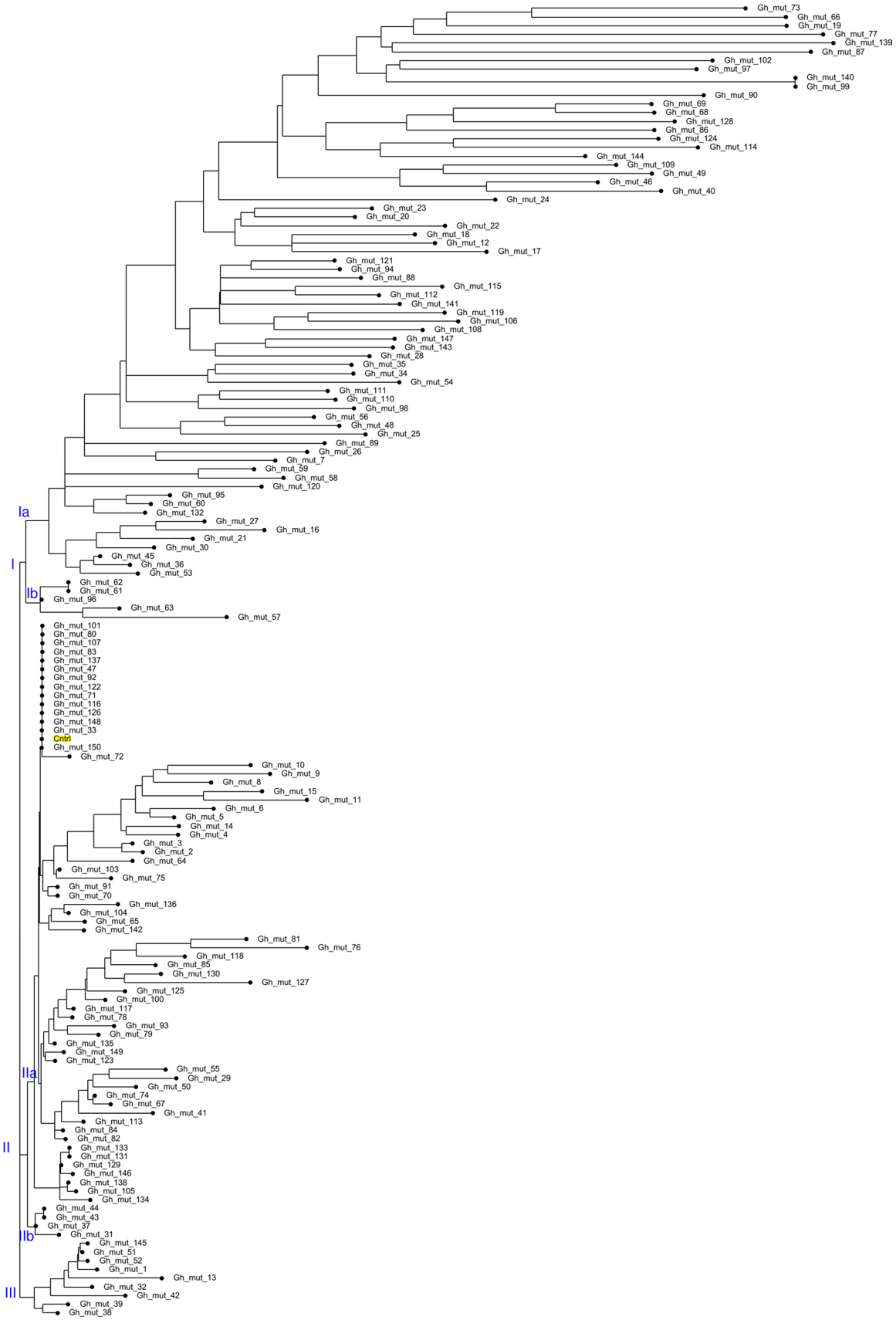


Fig. 4 RAPD profile of selected mutants along with the control. **(a)** OPW-05 primer, **(b)** OPE-20 primer. Absence of RAPD band has been shown by red arrow, (M = DNA ladder, W = Control)

