

SNP Markers and its Impact on Crop Improvement

Debasmita Das, Rukmini Mishra*

Department of Botany, School of Applied Sciences, Centurion University of Technology and Management, Odisha, INDIA.

Submission Date: 15-10-2021; Revision Date: 04-11-2021; Accepted Date: 30-12-2021

ABSTRACT

With the advent of new technologies, methods and trends, the global progression in agricultural world is quite fast paced. However, it seems that Indian agriculture is yet to catch up to the new age technologies considering the practice of cultivation using the conventional methods can be still seen extensive. SNPs marker aided selection can prove to be an advantage for Indian agronomy. The cost effectiveness, fast and accurate mining for suitable allele specific marker and rapid detection makes this sequence-based marker the best option for a vast and versatile land like India, where its diverse topographic, geographic and climatic zones require equally distinct characteristics in crop genome to endure the environmental stress. The present review article summarizes the development of SNP markers and its applications in crop improvement.

Key words: Crop Improvement, Marker Assisted Selection (MAS), Molecular Breeding, Molecular Marker, Sequence Based Marker, SNPs markers.

Correspondence:

Dr. Rukmini Mishra,

Department of Botany,
School of Applied
Sciences, Centurion
University of Technology
and Management,
Odisha, INDIA.

Phone no: +91-7077320293

Email: rukmini.mishra@
cutm.ac.in

INTRODUCTION

Agriculture has been the back bone of the Indian economy and society since Indus valley civilization. Agronomy is a vital sector for India contributing 16.5% of total GDP,^[1] while accounts 42.6 % of the total workforce of the nation.^[2] But with the continuous population expansion the amount of arable land is gradually decreasing, which is starting to affect overall quantity and quality of the crop production. Simultaneously, Indian agronomic practices are not for only to fill its own plate but also to export agricultural products to other countries. In fiscal year 2020-21, India exported agricultural goods worth \$41.25 billion, making itself an Agri-exporter giant universal.^[3] With the gradual development of science and technology in areas like, genetics, biotechnology and agriculture, agricultural practices have become highly sophisticated. During the green revolution period, the rise of

genetically modified, high yielding crop variants can be seen. This new technology not only was able to secure the overall quality and quantity of the production but also made the variants immune to abiotic and biotic stress increasing the production yields.

While the higher yielding variants are gaining popularity at an increasing rate, it is also gradually replacing the former original parent varieties. This results in easy availability of higher yielding cultivars for research, seed production and cultivation, while the wild parent cultivars are on the verge of extinction. The green revolution led the focus of agricultural research towards developing higher yielding cultivars. Still, it is a common occurrence for these improve varieties to be susceptible to many diseases. While the older wild cultivars may be not as much high yielding as our expectation, but they can be a significant source of disease resistant genes.^[4] Thus, the need to conserve and protect the genotypes of these wild variants has emerged as of paramount importance.

In an instance like this, molecular markers can prove to be an indispensable advantage. Essentially, molecular markers are fragments of DNA which are used as an indicator for presence or absence of specific trait expressing allele inside a plant genome. In case of chromosomal crossover, during, sexual reproduction, sequences of parental DNA which are close to each

SCAN QR CODE TO VIEW ONLINE



www.ajbls.com

DOI: 10.5530/ajbls.2021.10.72

other are likely to be inherited together. This inclination of genes to be inherited together with the genes that are nearby is called genetic linkage. Genetic linkages are considered as an important exclusion from mendelian law of independent assortment, which states that genes responsible for different phenotypic characteristics are inherited independent of one another.^[5] However, it is only true for markers in different chromosome.

Molecular markers often signify the presence of point polymorphisms within a genome. These polymorphisms can be in the form of deletion, insertion, duplication, point mutations or translocation.^[6] Although they normally don't affect the expression of genes.^[7] Molecular markers, are linked to the genes carrying traits of interest, and used not only for detecting the presence or absence of the genes but also help in identification and recognition of its precise location within a pool of unfamiliar DNA.

Genetic framework exhibits not only the historical composition of the demographic adaptive potential but the extensive tenacity of the population against the geographic and environmental factors. With the help of molecular markers regular crops can be engineered to suit a particular climatic geographic region with more yielding capacity and other desirable traits like disease, insect resistance, larger reproductive organs, shorter flowering and maturing period etc., to make crop production more efficient and of superior quality. Proper management of limited arable land with the suitable application of today's science and technology and informative assessment of the cultivated crop ultimately proceed towards a higher margin of production. The lower number of adverse variables due to more efficient crop management makes agriculture emerges as a more economically beneficial trade, simultaneously straitening the backs of majority of the Indian populace who depends upon agronomy for their livelihood.

However, there is still a significance difference in number of studies conducted in case of animal and human genome analysis than plant genome including some major crop lines which comprises parallel analysis of a large-scale library of SNPs. These libraries aid with association genetic studies for a particular trait like, disease resistant, stress tolerance or maturation timing of flowers or fruits etc.

In this paper, we discuss about molecular markers, specifically various application of SNP marker assisted breeding and study some of the most significant cases where mining for SNPs and its' analysis have paved the path for developing improved variety of crops.

Molecular markers

Based on the method of detection, molecular markers can be categorized as either hybridization based, PCR based and sequence based markers.

Based on nucleic acid hybridization (non-PCR based)

In this method the targeted DNA of a specific genome is isolated, cloned and hybridized with the genomic DNA which can be identified within the genome. Restriction fragment length polymorphism (RFLP) was the first and quite popular method based on hybridization.^[8] The core principle of this method was based on the variation found in the length of DNA fragments produced by specific endonuclease. However, although the detection of RFLP molecular markers was quite popular, it proved to be low throughput, rather expensive, time consuming and labor intensive in nature.^[9] Besides these markers were not suitable for automation due to requiring radioactivity and a large amount of DNA for the analysis, which was making these markers ultimately antiquated.

