

# Identification of functional SNPs in genes and their effects on plant phenotypes

Md. Amdadul Huq · Shahina Akter · Ill Sup Nou · Hoy Taek Kim · Yu Jin Jung · Kwon Kyoo Kang

Received: 18 March 2016 / Revised: 24 March 2016 / Accepted: 25 March 2016

© Korean Society for Plant Biotechnology

**Abstract** Single nucleotide polymorphism (SNP) is an abundant form of genetic variation within individuals of species. DNA polymorphism can arise throughout the whole genome at different frequencies in different species. SNP may cause phenotypic diversity among individuals, such as individuals with different color of plants or fruits, fruit size, ripening, flowering time adaptation, quality of crops, grain yields, or tolerance to various abiotic and biotic factors. SNP may result in changes in amino acids in the exon of a gene (asynonymous). SNP can also be silent (present in coding region but synonymous). It may simply occur in the noncoding regions without having any effect. SNP may influence the promoter activity for gene expression and finally produce functional protein through transcription. Therefore, the identification of functional SNP in genes and analysis of their effects on phenotype may lead to better understanding of their impact on gene function for varietal improvement. In this mini-review, we focused on evidences revealing the role of functional SNPs in genes and their phenotypic effects for the purpose of crop improvements.

**Keywords** Functional SNPs, Genetic diversity, Phenotypic variation, Biotic and abiotic stresses

## Introduction

Crop plants are very important for human being, therefore different strategies are using for their improvement accordance to current demands. Among these strategies, plant breeding program is a natural way of variety development. During breeding programs, a lot of genetic variations are arisen, which are corresponding to the phenotypes; such as quality of crops, grain yields, different colors of plants or fruits, size of fruits, and tolerance to various biotic and abiotic stresses (Vidal et al. 2012; Jang et al. 2015). Genetic diversity is also generated in different crop species through domestication of the same species in different geographical regions. The most common form of genomic variation is single nucleotide variation in the genome within the individuals. Analysis of DNA variation through DNA sequencing of a target gene regulating phenotypes is a good way to identify causal genes for the traits. The recent advances in sequencing technology are giving great opportunity for plant breeders to find out genetic diversity in different breeding populations, especially for the discovery of functional SNP (single nucleotide polymorphism) in causal genes and development of SNP markers, which are associated with diverse agronomic traits in crops (Vidal et al. 2012). Most of the crop plants contain high nutritional value, which provides some particular nutrients that have high impact to maintain healthy human body. These nutrients may vary largely depending on growing conditions, varieties and mutations in functional genes (Schreiber et al. 2014).

Sequencing of many crop plant genomes is already completed, which was a major milestone for plant research (Huq et al. 2016). Reference genome sequence is essential for measuring genetic polymorphisms among individuals of same species. In order to identify the sequence diversity within crop species like rice, potato, tomato, maize, etc., a lot of resequencing data are now available (Causse et al. 2013; Chen et al. 2014; Xu et al. 2014; Chung et al. 2014). These data contributed to evidence suggesting that during process of domestication,

---

M. A. Huq · S. Akter  
Department of Horticulture, Hankyong National University,  
Ansung City, Gyeonggi-do, 17579, Republic of Korea

Y. J. Yu (✉) · K. K. Kang (✉)  
Department of Horticulture, Hankyong National University,  
Ansung City, Gyeonggi-do, 17579, Republic of Korea  
Department of Horticulture, Hankyong National University,  
Ansung City, Gyeonggi-do, 17579, Republic of Korea  
e-mail: [kykang@hknu.ac.kr](mailto:kykang@hknu.ac.kr), [yuyu1216@hknu.ac.kr](mailto:yuyu1216@hknu.ac.kr)

I. S. Nou · H. T. Kim  
Department of Horticulture, Suncheon National University, 255,  
Jungang-ro, Suncheon, Jeonam-do, 57922, Korea

mutation, multiplication, selection breeding and exchange of cultivars, a huge number of polymorphisms were spontaneously or artificially generated in the genome of different individuals of same species. These changes in genome can alter the functions of important genes and ultimately make the phenotypic variations in plants (Vidal et al. 2012; Shi et al. 2015; Shirasawa et al. 2016). The most abundant DNA polymorphisms in the genome sequences are SNPs and are thought to play a major role in the induction of phenotypic variations. There are many reports about the gene specific or genome-wide functional SNP discovery in different breeding varieties or lines, which are associated with different phenotypic changes (Kharabian-Masouleh et al. 2012; Kumar et al. 2014; Jang et al. 2015). In this paper, we focused on the evidences revealing the role of functional SNPs of genes and their phenotype effects for crop improvements.

### SNP and its significance

SNP is a variation at a single position in DNA sequence among individuals of same species. In short, SNP is the polymorphism occurring within DNA samples with difference at single base. SNPs are the most common DNA polymorphisms in genome sequences of human, animals, and plants and they are thought to play a major role in the induction of phenotypic variations. According to international SNP map working group, human genome sequence contains 1.42 million SNPs and average one SNP per 1.9 kb (Sachidanandam et al. 2001). Also in plants, SNP polymorphisms are found in high density across the genome (Ching et al. 2002). In Nipponbare rice genome, 0.64 SNP was found per one kb (Jeong et al. 2013), while in tomato average 6.1 SNP per one kb was observed in the whole genome (Kim et al. 2014).

Different DNA markers are widely used for analysis of genetic diversity of plants, their evolutionary studies, association mapping as well as diagnostics, fingerprinting, and breeding applications. Among all DNA markers, SNPs are the most abundant and robust, feasible for automated high-throughput genotyping, and available for multiple assay options using different technology platforms to meet the demand for genetic studies and molecular breeding in crop plants (Steemers and Gunderson 2007; Alkan and Eichler 2011). In recent years, SNPs have gained much interest in the scientific and breeding community that could be used as potential genetic markers, which may be identified effectively in every gene (Rafalski 2002). SNPs also can identify the genomic diversity of species to demonstrate the speciation and evolution, and associate genomic variations with phenotypic traits (McNally et al. 2009). The major applications of SNP are described shortly.

