Mutation Breeding

9

Reinforcing Mutation Breeding by Augmenting Genomic Resources in Pulse Crops

J. Souframanien*[1], P. Dhanasekar[1], V. J. Dhole[1], Swapnonil Banerjee[2] and L. Srinivas[2]
[1]*Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre (BARC), Trombay-400085, INDIA*
[2]*Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre (V), Visakhapatnam-531011, INDIA*

Maruca pod borer tolerant cowpea mutant Mr2

ABSTRACT

Pulse crops, with high protein content, are valuable components of Indian agriculture as they can satisfy the nutritional demands of our growing population. However, these crops with narrow genetic diversities, have received comparatively less research attention. Mutation breeding has been successful in the genetic improvement of these orphan, yet important pulse crops such as pigeonpea, mungbean, black gram and cowpea. Also, induced mutagenesis has the potential to enrich the genetic resources in these crops, equipping the different crop improvement programs with arsenal to meet the challenges of climate change. Concurrently, augmenting the genomic resources is equally significant for accelerated and targeted breeding of pulse crops.

KEYWORDS: *Genetics, Genomics, Mutation breeding, Next-generation sequencing, Pulses, SSRs, SNPs, Whole genome*

Introduction

Pulse crops, the cornerstone of Indian agriculture, are the primary source of dietary proteins for the predominantly vegetarian population of our country. These crops play an immense role in enhancing and ensuring nutritional security by complementing carbohydrate-rich cereals-based diets[1]. India predominantly cultivates chickpeas, pigeonpeas, mungbeans, black gram, cowpeas, lentils, and various beans. Despite being the leading global producer, consumer, and processor of pulses, India's reliance on net imports is concerning [2]. Mutation breeding is a GMO-free [3] complementary method to the narrow genetic base limited conventional breeding for inducing novel variations and improving yields of pulse crops.

Traditionally, ionizing radiations such as X-rays, gamma rays, beta particles, and fast and thermal neutrons were employed to induce mutations in these crops. However, new energy sources like electron-, proton-, and heavy-ion beams are increasingly adding new perspectives to mutation breeding.Though induced mutation alone or in combination with conventional breeding has the potential to create variations, the availability of genomic resources profoundly influences the pace of accelerated genetic crop improvement. The advent of next-generation sequencing (NGS) technologies has resulted in the development of extensive molecular resources, including transcriptome sequence data, genetic and physical maps, and molecular markers, enabling trait mapping and marker-assisted breeding faster and more reliable. For fast-tracking pulse crop improvement, it is imperative to use radiation to broaden the variation and concurrently develop exhaustive genomic resources.

**Author for Correspondence: J. Souframanien*
E-mail: souf@barc.gov.in

Materials and Methods

Induced mutagenesis

Thousands of genetically pure, uniform, dry seeds (12% moisture content) of pulse crops after exposure to predetermined doses of ^{60}Co gamma rays were used to raise M_1 and M_2 generations. Variants with putative mutations for desirable traits, identified in the M_2 generation, were ascertained for their stability and uniformity (homozygosity) in the later generations M_3–M_5. If found to be agronomically superior, these were directly evaluated for their potential release as varieties. Conversely, the poor mutants were improved through recombination breeding before entering into varietal trials. In either case, the promising mutants or their derivatives underwent a battery of rigorous testing with other competitors and check varieties (in station to multi-location and adaptive varietal trials) for different parameters such as yield, nutrition, disease, insect pest and nematodes resistance, in coordination with various State Agricultural Universities and or All India Coordinated Research Project on Pulses of the Indian Council of Agricultural Research. Subsequently, potential candidates that successfully emerged from these meticulous testing were released and notified through the Gazette for commercial cultivation by the Ministry of Agriculture and Farmers Welfare.