29.29%, which was higher than the control (Table 4). The maximum CV was recorded for the leaf size (29.29%) followed by number of branch/plant (29.27%), seed weight/plant (27.57%), biological yield/plant (27.25%), stem circumference (27.18%), lint weight/plant (28.17%), number of boll/plant (26.54%), number of node/plant (23%). The other traits were reported less than 20% (Table 4). The increased CV indicated a broad genetic base and heterogeneous nature of the population. The mean values revealed that most of the traits were significantly ($p \leq 0.05$) varied from control except internodal distance and test weight (Table 4). Thus, M_3 population has shown significant differences against the control for the recorded traits and suggested induced quantitative variation as previously reported in the mutant population of chili (Hasan et al. 2020), chickpea (Tiliouine et al. 2018) and lentil (Laskar and Khan 2017). Thus, quantitative variation of the population over the control would be of direct interest to plant breeders for the selection of superior line/s.

RAPD analysis

Sometimes, morphological variations cannot be observed due to gene redundancy and silent mutation. In such conditions, RAPD profiling for the small set of mutant with the control helps in detection of induced mutation at the DNA level at the primary stage. In the previous reports, such hidden variation has been explored by RAPD fingerprinting in wheat (Chen et al. 2012) and torch ginger (Yunus et al. 2013). In this study,14 RAPD primers were selected that gave 92 clear scorable bands (Table 2). The numbers of bands per primer ranged from 4 (OPZ-13) to 11 (OPG-02) with an average of 6.57 bands (Table 2). Of these, 68 (75.1%) bands were found to be polymorphic between the control and selected M_3 plants (Table 2). The percentage of polymorphism varied from 33.3% (OPG-01) to 87.5% (OPE-20 and OPA-13) (Table 2). A representative gel electrophoresis image of the primer OPW-05 and OPE-20 are presented in Fig. 4a, b that showed a clear loss of bands in the number of amplified DNA fragment among M_3 plants than the control. These polymorphic bands suggested induced mutation which may be either point mutation or small indels (insertion or deletion). In other studies, similar efforts have been made for sorghum (Taryonoand and Cahyaningrum, 2011) and *J. curcas* (Dhakshanamoorthy et al. 2013) by using six and twenty-five RAPD primers and reported comparable percentage of polymorphic bands, respectively. However, Aslam et al. (2017) reported a lower percentage of polymorphism among *Capsicum annuum* mutants. According to Raj et al. (2014), cluster analysis is a productive approach in diversity analysis that uses obtained numerical data of fingerprint and calculates distances between every pair of entities. In the present work, pair-wise genetic distance was calculated from the obtained binary data that ranged from 0.0 to 0.99 with an average value of 0.50 ± 0.07 (data not shown). The average genetic dissimilarity of each plant with other plants varied from 0.29 ± 0.29 (Control) to 1.00 ± 0.00 (Gh_mut_139). The highest degree of genetic dissimilarity was observed in Gh_mut_139 followed by Gh_mut_77 and Gh_mut_87 among the others and considered most divergent. The NJ dendrogram was created to better understand the similarity and dissimilarity among 150 M_3 plants and grouped into three major clusters namely I, II and III (Fig. 5). Cluster I and II had further sub-divided into two sub-clusters as Ia, Ib and IIa, IIb, respectively. Sub-cluster IIa was found to be the largest with 67 plants while sub-cluster IIb was lowest with 4 plants. The control plant was