### SNP for genetic mapping

Genetic map refers to the arrangement of genes, identification of the locus of a gene and measurement of distances between genes. Construction of genetic maps are essential tools in plant breeding for genetic improvement as they are able to identify the gene location and quantitative trait loci (QTL), as well as crucial tools for genome sequence assembly and comparative genomic analysis and map based cloning. Biallelic nature of SNP, their high abundance in genome, uniform genome distribution and cost effectiveness (Ganal et al. 2009) make them an ideal marker for constructing new genetic maps compared to other genetic markers, which are often multiallelic (Kruglyak 1997). Therefore, SNP-based genetic maps have been developed in many economically important agricultural species such as cucumber (Wei et al. 2014), rice (Xie et al. 2010), maize (Buckler et al. 2009), apple (Sun et al. 2015), soybean (Akond et al. 2013), cotton (Byers et al. 2012), *Brassica* (Li et al. 2009) etc. SNPs also have been considered as one of the ideal marker for genome wide association mapping, which have led to the discovery of thousands to millions of SNPs in last few years and made it possible to produce genome-wide haplotypes of large numbers of genotypes. In many plant species, such type of studies were reported like *Arabidopsis* (Aranzana et al. 2005), rice (Huang et al. 2010), maize (Poland et al. 2011), barley (Pasam et al. 2012) etc.

### SNP for evolutionary studies

SNPs can be used for evolutionary studies of genome that can reveal about population history, how breeding system and selection affect variation at genetic level. Because, generally SNP is used for study of sequence variation among species and such type of variations are present at all levels of evolution and ultimately SNP can provide an understanding of how modern genome has evolved. The commonly used markers for evolutionary studies are SSRs (simple sequence repeats) and mitochondrial DNA which may be misinterpreted due to homoplasmy (Morin et al. 2004). It is possible to avoid this problem by using SNP markers that represent single base nucleotide substitutions (Vignal et al. 2002). Many successful reports are already published about the use of SNPs to study the evolution of genes such as WAG-2 (wheat AG-2) in wheat (Wei et al. 2011).

### Techniques for SNP genotyping

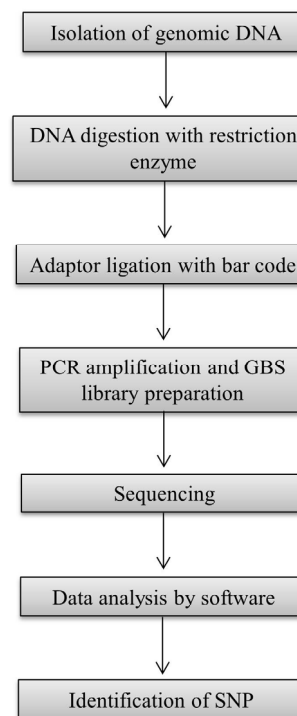
A large number of techniques have been developed for the identification of SNP polymorphisms in plants. Selection of

the technique depends on the cost, time, availability, reliability factors. There are many reports that described the different methodologies of SNP genotyping (Gut 2001; Kumar et al. 2012). From all of these methodologies, direct DNA sequencing technologies are considered as the most used and benefited for SNP identification.

Sequencing-based techniques were first invented at 1977 through Sanger method which depends on a combination of deoxy- and dideoxy-labeled chain terminator nucleotides (Sanger et al. 1977a). In the same year, the first complete genome of bacteriophage phi X174 was sequenced by this method (Sanger et al. 1977b). But in the last decade, several NGS (next generation sequencing) technologies (Roche/454, Illumina, SOLiD) have outperformed Sanger-based sequencing in throughput and overall cost (Kircher and Kelso 2010). With a throughput of hundreds of millions to several billions of bases per run, NGS are able to identify many SNPs in a species at much lower cost in a short time (Mardis 2007). Identification of SNP using NGS is reported in different plants such as *Arabidopsis* (Zhang and Borevitz 2009), rice (McNally et al. 2009), potato (Hamilton et al. 2011), eggplant (Barchi et al. 2011), maize (Jones et al. 2009), wheat (Allen et al. 2011), barley (Waugh et al. 2009), cotton (Byers et al. 2012), common beans (Cortés et al. 2011), soybean (Hyten et al. 2010), oat (Oliver et al. 2011) etc. In order to identify functional SNPs, first, need to prepare the genomic library through DNA fragmentation and in-vitro adaptor ligation, then clonal amplification by PCR, sequencing, data analysis and identification of SNP using software. For sequencing, different companies use their own technology, such as Roche/454 uses pyrosequencing protocol, SOLiD platform uses sequencing by ligation protocol and Illumina technology uses sequencing by synthesis protocol.

### Genotyping by sequencing (GBS)

Most recently a new method has been derived for SNP genotyping using illumina NGS platform to reduce the cost for DNA sequencing, is known as GBS which was developed in 2011 (Elshire et al. 2011). GBS is a sequencing by synthesis strategy. GBS system is becoming increasingly important, effective and unique tool for SNP identification in plant species because of its low cost, reduced sample handling, no size fractionation, fewer PCR and purification steps, no reference sequence limits, efficient barcoding and easiness to scale up (Davey et al. 2011). A schematic representation of GBS technology for SNP discovery from plants was shown in Figure 1. GBS is an ideal method for SNP genotyping in plants from single gene markers to whole genome profiling