Developing genomic resources

Next-generation sequencing-based Illumina short reads and the third generation Oxford Nanopore sequencing-based long reads of the whole genome sequences of the black gram cultivar 'Pu31' and that of cowpea cultivar 'CPD103' were individually hybrid assembled de novo using the MaSuRCA (Maryland Super Read Cabog Assembler) genome assembler. The completeness of the assembled genomes was determined through BUSCO analysis. These assemblies were subsequently

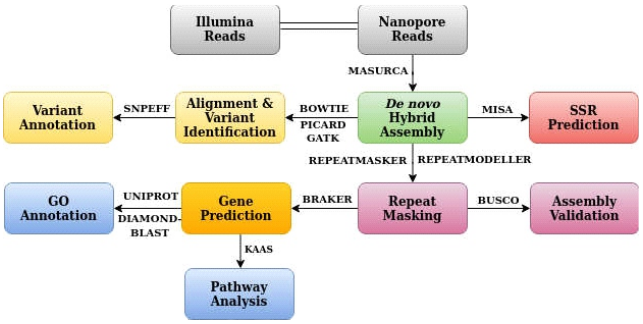

Fig.1: Whole genome sequencing bioinformatics workflow.

used for downstream analysis and also as reference genomes for resequencing of some induced mutants in these crops. While the entire set of genes and the encoded proteins across the genomes were predicted using BRAKER tool, the SSR motifs were deciphered using the MISA tool. The genome analysis tool kits, BOWTIE and PICARD, were employed to identify single base substitutions and small insertion or deletion variants (Indels) in the mutants as against the parental genomes. The gene products were subjected to gene ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for assigning gene ontology terms to describe their biological functions, processes, and cellular components and to map genes to complex interaction networks as per the KEGG database, respectively. The entire bioinformatic workflow is depicted in Fig.1. The whole genome sequences were also mined in silico for R-genes, transcriptionally active proteins and protein kinases involved in regulating gene expressions under a myriad of biotic and abiotic stresses. In addition, RNA-seq based transcriptome profiling enabled development of genic-SSRs and SNPs and treatment context-based identification of differentially expressed genes.

Results and Discussion

Induced mutagenesis

Of the 24 mutant varieties of pulse crops developed through mutation breeding at this Centre, five of them (three in black gram and one each in mungbean and cowpea) have been released in the last five years for commercial cultivation

Fig.2: Maruca pod borer tolerant cowpea mutant Mr2.

Fig.2: Recent varieties of Trombay pulse crops suitable for summer season. (a) Mungbean: TRCRM-147, (b) Cowpea: TC-901.

(Table 1, Fig.2). The cowpea variety, TC-901, is a gamma ray-induced direct mutant [4], while the rest of the four varieties are mutant derivatives involving mutant x cultivar or cultivar x mutant crosses (Table 1). The success of any mutation breeding program is largely dependent on the selection of appropriate parental genotype for mutation induction, appropriate dose for maximizing mutation recovery, size of the mutagenized population in M_1 and subsequent generations, and proper screening techniques for rapid and cost-effective identification of mutants [1]. A series of radio-sensitivity assays are conducted for determining the genotype-specific optimal dose prior to embarking on mutation breeding experiments. Optimal doses causing less than 50% lethality eg. LD_{20} or LD_{30}, or less than 50% growth reduction (GR_{50}) are understandably more desirable in autogamous plants like pulses that ensure maximal recovery of useful mutations with least mortality-like unintended damages. Successful outcomes of mutation breeding programs in pulse crops suggest that doses in the range of 100–200 Gy, 200–300 Gy and 300–400 Gy, are effective in pigeonpeas, cowpeas, and in mungbeans/urdbeans, respectively. Effective doses of electron beams were also in close proximity to that of gamma rays in different pulse crops (mungbean:500 Gy; urdbean:400 Gy; and cowpea:270 Gy) [5]. New mutant varieties have been carefully compared to competitors and check varieties, and have undergone extensive multi-location and adaptive varietal trials lasting almost a decade to ensure their success in the field. Thanks to this process, the farmers now have access to a range of high-quality, disease-resistant mutant pulse varieties that have improved yields and supported sustainable agriculture practices. As perceived in the present research outcome, the induced mutants often by themselves are not appropriate for varietal release, but have the potential to be more productive when used in cross breeding [6]. The contribution of induced-mutagenesis in enhancing the genetic resources can also be appreciated from the wide range of novel mutants identified in these pulse crops. For instance, the unavailability of useful resistance sources against *Maruca* pod borer in the cowpea gene pool was partially overcome by identifying tolerant mutants (Fig.3) (major accomplishment of BARC under the FAO/IAEA ongoing CRP project D22006) [7]. The induced mutants also serve as genetic treasures for identifying candidate genes and for exploring the intricacies of developmental biology in pulse crops.

