

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,300

Open access books available

130,000

International authors and editors

155M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



Single Nucleotide Polymorphisms: A Modern Tool to Screen Plants for Desirable Traits

Lovina I. Udoh, Willie Peggy Obaseoji and Chiebuka Uzoabo

Abstract

Single nucleotide polymorphism (SNP) represent a change in a single nucleotide within the genome. This can alter the phenotype of an individual within the same species if it occurs in a coding region of the gene. The change in nucleotide can produce desirable characteristic in plants and can become an object for selection. New SNPs have been discovered and subsequently converted to molecular markers using various non-gel based and next generation sequencing platforms. Considering that SNP markers are based on target genes, its abundance in the genome, high automation and multiplexability, has made it a marker of choice and an effective tool for screening plant germplasm for desirable traits. This chapter considers SNP as molecular marker, their discovery and different SNP genotyping methods was documented. A few case studies of SNP as allele specific markers and their association with traits of interest was considered. Thus, highlighting their efficacy as useful tool for marker assisted selection and plant germplasms screening.

Keywords: single nucleotide polymorphisms, molecular breeding, plant germplasm screening, molecular marker, marker assisted selection

1. Introduction

Plant breeders usually screen a large number of plants for traits of economic value as determined by the breeding goal which may include breeding for resistance, biofortification to increase some micronutrients and gene pyramiding. The larger volume of plants screened at early stages of a breeding program can be laborious, capital intensive and time consuming. Germplasm screening is usually an initial step for a number of breeding programs. The aim of screening large collection of plants is for narrowing on those with desired characteristic for advancement to the next stage. It is extremely important to get it right from the beginning in order to meet breeding objectives for crop improvement.

Recently, molecular breeding has brought about a revolution in plant breeding and has been widely applied in plant improvement programs. Molecular breeding also called marker assisted breeding or marker assisted selection (MAS) is the method of using molecular or DNA markers that are closely linked to a phenotype to aid selection for such trait in a breeding scheme. A number of molecular markers have been developed and successfully used in selection due to their association with a phenotype of interest in different plant species. Thus, they have been applied in plant germplasm screening for desirable traits. Because selection is based on target

genes, they provide a higher rate of accuracy during screening, reduce time and labor, therefore bringing about reduced cost.

Several types of molecular markers have been applied to different areas of plant breeding, firstly developed markers include Restriction fragment length polymorphism (RFLPs) and Randomly amplified polymorphic DNA (RAPDs). These gave way to the popular techniques Amplified fragment polymorphism (AFLP) and Simple sequence repeats (SSRs) due to their ease in detection and automation. Simple sequence repeats have been extensively applied to screen for resistant germplasm, biotic and abiotic stress and variety identification in potato, groundnut rice [1–3]. SSR markers are generally PCR-based, their technicality is simple and relatively cheap. The disadvantage of these markers is that they require polyacrylamide gel electrophoresis to achieve a high resolution of allele fragments, which is laborious to perform. They give information about a single locus per assay although multiplexing of more than two markers is possible, it is relatively expensive. The continued automation in sequencing technology has resulted in a shift from first generation DNA based markers to the use of functional and gene targeted single nucleotide polymorphism (SNP) markers. Therefore, this chapter covers ways of discovering new SNPs and their conversion to molecular marker. Also, different methods of SNPs genotyping and their application in some aspects of crop improvement are documented. The discovery of SNP markers is cost effective, it is highly multiplexable and they are availability of high throughput technologies for SNP genotyping. Because SNPs are based on target genes, they are highly reliable in cultivar identification and germplasm screening.

2. Single nucleotide polymorphism markers

Single nucleotide polymorphism is a change in a single base pair at a specific locus involving two alleles where the rare allele frequency is >1% (**Figure 1**). It is an individual nucleotide base difference between two DNA sequences. It represents the site where the DNA sequence shows a difference by one base. They are categorized based on nucleotide substitution as transition which is an interchange of pyrimidines (C/T) and purines (A/G) or transversion which is an interchange of purine base for pyrimidine (G/C, A/T, A/C and G/T). A nucleotide base represents the basic unit of inheritance therefore SNPs present a powerful tool as molecular