**Fig. 1** Overview of SNP discovery in plants through genotyping by sequencing (GBS) system

(Poland and Rife 2012). GBS experiments were needed to do isolation of genomic DNA from plant materials, then quantification and normalization, digestion with appropriate restriction enzyme, then ligate the adapter at both end of digested DNA with a bar coding (BC) region in adapter 1, following PCR amplification and sequencing. Finally, bioinformatic analysis of sequencing data is carried out and find out the SNPs (Fig. 2). Compared to other methods, GBS is a considerably less complicated, fragmentation and ligation of appropriate adapters are more straightforward, single-well digestion of genomic DNA, and fewer DNA purification steps make it easy. Moreover, GBS method avoids the separation step of fragments by size resulting in reduced sample handling and ultimately become cost effective. The low cost of GBS system makes it a powerful tool for SNP genotyping in a variety of crop species and populations as well as other plants. GBS has been shown as a valid tool for genomic diversity studies (Fu and Peterson 2011; Lu et al. 2013; Fu et al. 2014), which is already able to prove itself as an excellent system for SNP identification in plant breeding programs even in the absence of reference genome sequences or without any previous information about DNA polymorphism. Available reference genome makes easy to data analysis and identification of SNPs, but it is not essential in GBS system, which is a great advantage to plant breeders for crop improvement programs. Many reports already published about the use of GBS system for genetic

|                    |                                      |
|--------------------|--------------------------------------|
| Read 1             | ATGACGGTTG <b>T</b> ACAATTTGCGAC     |
| Read 2             | GACGGTTG <b>C</b> ACAATTTGCGACGT     |
| Read 3             | CGATGACGGTTG <b>T</b> ACAATTTGCGA    |
| Reference sequence | CGATGACGGTTG <b>C</b> ACAATTTGCGACGT |

**Fig. 2** Data analysis for SNP identification. Reads are aligned to reference sequence to find differences between the reference genome and newly sequenced genome. This concept is taken from Kumar et al. (2014) with modification

analysis, marker development and high throughput SNP genotyping of various crops such as rice, wheat, yellow mustard, rapeseed, lupin, lettuce, switchgrass, soybean, maize, etc. (Poland et al. 2012; Fu et al. 2014; Spindel et al. 2013; Truong et al. 2012; Lu et al. 2013; Sonah et al. 2013).

## SNP polymorphisms in different crop plants

### Rice

Rice is the main food for more than half of the world's population. The complete genome sequencing of rice in 2002 using bacterial artificial chromosomes (BAC) based approach was a major milestone for rice genomic research. In which genome size was 389 Mb, approximately three times larger than the model plant *Arabidopsis* and contains total of 37,544 non-transposable element related protein coding sequences (Yu et al. 2002; Goff et al. 2002; International Rice Genome 2005). After that a lot of genome resequencing data of rice are available that showed the high sequence diversity especially single nucleotide polymorphisms. For rice SNP genotyping, several high throughput array-based genotyping platforms have been developed which were considered critically important for dissecting phenotype-genotype associations in rice (McCouch et al. 2010; Tung et al. 2010; Zhao et al. 2011). Yu et al. (2014) identified more than four millions SNPs from around 500 rice landraces. Jeong et al. (2013) generated a total of  $1,165 \times 10^6$  raw reads and detected 1,154,063 DNA polymorphisms between the Korean rice accessions and Nipponbare. In average 0.64 SNP was found per one kb of Nipponbare genome, while Dongjin (Korean rice accession) genome contains a lower number of SNP (0.45 SNP/kb). Chen et al. (2014) resequenced 801 rice varieties and screened more than 10,000,000 SNP loci. Huang et al. (2009) analyzed and detected a total of 1,226,791 SNPs between indica cv. "9311" and japonica cv. "Nipponbare" that was average 3.2 SNPs/kb. Also, Parida et al. (2012) identified and validated SNPs in biotic and abiotic stress-responsive rice genes and determined the population structure in rice.

### Wheat

Wheat is one of the top three staple grains in the world, along with rice and maize whose genome size is around 17 Gb. The international wheat genome sequencing consortium revealed a chromosome-based draft genome sequence of hexaploid bread wheat in 2014 (The International Wheat Genome Sequencing Consortium 2014). The modern cultivated wheat also known as bread wheat (*Triticum aestivum* L.) with allohexaploid genome (AABBDD) is created by natural hybridization and spontaneous chromosome doubling between the tetraploid *Triticum turgidum* (AABB genome) and diploid *Aegilops tauschii* (DD genome) (Dubcovsky and Dvorak 2007). Allen et al. (2011) identified SNPs from different varieties of wheat and found 3251 SNPs in Cadenza, 2944 SNPs in Rialto, 3492 SNPs in Savannah and 3792 SNPs in Recital variety where Avalon variety was used as reference. Wang et al. (2014) discovered a total of 46,977 SNPs from eight mapping populations of wheat.

### Maize

Maize is the most produced cereal crop in the world which whole genome was first sequenced at 2009. The genome size of maize is 2.3 Gb with more than predicted 32,000 genes (Schnable et al. 2009). DNA sequence diversity in maize populations is more than human. Tenaillon et al. (2001) measured the sequence diversity in 21 loci distributed along chromosome 1 of maize. They sequenced from 25 inbred lines and data indicated that the maize has an average one SNP per 104 bases between two randomly sampled sequences that was higher than human or *Drosophila melanogaster*. Xu et al. (2014) identified SNPs from resequencing results of 15 inbred lines against B73 reference genome. A total of 6,385,011 SNPs were identified from 15 inbred lines. Chromosome 1 contains highest number of SNP (2,511,910) than other chromosomes of maize and that was 8.34 SNPs per Kb. Jones et al. (2009) obtained 1,088 loci from public sequencing data of 60 inbred lines and found total 9,194 SNPs that was average one SNP per 43 bases. Kumar et al. (2014) selectively amplified and sequenced four root genes (*Rtcl*, *Rth3*, *Rum1* and *Rul*) from 74 maize inbred lines and found DNA polymorphisms. They sequenced 2386 bases across four candidate genes involved in root development, resulting in 78 SNPs and SNP frequency was one per 31 bases. In another study, 383,145 SNPs were identified from 21 diverse inbred maize lines. These single nucleotide polymorphisms have the potential to broaden functional diversity and generate phenotypic variation in populations that may lead to new adaptations and

the modification of important agronomic traits (Muraya et al. 2015).