Development of Genomic Resources

Draft whole genome sequence

A draft whole genome sequence of blackgram cultivar Pant U-31 was constructed for the first time by employing hybrid genome assembly based on Illumina and Nanopore reads with a total sequencing depth of ~148X. The final de novo whole genome spanned 475 Mb with a maximum scaffold

Table 1: Trombay mutant varieties notified in the last five years in different pulse crops.

Sr. No.	Mutant variety	Mutant/Mutant derivative	Year of release	Released for	Important attributes	Collaborating Institute
	Blackgram					
1	TJU-130	Cultivar x mutant	2023	MP	YMD resistant, Summer suitable	JNKVV, Jabalpur
2	TJU-339	Mutant x cultivar	2023	MP	YMD resistant, Summer suitable	JNKVV, Jabalpur
3	TRCRU-22	Mutant x cultivar	2023	Karnataka	Medium large seeds, YMD resistant, summer suitable	UAS, Raichur
	Mungbean					
4	TRCRM-147	Mutant x cultivar	2023	Karnataka	Large seed size, YMD resistant, Summer suitable	UAS, Raichur
	Cowpea					
5	TC-901	Direct mutant	2018	North Zone	Mosaic and root-rot resistant, early, summer suitable	IIPR-Kanpur

length of 6.3 Mb and scaffold N50 of 1.42 Mb[8]. Along similar lines, whole genome assembly of cowpea cultivar CPD103 resulted in an estimated genome size of ~377 Mb with a maximum scaffold length of 1.6 Mb and scaffold N50 of 26.7 Kb at a sequencing depth of ~120X [9]. Next generation sequencing reads are too short to resolve plant genomes abundant with repetitive elements resulting in assembly ambiguity [10]. Therefore, in the present study, a hybrid assembly approach involving the third generation Nanopore sequencing technique that is capable of delivering several kilobases of long reads was integrated with more accurate Illumina short reads to construct scaffold level reference genome assembly. The completeness of the genome assembly was gauged using the BUSCO analysis wherein 96.8% and 90.0% of complete genes in respect of black gram and cowpea, respectively, were recovered suggesting robustness of the assembly. About 42,115 genes were predicted with a mean coding length of 1131 bp in black gram as against 65,708 genes with a mean coding length of 915.7 bp in case of cowpea. The cowpea coding sequences varied between 42 bp and 15.2 Kb. Roughly, 80.9% and 85.4% of the predicted genes in black gram and cowpea, respectively, could be annotated (Fig.4).

Transposable elements (TEs)

In black gram, almost half of the assembled genome comprised of repetitive elements, majorly retrotransposons (RNA, 47.3% of genome) and minority of DNA transposons (2.29% of genome). But unlike the other related *Vigna* species, black gram housed more members of *Copia* super family

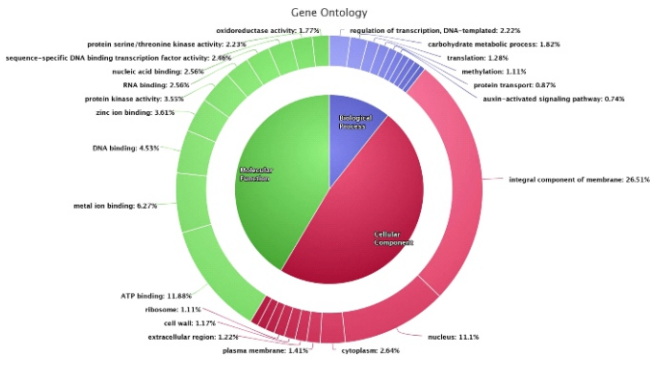

Fig.4: Gene ontology chart of Vigna mungo [8].

