

Allele Mining for Smart Breeding

Delma Ulahannan¹, Dr. S. R. Patel¹ and Dr. R. K. Kalaria²

¹Department of Genetics and Plant Breeding, College of Agriculture, Navsari Agricultural University, Bharuch, Gujarat

²Bioinformatics section, ASPEE Shakilam Biotechnology Institute, Navsari Agricultural University, Surat, Gujarat

Received: April, 2025; Accepted: May, 2025; Published: July, 2025

INTRODUCTION

Allele mining is an approach that is used to dissect naturally occurring allelic variations or suitable alleles of a candidate gene those controlling key agronomic traits which has potential in crop improvement. It is a research field aimed at identifying allelic variation of relevant traits within genetic resource collections. It is a finding of superior alleles from the natural population. Enormous sequence information is available

in public databases as a result of sequencing of diverse crop genomes. It is important to use this genomic information for the identification and isolation of novel and superior alleles of agronomically important genes from crop gene pools to suitably deploy for the development of improved cultivars. It helps in tracing the evolution of alleles, identification of new haplotypes and development of allele-specific markers for use in marker-assisted selection.

APPROACHES FOR ALLELE MINING

Eco-TILLING: Technique used in allele mining to detect natural genetic variations or mutations in genes of interest within a population. It combines the principles of TILLING (Targeting Induced Local Lesions in Genomes) with evolutionary studies.

Sequence based allele mining: Powerful approach used to identify and characterize novel

CASE STUDIES

Bhullar *et al.* (2010) undertaken a large-scale molecular allele mining to isolate new alleles of the powdery mildew resistance gene *Pm3* from wheat gene bank accessions. The search for new *Pm3* alleles were carried out on a geographically diverse set of 733 wheat accessions originating from 20 countries. *Pm3* specific molecular tools as well as classical pathogenicity tests were used to characterize the accessions. Two new functional alleles *Pm3-4650* and *Pm3-2816* were identified out of the eight newly cloned *Pm3* sequences. Thus, the set of functional *Pm3* alleles family now includes 17 genes, making it one of the largest allelic series of plant resistance genes.

Karikari *et al.* (2020) analyzed the genetic base

alleles from genomic or transcriptomic data using various bioinformatics tools.

Association mapping-based allele mining: Efficient method that combines association mapping with allele mining to identify genetic variants linked to specific traits.

of 163 soybean genotypes from three continents *viz.* Africa, America and Asia using 68 trait-linked simple sequence repeats (SSR) markers. Association mapping Model-based population structure analysis also divided them into 4 subpopulations. The results indicated that soybean germplasm has moved from Asia to America, and from America to Africa. Based on the association mapping, a total of 21 SSR markers showed significant association with days to flowering (DoF) and 100-seed weight (HSW). Two markers Satt365 and Satt581 located on chromosome No.6 and No.10, respectively, showed pleiotropic effect or linkage on both traits. The three studies on diverse genotypes, polymorphic SSR markers

and desirable alleles for DoF and HSW are useful in future breeding programs to improve quantitative traits.

Singh *et al.* (2024) re-sequenced 399 genotypes from 3K-RG (Rice Genome) and mined to identify the superior haplotypes for 95 drought-responsive candidate genes. Candidate gene-based association analysis revealed 69 marker-trait associations (MTAs) in 16 genes for single plant yield (SPY) under drought stress. Hapl-pheno analysis of these 16 genes, they identified seven genes *viz.*, *OsGSK1*-H4, *OsDSR2*-H3, *OsDIL1*-H22, *OsDREB1C*-H3, *ASR3*-H88, *DSM3*-H4 and *ZFP182*-H4 with superior haplotypes which were associated with the higher SPY under drought stress. These lines carrying the superior haplotypes can be used as potential donors in haplotype-based smart breeding to develop high-yielding drought-tolerant rice varieties.