Gh_mut_73
Gh_mut_66
Gh_mut_19
Gh_mut_77
Gh_mut_139
Gh_mut_87
Gh_mut_102
Gh_mut_97
Gh_mut_140
Gh_mut_99
Gh_mut_90
Gh_mut_69
Gh_mut_68
Gh_mut_128
Gh_mut_86
Gh_mut_124
Gh_mut_114
Gh_mut_144
Gh_mut_109
Gh_mut_49
Gh_mut_46
Gh_mut_40
Gh_mut_24
Gh_mut_23
Gh_mut_20
Gh_mut_22
Gh_mut_18
Gh_mut_12
Gh_mut_17
Gh_mut_121
Gh_mut_94
Gh_mut_88
Gh_mut_115
Gh_mut_112
Gh_mut_141
Gh_mut_119
Gh_mut_106
Gh_mut_108
Gh_mut_147
Gh_mut_143
Gh_mut_28
Gh_mut_35
Gh_mut_34
Gh_mut_54
Gh_mut_111
Gh_mut_110
Gh_mut_98
Gh_mut_56
Gh_mut_48
Gh_mut_25
Gh_mut_89
Gh_mut_26
Gh_mut_7
Gh_mut_59
Gh_mut_58
Gh_mut_120
Gh_mut_95
Gh_mut_60
Gh_mut_132
Gh_mut_27
Gh_mut_16
Gh_mut_21
Gh_mut_30
Gh_mut_45
Gh_mut_36
Gh_mut_53
Gh_mut_62
Gh_mut_61
Gh_mut_96
Gh_mut_63
Gh_mut_57
Gh_mut_101
Gh_mut_80
Gh_mut_107
Gh_mut_83
Gh_mut_137
Gh_mut_47
Gh_mut_92
Gh_mut_122
Gh_mut_71
Gh_mut_116
Gh_mut_126
Gh_mut_148
Gh_mut_33
Cnlrl
Gh_mut_150
Gh_mut_72
Gh_mut_10
Gh_mut_9
Gh_mut_8
Gh_mut_15
Gh_mut_11
Gh_mut_6
Gh_mut_5
Gh_mut_14
Gh_mut_4
Gh_mut_3
Gh_mut_2
Gh_mut_64
Gh_mut_103
Gh_mut_75
Gh_mut_91
Gh_mut_70
Gh_mut_136
Gh_mut_104
Gh_mut_65
Gh_mut_142
Gh_mut_81
Gh_mut_76
Gh_mut_118
Gh_mut_85
Gh_mut_130
Gh_mut_127
Gh_mut_125
Gh_mut_100
Gh_mut_117
Gh_mut_78
Gh_mut_93
Gh_mut_79
Gh_mut_135
Gh_mut_149
Gh_mut_123
Gh_mut_55
Gh_mut_29
Gh_mut_50
Gh_mut_74
Gh_mut_67
Gh_mut_41
Gh_mut_113
Gh_mut_84
Gh_mut_82
Gh_mut_133
Gh_mut_131
Gh_mut_129
Gh_mut_146
Gh_mut_138
Gh_mut_105
Gh_mut_134
Gh_mut_44
Gh_mut_43
Gh_mut_37
Gh_mut_31
Gh_mut_145
Gh_mut_51
Gh_mut_52
Gh_mut_1
Gh_mut_13
Gh_mut_32
Gh_mut_42
Gh_mut_39
Gh_mut_38
Ia
I
Ib
IIa
II
IIb
III

◂**Fig. 5** Dendrogram of 150 M_3 plants with the control have built from 14 RAPD markers based on Jaccard's coefficient using NJ clustering method. Control plant is highlighted by yellow color. The abbreviation Gh_mut = *G. herbaceum*_mutant whereas digits indicates ID of the M_3 plant that used in RAPD analysis

clustered in IIa. Sub-cluster Ia and IIb had computed 66 and 5 plants, respectively. Cluster III was represented by 9 plants. In addition, PCA analysis also differentiated all the 150 M_3 plants in to three major clusters almost similar to the NJ dendrogram (Fig. 6). Thus, both NJ dendrogram and PCA analysis of the present study have shown distinct clusters and subclusters due to the accumulation of induced mutations among M_3 plants. As a result, current study gives information about the quality of the developed mutant resource that can be used in both functional genomics study and cotton breeding programs.

Conclusion

The aim of the current study was to utilize EMS mutagenesis and establish a saturated mutant population. A set of 5473 fertile M_3 plants has been established and shown to possess considerable degree of phenotypic as well as genotypic polymorphism. These plants will serve as a potential resource for forward and reverse genetic studies. Further, phenotypically diverse plants will allow exploring causal point mutation for functional validation of genes responsible for growth and development as well as marker-assisted selection.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42535-022-00385-4.