Based on PCR amplification (PCR based approaches)

PCR based markers revolutionized the field of molecular biotechnology. They are considered as the second generation of genetic markers.^[10] The basic principle of this techniques is based on detection of polymorphism in nucleotide sequences by PCR amplification of the sample DNAs. They can be broadly distributed into two variants: Locus non-specific markers e.g., random amplified polymorphic DNA (RAPD); amplified fragment length polymorphism (AFLP), and locus specific markers e.g., simple sequence repeats (SSR); single nucleotide polymorphism (SNP). RAPD, AFLP, SSR are some of the new generations of PCR based markers which emerged in the beginning of the nineties and still remained as majorly used markers within the scientific community. While selecting an ideal genetic marker, codominance, with higher level of polymorphism detection, high reproductivity and expansive distribution across the genome are some of the crucial elements that should be kept in mind.^[11] While RAPD can conduct identification of polymorphic locus in several sample region of a specific genome, some significant variables like, the quantity and quality of DNA, PCR buffer, taq polymerase as well as annealing temperature affect the reproductivity of these markers.^[12] Anonymity and very low reproducibility proved to be some of the major limitation of RAPD.^[13] However, despite being anonymous markers,

AFLPs not only exhibit significantly high reproducibility but also high sensitivity. Although AFLPs are still used in case of unavailability of reference genome, they are not quite popular or used expansively due to the long and laborious identification process as well as inability to automate. Unsurprisingly, with the emergence of SSR markers they are pronounced as “markers of choice”.^[14] SSRs not only show higher reproductivity, higher polymorphous, and more flexible to automation, they are not unspecified. However, the cost of identification is quite expensive. SSRs markers remained quite popular throughout late 90s and the beginning of 21st century but their dominion over molecular research community came to a halt with the discovery of SNPs markers.

Sequence based markers

SNPs markers were first discovered in human genome. Due to their biallelic nature they are less polymorphic than SSRs markers while being more expensive than the later. But they are in copiousness amount within genome while being ultra-high-throughput and amenable to automation.^[15] Even though SNPs are generally more common in non-coding region of the genome, sometimes their presence in coding region develops either non-synonymous mutation resulting as change in amino acid sequence^[16] or synonymous mutation. Although synonymous mutation keeps the

amino acid sequence unchanged, it has the ability to alter phenotypic characters by changing mRNA splicing.^[17] Analysis of SNPs marker genotypes are broadly based on some basic procedures like extension of primers or invasive cleavage, ligation of oligonucleotide and allele-specific hybridization.^[18] Since their emergence, SNPs are used in various crops but the crops having simple genome are quite easy and upfront process, plants with complex genome are more problematic to map due to presence of highly repetitive sequences. But recent uprising of Next Generation Sequencing (NGS) has eliminated the issue by avoiding the highly repetitive area of the genome while mining for SNPs at a very low cost. This method proved to be quite popular and was applied successfully in genomes of various plants.

Due to their ultra-high throughput nature, SNPs are multipurposed. However, most importantly, they are used in rapid identification of variations in different crops and building ultra-high-density maps of genomes.^[19] A comprehensive comparison between major molecular markers is shown in Table 1.

Molecular Markers and Crop Improvement

Crop improvement, although an ancient concept generally refers to engineering or alteration of the plants for the greater good of humanity. It is deemed that agronomic practices developed as a part of cultural

Table 1: A comparison between major molecular markers.

Feature and description	RFLP	RAPD	AFLP	SSR	SNP
Expression/inheritance	Co-dominant	Dominant	Dominant / co-dominant	Co-dominant	Co-dominant
Primary application	Genetics	Diversity	Diversity and genetics	All purposes	All purposes
Genomic coverage	Low copy coding region	Whole genome	Whole genome	Whole genome	Whole genome
Genomic abundance	High	High	High	Moderate to high	Very high
Type of polymorphism	Variable length of restriction digest fragments	Random presence/absence of primer sites in DNA	Variable length of restriction digest fragments	Variable numbers of short polynucleotide repeats	Single-nucleotide polymorphism
Level of polymorphism	Moderate	High	High	High	High
Number of loci	Small (<1,000)	Small (<1,000)	Moderate (1,000s)	High (1,000s – 10,000s)	Very high (>100,000)
Reproducibility/reliability	High	Low	High	High	High
Genotyping throughput	Low	Low	High	High	High
Cost per analysis	High	Low	Moderate	Low	Low
Ease of automation	Low	Moderate	Moderate to high	High	High
Ease of use	Not easy	Easy	Moderate	Easy	Easy
Time demanding	High	Low	Moderate	Low	Low

evolution naturally, when humans started domesticating wild plants for food instead of hunting and scavenging for food. From the beginning of civilization, humans have been adopting various techniques and methods to improve the quality and quantity of their cultivation and cultivars, but until recently, a few decades ago, a significant advancement can be seen. According to Lawrence Bogorad, an American plant molecular biologist, pioneer of photosynthesis research, the improvement of crops includes two major techniques, selection and breeding.^[20]

Selection is the process of opting for cultivars having advantageous traits for cultivation. While in case of breeding, the breeder crosses two varieties of cultivars with advantageous traits to produce offspring of desired characteristics. Even with establishment of physical basis of inheritance during early 1900s, the process was still unpredictable.^[20] But with time, due to the progression of science and technology, discovery of genes, genomic expressions, cell division and most importantly Mendelian genetics and crossing methods with ratio for the phenotypic traits have cleared the path for formulating new improved varieties with a specific gene pool only containing desired characteristics.