### Barley

The entire genome of barley was first sequenced at 2012 and the total genome size was around 5.1 Gb, containing 79,379 transcript clusters, including 26,159 high-confidence genes (Mayer et al. 2012). Xia et al. (2013) investigated SNPs in small heat shock protein 17.8 (*HSP17.8*) across 210 barley accessions and discovered eleven SNPs including 10 from the coding region which are deleterious for *HSP17.8* gene function. In another study, Clark et al. (2003) reported the effect of single nucleotide polymorphisms on the functional properties of the  $\beta$ -amylase of barley. They found three SNPs in coding region of  $\beta$ -amylase (*bmy1*) from a malt (Morex) and a feed (Stephoe) barley that caused differences in the amino acid sequences. Rostoks et al. (2005) identified SNPs by resequencing unigene fragments from eight diverse accessions of barley and observed the SNP frequency in 877 unigenes was 1 per 200 bases.

### Soybean

The reference genome sequence of soybean is available from 2010 which make it easy to identify the DNA polymorphisms among soybean populations. The genome size is approximately 1.1 Gb with 46,430 protein coding genes (Schmutz et al. 2010). Lee et al. (2015) identified more than four millions high quality SNPs by resequencing 16 soybean accessions. Chung et al. (2014) obtained 3,871,469 high quality SNPs by resequencing of 10 cultivated and 6 wild soybean accessions after mapping reads for each accession to the reference genome sequence. Genic regions contain 20.4% (788,809 SNPs) SNPs and rest of the SNPs were located in the intergenic regions. Jang et al. (2015) discovered a single nucleotide polymorphism in an *endo-1,4- $\beta$ -Glucanase* gene of soybean that altered the amino acid sequence and possibly reducing or eliminating its affinity for substrates in permeable cultivars. Vidal et al. (2012) found more than 6,000 SNPs in drought stress related genes from two contrasting cultivars of soybean, sensitive (BR 16) and tolerant (Embrapa 48). Among these SNPs, 165 are related to tolerance to abiotic stresses. Shi et al. (2015) identified three functional SNP in soybean (two for *Rhg1* locus and one for *Rhg4* locus) which are responsible for soybean cyst nematode resistance. In another study, Lam et al. (2010) re-sequenced a total of 17 wild and 14 cultivated soybean genomes and discovered a set of 205,614 tag SNPs that may be useful for QTL mapping and association studies.

They also concluded that the allelic diversity in wild soybeans is higher than cultivated soybeans. From another study, 209,903 SNPs were found from several soybean accessions. The average distance between adjacent SNPs was 4.5 kb (Song et al. 2013). Zhou et al. (2015) obtained 9,790,744 single nucleotide polymorphisms (SNPs) by resequencing 302 wild and cultivated soybean accessions after mapping against the soybean reference genome which make easy to identify the multiple loci and genes for important agronomic traits.

### Potato

Potato genome sequencing consortium first revealed the entire genome sequence of potato at 2011 that was 850 Mb in size. Hamilton et al. (2011) discovered 575,340 SNPs by sequencing normalized cDNA prepared from three commercial potato cultivars (Atlantic, Premier Russet, and Snowden). 230 SNPs were found in *Allene Oxide Synthase 2* gene of 184 tetraploid potato individuals which are associated with field resistance to late blight in populations of tetraploid potato cultivars (Pajeroska-Mukhtar et al. 2009). Uitdewilligen et al. (2013) sequenced 807 target genes from 83 tetraploid potato cultivars using genotyping by sequencing technique, and finally obtained 129,156 sequence variants where SNP density was 1 per 24 bases in exons and 1 per 15 bases in introns.

### Tomato

The complete genome of tomato has been sequenced and assembled by tomato genome consortium at 2012 which is enabling the identification of genome-wide SNPs and considered as a model for genomic research in *Solanaceae*, as well as for studying crop breeding (Tomato Genome Consortium 2012; Kim et al. 2014). The total genome size of cultivated tomato (*Solanum lycopersicum*) is approximately 950 Mb and a total of 34,727 protein coding genes in tomato genome were predicted by the international tomato annotation group (ITAG) (Tomato Genome Consortium 2012, <http://www.uk-sol.org>). According to Causse et al. (2013), genome sequencing of the eight tomato lines including *S. cerasiforme* and *S. lycopersicum* yielded a total of 4,290,679 unique SNPs when comparing each genome separately to the reference sequence. Total number of SNP varied widely from one line to another and also varied largely between different chromosomes. Chromosomes 4, 5, 7, 8, 9 and 11 contain highest number of SNPs and that is more than 350,000 SNPs per chromosome. On the other hand, chromosomes 1, 6, and 10 contain lowest number of SNPs (less than 150,000 unique SNPs per chromosome) (Causse et

al. 2013). Kim et al. (2014) discovered 4,680,647 putative SNPs from two accessions of *S. pimpinellifolium* by comparing with reference, of which 89.9% (4,210,454) were homo and 10.1% (470,193) were hetero-type SNPs. The total number of SNP and the density of SNP in different chromosomes also varied widely. An average 6.1 SNPs/kb was observed in the whole genome. In another study, around 1.5 million SNPs were identified from each resequencing data of six tomato lines by mapping onto tomato reference genome. They also identified nine SNP loci that were significantly associated with eight morphological traits (Shirasawa et al. 2013). The 100 tomato genome sequencing consortium (2014) reported that the SNP frequency in tomato genome is significantly higher in intergenic regions ( $89.47 \pm 3.03\%$ ) than in genic regions ( $7.55 \pm 2.19\%$  in introns and  $2.33 \pm 0.68\%$  in exons) for all accessions of cultivated tomato (*Solanum lycopersicum*). Also, many studies are reported about the SNP discovery in tomato and the role of these SNPs in gene function and development of agronomic traits (Shirasawa et al. 2016; Shirasawa et al. 2013; Hirakawa et al. 2013; Hamilton et al. 2012).