(34.5%) of LTR retrotransposons as against the Gypsy elements (13.4%). In cowpea also the retrotransposons (91.1%) were the major TEs with DNA transposons forming a minor component (4.4%). But among the LTR retrotransposons, the Gypsy superfamily predominated compared to the *Copia* elements. TEs, the potential resources of phenotypic variation and plasticity, are the major drivers of genome expansion [8]. They aid crop improvement programs by serving as molecular markers within or are closely linked to various QTLs as well as defense-related genes. They have been widely used for diversity studies and trait mapping [11].

Identification of genomic SSR motifs

From the whole genome sequencing of black gram, a total of 166,014 SSRs, including 65,180 compound SSRs, were identified and primer pairs for 34,816 SSRs were designed. Among 1,66,014 SSRs (excluding mono nucleotide repeats) identified, the proportions of dinucleotide repeats were higher (38.1%) than the other repeats in *V. mungo* [8]. Similarly, in cowpea, SSRs numbering 1,73,866 were identified from 16,063 sequences including 23,082 compound SSRs. Among these SSRs, barring the mono nucleotide repeats, tetra nucleotide repeats predominated (19.5%) in contrast to the dinucleotide repeats (15.6%). Pulses exhibited variation with respect to the predominance of particular nucleotide repeats. Mungbean showed prevalence of dinucleotide repeats, while in cluster beans tetra nucleotide repeats prevailed (19.9%), whereas common beans exhibited preponderance of trinucleotide repeats. Proportion of tri-, tetra-, penta-, and hexa-nucleotide SSRs in black gram were more or less same in comparison to *V. radiata* (24.6%, 2.5%, 1.2%, 0.2%) but lower than *V. marina* (49%, 3%, 7%, 5%), except for the tetra-nucleotide repeats. In cowpea, the corresponding proportions were 5.9%, 19.5%, 0.12% and 0.001%. These SSR markers are valuable for genomic mapping, DNA fingerprinting, and marker-assisted selection in plant species and pulses in particular [12].

Discovery of R-gene domains

In black gram, of the 33,959 annotated gene products, 1659 proteins showed the presence of R-gene (resistance gene) related domains, among which the predominant KIN class (905) was followed by RLK (239) and RLP classes (138) [8]. In contrast, a total of 2188 R-gene domain containing proteins were identified in cowpea, which also exhibited

predominance of KIN class (855) genes followed by the transmembrane receptors, RLKs (258) and RLPs (238). The classic nucleotide binding sites (NBS)-harbouring genes numbered 142 in black gram, which was in sharp contrast to 392 in cowpea. Only twenty-seven proteins were found to represent the class of cytoplasmic proteins (CNL and TNL) in black gram that was distinctly differing from 183 proteins in cowpea. These R-genes could be exploited in pyramiding resistance genes in new cultivars favouring effective and environment friendly approach of plant disease control [8].

KEGG pathway analysis

KEGG analysis identified 16,404 unique pathways in black gram of which majority of the genes were grouped into protein families (genetic information processing, signalling and cellular processes, and metabolism) followed by carbohydrate metabolism and transcription-related activities. In cowpea, 17,283 unique KEGG pathways were identified. Like in black gram, majority of them were grouped into protein families, followed by carbohydrate metabolism and transcription activities. KEGG pathway analysis provides a comprehensive overview of biological pathways enabling scientists to identify key molecules and pathways involved in specific biological and cellular processes. They could be applied to different organisms that share orthologous genes and help uncover hidden features in large-scale biological data, such as cellular and organism-level functions [13].

Orthologous gene comparison

Orthologous gene comparison study involving genes from black gram, mungbean, cowpea and adzuki bean showed 19,095 gene clusters to be shared by all four species, while 1970 clusters were specific to black gram. Cowpea exhibited 789 unique clusters among the four species, while it shared 20,091 orthologous genes with black gram (Fig.5). Such orthologous gene study aids in comparative genomics, phylogenetics, protein function annotation, understanding genome rearrangements, providing hypotheses about gene function across different species. In, it also provides insights into evolutionary relationship and gene family dynamics [14, 15].