Norman E. Borlaug, the father of green revolution once said, "Food is the moral right of all who are born into this world. Without food, man can live at most but a few weeks; without it, all other components of social justice are meaningless." With the rapid increase in population and more and more mouth to feed, food production is carried out expansively. It is estimated that roughly half of global habitable land mass is used for agricultural purposes.^[21] But at the same time, land mass is not increasing with respect to the populace, demanding production of more food from even less amount of arable land mass. Better, higher yielding crops with better nutrient quotient is the only answer for this problem. Marker assisted molecular breeding can be a crucial tool for designing and producing these kind of elite crop lines.

The traditional or regional cultivars having undergone natural selection process are unique to regional ecotype due to isolation from other populace of same species. They are heterogenous and adapted to local climate and its pathogen. Hence, they have accumulated substantial resistance alleles in their genome. However, they are susceptible to many undesired alleles i.e., low yielding, low fertility, susceptibility to lodging or low winter hardiness, susceptibility to draught or flooding etc. Similarly, in modern higher yielding cultivars pathogens resistance is comparatively less than their ancestry.^[22] Although they show significant tolerance to drought and

flooding. Traditional agronomy practices simply cannot serve the energy requirement of the world. Moreover, the involvement of various chemicals like, fungicide, herbicide, pesticides, and fertilizers not only prove to be major pollutants but they simultaneously destroy the agro ecosystem. The haphazard and extensive use of chemical ensued in an uprising of resistant pests' development and residual toxicity. Consequently, development of high yielding resistant cultivars with adequate desired characteristics has become a crucial requisite. In order to achieve this, a deep knowledge of genome and its expression, genetic diversity, genome mapping is needed.

Development of a new variety usually takes almost 25 years through convectional breeding. Although application of biotechnology can significantly reduce the duration to almost 7 to 8 years. Molecular breeding implicates genetic manipulation of DNA to engineer a specific plant or animal genome with only desired characteristics. Although there are several ways to manipulate genetic compositions like, genetic engineering, genomic selection, and molecular marker-assisted selection.^[23] Molecular marker-assisted selection is more extensively used. The practice of selection of plants by using linked markers is defined as marker assisted selection (MAS), where molecular markers are applied as a tool for selection of plants with desired trait by cultivators, researcher or breeders.^[24,25] The conventional method of selection involves physically growing the plants and waiting for its maturation in order to closely observes the recessive trait and surpasses it. But with marker assisted method not only the selection process has become more immediate and cost effective, it saves much energy and labor required to maintain a harvest.

Marker assisted selection helps in efficient screening procedure for selection of superior traits (Figure 1). Not only that it can also help in segregation of characters which are generally expressed at a later stage in plants like male sterility, photoperiod sensitivity, and grain quality. Similarly, characteristics like, tolerance to environmental stress (i.e., draught, flooding's, soil nutrient and mineral deficiency or toxicity, salt stress or disease tolerant etc.) which are far more expensive, extremely time consuming or difficult to examine can be easily screened if the identity and the location of relative markers are known within a single or multiple samples simultaneously. Moreover, identification of heterozygous can be done among homozygous without having to proceed through time consuming and laborious progeny testing.^[26]

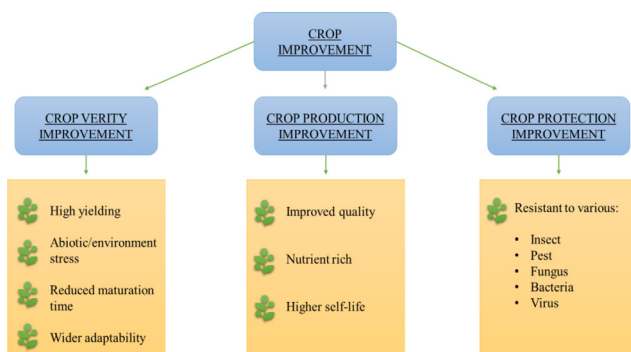


Figure 1: Possible applications of crop improvement.

Role of SNP markers in Crop Improvement

Genetic linkage analysis is a very crucial part of the studies to map genome. The principal of this linkage analysis is structured upon the concept of genetic recombination during meiosis.^[27] Genome mapping includes genotyping of individuals of isolated populaces. Similarly, development of genetic linkage maps with respect to the said genotypes while abundance presence of DNA markers in the genome is equally important. With development of NGS technologies and the availability of hefty amount of genetic sequence information, mapping for SNPs markers has increased significantly. Due to their abundance in genome and providing highest resolution for maps SNPs genotyping is primarily used for discovery of genes or QTLs.

SNP genotyping is a valuable tool for gene mapping, map-based cloning, and marker assisted selection (MAS) in crops.^[28] In a dated study on rice plants, it is shown that, during quantitative trait locus analysis (QTL) for yield and it's three other component traits, i.e., grain weight, number of tillers per plants and number of grains per panicle, ultra-high throughput SNPs mapping demonstrated a more powerful resolution map comparative to RFLP/SSR based QTL analysis.^[29] Maize plant has a complex genome with highly repetitive sequences. Specifically, the trait responsible for flowering time of the plant also controls the adaptation of the plant to its surrounding environment. After a dissection for time of flowering across nearly one million plants in eight different environments was carried out, it is found by the authors that numerous small sections of additive QTLs control the time of flowering across different environment instead of a single large-effect QTLs, which are relative among the family.^[30]

Similarly, another study revealed that the hypothesis of multiple loci of small effects controlling the expression of disease resistance in maize plant is in accordance to the results of the study. The researchers mapped approx. 5000 inbred genomes for quantitative resistance against

northern late blight disease of maize plant. 1.6 million SNPs were used to identify 29 loci of quantitative traits with multiple alleles.^[31] Wheat is one of the major staple foods of the world. Lately SNPs mapping assay was performed on several tetraploids and hexaploids wheat lines to identify nearly 96 SNPs loci.^[32] Several studies have been conducted the genomic mapping for SNPs marker discovery in wheat plants.^[33-35]