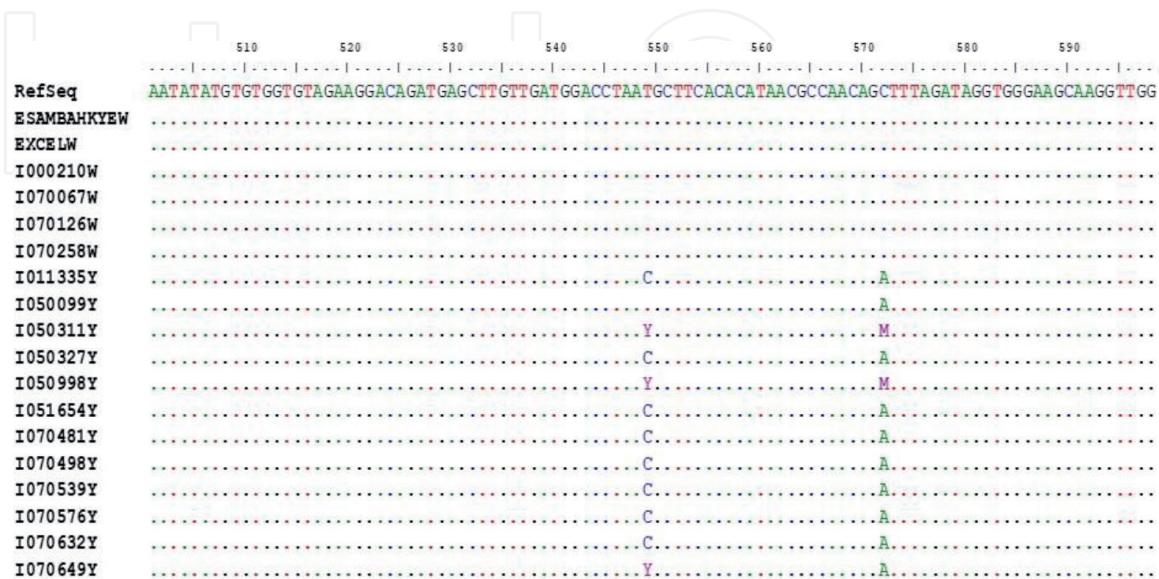


Figure 1. Multiple sequence alignment of nucleotide sequences to a reference sequence (RefSeq) reveal SNPs of PSY2 gene (Y)C/T and (M)A/C in yellow and white root accessions of cassava. Consensus to reference sequence are plotted as dots.

marker. SNPs which can be categorized as insertion or deletions, may be found in the coding or non-coding sequences of crop plants. Some individuals of a species may be heterozygous at a SNP locus or ambiguous as seen in **Figure 1** (Y refers to C and T nucleotides being present at one locus), this means they possess both nucleotides on the same position on that gene. Such individuals will display an intermediate phenotype when screened phenotypically.

When SNPs or insertion/deletions found in coding sequences result in non-conservative amino acid changes, it can cause variation in the phenotype of individuals of same species (**Figure 2**). But the translation of non-coding sequences containing SNPs that results in conservative amino acid change will not significantly affect the phenotype of the individuals. Considering the figures above, **Figure 1** had SNP present at two different positions (549 and 572) for C/T and A/C nucleotides of the PSY2 gene in cassava, respectively. Upon translation to amino acid as seen in **Figure 2**, the SNP at position 549 was synonymous because it caused no amino acid changes therefore may not be an effective marker in marker assisted selection with respect to yellow and white cassava roots. But the SNP located at position 572 of the gene caused non-synonymous changes in their amino acid sequence as individuals with the A nucleotide gave Alanine while those with the C nucleotide gave aspartic acid. Individuals with white root carrying the C nucleotide gave Alanine while those with yellow root carrying the A nucleotide gave Aspartic acid. Thus, this SNP can be effectively utilized as molecular marker in selecting for root color of cassava even before it gets to the stage of developing roots for phenotypic evaluation. The presence of SNPs in regulatory and coding regions of genes can cause significant phenotypic effect on function of protein and how genes are expressed. This permits the association of genotypic and phenotypic variations which has been successfully exploited in cultivar identification and genetic diversity analysis. Though the association of traits of economic importance may be less than 100%, they can still be successfully utilized in marker assisted selection and gene isolation. Given its precision in germplasm identification, it can also be an efficient tool for plant germplasm screening.

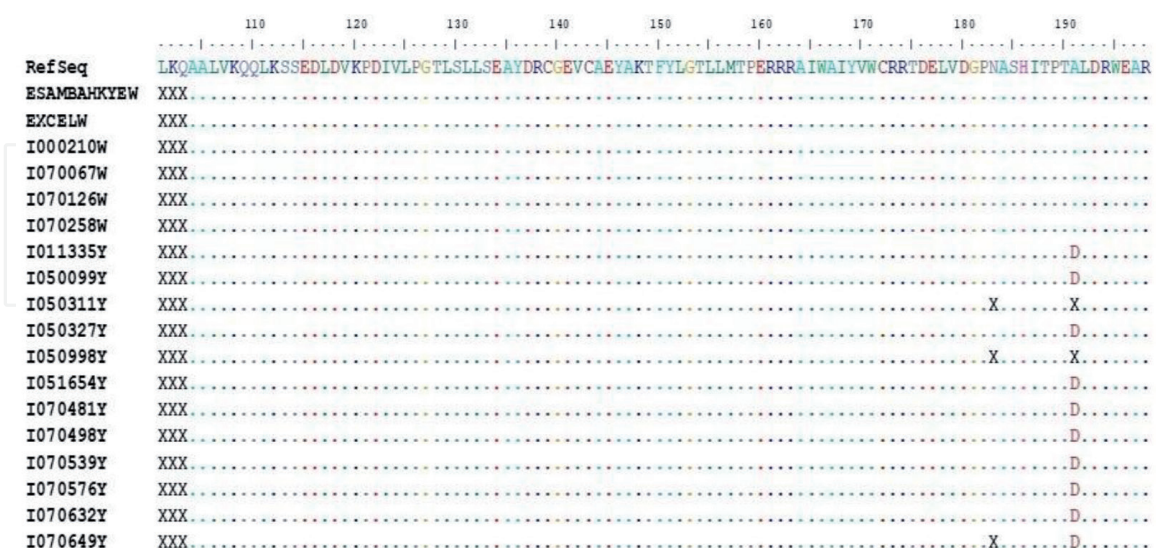


Figure 2. Multiple sequence alignment of amino acid sequences showing a non-synonymous mutation at the position of the A/C nucleotide polymorphism of PSY2 gene in cassava.

The translation of individuals with the A nucleotide gave Alanine (A) while those with the C nucleotide gave Aspartic acid (D), X shows ambiguous position. Consensus sequences are plotted as dot with reference to the reference sequence (RefSeq).