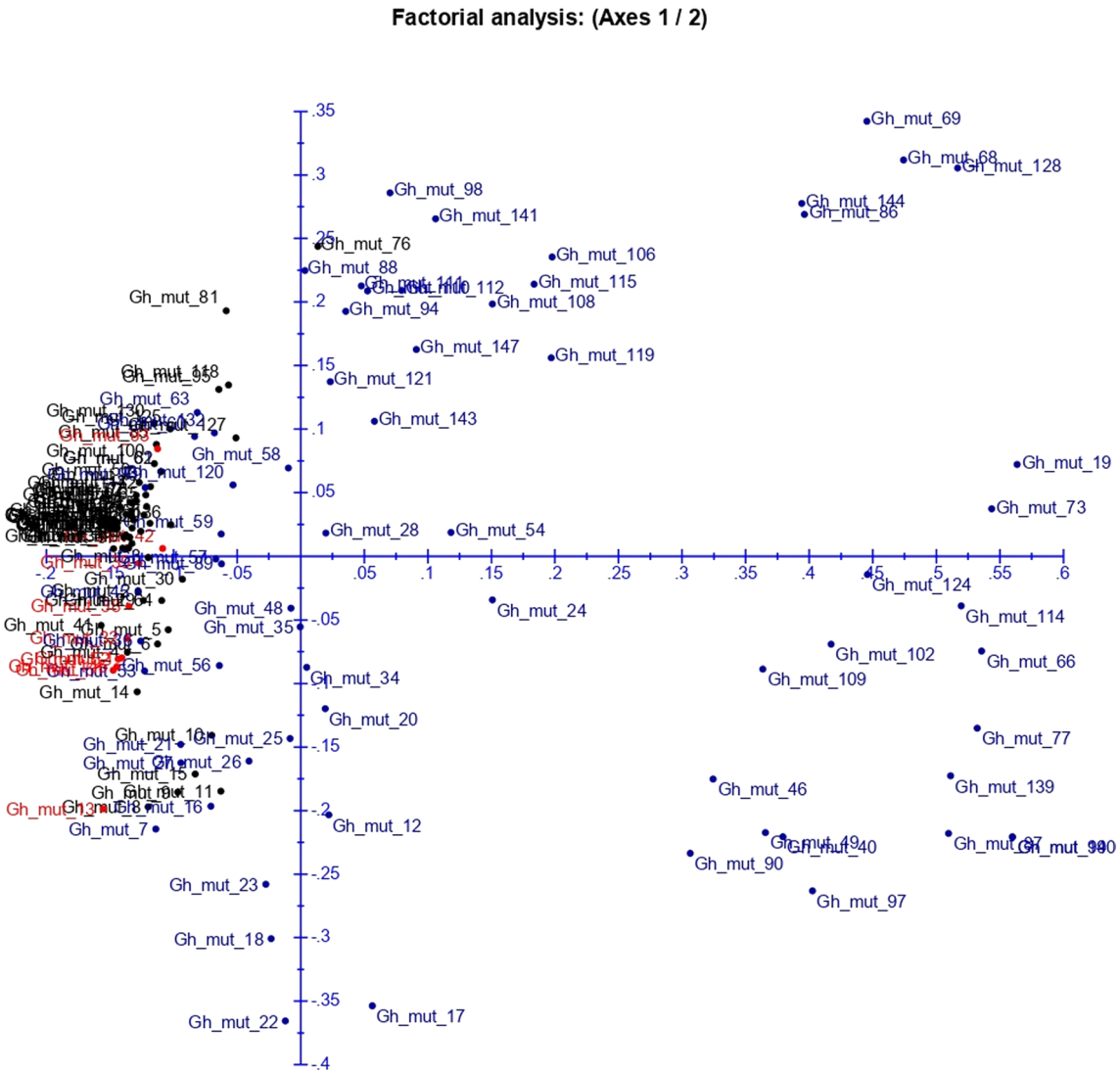


Fig. 6 Two-dimensional plot of principal component analysis (PCA) for genetic relationship among 150 M_3 plants of *G. herbaceum* along with the control, (Cluster I in blue color, Cluster II in black color and Cluster III in red color)

Acknowledgements The authors thank to the director CSIR-NBRI and MGM for providing the facilities to carry out the present study. UK thanks to CSIR for the Fellowship (Junior Research Fellowship & Senior Research Fellowship).

Funding The present study was funded by Council of Scientific & Industrial Research (CSIR) through a project "Genomics of medicinal plants and agronomically important traits (PlaGen)" (BSC 0107). The institutional assigned manuscript number is "CSIR-NBRI_MS/2020/06/03".