In case of soybean plant, mapping for SNPs marker was carried out by different studies. The main aim of these studies was to construct a higher resolution map for the targeted locus to enhance the disease resistivity and cost efficiency of the production. The detection of precise location of the soybean aphid resistant genes *Rag1* and *Rag2* were carried out through fine genome mapping.^[36,37] Similarly, another study was carried out for SNPs marker tight linked QTLs identification against the southern root-knot nematode.^[38] SNPs markers were developed in canola plant after identification of single nucleotide mutation for *fad2* and *fad3* genes in the plant genome. However, marker assisted selection of those genes for characteristic introgression and breeding was facilitated only after the development of allele specific PCR assays.^[39]

In association with *Rnh.7H*, a resistant gene against the covered smut disease in barley plant, SNPs markers were developed. The researchers have conducted the identification through high resolution melting technique.^[40] In case of sugar beet, SNPs assisted mapping for QTL regarding beet necrotic yellow vein virus was conducted. Anchor linkage marker for *Rz4*^[41] and *Rz5*^[42] were developed. Molecular marker discovery is fast paced field of agronomy where emergence of new analytical technologies is frequent and becoming more and more sophisticated and efficient in nature. While above discussed cases were dated from nearly a decade ago, more current studies are being conducted with the availability of prior information and novel cutting-edge technologies.

Recently, a study for genome-wide SNPs mapping in Darjeeling tea plants had reported 54,206 high-quality SNP markers across 15 chromosomes. In an Illumina HiSeq X10 platform the application of double-digest restriction-site associated DNA based paired-end sequencing revealed about genomic association among some specified genes of high impact varieties of Darjeeling tea and their efficient agronomic and biochemical variables.^[43] In another study, a selection of 368 inbred cultivars of maize plant was undergone through an extensive quantitative trait loci (QTL) analysis in order to detect kompetitive allele specific PCR (KASP) SNP markers by high-throughput RNA sequencing.

On the basis of a few conventional requirements like Polymorphism information content (PIC) value ≥ 0.4 , bi-allelic and conserved primer sequences, unique genomic region etc. 71,311 KASP SNPs were located out of total 2,948,985 SNPs markers. These markers were functionally annotated to over 52 genes which also include genes responsible for most of the primary and secondary metabolic pathway of the plant.^[44]

In case of sweet potato plants (*Ipomoea batatas*), the transgenic variant, exhibit a distinct light orange/ yellow flesh colour of the storage roots due to the orange gene (IbOr-R96H). The original wild parent variant of the transgenic plant also shows the presence of a distinct wild-type IbOr gene (IbOr-WT).^[45] However, a recent study demonstrates that the said gene carries a unique single nucleotide polymorphism located in the 96th amino acid positions whose overexpression also results in carotenoid accumulation in the storage roots as well as a well-developed heat stress tolerance.^[46] Rice is a well-established staple food crop. However, as it is discussed above SNPs analysis poses as a significant tool to be used in the marker assisted selection (MAS) in case of rice cultivation. A novel SNPs marker was recently discovered which is linked to the gene responsible for resistance against the narrow brown leaf spot (NBLs) disease caused by a fungal pathogen named *Cercospora janseana*. This study revealed that upon the resistant analysis under natural conditions for 3 years the recombinant inbred line of the rice cultivar showed 81.4% of phenotypic variation caused by a novel solitary large-effect QTL named CRSP-2.1. Additionally, 13 more SNPs markers were identified which are responsible to haplotype diversity in the current rice cultivars of U.S. rice germplasm.^[47]

In rice cultivation of temperate and tropical regions Bacterial blight (BB), is another major disease, which is caused by *Xanthomonas oryzae* (*Xoo*). But due to specific interaction between the resistance gene and the corresponding virulence gene of the pathogen, new strains of pathogen are rapidly evolving with the capacity to overcome the installed resistant. An extensive genome-wide association study (GWAS) reveals the presence of 11 novel QTLs in 4th, 5th, 11th, 12th chromosomes which are related to BB resistance.^[48] Similarly, Rice blast, caused by *Magnaporthe oryzae*, is another major disease known to wreak havoc on the paddy fields. During a topical analysis 23 genomic variants of rice were screened for resistant against the disease. It was revealed that the SNPs linked to *Pi54* and *Pita* genes not only were responsible for resistance against the rice blast disease, but also responsible for 11 other phenotypic variation and two haplotypings.^[49]

SNPs genotyping is proved to be useful for not only developing disease resistant cultivars in rice crop but also aids in intentional breeding for improved plant architecture. An ideal tiller angle is a crucial factor for attaining high yield which is controlled by the TAC₁ (*Tiller Angle Controlling*) gene. A single nucleotide polymorphism on the 4th intron 3' splicing point in the said gene is found to be linked to the variation found in the architecture of the tiller angle in the plant. A fluorescent functional molecular marker, PM-TAC1 was developed via the penta-primer amplification refractory mutation system Based on this SNPs variation analysis.^[50]

In another investigation to find the best suited wheat SNP arrays, various current SNP arrays were evaluated on the basis of identity, alignment length, mismatches, bit score and e-value ($< 1e-20$). Among the 9K, 15K, 35K, 55K, 90K, 660K, 820K arrays, 660K SNP arrays found to be the finest pick considering reliable application, distribution, density, associated genes, heterozygosity and most importantly number, application of SNPs marker present in genome and cost effectiveness for marker assisted selection (MAS).^[51] Equally, with the progression of technology, nowadays SNPs variation can also be used for high throughput DNA barcoding to identify commercial tomato cultivars with a remarkable 80.0–93.6 % identification rate. In a recent analysis it was found that among the commercial 48 F₁ variants of tomatoes, five major SNPs markers subsets specified as 192, 96, 48, 24, 12 markers are highly effective for DNA barcoding system for specification of commercial tomato cultivars.^[52]

Application of SNPs markers in QTL mapping a genotypic study is very vast. While very few of these studies are mentioned previously, much superior work is already done in this field. Still even more contemporary approaches to analyze and identify SNPs and implication of the knowledge in dissection of beneficial traits have been published recently while this review was still being drafted.