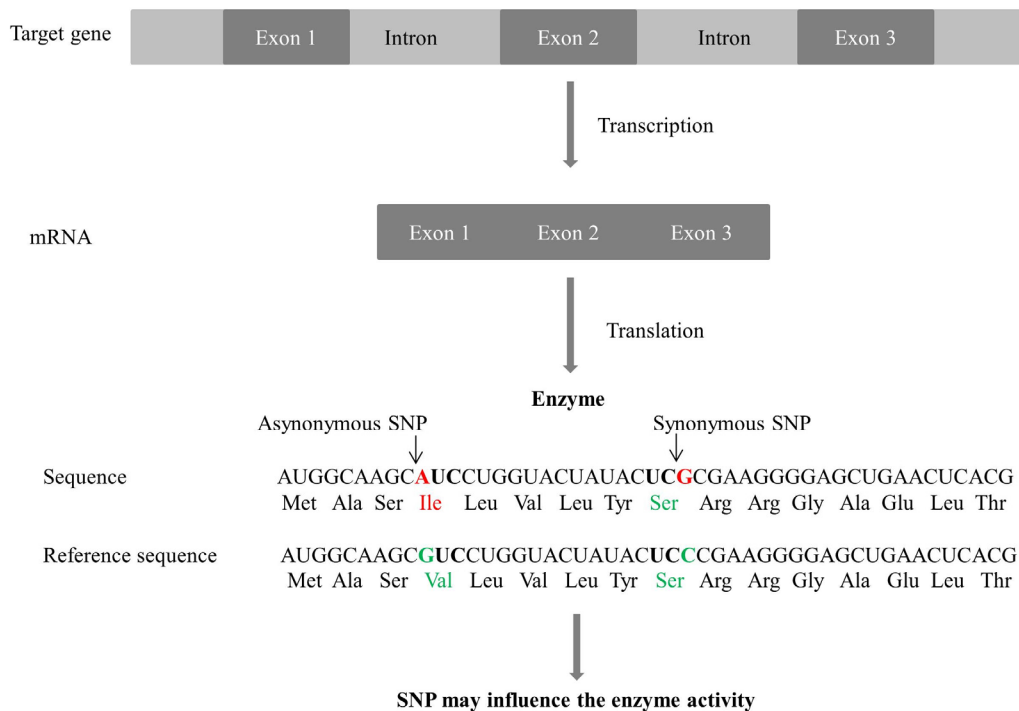
#### Other crops

There are so many other crop plants whose full genome sequence have been completed such as grape (Velasco et al.

2007), cucumber (Huang et al. 2009), apple (Velasco et al. 2010), banana (Hont et al. 2012), oil palm (Singh et al. 2013), eggplant (Hirakawa et al. 2014) etc. These reference genome sequences help the plant breeders to discover SNP among different cultivars or breeding lines which facilitate the development and selection of improved crop varieties.

#### Effect of SNPs on gene function

Single Nucleotide Polymorphism may influence the promoter activity for gene expression, transcriptional and translational efficiency (LeVan et al. 2001). Therefore, they may be responsible for phenotypic variations among individuals for improving of agronomical traits. A gene contains two parts, exon and intron. Intron is removed during post transcriptional modification but the exons are finally translated into amino acid sequence and produce enzyme. So, the SNP in the exon part (coding region) is most important because they can affect the gene function. SNPs in the coding region are of two types, synonymous and asynonymous SNPs. Synonymous SNPs do not affect the amino acid sequence but asynonymous SNPs change the amino acid sequence of protein and may influence the enzyme activity (Fig. 3). There are many reports about the effect of SNP on gene function in different crop plants. One study conducted by Schreiber et al. (2014) and identified SNPs in plastidic starch phosphorylase *PHO1* gene of potato



**Fig. 3** A schematic representation of the role of SNP in gene function that can influence enzyme activity by changing amino acids. Met, Methionine; Ala, Alanine; Ser, Serine; Ile, Isoleucine; Leu, Leucine, Val, Valine; Tyr, Tyrosine; Arg, Arginine; Gly, Glycine; Glu, Glutamic acid and Thr, Threonine. This concept is taken from Jang et al. (2015) with modification

that changed some amino acids. This change might cause the reduced enzyme activity and decreases the starch breakdown. The ultimate result of that is to increase starch and decrease sugar in tuber of potato. Potato with decreased reducing sugars has positive effects on the quality of processed products such as chips, French fries etc. Fridman et al. (2004) discovered an SNP in an invertase gene of tomato that changed an amino acid near the catalytic site of the invertase crystal, affecting enzyme activity, which was responsible for sugar yield. In rice, 66 functional SNPs were discovered in exonic regions from 18 genes involved in starch synthesis. A novel SNP was reported in Glucose-6-Phosphate Translocator 1 (*GPT1*) gene at the position 1188 of *GPT1* gene that alters amino acid associated with amylose and resistant-retrograded starch content (Kharabian-Masouleh et al. 2012). Clark et al. (2003) examined the effect of SNP on the activity of  $\beta$ -amylase (*bmy1*) gene from barley and found that SNPs have high effect on the enzyme activity. They found only three SNPs between Morex and Steptoe that altered the three amino acids at the positions 115, 165, and 430, which made 67% more *bmy1* activity of Steptoe than the *bmy1* from Morex. From this study, they concluded that SNPs in the coding region of *bmy1* gene greatly affect the activity of  $\beta$ -amylase enzyme. Kumar et al. (2014) identified SNPs from several genes *Rtcl* (rootless concerning crown and seminal roots like protein), *Rth3* (roothairless 3), *Rum1* (rootless with undetectable meristems 1), and *Rul1* (*Rum1*-like 1) that involved in maize root development and observed that these polymorphisms were significantly associated with seedling root traits in maize and suggested that the SNPs existing in the examined genes can be used to improve the quality of maize root. Jang et al. (2015) developed a new cultivar of soybean, Tachinagaha through the insertion of a quantitative trait locus, *qHS1* that is responsible for hard seed from the impermeable (hard-seeded), wild soybean (*G. soja*) into the permeable cultivar Kariyutaka. The seed coat of resulted new cultivar Tachinagaha was more rigid than its parent cultivar due to increasing amount of  $\beta$ -1,4-glucans in the outer layer of the seed coat. The *qHS1* locus encoded an endo-1,4- $\beta$ -glucanase and sequencing results revealed one SNP in endo-1,4- $\beta$ -glucanase gene that altered an amino acid, effecting on enzyme activity and increasing the amount of  $\beta$ -1,4-glucans, resulting rigid, impermeable seed in new cultivar. 165 functional SNPs were identified from 127 abiotic stress related genes of soybean (*Glycine max*) and established that these SNPs play an important role in tolerance to drought stress (Vidal et al. 2012). In another study, Xia et al. (2013) discovered four functional SNPs in *HSP17.8* gene of different barley accessions that control some agronomic traits in barley. Pajeroska-Mukhtar et al. (2009)