Identification of genic-SSRs and SNPs

Transcriptome sequencing using NGS technologies allows quick and inexpensive identification of SNPs/SSR within the coding regions. These are more likely to be related with various biological functions. In cultivated blackgram (TU94-2), towards developing genic-SSR and SNP markers, a

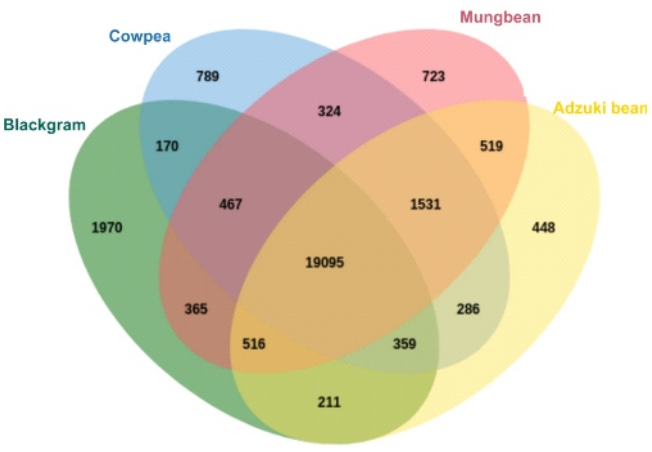

Fig.5: Venn diagram showing shared orthologous gene clusters among black gram, cowpea, mungbean and adzuki bean [8].

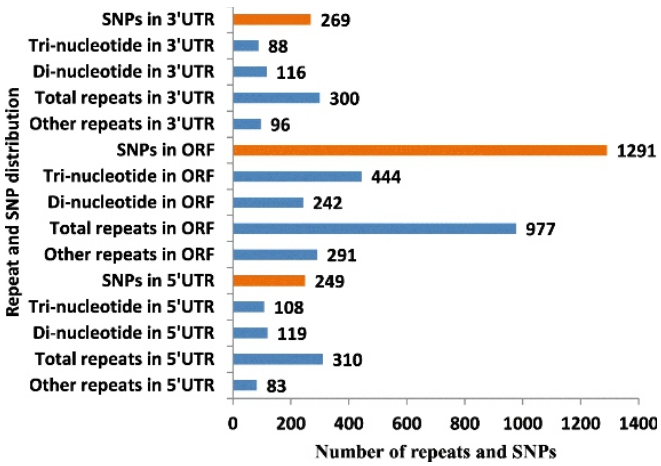

Fig.6: Frequency and distribution of SSRs and SNPs in coding sequence and untranslated region (UTRs) of blackgram TCS [17].

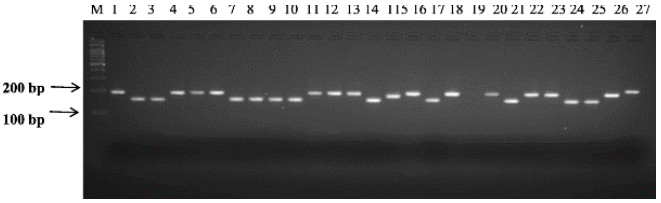

Fig.7: PCR amplification of a genic-SSR marker in 27 blackgram genotypes [17].