Defense mechanisms against various diseases are a very popular choice to explore. In a new genome-wide association analysis to identify the specific region of the chromosome, responsible for resistance to leaf rust and stripe rust in seedlings and adult plant stages, KASP markers were identified and validated using 90K SNP array. After analyzing 268 diverse wheat lines 22 known resistance and 18 potentially novel loci were identified which are claimed to have the ability to explain about 4.6 to 25.2% of the phenotypic variation.^[12] In another study in the same foliar stripe rust disease of wheat, nine resistance quantitative trait loci (QTL) were mapped for resistance against the disease using the same

90K single nucleotide polymorphisms array on 137 lines of recombinant inbred lines (RILs).^[53] Similarly, in another study on resistance against common scab disease of potato caused by *Streptomyces* sp. an genome-wide association study (GWAS) was conducted among 165 genotypes of potato to establish a novel resistance QTL on 640 kb interval between 0.43 Mb and 1.07 Mb in chromosome 1 in potato.^[54] The Mesoamerican evergreen crop *Theobroma cacao* L. may be one of the major economically beneficial crops. However, frosty pod and black pod rots are major preparators of the severe losses which have significantly decreased the production of cacao beans. In a recent research the is conformed that seven QTLs are responsible for resistance against the frosty pod rust (FPR), and black pod rust (BPR) disease. They are reported on 2, 7, 8 chromosomes for FPR and on chromosome 4, 8, 9 for BPR. Furthermore, another 8 QTLs are acknowledged to be resistant to FPR (in chromosome 4,9,10) and BPR (chromosome 2).^[55] Limiting agricultural land mass with respect to the growing population and developments forces us to choose a more sustainable approach in agriculture. Genetically disease resistant crops pave the way for more sustainable application of SNPs marker assisted breeding where the focus remain on making the plants sustain themselves without external chemical agents against diseases and pathogens. But still there are many crop diseases like Common Scab of Potato that although have many sources of resistant but still don't have much highly resistant germplasm identified till now. More research projects should be designed towards the in-depth detection and analysis. Just like genetically resistant against diseases or biotic stress, genetic resistance against environmental stress is another major aspect of molecular breeding. It defines behavioral or mechanical responses to an environmental stimulus like salinity, drought, flood, extremes in temperature, heavy metals, radiation etc.

In case of oil palm (*Elaeis guineensis*) a topical research explained that the production of fresh fruit bunch (FFB) is negatively affected by drought stress. The double digest restriction amplified DNA (ddRAD) genotyping was conducted to identify the SNP marker associated with drought tolerance gene. A total of 538k SNPs were identified, from which 58 SNPs associations cite across 21 genes were considered optimal.^[56] Similarly, a KASPar SNP genetic mapping of cassava suggests the presence of 27 QTL associations against drought tolerance.^[57]

Likewise, salinity stress is another factor which can be mitigated through SNPs marker mining and mapping. In Barley, salt tolerance can be a deciding factor for

seed germination and growth of seedlings. A genome-wide association scan (GWAS) was carried out using 9K SNPs to reveal 80 associated SNPs with significant role in *Squamosa promoter-binding-like protein 6* at chromosome 5H of Barley. It essentially promotes seedling growth even with saline condition.^[58]

With the advancement of the next generation sequencing, it is estimated that more and more attention will be focused on the acquiring knowledge about the genotypes of known plants. Genetic diversity explains the importance of variation within the population and relationships in-between the variants. The knowledge gained from genotyping is crucial for crossing, hybrid experiments, assigning individuals into heterotic groups, inbred line identification and most importantly germplasm conservation for preserving valuable traits of the wild and ancient varieties. Moreover, introgression of novel alleles into already establishes exclusive verities require thorough identification of the said genome in order to eliminate any undesirable characteristics.

A recent genome wide diversity study in rice led the authors to discover a miRNA-regulated gene named CYP704A3 which is responsible for seed length.^[59] Similarly, in maize, the diversity studies using SNPs across 21 loci of chromosome 1 helped to understand the relation pattern between different species of maize.^[60] Formerly, molecular markers were used to analyze the genetic variation in-between populations through AFLPs, SSRs, or isozymes methods.^[61] However, the analysis could be conducted on a specific regions or genes of the genome, which was very laborious and time-consuming process. With the emergence of SNPs along with NGS technology, now, a whole genome wide analysis can be carried out. Simultaneous comparison of genotypic of both ancient and recent cultivars can unravel the molecular mechanism of evolution between genomes. A few years back, the researchers were able to identify SNPs marker in the region of regulation on the 5' end of the vital *qSH1* gene which is related to shattering in some subspecies of rice. This indicated that seed shattering was intentionally selected from wild varieties, thus proving to be a major turning event in the process of domestication of rice plants.^[62] Likewise, SNPs were used to understand the molecular basis of the genes like WAG-2 among wheat and its relatives.^[63] As it is known that chloroplast and nuclear genes are prolific resources of information regarding phylogenetics. SNPs mapping can be applied to observe the distinction and extent of resemblances between diverse sequences to establish phylogenetical and evolutionary association among an extensive collection of varieties of plants.