Declarations

Conflict of interest The authors have no conflict of interest.

References

Abid MA, Wang P, Zhu T, Liang C, Meng Z, Malik W, Guo S, Zhang R (2020) Construction of *Gossypium barbadense* mutant library provides genetic resources for cotton germplasm improvement. Int J Mol Sci 21(18):6505. https://doi.org/10.3390/ijms21186505

Aslam U, Cheema H, Ahmad S, Khan IA, Malik W, Khan AA (2016) COTIP: cotton TILLING platform, a resource for plant improvement and reverse genetic studies. Front Plant Sci 7:1863. https://doi.org/10.3389/fpls.2016.01863

Aslam R, Bhat TM, Choudhary S, Ansari MYK, Shahwar D (2017) Estimation of genetic variability, mutagenic effectiveness and efficiency in M_2 flower mutant lines of *Capsicum annuum* L. treated with caffeine and their analysis through RAPD markers. J King Saud Univ Sci 29(3):274–283. https://doi.org/10.1016/j.jksus.2016.04.008

Bechere E, Auld DL, Dotray P, Kebede H (2010) Registration of four upland cotton (*Gossypium hirsutum* L.) genetic stock mutants with tolerance to Imazamox. J Plant Regist 4(2):155–158. https://doi.org/10.3198/jpr2009.08.0446crgs

Brown N, Smith CW, Hague S, Auld D, Hequet E, Joy K, Jones D (2015) Within-boll yield characteristics and their correlation with fiber quality parameters following mutagenesis of upland cotton, TAM 94L–25. Crop Sci 55(4):1513–1523. https://doi.org/10.2135/cropsci2014.06.0442

Chantreau M, Grec S, Gutierrez L, Dalmais M, Pineau C, Demailly H, Paysant-Leroux C, Tavernier R, Trouve JP, Chatterjee M, Guillot X, Brunaud V, Chabbert B, Wuytswinkel O, Bendahmane A, Thomasset B, Hawkins S (2013) PT-Flax (phenotyping and TILLING of flax): development of a flax (*Linum usitatissimum* L.) mutant population and TILLING platform for forward and reverse genetics. BMC Plant Biol 13(1):159. https://doi.org/10.1186/1471-2229-13-159

Chawade A, Sikora P, Bräutigam M, Larsson M, Vivekanand V, Nakash MA, Chen T, Olsson O (2010) Development and characterization of an oat TILLING population and identification of mutations in lignin and beta glucan biosynthesis genes. BMC Plant Biol 10:86. https://doi.org/10.1186/1471-2229-10-86

Chen L, Huang L, Min D, Phillips A, Wang S, Madgwick PJ, Parry MAJ, Hu YG (2012) Development and characterization of a new TILLING population of common bread wheat (*Triticum aestivum* L.). PLoS One 7(7):e41570. https://doi.org/10.1371/journal.pone.0041570

Dahmani-Mardas F, Troadec C, Boualem A, Leveˆque S, Alsadon AA, Aldoss AA, Dogimont C, Bendahmane A (2010) Engineering melon plants with improved fruit shelf life using the TILLING approach. PLoS ONE 5(12):e15776. https://doi.org/10.1371/journal.pone.0015776

Danylchenko O, Sorochinsky B (2005) Use of RAPD assay for the detection of mutation changes in plant DNA induced by UV-B and γ-rays. BMC Plant Biol 5:S9. https://doi.org/10.1186/1471-2229-5-S1-S9

Dhakshanamoorthy D, Selvaraj R, Chidambaram A (2013) Induced mutagenesis in *Jatropha curcas* L. using Ethyl methanesulphonate (EMS) and assessment of DNA polymorphism through RAPD markers. J Crop Sci Biotechnol 16(3):201–207. https://doi.org/10.1016/j.crvi.2010.11.004

Espina MJ, Ahmed C, Bernardini A, Adeleke E, Yadegari Z, Arelli P, Pantalone V, Taheri A (2018) Development and phenotypic screening of an Ethyl methanesulfonate mutant population in soybean. Front Plant Sci 9:394. https://doi.org/10.3389/fpls.2018.00394