Single nucleotide polymorphism is now a highly preferred genetic marker due to the increase in amount of sequence information and the determination of gene function due to genomic research. Their widespread abundance in the genome and the development of new SNP genotyping platforms has made them the preferred marker for plant germplasm screening or characterization and identification of functional genes for traits of interest [4–6]. SNPs are easily automated with high throughput techniques and are being used for large segregating populations.

A lot of techniques and methods have been applied to identify SNPs and use those that can successfully discriminate between traits of interest for marker development. SNP markers can be identified by carrying out locus specific PCR. Here, locus specific PCR primers are synthesized from genomic sequences that are known and available in the public databases or previously sequenced data. The primers are used to amplify DNA samples from several individuals of a plant species. The resultant PCR amplicons are sequenced and aligned. Alignment is searched for availability of SNPs which are base changes within a particular locus. Depending on the informativeness of such SNP after characterization, it can be further evaluated for its effectiveness as a marker for germplasm screening or marker assisted selection. This method of SNP discovery can only be used if there is an existing information concerning the sequence to be amplified. This method was used by Udoh et al. [7] to identify SNPs causing non-synonymous changes in amino acid of phytoene synthase gene in cassava linked to expression of yellow color in roots of some cassava varieties. Also, Harjes et al. [8] identified SNPs in the regulatory regions of lycopene epsilon cyclase genes causing accumulation of carotenoids in maize.

The availability of whole genome sequences and expressed sequence tags (ESTs) databases has allowed for non-gel-based approach to SNP discovery. Unigenes or EST sequences of interest can be analyzed de novo or exported to other convenient computer software programs for alignment and SNP searches. Alignment of genomic sequences may identify SNPs in both coding and non-coding regions of the genome but ESTs are preferred because they are coding sequences and SNPs identified here can affect gene expression thus can be evaluated further for downstream applications. This approach to discover SNP is relatively easy and cost effective although the authenticity of sequences used may not be guaranteed because they were mined from public databases.

Also, high throughput automated next generation sequencing (NGS) platforms such as Illumina Genome Analyzer, Roche/454 FLX and ABI SOLiD can generate lots of SNPs when used for whole genome sequencing, RNA sequencing, methylated DNA sequencing and exome capture procedures. SNPs generated through these platforms can be between different varieties of plant species or between the same unigenes. Although these platforms are relatively expensive to utilize, but prices are gradually easing with increasing patronage. This method has been used to discover thousands of good quality SNPs in four pea recombinant inbred lines [9]. Nevertheless, limitations exist with regards to accuracy, sensitivity and reproducibility of reads generated. A major concern using the NGS platform is the need to use very good assemblers to organize reads for SNP calling; examples of some assemblers include Genome Analysis Toolkit (GATK) [10], SOAPsnp [11, 12] and freebayes [13]. Different SNP callers have been compared in searching for a more versatile tool [14–18]. In a study for SNP discovery using RNA-sequence data, a combination of SNP callers Trinity-GATK gave 100% accuracy in peach and mandarin RNA-sequencing [19].

The versatility of SNPs has also led to their widespread use in phylogenetics to study the relatedness of organisms through the use of molecular sequencing data resulting in the identification and accurate classification of organisms. It has

also been applied in phytogeography for determining the distribution of plant species. A major advantage of the single-base resolution of SNPs is that it allows better detection of 'perfect' markers, which are causally linked to agronomic traits. Another high-throughput method used for detecting SNPs is the GBS (genotyping by sequencing) which utilizes a range of techniques including those of reduced-representation sequencing and whole genome resequencing. Generally, this method identifies SNPs that are broadly distributed throughout the genome of organisms by fragmenting the genome using restriction enzymes, fragments are ligated to adapters and amplified. Amplified products are sequenced and aligned to a reference genome to call SNPs. GBS is gradually leading a transition from population genetics to population genomics, so that high-throughput marker recognition in plant population is affordable. A lot of commercial crops have been studied using GBS to aid breeding processes in Rice [20–22], Maize [23–25], potato [26–28]. Although GBS was initially developed as a reduced-representation sequencing (RRS) approach using restriction enzymes to decrease genome complexity before sequencing [29–32]. Whole genome re-sequencing approaches was applied to allow higher genomic resolution. Since the creation of GBS, it has undergone continuous development, based on reduced-representation sequencing or whole genome resequencing methods. Combined with phenotypic data, GBS procedures provide a powerful basis for rapid mapping and identification of SNPs in genes underlying agronomic traits, which can then be utilized as efficient molecular markers for crop germplasm improvement.

Notwithstanding, the fact that a lot of next generation SNP genotyping techniques have been developed, they are all within the same bracket with regards to limitations of cost, complexity and accuracy. Important quantitative trait loci and SNPs associated with desirable agronomic traits have been employed to improve productivity of crops. Whole genome resequencing of *G.max* and *G. soja* by Ramakrishna et al. [33], identified SNPs and InDels and seven genes that hold a probable role in the determination of seed permeability. The expression differences of these genes at different stages of water imbibition was analyzed and two genes were identified that revealed preliminary, but a relevant association with soybean seed permeability trait. Genome-wide association study was performed by Do et al. [34] to map genomic regions for salt tolerance in a diverse panel of 305 soybean accessions using a SNP dataset derived from the SoySNP50K iSelect BeadChip [35]. The analysis revealed a major locus for salt tolerance on Chr. 3 confirmed by a number of significant SNPs, of which three gene-based SNP markers, Salt-20, Salt14056 and Salt11655, had the highest association with the studied trait.