CONCLUSION

With the advancement of NGS technology, molecular marker assisted crop improvement is swiftly gaining popularity. Not only has this technique made discovery of SNPs marker more cost effective, but far more accurate and easier to access. As discussed above, there are ample numbers of SNPs available across various genomes for genomic reference, even in case of complex genome. Still there are many plant genomes yet to be genotyped in order to fill the knowledge gap regarding non coding regions and repetitive regions. During past decades, there has been a tremendous growth in molecular genetics considering development of the first and second-generation sequencing methods. Not only sequencing for an entire genome has become so affordable but provides higher quality analysis too. Although improvement in crops through molecular marker is not a new concept, it has yet to be used extensively in Indian agricultural arena. DNA based molecular markers gives us a better understanding of the molecular mechanics of genomics, expression of quantitative characters, evolutionary co-relations of crop plants, which plays a crucial role in understanding the demand of the populaces and crops engineered to meet them. Among all the markers as established “marker of choice” SNPs markers demonstrates a promising future with respect to the vast possibility of its application. Moreover, invention and implementation of new techniques regarding agricultural biotech not only aids to the overall wellbeing of both economy and vitality of the producer and the consumers as well as the nation but it also paves ways towards the emergence of more sophisticated vital technology for the benefit of the humankind.

AUTHOR CONTRIBUTIONS

RM and DD conceived the idea and DD wrote the manuscript. RM critically reviewed the manuscript. Both the authors have approved it.

ACKNOWLEDGMENTS

Authors are thankful to the President, Centurion University of Technology and Management, Odisha for his constant help and support. We apologize to the authors for the omission of their work that could not be included in this paper due to space constraints.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Online FE. India economic survey 2018: farmers gain as agriculture mechanisation speeds up, but more R&D needed; 2018 Jan 29. Financial Express [internet] [cited Aug 5 2021]. Available from: <https://www.financialexpress.com/budget/india-economic-survey-2018-for-farmers-agriculture-gdp-msp/1034266/>.
2. India - Distribution of the workforce across economic sectors 2019 [internet]. Statista.com [cited Aug 5 2021]. Available from: <https://www.statista.com/statistics/271320/distribution-of-the-workforce-across-economic-sectors-in-india/>.
3. Nandi S. India's agriculture exports jump 17.34% to \$41.25 billion in FY21 [internet]; 2021. Available from: Business-standard.com [cited Aug 5 2021]. Available from: https://www.business-standard.com/article/economy-policy/india-s-agriculture-exports-jump-17-34-to-41-25-billion-in-fy21-121061001623_1.html.
4. Villa TCC, Maxted N, Scholten M, Ford-Lloyd B. Defining and identifying crop landraces. *Plant Genet Resour.* 2005;3(3):373-84. doi: 10.1079/PGR200591.
5. Campbell NA Markl J, editor. *Biologie*. 6th ed. Heidelberg, Germany: Spektrum Akademischer Verlag; 2003.
6. Nadeem MA, Nawaz MA, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, et al. DNA molecular markers in plant breeding: Current status and recent advancements in genomic selection and genome editing. *Biotechnol Biotechnol Equip.* 2018;32(2):261-85. doi: 10.1080/13102818.2017.1400401.
7. Mondini L, Noorani A, Pagnotta M. Assessing plant genetic diversity by molecular tools. *Diversity.* 2009;1(1):19-35. doi: 10.3390/d1010019.
8. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet.* 1980;32(3):314-31. PMID 6247908.
9. Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica.* 2005;142(1-2):169-96. doi: 10.1007/s10681-005-1681-5.
10. Kole C, Abbott AG, editors. *Principles and practices of plant genomics*, vol. 1. Genome mapping. Enfield, CT: Science; 2008.
11. Brettschneider R. RFLP analysis. In: *Molecular tools for screening biodiversity*. Dordrecht: Springer Netherlands; 1998. p. 85-95.
12. Wolff K, Schoen ED, Rijn JP. Optimizing the generation of random amplified polymorphic DNAs in chrysanthemum. *Theor Appl Genet.* 1993;86(8):1033-7. doi: 10.1007/BF00211058, PMID 24194014.
13. Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, Van de Wiel C, et al. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol Breed.* 1997;3(5):381-90. doi: 10.1023/A:1009612517139.
14. Powell W, Machray GC, Provan J. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* 1996;1(7):215-22. doi: 10.1016/S1360-1385(96)86898-0.
15. Rafalski JA. Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant Sci.* 2002;162(3):329-33. doi: 10.1016/S0168-9452(01)00587-8.
16. Sunyaev S, Hanke J, Aydin A, Wirkner U, Zastrow I, Reich J, et al. Prediction of non-synonymous single nucleotide polymorphisms in human disease-associated genes. *J Mol Med (Berl).* 1999;77(11):754-60. doi: 10.1007/s001099900059, PMID 10619435.
17. Richard I, Beckmann JS. How neutral are synonymous codon mutations? *Nat Genet.* 1995;10(3):259. doi: 10.1038/ng0795-259, PMID 7670461.
18. Sobrino B, Brión M, Carracedo A. SNPs in forensic genetics: A review on SNP typing methodologies. *Forensic Sci Int.* 2005;154(2-3):181-94. doi: 10.1016/j.forsciint.2004.10.020, PMID 16182964.
19. Govindaraj M, Vetriventhan M, Srinivasan M. Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. *Genet Res Int.* 2015;2015:431487. doi: 10.1155/2015/431487.
20. National Research Council. (US) Board on Agriculture. *Crop Improv.* 1984.
21. EC, Klein Goldewijk K, Siebert S, Lightman D, Ramankutty N. Anthropogenic transformation of the biomes, 1700-2000: Anthropogenic transformation of the biomes. *Glob Ecol Biogeogr.* 2010;no-no. doi: 10.1111/j.1466-8238.2010.00540.x.