total of 48,291 transcript contigs (TCS) were searched for SSRs and 1,840 SSRs were identified in 1,572 TCS with an average frequency of one SSR per 11.9 kb. The tri-nucleotide repeats were most abundant (35%), followed by di-nucleotide repeats (32%). PCR primer pairs were successfully designed for 933 SSR loci. Sequences analyses indicate that about 64.4% and 35.6% of the SSR motifs were present in the coding sequences (CDS) and untranslated regions (UTRs), respectively (Fig.6) [16,17]. Similarly in the wild blackgram accession (*Vigna mungo* var. *silvestris*), out of 40,178 TCS, 1621 SSRs were identified in 1339 (3.3%) TCS, with an average frequency of one SSR per 11.1 kb. Tri-nucleotide repeats were found to be most abundant (646,39.9%) followed by di-nucleotide (490,30.2%), which together constituted 70.1% of the identified SSRs. Of the 1621 SSR motifs that were identified, PCR primer pairs were successfully designed for 1171 SSR loci. Of the total of 1844 SNPs that were identified in TW samples, 17 SNPs were heterozygous and 1828 SNPs were homozygous. Out of 1844 SNPs identified, PCR primer pairs were successfully designed for 1749 SNP loci and validated in a set of black gram genotypes (Fig.7) [17]. In cowpea, a total of 41,506 transcript contigs (TCS) were searched for SSRs and 3878 SSRs were identified in 3315 TCS with an average frequency of one SSR per 9.34 kb in cultivar CPD103. The tri-nucleotide repeats were most abundant (54.3%), followed by tetra-nucleotide repeats (21.5%). The most frequent motif type was A/T, which was followed by GAA/CTT. The availability of such a large number of sequence-based markers allows genetic diversity analysis, linkage mapping, comparative genomics, and association studies [17].

Nature of Gamma Rays-induced Stable Mutations at Whole Genome Level

Gamma rays induced single base substitutions and indels in three cowpea mutants (M_6 generation) (disease resistant, large seed size and small seed size mutants) as deduced through Illumina based next-generation

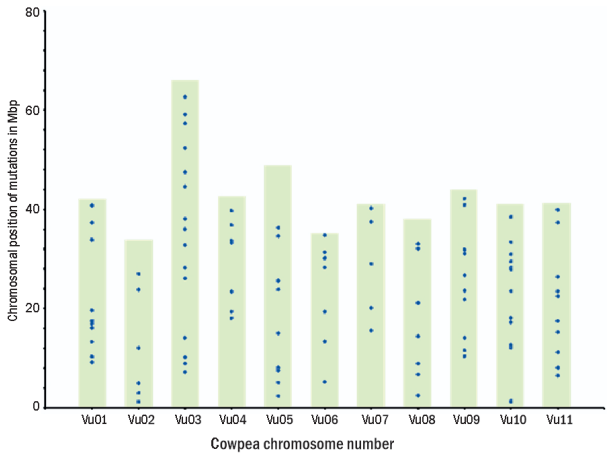

Fig.8: Chromosomal distribution of induced mutations (SBSs and indels) in gamma ray-induced cowpea mutants (M_6) [19] .

resequencing. A relatively higher frequency (88.9%) of single base substitutions (SBSs) with an average transition to transversion ratio (Ti/Tv) of 3.51 was observed [9]. A>G transitions, including its complementary T>C transitions, predominated the transition mutations and all four types of transversion mutations were detected. Among the indels (11%), small insertions (6.3%) were relatively more prominent than small deletions (4.8%). Single-base indels (involving A/T bases) showed preponderance, but indels of up to three bases were also detected in low proportions. Scattered across all the 11 chromosomes, 5.61±0.67 SBSs and 0.7±0.15 indels were induced per chromosome on an average (Fig.8). Only a fraction of SBSs (19.45%) and indels (20.2%) potentially altered the encoded amino acids/peptides, which concurs with previous reports [18]. The mutation rate (1.4×10^{-7} per bp) induced by gamma rays in the present study is about 0.6× of that in rice ($2.31 \pm 1.5 \times 10^{-7}$) [19], probably due to genomic differences.

Impact of Mutant Pulse Crop Varieties in India

Indian farmers are increasingly adopting Trombay mutant pulse crop varieties due to their superior yields, which are leading to higher on-farm incomes. One of the earliest black gram mutant varieties, TAU-1, released in 1985, is still highly coveted among farmers in Maharashtra and remains the mainstay of black gram cultivation in the state. TU-40 black gram variety is gaining popularity in the southern states, especially in the rice fallow niches. Pigeonpea varieties, TJT-501 and TT-401, are also incredibly popular, with the former being grown in almost half of the pigeonpea growing areas in Madhya Pradesh, while the latter, released for the central zone, is making inroads into the southern states with yields as high as 2500 kg/ha. The Maharashtra farmers cultivating PKV-TARA are reaping record pigeonpea yields under drip irrigation. The mungbean variety TMB-37 released for North East Plain Zone is garnering huge demands in the country for its earliness and summer suitability, and is re-released in the Punjab state. Likewise, the powdery mildew and *Corynespora* leaf spot disease-resistant varieties, TM-96-2 and TM-2000-2, are bringing rice fallow areas under mungbean cultivation in Andhra Pradesh and Chhattisgarh, respectively [5]. The recently released pulse varieties, TC-901 (cowpea), TRCRM-147 (mungbean) and TRCRU22, TJU339 and TJU-130 (blackgram), with resistance to yellow mosaic disease are contributing to area expansion under summer cultivation. These Trombay mutant pulse varieties, enjoying the sustained patronage of the Indian farmers, are instrumental in boosting the productivity and nutritional security of the country and now fulfill more than 3% of the national breeder seed demand of pulse crops in the country.