identified SNP polymorphisms in *StAOS2* (*Allene Oxide Synthase 2*) gene, encodes an enzyme involve in the defense signaling pathway in potato. Among these SNPs, two SNPs at the *StAOS2* locus, *StAOS2\_snp691* and *StAOS2\_snp692* are involved with increased tolerance to late blight disease of potato. Hirakawa et al. (2013) examined the genome-wide SNP in tomato and their effect on gene functions. They found that when the SNPs are located on the functional sites of candidate genes could directly affect the gene expressions and protein functions, also might be related with phenotypic differences among tomato lines. Parida et al. (2012) investigated SNPs in stress-responsive rice genes and assessed the functional and adaptive significance of the validated SNPs in biotic and abiotic stress tolerance in rice.

## Conclusion

As SNPs can change the amino acid that might affect the enzyme activity, so the study of functional SNPs is very important regarding crop improvements. It is important to know the location of SNP in the genome because if the SNP is present in the coding region can highly affect the activity and thermostability level of the enzyme. Sometimes it is also depends on the substituted amino acid positions because some amino acid controls the activity of enzyme. Recent technological advances make it easy to find out functional SNP from various breeding lines which could be used for crop improvements. The success stories indicate that SNPs in the functional parts of the gene may control the level of biotic and abiotic stresses and may develop various abiotic and biotic stress tolerance crop varieties through modifying enzyme activity.

## Acknowledgements

This research was supported by Golden Seed Project (Center for Horticultural Seed Development, No. 213003-04-4-SBC10), by a research grant of the iPET, Ministry of Food, Agriculture, Forestry and Fisheries, Republic of Korea

## References

- Akond M, Liu S, Schoener L, Anderson JA, Kantartzi SK, Meksem K, Song Q, Wang D, Wen Z, Lightfoot DA, Kassem MA (2013) A SNP-based genetic linkage map of soybean using the SoySNP6K illumina Infinium beadchip genotyping array. *Journal of Plant Genome Sciences* 1:80-89
- Alexandra M, Allen AM, Barker GLA, Berry ST, Coghill JA,



- Gwilliam R, Kirby S, Robinson P, Brenchley RC, D'Amore R, McKenzie N, Waite D, Hall A, Bevan M, Hall N, Edwards KJ (2011) Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant Biotechnol J* 9:1086-1099
- Alkan C, Coe BP, Eichler EE (2011) Genome structural variation discovery and genotyping. *Nat Rev Genet* 12:363-376
- Allen AM, Barker GL, Berry ST, Coghill JA, Gwilliam R, Kirby S, Robinson P, Brenchley RC, D'Amore R, McKenzie N, Waite D, Hall A, Bevan M, Hall N, Edwards KJ (2011) Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant Biotechnol J* 9:1086-1099
- Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C, Toomajian C, Traw B, Zheng H, Bergelson J, Dean C, Marjoram P, Nordborg M (2005) Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. *PLoS Genet* 1:e60
- Barchi L, Lanteri S, Portis E, Acquadro A, Valè G, Toppino L, Rotino GL (2011) Identification of SNP and SSR markers in eggplant using RAD tag sequencing. *BMC Genomics* 12:304
- Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC (2009) The genetic architecture of maize flowering time. *Science* 325:714-718
- Byers RL, Harker DB, Yourstone SM, Maughan PJ, Udall JA (2012) Development and mapping of SNP assays in allotetraploid cotton. *Theor Appl Genet* 124:1201-1214
- Causse M, Desplat N, Pascual L, Paslier MCL, Sauvage C, Bauchet G, Bérard A, Bounon R, Tchoumakov M, Brunel D, Bouchet JP (2013) Whole genome resequencing in tomato reveals variation associated with introgression and breeding events. *BMC Genomics* 14:791
- Chen H, Xie W, He H, Yu H, Chen W, Li J, Yu R, Yao Y, Zhang W, He Y, Tang X, Zhou F, Deng XW, Zhang Q (2014) A high-density SNP genotyping array for rice biology and molecular breeding. *Mol Plant* 7:541-553
- Ching A, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski AJ (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet* 3:19
- Chung WH, Jeong N, Kim J, Lee WK, Lee YG, Lee SH, Yoon WC, Kim JH, Choi IY, Choi HK, Moon JK, Kim N, Jeong SC (2014) Population structure and domestication revealed by high-depth resequencing of Korean cultivated and wild soybean genomes. *DNA Research* 21:153-167
- Clark SE, Hayes PM, Henson CA (2003) Effects of single nucleotide polymorphisms in  $\beta$ -amylase I alleles from barley on functional properties of the enzymes. *Plant Physiol Bioch* 41:798-804
- Cortés AJ, Chavarro MC, Blair MW (2011) SNP marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 123:827-845
- Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* 12:499-510
- Dubcovsky, J. and Dvorak, J. (2007) Genome plasticity: a key factor in the success of polyploid wheat under domestication. *Science*, 316, 1862-1866.
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379
- Fridman E, Carrari F, Liu YS, Fernie AR, Zamir D (2004) Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* 305:1786-9
- Fu YB, Cheng B, Peterson GW (2014) Genetic diversity analysis of yellow mustard (*Sinapis alba* L.) germplasm based on genotyping by sequencing. *Genet Resour Crop Evol* 61:579-594
- Fu YB, Peterson GW (2011) Genetic diversity analysis with 454 pyrosequencing and genomic reduction confirmed the eastern and western division in the cultivated barley gene pool. *Plant Genome* 4:226-237
- Ganal MW, Altmann T, Röder MS (2009) SNP identification in crop plants. *Curr Opin Plant Biol* 12:211-217
- Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science* 296:92-100
- Gut IG (2001) Automation in genotyping of single nucleotide polymorphisms. *Hum Mutat* 17:475-92
- Hamilton JP, Hansey CN, Whitty BR, Stoffel K, Massa AN, Deynze AV, Jong WSD, Douches DS, Buell CR (2011) Single nucleotide polymorphism discovery in elite north american potato germplasm. *BMC Genomics* 12:302
- Hamilton JP, Sim SC, Stoffel K, Deynze AV, Buell CR, Francis DM (2012) Single nucleotide polymorphism discovery in cultivated tomato via sequencing by synthesis. *Plant Genome* 5:17-29
- Hirakawa H, Shirasawa K, Miyatake K, Nunome T, Negoro S, Ohyama A, Yamaguchi H, Sato S, Isobe S, Tabata S, Fukuoka H (2014) Draft genome sequence of eggplant (*Solanum melongena* L.): the representative Solanum species indigenous to the old world. *DNA Res* 2:649-660
- Hirakawa H, Shirasawa K, Ohyama A, Fukuoka H, Aoki K, Rothan C, Sato S, Isobe S, Tabata S (2013) Genome-wide SNP genotyping to infer the effects on gene functions in tomato. *DNA Res* 20:221-233
- Hont AD, Denoeud F, Aury JM, Baurens FC, Carreel F, Garsmeur O, Noel B, Bocs S, Droc G, Rouard M (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488:213-217
- Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas WJ, Wang X, Xie B, Ni P (2009) The genome of the cucumber, *Cucumis sativus* L. *Nat Genet* 41:1275-1281
- Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D, Weng Q, Huang T, Dong G, Sang T, Han B (2009) High-throughput genotyping by whole-genome resequencing.