In *Arachis* species Clevenger et al. [36] re-sequenced 20 genotypes and selected genome-wide SNPs to develop large-scale SNP genotyping array which will be very useful for further genetic and breeding applications in *Arachis*. In Maize, Unterseer et al. [37] developed a high-density maize SNP array comprising 616,201 variants (SNPs and indels) and used to design commercially available Affymetrix® Axiom® Maize Genotyping Array. The array is composed of 609,442 SNPs and 6759 indels. Among these were 116,224 variants in coding regions and 45,655 SNPs of the Illumina® MaizeSNP50 BeadChip for study comparison. The Array although optimized for European and American climate is suitable for a broad range of applications because of the stringent quality filter criteria implemented. Cereals like *Zea mays* and *Oryza sativa* have extensively been studied for SNP diversity using diverse germplasm.

2.1 Single nucleotide polymorphism genotyping methods

Single nucleotide polymorphism is an individual nucleotide base difference between two DNA sequences. When SNPs occur within a gene, they may play a

more direct role on the trait by affecting the gene's function and such SNPs can be exploited as molecular markers. Molecular markers enable precise identification of genotypes without the confounding effect of the environment [38], because selection is based on molecular determination and not the morphological expressions observed. A more informative marker gives a high polymorphic information content result [39]. SNP markers for chickpea and pigeon pea were evaluated and found to show 100% consistency and polymorphic information content values between 0.02 to 0.5 [22, 40]. SNP markers can be used for association studies, conservation genetics, germplasm screening or characterization, genetic diversity analysis and are fast becoming the preferred marker system in marker assisted breeding programs.

In the last 10 years, the rapid transformation in sequencing technology have enormously affected crop genotyping procedures. These new procedures enhanced rapid, high-throughput genotyping of whole crop population and gives opportunity to advance use of molecular tools in plant breeding. There is an urgent need in crop improvement programs to speed up crop production through marker assisted selection or introduce alleles that confers plants with resistance to pest and diseases, abiotic stress adaptation and high yield potential. Elite cultivars, store very useful genetic information that needs to be introgressed. Molecular marker approaches have been used in analyzing and identifying alleles associated with desirable agronomic traits in diverse germplasm pool of legumes and cereals [41].

Some SNPs genotyping methods that are easy to use and accurate and can specifically genotype SNP markers at specific loci for a collection of plant population are presented below.

2.1.1 Tetra ARMS allele specific PCR

Tetra ARMS (tetra-primer amplification refractory mutation system) allele specific PCR is a versatile, rapid and economical SNP detection tool. Other contemporary SNP genotyping tools include allele specific PCR, high resolution melting analysis, PCR single stranded conformation polymorphism, PCR-primer introduced restriction analysis and real-time PCR-based genotyping. It involves a single PCR step followed by gel electrophoresis. Tetra ARMS allele specific PCR utilizes four primers including outer forward, outer reverse, inner forward and inner reverse primers. The outer forward or outer reverse primer combination generates the outer fragment of the SNP locus and acts as an internal control for the PCR. The inner forward or outer reverse and outer forward or inner reverse primer combination yield allele-specific amplicons depending on the genotype of the sample used. The placing of the inner primers is not the same as those of the corresponding outer primer to produce amplicons with different sizes and easily visualized on gel and distinction is made accordingly [42–44].

A study by Ehnert et al. [45] using tetra ARMS allele specific PCR method described three common single nucleotide polymorphisms in the *PADI4* gene involved in diverse post-translational modifications of proteins in eukaryotes. The SNPs which are thought to affect PAD4 expression activity are rs874881, rs11203366 and rs11203367. Hypercitrullination as a result of increased PAD expression or activity, is associated with autoimmune diseases like rheumatoid arthritis, lupus, Alzheimer's disease, ulcerative colitis, multiple sclerosis, and certain cancers. SNP markers identified by genotyping-by-sequencing were used to distinguish four varieties of sweet potato through tetra-primer ARMS-PCR method. It was shown that three variety-specific fragments (164 bp and 241 bp of SNP 04-27457768 and 292 bp of SNP 03-16195623) were amplified in the

'Beniharuka', 'Pungwonmi', and 'Annobeni' sweet potato varieties [43]. Some SNP markers developed by Angiolillo et al. [46] was used to resolve the issue of nomenclature for 65 olive samples as the markers were able to discriminate 77% of the olive cultivars. SNP markers developed in this study were used to assess the genetic variability and diversity of widely cultivated olive cultivars important for oil production. To show the reliability of the tetra-primer ARMS-PCR technique and its potential for use in low- to moderate-throughput situations Chiapparino et al. [47] unambiguously assayed five SNPs in a set of 132 varieties of cultivated barley.

With this technique, there is almost always a need for troubleshooting to standardize the procedure especially at initial steps of the protocol to adapt it to the genotype investigated. This really reduced its wide spread application, despite the fact that it is economical and precise in SNP genotyping. In order to improve the ARMS-PCR procedure, several modifications have been suggested to optimize its usage. Two improvements were suggested by Tanha et al. [42]; one is to equalize outer primer and inner primer strength by adding a mismatch at 2 positions of outer primers and the second is to equal annealing temperature which should be a little higher than melting temperature. This resulted in the improvement of expected result and specificity. Another study by Alyethodi et al. [44] suggest that the use of Strand displacement polymerase rather than conventional Taq polymerase resulted in the generation of amplicons by 25 cycles in the PCR reaction while Taq polymerase needed a minimum of 35 cycles. Also, reaction with Strand displacement polymerase did not require PCR enhancers like dimethyl sulfoxide, thus it was time saving and efficient.