Conclusion

The accomplishments in mutation breeding of pulse crops at this Centre underscores the significance of the physical mutagens such as gamma rays in broadening the genetic variability of pulse crops and their contribution in sustained enhancement of the productivity to meet the growing demands and emerging challenges of climate change. The impact of mutant varieties demonstrates the immense potential of induced mutagenesis in rescuing the conventional breeding and contributing towards an environmentally sustainable agriculture. The rich genomic data generated through the molecular studies provide valuable resources for accelerated genetic enhancement of pulse crops through marker assisted breeding and enables adoption of newer genome editing tools like CRISPR-Cas in these important 'orphan' crops.

References

[1] Souframanien J, Dhanasekar P. In: Penna S, Jain SM (eds) Mutation Breeding for Sustainable Food Production and Climate Resilience. Springer, 2023, p 445–485.

[2] Dixit GP, Srivastava AK, Ali H, Indian Farming, 2024, 74(2), 03-06.

[3] Parry MAJ, Madgwick PJ, Bayon C, et al., J Exp Bot, 2009, 60(10), 2817-2825.

[4] Dhanasekar P, Reddy KS, In: Abstracts of FAO/IAEA International symposium on Plant mutation breeding and biotechnology, Vienna, 27-31 August, 2018, IAEA-CN-263-170

[5] Souframanien J, Saha AJ, Dhole VJ, et al., IANCAS Bull, 2020, XV, 71–80.

[6] Pawar SE, Manjaya JG, Souframanien J et al., Paper presented in DAE-BRNS Symposium on the use of Nuclear Techniques in crop improvement Dec 6-8, 2000, Mumbai, p 170-174.

[7] Dhanasekar P, Souframanien J, Dhole VJ et al., In: Souvenir of International Conference on Pulses: Smart crops for Agricultural Sustainability and Nutritional Security, New Delhi, India, 10-12 February, 2023, S6.46, p 519.

[8] Souframanien J, Raizada A, Dhanasekar P, et al., Sci Rep, 2021, 11, 11247.

[9] Dhanasekar P, Souframanien J, Int J Radiat Biol, 2024 doi:10.1080/09553002.2024.2345087

[10] Wang Y, Zhao Y, Bollas A, et al., Nat Biotechnol, 2021, 39(11), 1348–1365.

[11] Arvas YE, Abed MM, Zaki QA, et al., IOP Conf. Ser.: Earth Environ Sci, 2021, 761, 012031.

[12] Panigrahi KK, Panigrahi P, Mohanty A, et al., Genetika, 2020, 52(3), 1161-1179.

[13] David C, Kondylakis H., Information, 2024, 15(1), 56.

[14] Altenhoff AM, DessimozC, PLoS Comput Biol., 2019, 5(1), e1000262.

[15] Hu Y, Flockhart I, Vinayagam A et al., BMC Bioinform., 2011, 12, 357.

[16] Souframanien J, Reddy KS, PLoS ONE, 2015, 10(6), e0128748.

[17] Raizada A, Souframanien J. BMC Plant Biol, 2019, 19, 358.

[18] Yang G, Luo W, Zhang J, et al., Front Plant Sci, 2019, 10, 1514.

[19] Li F, Shimizu A, Nishio T et al., G3: Genes Genomes Genet, 9 (1), 3743–3751.