22. Samal KC, Rout GR. Genetic improvement of vegetables using transgenic technology. In: Genetic engineering of horticultural crops. Elsevier; 2018;193-224.
23. Ribaut JM, De Vicente MC, Delannay X. Molecular breeding in developing countries: Challenges and perspectives. *Curr Opin Plant Biol.* 2010;13(2):213-8. doi: 10.1016/j.pbi.2009.12.011, PMID 20106715.
24. Beckmann JS, Soller M. Restriction fragment length polymorphisms in genetic improvement: Methodologies, mapping and costs. *Theor Appl Genet.* 1983;67(1):35-43. doi: 10.1007/BF00303919, PMID 24258478.
25. Darvasi A, Soller M. Optimum spacing of genetic markers for determining linkage between marker loci and quantitative trait loci. *Theor Appl Genet.* 1994;89(2-3):351-7. doi: 10.1007/BF00225166, PMID 24177853.
26. Varshney A, Mohapatra T, Sharma RP. Molecular mapping and marker assisted selection of traits for crop improvement. In: Plant biotechnology and molecular markers. Dordrecht: Kluwer Academic Publishers; 2006. p. 289-330.
27. Tanksley SD. Mapping polygenes. *Annu Rev Genet.* 1993;27(1):205-33. doi: 10.1146/annurev.ge.27.120193.001225, PMID 8122902.
28. Mammadov J, Aggarwal R, Buyyarapu R, Kumpatla S. SNP markers and their impact on plant breeding. *Int J Plant Genomics.* 2012;2012:728398. doi: 10.1155/2012/728398.
29. Yu H, Xie W, Wang J, Xing Y, Xu C, Li X, *et al.* Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. *PLOS ONE.* 2011;6(3):e17595. doi: 10.1371/journal.pone.0017595, PMID 21390234.
30. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, *et al.* The genetic architecture of maize flowering time. *Science.* 2009;325(5941):714-8. doi: 10.1126/science.1174276, PMID 19661422.
31. Poland JA, Bradbury PJ, Buckler ES, Nelson RJ. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc Natl Acad Sci U S A.* 2011;108(17):6893-8. doi: 10.1073/pnas.1010894108, PMID 21482771.
32. Akhunov E, Nicolet C, Dvorak J. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina Golden Gate assay. *Theor Appl Genet.* 2009;119(3):507-17. doi: 10.1007/s00122-009-1059-5, PMID 19449174.
33. Allen AM, Barker GLA, Berry ST, Coghill JA, Gwilliam R, Kirby S, *et al.* Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant Biotechnol J.* 2011;9(9):1086-99. doi: 10.1111/j.1467-7652.2011.00628.x, PMID 21627760.
34. Brito LF, Kijas JW, Ventura RV, Sargolzaei M, Porto-Neto LR, Cánovas A, *et al.* Genetic diversity and signatures of selection in various goat breeds revealed by genome-wide SNP markers. *BMC Genomics.* 2017;18(1):229. doi: 10.1186/s12864-017-3610-0, PMID 28288562.
35. Winfield MO, Wilkinson PA, Allen AM, Barker GLA, Coghill JA, BurrIDGE A, *et al.* Targeted re-sequencing of the allohexaploid wheat exome. *Plant Biotechnol J.* 2012;10(6):733-42. doi: 10.1111/j.1467-7652.2012.00713.x, PMID 22703335.
36. Kim KS, Hill CB, Hartman GL, Hyten DL, Hudson ME, Diers BW. Fine mapping of the soybean aphid-resistance gene Rag2 in soybean PI 200538. *Theor Appl Genet.* 2010;121(3):599-610. doi: 10.1007/s00122-010-1333-6, PMID 20454773.
37. Kim KS, Bellendir S, Hudson KA, Hill CB, Hartman GL, Hyten DL, *et al.* Fine mapping the soybean aphid resistance gene Rag1 in soybean. *Theor Appl Genet.* 2010;120(5):1063-71. doi: 10.1007/s00122-009-1234-8.
38. Ha BK, Hussey RS, Boerma HR. Development of SNP assays for marker-assisted selection of two southern root-knot nematode resistance QTL in soybean. *Crop Sci.* 2007;47(S2);Suppl 2:S-73-S-82. doi: 10.2135/cropsci2006.10.0660tpg.
39. Hu X, Sullivan-Gilbert M, Gupta M, Thompson SA. Mapping of the loci controlling oleic and linolenic acid contents and development of fad2 and fad3 allele-specific markers in canola (*Brassica napus* L.). *Theor Appl Genet.* 2006;113(3):497-507. doi: 10.1007/s00122-006-0315-1, PMID 16767448.
40. Lehmsiek A, Sutherland MW, McNamara RB. The use of high resolution melting (HRM) to map single nucleotide polymorphism markers linked to a covered smut resistance gene in barley. *Theor Appl Genet.* 2008;117(5):721-8. doi: 10.1007/s00122-008-0813-4, PMID 18553067.
41. Grimmer MK, Trybush S, Hanley S, Francis SA, Karp A, Asher MJC. An anchored linkage map for sugar beet based on AFLP, SNP and RAPD markers and QTL mapping of a new source of resistance to Beet necrotic yellow vein virus. *Theor Appl Genet.* 2007;114(7):1151-60. doi: 10.1007/s00122-007-0507-3, PMID 17333102.
42. Grimmer MK, Kraft T, Francis SA, Asher MJC. QTL mapping of BNYVV resistance from the WB258 source in sugar beet. *Plant Breed.* 2008;127(6):650-2. doi: 10.1111/j.1439-0523.2008.01539.x.
43. Hazra A, Kumar R, Sengupta C, Das S. Genome-wide SNP discovery from Darjeeling tea cultivars - their functional impacts and application toward population structure and trait associations. *Genomics.* 2021;113(1 Pt 1):66-78. doi: 10.1016/j.ygeno.2020.11.028, PMID 33276009.
44. Chen Z, Tang D, Ni J, Li P, Wang L, Zhou J, *et al.* Development of genic KASP SNP markers from RNA-Seq data for map-based cloning and marker-assisted selection in maize. *BMC Plant Biol.* 2021;21(1):157. doi: 10.1186/s12870-021-02932-8, PMID 33771110.
45. Kim SH, Ahn YO, Ahn MJ, Jeong JC, Lee HS, Kwak SS. Cloning and characterization of an Orange gene that increases carotenoid accumulation and salt stress tolerance in transgenic sweetpotato cultures. *Plant Physiol Biochem.* 2013;70:445-54. doi: 10.1016/j.plaphy.2013.06.011, PMID 23835362.
46. Kim SE, Lee CJ, Park SU, Lim YH, Park WS, Kim HJ, *et al.* Overexpression of the golden SNP-carrying orange gene enhances carotenoid accumulation and heat stress tolerance in Sweetpotato plants. *Antioxidants (Basel).* 2021;10(1):51. doi: 10.3390/antiox10010051, PMID 33406723.
47. Addison CK, Angira B, Cerioli T, Groth DE, Richards JK, Linscombe SD, *et al.* Identification and mapping of a novel resistance gene to the rice pathogen, *Cercospora janseana*. *Theor Appl Genet.* 2021;134(7):2221-34. doi: 10.1007/s00122-021-03821-2, PMID 33825949.
48. Lu J, Wang C, Zeng D, Li J, Shi X, Shi Y, *et al.* Genome-wide association study dissects resistance loci against bacterial blight in a diverse rice panel from the 3000. *Rice, NY: Rice Genomes Project;* 2021. p. 14(1):22.
49. Mishra A, Wickneswari R, Bhuiyan MAR, Jena KK, Abd Aziz Shamsudin N. N. Broad spectrum blast resistance alleles in newly developed Malaysian rice (*Oryza sativa* L.) genotypes. *Euphytica.* 2021;217(1). doi: 10.1007/s10681-020-02738-z.
50. Gao J, Liang H, Huang J, Qing D, Wu H, Zhou W, *et al.* Development of the PARMS marker of the TAC1 gene and its utilization in rice plant architecture breeding. *Euphytica.* 2021;217(3). doi: 10.1007/s10681-020-02747-y.
51. Sun C, Dong Z, Zhao L, Ren Y, Zhang N, Chen F. The Wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat. *Plant Biotechnol J.* 2020;18(6):1354-60. doi: 10.1111/pbi.13361, PMID 32065714.
52. Kim M, Jung JK, Shim EJ, Chung SM, Park Y, Lee GP, *et al.* Genome-wide SNP discovery and core marker sets for DNA barcoding and variety identification in commercial tomato cultivars. *Scientia Horticulturae.* 2021;276. doi: 10.1016/j.scienta.2020.109734, PMID 109734.
53. Zhou X, Zhong X, Roter J, Li X, Yao Q, Yan J, *et al.* Genome-wide mapping of loci for adult-plant resistance to stripe rust in durum wheat Svevo using the 90K SNP array. *Plant Dis.* 2021;105(4):879-88. doi: 10.1094/PDIS-09-20-1933-RE, PMID 33141640.
54. Koizumi E, Igarashi T, Tsuyama M, Ogawa K, Asano K, Kobayashi A, *et al.* Association of genome-wide SNP markers with resistance to common scab of potato. *Am J Potato Res.* 2021;98(2):149-56. doi: 10.1007/s12230-021-09827-2.
55. Gutiérrez OA, Puig AS, Phillips-Mora W, Bailey BA, Ali SS, Mockaitis K, *et al.* SNP markers associated with resistance to frosty pod and black pod rot diseases in an F1 population of *Theobroma cacao* L. *Tree Genet Genomes.* 2021;17(3). doi: 10.1007/s11295-021-01507-w.
56. Yono D, Nugroho YA, Tanjung ZA, Utomo C, Liwang T. Genome wide SNP marker identification associated with drought tolerance in oil palm. *Bio Diversitas.* 2021;22(6). doi: 10.13057/biodiv/d220616.
57. Ewa F, Asiwe JNA, Okogbenin E, Ogbonna AC, Egesi C. KASPar SNP genetic map of cassava for QTL discovery of productivity traits in moderate drought stress environment in Africa. *Sci Rep.* 2021;11(1):11268. doi: 10.1038/s41598-021-90131-8. PMID 34050196.
58. Thabet SG, Moursi YS, Sallam A, Karam MA, Alqudah AM. Genetic associations uncover candidate SNP markers and genes associated with salt tolerance during seedling developmental phase in barley. *Environ Exp Bot.* 2021;188. doi: 10.1016/j.envexpbot.2021.104499, PMID 104499.