Gunapati S, Naresh R, Ranjan S, Nigam D, Hans A, Verma PC, Gadre R, Pathre UV, Sane AP, Sane VA (2016) Expression of *GhNAC2* from *G. herbaceum*, improves root growth and imparts tolerance to drought in transgenic cotton and *Arabidopsis*. Sci Rep 6:24978. https://doi.org/10.1038/srep24978

Hammer Ø, Harper D, Ryan P (2001) PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron 4(1):9

Hasan N, Choudhry S, Laskar RA (2020) Studies on qualitative and quantitative characters of mutagenised chili populations induced through MMS and EMS. Vegetos 33(4):793–799. https://doi.org/10.1007/s42535-020-00164-z

Jena SN, Srivastava A, Rai KM, Ranjan A, Singh SK, Nisar T, Srivastava M, Bag SK, Mantri S, Asif MH, Yadav HK, Tuli R, Sawant SV (2012) Development and characterization of genomic and expressed SSRs for levant cotton (*Gossypium herbaceum* L.). Theor Appl Genet 124(3):565–576. https://doi.org/10.1007/s00122-011-1729-y

Khan A, Pan X, Najeeb U, Tan DKY, Fahad S, Zahoor R, Luo H (2018) Coping with drought: stress and adaptive mechanisms, and management through cultural and molecular alternatives in cotton as vital constituents for plant stress resilience and fitness. Biol Res. https://doi.org/10.1186/s40659-018-0198-z

Krupa-Małkiewicz M, Bienias A (2018) BSA and molecular markers screening for salt stress tolerant mutant of Petunia obtained in *in vitro* culture. Ciência Rural. https://doi.org/10.1590/0103-8478cr20170042

Kulkarni VN, Khadi BM, Maralappanavar MS, Deshapande LA, Narayanan S (2009) The worldwide gene pools of *Gossypium arboreum* L. and *G. herbaceum* L., and their improvement. Genetics and Genomics of Cotton pp 69–97. https://doi.org/10.1007/978-0-387-70810-2_4

Kumar AP, Boualem A, Bhattacharya A, Parikh S, Desai N, Zambelli A, Leon A, Chatterjee M, Bendahmane A (2013) SMART–Sunflower mutant population and reverse genetic tool for crop improvement. BMC Plant Biol 13:38. https://doi.org/10.1186/1471-2229-13-38

Kumar U, Sawant SV, Yadav HK (2021) Exploring genetic variability in Ethyl methanesulfonate mediated mutant population of Wagad cultivar of *Gossypium herbaceum*. J Environ Biol 42(3):589–596. https://doi.org/10.22438/jeb/42/3/MRN-1399

Kurowska M, Daszkowska-Golec A, Gruszka D, Marzec M, Szurman M, Szarejko I, Maluszynski M (2011) TILLING-a shortcut in functional genomics. J Appl Genet 52(4):371–390. https://doi.org/10.1007/s13353-011-0061-1

Laskar RA, Khan S (2017) Assessment on induced genetic variability and divergence in the mutagenized lentil populations of microsperma and macrosperma cultivars developed using physical and chemical mutagenesis. PLoS ONE 12(9):e0184598. https://doi.org/10.1371/journal.pone.0184598

Lian X, Liu Y, Guo H, Fan Y, Wu J, Guo H, Gou Z (2020) Ethyl methanesulfonate mutant library construction in *Gossypium hirsutum* L. for allotetraploid functional genomics and germplasm innovation. Plant J 103(2):858–868. https://doi.org/10.1111/tpj.14755

Maghuly F, Pabinger S, Krainer J, Laimer M (2018) The pattern and distribution of induced mutations in *J. curcas* using reduced representation sequencing. Front Plant Sci 9:524. https://doi.org/10.3389/fpls.2018.00524

Mardas FD, Troadec C, Boualem A, Leveque S, Alsadon AA, Aldoss AA, Dogimont C, Bendahmane A (2010) Engineering melon plants with improved fruit shelf life using the TILLING approach. PLoS ONE 5(12):e15776. https://doi.org/10.1371/journal.pone.0015776