2.1.2 KASP assay for SNPs genotyping

Another robust and easy to use SNP genotyping method is Kompetitive allele-specific PCR (KASP) genotyping assay based on competitive allele specific polymerase chain reaction, developed by LGC genomics (www.lgcgroup.com). It is widely applied in plant breeding because of its reduced cost in genotyping large number of samples. It allows for biallelic scoring of SNPs and insertion/deletions at specific loci and can be conveniently used for small number of SNPs. Here, the SNP-specific KASP assay mix and the universal KASP master mix are added to the DNA samples, followed by thermal cycling. The bi-allelic discrimination is carried out by competitive binding of two allele-specific forward primers. Each of the primers are labeled with fluorescence resonant energy transfer cassettes a FAM dye and an HEX dye [48, 49].

This method is a PCR-based homogenous fluorescent SNP genotyping set up which is cost-effective to run and more reliable than other SNP genotyping techniques. Since the introduction of KASP, it has been developed and used to genotype rice, wheat, soybean, cucumber, chickpeas and many other crops. It has been employed in the enhancement and production of efficient markers in Chinese cabbage. The authors re-sequenced 4 Chinese cabbage and carried out SNP survey in the genome. They established KASP-SNP resource and converted 258 SNP variations into KASP molecular markers. These molecular markers discovered in Chinese cabbage will be invaluable for germplasm identification and cabbage research around the world [50]. Also, Khanal et al. [51] reported flanking sequences of 162 putative SNPs, none of them have been previously evaluated to determine whether they performed as intended. Therefore, a subset of 31 putative SNPs that represent the entire nematode genome were designed to form a residual emission fluorescence KASP.

With KASP primers, biallelic scoring of SNPs at specific loci is possible. Cotton (*Gossypium hirsutum*) leaf was mapped using the underlying gene of okra leaf using KASP assay. The sequences of okra leaf gene, GhOKRA, has been link to other plant species and is involved in regulating leaf morphology in plants. SNP markers located on the okra gene (GhOKRA) was successfully establish to be the best candidate gene responsible for okra leaf trait in upland cotton [52]. SNP marker analysis can be used for genetic diversity analysis, create genetic maps, and marker assisted selection of crops. To use these technologies, one must first identify and validate putative SNPs. Many technologies exist for SNP genotyping but KASP performs well when it comes to adaptability, efficiency and cost-effectiveness. It is efficient in determining the alleles at specific locus within genomic DNA [53].

For crop improvement purposes, the maize breeders at International Maize and Wheat Improvement Center developed 16 marker-assisted recurrent select (MARS) populations. The parents of these MARS populations were initially genotyped along with over 450 maize inbred and advanced breeding lines using the GoldenGate assay [54].

2.2 Application of SNP in crop improvement

SNP markers are used for the improvement of crops in a number of ways which include to select disease resistant crops, high yielding varieties, plants that can withstand biotic and abiotic stress and many more. The development of SNP markers has become a regular process, especially for crops with reference genome and this new order has influenced the application of SNP markers in plant breeding. For more than two decades, researchers experienced so much throwback in a bid to develop markers linked to discovered QTLs. However, the revolution in sequencing technology, brought about easy identification of SNP markers underlying genes in a QTL. SNP markers were used to characterize natural variation of sorghum grain nutrients composition in a global sorghum panel and genome wide association study was use to map QTL responsible for this variation. It was discovered by Rhodes et al. [55] that protein, fat and starch all had strong correlation across years, but protein was the most significant. Also, protein had the highest narrow sense-heritability. Further investigation showed that there is a strong negative correlation between starch and protein and fat, and strong positive correlation between protein and fat.

SNP markers have been frequently used in marker assisted selection due to its abundance in the genes of all species. In breeding for resistance to root-knot nematode (*Melodogyne incognita*) that infects soybean by Dubiela et al. [56], SNP markers were identified for *M. incognita* in soybean using a microarray panel. The markers were used to identify susceptible soybean which was confirmed by phenotypical evaluation of plants on the field. This marker assisted selection protocol helped in the identification of resistant varieties. Also, Potato blight is considered as one of the most destructive disease of potato, which is caused by a fungus, *Phytophthora infestans*. Globally, there are 2 dominant mating type strains: type A1 and A2. SNP markers were used to determine potato late blight of susceptible and resistant genotypes. Using genome wide association studies, Nay et al. [57] identified SNP markers that co-segregated with resistance loci for Angular leaf spot (ALS) in common bean and haplotypes. The discovered markers will increase breeding efficiency for ALS resistance and allow researchers to react faster to future changes in pathogen pressure and composition.

The allelic, high number of loci that can be multiplexed and possibility of automation of SNP markers makes them very useful for cultivar identification.

Grapevine cultivar identification was carried out using over 300 SNPs in its genome. Re-sequencing method was used in the selection of 11 genotypes, 48 SNPs spread across all grapevine chromosome providing enough information content for genetic cultivar identification [58].

The quality of a crop is highly dependent on the number of micronutrients it contains. Genome wide study was used to identify SNPs associated with micronutrient (Fe, Se and Zinc) concentration in pea (*Pisum sativum*). The SNP was very useful in identification of seed mineral concentration in pea and the loci were mapped. For Fe concentration, 5 SNPs were identified and each marker was distinct phenotypically. Markers identified for Zn was 5 chromosomal and 3 non-chromosomal SNP markers while eight was identified for Se. It was also stated by Dissanayaka et al. [59] that genome wide association studies can be effectively used for reliable marker assisted selection scheme.