Marla *et al.* (2025) validated the Segaolane *dw3* loss-of-function variant in sorghum, resulting from a 137 base pair deletion in the third exon, to suppress revertant production. In nested association mapping, they crossed RTx430 X Segaolane and produced recombinant inbred line (RILs). The RIL with the *dw3*-ref allele (RTx430) produced revertants while no revertants were observed in RILs with the new Segaolane *dw3* allele. The availability of resequencing data enabled the designing of haplotype-based markers detecting the Segaolane stable *dw3* allele for marker-assisted trait introgression into elite sorghum breeding lines. This research mine identified new stable-dwarfing *dw3* alleles will be useful for smart

CONCLUSION

Allele mining plays an increasingly vital role in the development of improved crop varieties by searching the elite alleles and utilize them in smart breeding to meet the global growing demand for food while minimizing environmental stress impacts. Through the systematic exploration of genetic diversity within crop populations,

breeding to develop stable dwarf varieties in sorghum.

Pattnaik *et al.* (2025) employed standalone BLAST search, to identify genes involved in flowering regulation in pigeon pea. Protein sequences of six known flowering regulating genes from *Arabidopsis* and related plants were retrieved from the NCBI database. The entire set of protein sequences from pigeon pea was used as the database for comparison. The top hits with more than 30% identity and known conserved domains were considered true orthologs, resulting in the identification of six pigeon pea genes, *viz.*, *CcFrigida*, *CcFrigida Like1*, *CcFrigida Like2*, *CcFrigida Essential1*, *CcTerminal Flowering1* and *CcTerminal Flowering2*.

Thyssen *et al.* (2025) conducted studies in a set of 132 diverse cotton lines (*Gossypium hirsutum*), where the genomic DNA was extracted and sequenced by using Illumina HiSeq paired end short read. Genes near to *qA07-STR-1731* locus had been analyzed through transcriptome mining and identified one candidate gene having non-synonymous variation in expression at 8th & 16th days post anthesis (DPA). Variation in RNA expression in a diverse population at different developmental points was also indicative of a gene that potentially controls a phenotype. Candidate genes for traits like flame resistance that are likely due to the combination of many QTLs that could be benefited from this multi-mining approach. The candidate gene identified near to *qA07-STR-1731* locus could be useful for smart breeding to improve cotton fiber quality.

allele mining has the potential to unlock valuable traits, such as powdery mildew disease resistance, drought tolerance and yield which are essential for ensuring food security and sustainable agriculture in a changing climate through designing smart breeding.

REFERENCES

1. Bhullar, N. K.; Zhang, Z.; Wicker, T.; Keller, B. (2010). Wheat gene bank accessions as a source of new alleles of the powdery mildew resistance gene Pm3: A large-scale allele mining project. *BMC Plant Biology*, **10**(1):88.
2. Karikari, B.; Bhat, J. A.; Denwar, N. N. and Zhao, T. (2020). Exploring the genetic base of the soybean germplasm from Africa, America and Asia as well as mining of beneficial allele for flowering and seed weight. *Biotech*, **10**(5):195.
3. Marla, S. R.; Olatoye, M.; Davis, M.; Otchere, V.; Sexton-Bowser, S.; Morris, G. P. and Felderhoff, T. (2025). Mining sorghum pangenome enabled identification of new dw3 alleles for breeding stable-dwarfing hybrids. *G3: Genes, Genomes, Genetics*, **15**(5):54.
4. Pattnaik, A.; Barik, M.; Mishra, R. and Mohanty, J. N. (2025). Identification and characterization of genes that regulate flowering in pigeon pea (*Cajanus cajan*): An in-silico exploration. *Plant Sci. Today*, **12**(1): 1-11.
5. Singh, P.; Sundaram, K. T.; Vinukonda, V. P.; Venkateshwarlu, C.; Paul, P. J.; Pahi, B.; and Sinha, P. (2024). Superior haplotypes of key drought-responsive genes reveal opportunities for the development of climate-resilient rice varieties. *Communications Biology*, **7**(1):89.
6. Thyssen, G. N.; Smith, W.; Naoumkina, M.; Pinnika, G.; Jenkins, J. N.; McCarty, J. C. and Fang, D. D. (2025). Allele and transcriptome mining in *Gossypium hirsutum* reveal variation in candidate genes at genetic loci affecting cotton fiber quality and textile flammability. *BMC Plant Biology*, **25**(1):305.