59. Tang W, Wu T, Ye J, Sun J, Jiang Y, Yu J, *et al.* Erratum to: SNP-based analysis of genetic diversity reveals important alleles associated with seed size in rice. *BMC Plant Biol.* 2016;16(1):128. doi: 10.1186/s12870-016-0801-9, PMID 27267593.
60. Tenailon MI, Sawkins MC, Anderson LK, Stack SM, Doebley J, Gaut BS. Patterns of diversity and recombination along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Genetics.* 2002;162(3):1401-13. doi: 10.1093/genetics/162.3.1401, PMID 12454083.
61. Vigouroux Y, Mitchell S, Matsuoka Y, Hamblin M, Kresovich S, Smith JSC, *et al.* An analysis of genetic diversity across the maize genome using microsatellites. *Genetics.* 2005;169(3):1617-30. doi: 10.1534/genetics.104.032086, PMID 15654118.
62. Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, *et al.* An SNP caused loss of seed shattering during rice domestication. *Science.* 2006;312(5778):1392-6. doi: 10.1126/science.1126410, PMID 16614172.
63. Wei S, Peng Z, Zhou Y, Yang Z, Wu K, Ouyang Z. Nucleotide diversity and molecular evolution of the WAG-2 gene in common wheat (*Triticum aestivum* L) and its relatives. *Genet Mol Biol.* 2011;34(4):606-15. doi: 10.1590/S1415-47572011000400013, PMID 22215965.

Cite this article: Das D, Mishra R. SNP Markers and its Impact on Crop Improvement. *Asian J Biol Life Sci.* 2021;10(3):539-48.