3. Conclusion

Germplasm screening is usually a first step in plant breeding programs. It aims to reduce the large collection of plants and narrow down on those that fit the breeding objectives in view. It is usually laborious and time consuming, but the use of molecular markers can substantially aid this process. Several molecular markers including simple sequence repeats (SSRs) have previously been utilized in plant germplasm screening but SNPs markers enable selection based on target genes that code for specific trait of interest.

Advances in sequencing technologies has given rise to SNP markers and now the markers of choice in genetic studies because they are robust, widely distributed throughout the genome of plants and highly multiplexable. SNPs represent difference in a single nucleotide in the genome and those that are linked to a phenotype as a result of nonsynonymous amino acid changes can be reliably used as molecular markers.

A number of techniques and methods have been applied to discover new SNPs including non-gel methods where SNPs are mined from multiple sequence alignment in databases. SNPs have also been generated using next generation sequencing platforms like Illumina genome analyzer and Roch/454 FLX. Discovered SNPs are developed into user-friendly SNPs markers and used for genotyping. A number of SNP markers have been validated for marker assisted selection. In a study by Burow et al. [60], SNP markers were developed from sequences of brown midrib (bmr) trait of sorghum and used to accurately identify bmr6 or bmr12 individuals at the seedling stage. This validation was for a group of sorghum germplasm and a genetic population. Also, fifteen KASP SNP markers for bmr6 and bmr12 were developed and used for allele discrimination to select bmr individuals. Another study by Khanal et al. [61] employed KASP SNPs to determine the genetic variability present in 26 isolates of *Rotylenchulus reniformis* a plant parasitic nematode of cotton and soybean, this will be of benefit in resistant breeding programs. Also, Udoh et al. [62] developed KASP SNP markers for phytoene synthase2 gene associated with carotenoids in cassava. The validation SNP makers explained most of the proportion of phenotypic variation for carotenoids in a genetic gain cassava population.

In Panax species, Nguyen et al. [30] identified 1128 SNPs in coding gene sequences and developed 18 SNP markers from the chloroplast genic coding sequence region that can be used to distinguish all the seven Panax species from each other. Because SNPs markers are based on target gene, they are a highly reliable tool in identification of cultivars and germplasm screening.

Acknowledgements

My profound gratitude goes to my husband and children for their understanding and support during the period of writing this chapter. God bless you all.

Conflict of interest

No conflict of interest.

Notes/thanks/other declarations

I thank my coauthors for accepting to contribute to the chapter. We are grateful to Almighty God for the grace of life.

Appendices and nomenclature

AFLP	Amplified fragment length polymorphism
ARM	Amplification refractory mutation system
GBS	Genotyping-by-sequencing
KASP	Competitive allele-specific polymerase chain reaction assay
MAS	Marker assisted selection
NGS	Next generation sequencing
PCR	Polymerase chain reaction
SNPs	Single nucleotide polymorphism
SSR	Simple sequence repeats
QTL	Quantitative trait loci

Author details

Lovina I. Udoh^{1*}, Willie Peggy Obaseoji² and Chiebuka Uzoebo³

¹ Department of Botany, Akwa Ibom State University, Ikot Akpaden, Uyo, Akwa Ibom State, Nigeria

² Department of Genetics and Biotechnology, University of Calabar, Calabar, Nigeria

³ Genomics Training Center Ukana Offot, Uyo, Akwa Ibom State, Nigeria

*Address all correspondence to: lnkantol@yahoo.com; lovinaudoh@aksu.edu.ng

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Slater AT, Schultz L, Lombardi M, et al. Screening for Resistance to PVY in Australian Potato Germplasm. 2020; 1-15.
- [2] Bhawar PC, Tiwari S, Tripathi MK, et al. Screening of Groundnut Germplasm for Foliar Fungal Diseases and Population Structure Analysis Using Gene Based SSR Markers. 2020; 39: 75-84.
- [3] Zhang K, Wu Z, Tang D, et al. Development and Identification of SSR Markers Associated with Starch Properties and β -Carotene Content in the Storage Root of Sweet Potato (*Ipomoea batatas* L.). *Front Plant Sci* 2016; 7: 1-21.
- [4] Gray IC, Campbell DA, Spurr NK. Single nucleotide polymorphisms as tools in human genetics. *Hum Mol Genet* 2000; 9: 2403-2408.
- [5] Zhou L, Vega FE, Tan H, et al. Developing Single Nucleotide Polymorphism (SNP) Markers for the Identification of Coffee Germplasm. *Trop Plant Biol* 2016; 9: 82-95.
- [6] Chen X, Sullivan PF. Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. *Pharmacogenomics J* 2003; 3: 77-96.
- [7] Udoh LI, Adesoye A, Gedil M. Identification and Molecular Analysis of Pro - vitamin A Carotenoid Genes in Cassava (*Manihot esculenta* Crantz). 2017; 8: 38-44.
- [8] Harjes CE, Rocheford TR, Bai L, et al. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 2008; 319: 330-3.
- [9] Boutet G, Carvalho SA, Falque M, et al. SNP discovery and genetic mapping using genotyping by sequencing of whole genome genomic DNA from a pea RIL population. *BMC Genomics*. Epub ahead of print 2016. DOI: 10.1186/s12864-016-2447-2.
- [10] Poplin R, Ruano-Rubio V, DePristo M, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* 2017; 201178.
- [11] Li R, Li Y, Kristiansen K, et al. SOAP: Short oligonucleotide alignment program. *Bioinformatics* 2008; 24: 713-714.
- [12] Li R, Li Y, Fang X, et al. SNP detection for massively parallel whole-genome resequencing. *Genome Res* 2009; 19: 1124-1132.
- [13] Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing.
- [14] Sandmann S, De Graaf AO, Karimi M, et al. Evaluating Variant Calling Tools for Non-Matched Next-Generation Sequencing Data. *Sci Rep* 2017; 7: 1-12.
- [15] Cornish A, Guda C. A Comparison of Variant Calling Pipelines Using Genome in a Bottle as a Reference. *Biomed Res Int*; 2015. Epub ahead of print 2015. DOI: 10.1155/2015/456479.
- [16] Liu F, Zhang Y, Zhang L, et al. Systematic comparative analysis of single-nucleotide variant detection methods from single-cell RNA sequencing data. *Genome Biol*; 20. Epub ahead of print 19 November 2019. DOI: 10.1186/s13059-019-1863-4.
- [17] Ni G, Strom TM, Pausch H, et al. Comparison among three variant callers and assessment of the accuracy of imputation from SNP array data to whole-genome sequence level in