Parekh MJ, Kumar S, Zala HN, Fougat RS, Patel CB, Bosamia TC, Kulkarni KS, Parihar A (2016) Development and validation of novel fiber relevant dbEST-SSR markers and their utility in revealing genetic diversity in diploid cotton (*Gossypium herbaceum* and *G. arboreum*). Ind Crops Prod 83:620–629. https://doi.org/10.1016/j.indcrop.2015.12.061

Parekh MJ, Kumar S, Fougat RS, Zala HN, Pandit RJ (2018) Transcriptomic profiling of developing fiber in levant cotton (*Gossypium herbaceum* L.). Funct and Integr Genomics 18(2):211–223. https://doi.org/10.1007/s10142-017-0586-4

Perrier X, Flori A, Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Genetic diversity of cultivated tropical plants. Enfield Science Publishers, Montpellier, pp 43–76

Raj A, Kumar S, Haq I, Kumar M (2014) Detection of tannery effluents induced DNA damage in mung bean by use of random amplified polymorphic DNA markers. ISRN Biotechnol. https://doi.org/10.1155/2014/727623

Ranjan A, Nigam D, Asif MH, Singh R, Ranjan S, Mantri S, Pandey N, Trivedi I, Rai KM, Jena SN, Koul B, Tuli R, Pathre UV, Sawant SV (2012) Genome wide expression profiling of two accession of *G. herbaceum* L. in response to drought. BMC genom 13:94. doi: https://doi.org/10.1186/1471-2164-13-94

Rawat N, Schoen A, Singh L, Mahlandt A, Wilson DL, Liu S, Lin G et al (2018) TILL-D: An *Aegilops tauschii* TILLING resource for wheat improvement. Front Plant Sci 9:1665. https://doi.org/10.3389/fpls.2018.01665

Sabetta W, Alba V, Blanco A, Montemurro C (2011) sunTILL: a TILLING resource for gene function analysis in sunflower. Plant Methods 7:20. https://doi.org/10.1186/1746-4811-7-20

Saito T, Ariizumi T, Okabe Y, Asamizu E, Hiwasa-Tanase K, Fukuda N, Mizoguchi T, Yamazaki Y, Aoki K, Ezura H (2011) TOMA-TOMA: a novel tomato mutant database distributing Micro-Tom mutant collections. Plant Cell Physiol 52(2):283–296. https://doi.org/10.1093/pcp/pcr004

Schreiber M, Barakate A, Uzrek N, Macaulay M, Sourdille A, Morris J, Hedley PE, Ramsay L, Waugh R (2019) A highly mutagenised barley (cv. Golden Promise) TILLING population coupled with strategies for screening-by-sequencing. Plant Methods 15(1):99. https://doi.org/10.1186/s13007-019-0486-9

Slade AJ, Mcguire C, Loeffler D, Mullenberg J, Skinner W, Fazio G, Holm A, Brandt KM, Steine MN, Goodstal JF (2012) Development of high amylose wheat through TILLING. BMC Plant Biol 12(10):1186. https://doi.org/10.1186/1471-2229-12-69

Taryono T, Cahyaningrum P (2011) The detection of mutational changes in Sorghum using RAPD. Indones J Biotechnol 16(1)

Tiliouine WA, Laouar M, Abdelguerfi A, Cieslak JJ, Jankuloski L, Till BJ (2018) Genetic variability induced by gamma rays and preliminary results of low-cost TILLING on M_2 generation of chickpea (*Cicer arietinum* L.). Front Plant Sci 9:156. https://doi.org/10.3389/fpls.2018.01568

Trivedi I, Ranjan A, Sharma YK, Sawant S (2012) The histone H1 variant accumulates in response to water stress in the drought tolerant genotype of *Gossypium herbaceum* L. Protein J 31(6):477–486. https://doi.org/10.1007/s10930-012-9425-6

Wang M, Wang Q, Wang B (2012) Identification and characterization of microRNAs in Asiatic cotton (*Gossypium arboreum* L.). PLoS One 7(4):e33696. https://doi.org/10.1371/journal.pone.0033696

Yunus MF, Aziz MA, Kadir MA, Daud SK, Rashid AB (2013) In vitro mutagenesis of *Etlingera elatior* (Jack) and early detection of mutation using RAPD markers. Turk J Biol 37:716–725. https://doi.org/10.3906/biy-1303-19