- chicken. *BMC Genomics*; 16. Epub ahead of print 21 October 2015. DOI: 10.1186/s12864-015-2059-2.
- [18] Bian X, Zhu B, Wang M, et al. Comparing the performance of selected variant callers using synthetic data and genome segmentation. *BMC Bioinformatics* 2018; 19: 429.
- [19] Zhao Y, Wang K, Wang W, et al. A high-throughput SNP discovery strategy for RNA-seq data. 2019; 1-10.
- [20] Yadav S, Sandhu N, Singh VK, et al. Genotyping-by-sequencing based QTL mapping for rice grain yield under reproductive stage drought stress tolerance. 2019; 1-12.
- [21] Ali J, Aslam UM, Tariq R, et al. Exploiting the Genomic Diversity of Rice (*Oryza sativa* L.): SNP-Typing in Introgression-Breeding Populations. 2018; 9: 1-10.
- [22] Hiremath PJ, Kumar A, Penmetsa RV, et al. Large-scale development of cost-effective SNP marker assays for diversity assessment and genetic mapping in chickpea and comparative mapping in legumes. *Plant Biotechnol J* 2012; 10: 716-32.
- [23] Wang N, Yuan Y, Wang H, et al. Applications of genotyping - by - sequencing (GBS) in maize genetics and breeding. *Sci Rep* 2020; 1-12.
- [24] Su C, Wang W, Gong S, et al. High Density Linkage Map Construction and Mapping of Yield Trait QTLs in Maize (*Zea mays*) Using the Genotyping-by-Sequencing (GBS) Technology Genome Wide Identification of SNPs Using. 2017; 8: 1-14.
- [25] Paulsmeyer MN, Brown PJ, Juvik JA. Discovery of Anthocyanin Acyltransferase1 (AAT1) in Maize Using Genotyping-by-Sequencing (GBS). 2018; 8: 3669-3678.
- [26] Technology FSS, Yu X, Zhang M, et al. An SNP-Based High-Density Genetic Linkage Map for Tetraploid Potato Using Specific Length Amplified.
- [27] SALIHU BZ, FALUSI OA, ADEPOJU AO, et al. Assessment of Genetic Diversity of Promising Castor Been (*Ricinus communis* L.) Genotypes in Nigeria. *Not Sci Biol* 2019; 11: 467-474.
- [28] Wadl PA, Olukolu BA, Branham SE, et al. Genetic Diversity and Population Structure of the USDA Sweetpotato (*Ipomoea batatas*) Germplasm Collections Using GBSpoly. 2018; 9: 1-13.
- [29] Elshire RJ, Glaubitz JC, Sun Q, et al. A Robust , Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. 2011; 6: 1-10.
- [30] Nguyen VB, Linh Giang VN, Waminal NE, et al. Comprehensive comparative analysis of chloroplast genomes from seven Panax species and development of an authentication system based on species-unique single nucleotide polymorphism markers. *J Ginseng Res* 2020; 44: 135-144.
- [31] Shaibu AS, Sneller C, Motagi BN, et al. Genome-Wide Detection of SNP Markers Associated with Four Physiological Traits in Groundnut (*Arachis hypogaea* L.) Mini Core Collection. *Agronomy* 2020; 10: 192.
- [32] Meger J, Ulaszewski B, Vendramin GG, et al. Using reduced representation libraries sequencing methods to identify cp DNA polymorphisms in European beech (*Fagus sylvatica* L).
- [33] Ramakrishna G, Kaur P, Nigam D, et al. Genome-wide identification and characterization of InDels and SNPs in *Glycine max* and *Glycine soja* for contrasting seed permeability traits. 2018; 1-15.

- [34] Do TD, Vuong TD, Dunn D, et al. Identification of new loci for salt tolerance in soybean by high-resolution genome-wide association mapping. 2019; 1-16.
- [35] Song Q, Hyten DL, Jia G, et al. Development and Evaluation of SoySNP50K, a High-Density Genotyping Array for Soybean. 2013; 8: 1-12.
- [36] Clevenger J, Chu Y, Chavarro C, et al. Genome-wide SNP Genotyping Resolves Signatures of Selection and Tetrasomic Recombination in Peanut. *Mol Plant* 2017; 10: 309-322.
- [37] Unterseer S, Bauer E, Haberer G, et al. A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. 2014; 1-15.
- [38] Meuwissen THE, Hayes BJ, Goddard ME. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps.
- [39] Species R, Lebedev VG, Subbotina NM, et al. Transferability and Polymorphism of SSR Markers Located in Flavonoid Pathway Genes in *Fragaria* and.
- [40] Saxena RK, Penmetsa RV, Upadhyaya HD, et al. Large-scale development of cost-effective single-nucleotide polymorphism marker assays for genetic mapping in pigeonpea and comparative mapping in legumes. *DNA Res* 2012; 19: 449-461.
- [41] Dwivedi SL, Scheben A, Edwards D, et al. Assessing and Exploiting Functional Diversity in Germplasm Pools to Enhance Abiotic Stress Adaptation and Yield in Cereals and Food Legumes. 2017; 8: 1-30.
- [42] Tanha HM, Naeini MM, Rahgozar S, et al. Modified Tetra-Primer ARMS PCR as a Single-Nucleotide Polymorphism Genotyping Tool. Epub ahead of print 2015. DOI: 10.1089/gtmb.2014.0289.
- [43] Park H, Kim S, Nie H, et al. Molecular identification of sweet potato accessions using ARMS-PCR based on SNPs. 2020; 2818: 124-130.
- [44] Alyethodi RR, Singh U, Kumar S, et al. RESEARCH NOTE T - ARMS PCR genotyping of SNP rs445709131 using thermostable strand displacement polymerase. *BMC Res Notes* 2018; 1-5.
- [45] Ehnert S, Linnemann C, Braun B, et al. One-Step ARMS-PCR for the Detection of SNPs — Using the Example of the PADI4 Gene.
- [46] Angiolillo A, Lucentini L, Pilla F, et al. SNP-based markers for discriminating olive (*Olea europaea* L.) cultivars. Epub ahead of print 2006. DOI: 10.1139/g06-068.
- [47] Chiapparino E, Lee D, Donini P. NOTE / NOTE Genotyping single nucleotide polymorphisms in barley by tetra-primer ARMS – PCR. Epub ahead of print 2014. DOI: 10.1139/g03-130.
- [48] Neelam K, Brown-Guedira G, Huang L. Development and validation of a breeder-friendly KASPar marker for wheat leaf rust resistance locus Lr21. *Mol Breed* 2013; 31: 233-237.
- [49] Allen AM, Barker GLA, Berry ST, et al. Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant Biotechnol J* 2011; 9: 1086-1099.
- [50] Yang S, Yu W, Wei X, et al. An extended KASP-SNP resource for molecular breeding in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *PLoS One* 2020; 15: 1-17.
- [51] Khanal C, Kularathna MT, Ray JD, et al. Single Nucleotide Polymorphism

Analysis Using KASP Assay Reveals Genetic Variability in *Rotylenchulus reniformis*. 2019; 1835-1842.

[52] Zhu QH, Zhang J, Liu D, et al. Integrated mapping and characterization of the gene underlying the okra leaf trait in *Gossypium hirsutum* L. *J Exp Bot* 2016; 67: 763-774.

[53] Graves H, Rayburn AL, Gonzalez-Hernandez JL, et al. Validating DNA Polymorphisms Using KASP Assay in Prairie Cordgrass (*Spartina pectinata* Link) Populations in the U.S. *Front Plant Sci*; 6. Epub ahead of print 2016. DOI: 10.3389/fpls.2015.01271.

[54] Semagn K, Babu R, Hearne S, et al. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. *Molecular Breeding* 2014; 33: 1-14.

[55] Rhodes DH, Hoffmann L, Rooney WL, et al. Genetic architecture of kernel composition in global sorghum germplasm. *BMC Genomics* 2017; 18: 1-8.

[56] Dubiela C, Montecelli TDN, Lazzari F, et al. Development and validation of SNP assays for the selection of resistance to *Meloidogyne incognita* in soybean. *Crop Breed Appl Biotechnol* 2019; 19: 102-109.

[57] Nay MM, Mukankusi CM, Studer B, et al. Haplotypes at the Phg-2 Locus Are Determining Pathotype-Specificity of Angular Leaf Spot Resistance in Common Bean. *Front Plant Sci* 2019; 10: 1-11.

[58] Cabezas JA, Ibáñez J, Lijavetzky D, et al. A 48 SNP set for grapevine cultivar identification. *BMC Plant Biol*; 11. Epub ahead of print 2011. DOI: 10.1186/1471-2229-11-153.

[59] Dissanayaka DN, Gali KK, Jha AB, et al. Genome-wide association study to identify single nucleotide polymorphisms associated with Fe, Zn, and Se concentration in field pea. *Crop Sci* 2020; 60: 2070-2084.

[60] Burow G, Chopra R, Sattler S, et al. Deployment of SNP (CAPS and KASP) markers for allelic discrimination and easy access to functional variants for brown midrib genes bmr6 and bmr12 in *Sorghum bicolor*. *Mol Breed* 2019; 39: 115.

[61] Khanal C, Kularathna MT, Ray JD, et al. Single Nucleotide Polymorphism Analysis Using KASP Assay Reveals Genetic Variability in *Rotylenchulus reniformis*. *Plant Dis* 2019; 103: 1835-1842.

[62] Udoh LI, Gedil M, Parkes EY, et al. Candidate gene sequencing and validation of SNP markers linked to carotenoid content in cassava (*Manihot esculenta* Crantz). *Mol Breed*; 37. Epub ahead of print 2017. DOI: 10.1007/s11032-017-0718-5.