- Genome Res 19:1068-1076
- Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet* 42:961-967
- Huq MA, Akter S, Jung YJ, Nou IS, Cho YG, Kang KK (2016) Genome sequencing, a milestone for genomic research and plant breeding. *Plant Breeding and Biotechnology* 4:29-39
- Hyten DL, Cannon SB, Song Q, Weeks N, Fickus EW, Shoemaker RC, Specht JE, Farmer AD, May GD, Cregan PB (2010). High-throughput SNP discovery through deep resequencing of a reduced representation library to anchor and orient scaffolds in the soybean whole genome sequence. *BMC Genomics* 11:38
- International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature* 436:793-800
- Jang SJ, Sato MI, Sato K, Jitsuyama Y, Fujino K, Mori H, Takahashi R, Benitez ER, Liu B, Yamada T, Abe J (2015) A single-nucleotide polymorphism in an *endo-1,4-β-glucanase* gene controls seed coat permeability in soybean. *Plos One* 10:e0128527
- Jeong IS, Yoon UH, Lee GS, Ji HS, Lee HJ, Han CD, Hahn JH, An G, Kim TH (2013) SNP-based analysis of genetic diversity in anther derived rice by whole genome sequencing. *Rice* 6:6 doi: 10.1186/1939-8433-6-6
- Jones E, Chu WC, Ayele M, Ho J, Bruggeman E, Yourstone K, Rafalski A, Smith OS, McMullen MD, Bezawada C, Warren J, Babayev J, Basu S, Smith S (2009) Development of single nucleotide polymorphism (SNP) markers for use in commercial maize (*Zea mays* L.) germplasm. *Mol Breeding* 24:165-176
- Kharabian-Masouleh A, Waters DLE, Reinke RF, Ward R, Henry RJ (2012) SNP in starch biosynthesis genes associated with nutritional and functional properties of rice. *Sci Rep* 2:557
- Kim JE, Oh SK, Lee JH, Lee BM, Jo SH (2014) Genome-wide SNP calling using next generation sequencing data in tomato. *Mol Cells* 37:36-42
- Kircher M, Kelso J (2010) High-throughput DNA sequencing—concepts and limitations. *BioEssays* 32:524-536
- Kruglyak L (1997) The use of a genetic map of biallelic markers in linkage studies. *Nat Genet* 17:21-24
- Kumar B, Abdel-Ghani AH, Pace J, Reyes-Matamoros J, Hochholdinger F, Lübberstedt T (2014) Association analysis of single nucleotide polymorphisms in candidate genes with root traits in maize (*Zea mays* L.) seedlings. *Plant Science* 224:9-19
- Kumar S, Banks TW, Cloutier S (2012) SNP discovery through next-generation sequencing and its applications. *Int J Plant Genomics* 2012:831460 doi:10.1155/2012/831460
- Lam HM, Xu X, Liu X, Chen W, Yang G, Wong FL, Li MW, He W, Qin N, Wang B, Li J, Jian M, Wang J, Shao G, Wang J, Sun SS, Zhang G (2010) Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat Genet* 42:1053-1059
- Lee YG, Jeong N, Kim JH, Lee K, Kim KH, Pirani A, Ha BK, Kang ST, Park BS, Moon JK, Kim N, Jeong SC (2015) Development, validation and genetic analysis of a large soybean SNP genotyping array. *Plant J* 81:625-636
- LeVan TD, Bloom JW, Bailey TJ, Karp CL, Halonen M, Martinez FD, Vercelli D (2001) A common single nucleotide polymorphism in the CD 14 promoter decreases the affinity of Sp protein binding and enhances transcriptional activity. *J Immunol* 167:5838-5844
- Li F, Kitashiba H, Inaba K, Nishio T (2009) A *Brassica rapa* linkage map of EST-based SNP markers for identification of candidate genes controlling flowering time and leaf morphological traits. *DNA Res* 16:311-323
- Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, Buckler ES, Costich DE (2013) Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genet* 9:e1003215
- Mardis ER (2007) The impact of next-generation sequencing technology on genetics. *Trends Genet* 24:133-141
- Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K, Close TJ, Wise RP, Stein N (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491:711-716
- McCouch SR, Zhao K, Wright M, Tung CW, Ebana K, Thomson M, Reynolds A, Wang D, DeClerck G, Ali ML (2010) Development of genome-wide SNP assays for rice. *Breed Sci* 60:524-535
- McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulata VJ, Zeller G, Clark RM, Hoen DR, Bureau TE (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc Natl Acad Sci USA* 106:12273-12278
- Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution* 19: 208-216
- Muraya MM, Schmutzer T, Ulpinnis C, Scholz U, Altmann T (2015) Targeted sequencing reveals large-scale sequence polymorphism in maize candidate genes for biomass production and composition. *PLoS One* 10:e0132120
- Oliver RE, Lazo GR, Lutz JD, Rubenfield MJ, Tinker NA, Anderson JM, Wisniewski Morehead NH, Adhikary D, Jellen EN (2011) Model SNP development for complex genomes based on hexaploid oat using high-throughput 454 sequencing technology. *BMC Genomics* 12:77
- Pajeroska-Mukhtar K, Stich B, Achenbach U, Ballvora A, Lübeck J, Strahwald J, Tacke E, Hofferbert HR, Ilarionova E, Bellin D, Walkemeier B, Basekow R, Kersten B, Gebhardt C (2009) Single nucleotide polymorphisms in the allene oxide synthase 2 gene are associated with field resistance to late blight in populations of tetraploid potato cultivars. *Genetics* 181:1115-1127
- Parida SK, Mukerji M, Singh AK, Singh NK, Mohapatra T (2012) SNPs in stress-responsive rice genes: validation, genotyping, functional relevance and population structure. *BMC Genomics* 13:426
- Pasam RK, Sharma R, Malosetti M, Eeuwijk FAV, Haseneyer G, Kilian B, Graner A (2012) Genome-wide association studies for agronomical traits in a world-wide spring barley collection. *BMC Plant Biology* 12:16

- Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S, Crossa J, Sanchez-Villeda H, Sorrells M, Jannink JL (2012) Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5:103-113
- Poland JA, Bradbury PJ, Buckler ES, Nelson RJ (2011) Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc Natl Acad Sci USA* 108:6893-6898
- Poland JA, Rife TW (2012) Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* 5:92-102
- Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. *Nature* 475:189-195
- Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. *Curr Opin Plant Biol* 5:94-100
- Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson JT, Wanamaker SI, Walia H, Rodriguez EM, Hedley PE, Liu H, Morris J, Close TJ, Marshall DF, Waugh R (2005) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Mol Genet Genomics* 274:515-27
- Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409:928-933
- Sanger F, Nicklen S, Coulson AR (1977a) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463-5467
- Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes CA, Hutchison CA, Slocumbe PM, Smith M (1977b) Nucleotide sequence of bacteriophage phiX174 DNA. *Nature* 265:687-695
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178-183
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112-1115
- Schreiber L, Nader-Nieto AC, Schönhals EM, Walkemeier B, Gebhardt C (2014) SNPs in genes functional in starch-sugar interconversion associate with natural variation of tuber starch and sugar content of potato (*Solanum tuberosum* L.). *G3 (Bethesda)* 4:1797-1811
- Shi Z, Liu S, Noe J, Arelli P, Meksem K, Li Z (2015) SNP identification and marker assay development for high-throughput selection of soybean cyst nematode resistance. *BMC Genomics* 16:314
- Shirasawa K, Fukuoka H, Matsunaga H, Kobayashi Y, Kobayashi I, Hirakawa H, Isobe S, Tabata S (2013) Genome-wide association studies using single nucleotide polymorphism markers developed by re-sequencing of the genomes of cultivated tomato. *DNA Res* 20:593-603
- Shirasawa K, Hirakawa H, Nunome T, Tabata S, Isobe S (2016) Genome-wide survey of artificial mutations induced by ethyl methanesulfonate and gamma rays in tomato. *Plant Biotechnol J* 14:51-60
- Singh R, Ong-Abdullah M, Low ETL, Manaf MAA, Rosli R, Nookiah R (2013) Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. *Nature* 500:335-339
- Sonah H, Bastien M, Iquiria E, Tardivel A, Légaré G, Boyle B, Normandeau É, Laroche J, Larose S, Jean M, Belzile F (2013) An improved genotyping by sequencing (GBS) approach offering increased ver-satility and efficiency of SNP discovery and genotyping. *PLoS One* 8:e54603
- Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, Cregan PB (2013) Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PLoS One* 8:e54985
- Spindel J, Wright M, Chen C, Cobb J, Gage J, Harrington S, Lorieux M, Ahmadi N, McCouch S (2013) Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theor Appl Genet* 126:2699-2716
- Steemers FJ, Gunderson KL (2007) Whole genome genotyping technologies on the bead array platform. *Biotechnol J* 2:41-49
- Sun R, Chang Y, Yang F, Wang Y, Li H, Zhao Y, Chen D, Wu T, Zhang X, Han Z (2015) A dense SNP genetic map constructed using restriction site-associated DNA sequencing enables detection of QTLs controlling apple fruit quality. *BMC Genomics* 16:747
- Tenaillon MI, Sawkins MC, Long AD, Gaut RL, Doebley JF, Gaut BS (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proc Natl Acad Sci USA* 98:9161-9166
- The 100 Tomato Genome Sequencing Consortium (2014) Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *Plant J* 80:136-148
- Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485:635-641
- Truong HT, Ramos AM, Yalcin F, de Ruiter M, van der Poel HJ, Huvenaars KH, Hogers RC, van Enckevort LJ, Janssen A, van Orsouw NJ, van Eijk MJ (2012) Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. *PLoS One* 7:e37565
- Tung CW, Zhao K, Wright MK, Ali ML, Jung J, Kimball J, Tyagi W, Thomson MJ, McNally K, Leung H (2010) Development of a research platform for dissecting phenotype-genotype associations in rice (*Oryza* spp.). *Rice* 3:205-217
- Uitdewilligen JG, Wolters AM, D'hoop BB, Borm TJ, Visser RG, van Eck HJ (2013) A next-generation sequencing method for genotyping by-sequencing of highly heterozygous autotetraploid potato. *PLoS ONE* 8:e62355
- Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D (2010) The genome of the domesticated apple (*Malus x domestica* Borkh.). *Nat Genet* 42:833-839
- Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, FitzGerald LM, Vezzulli S, Reid J (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS One* 2:e1326

- Vidal RO, Nascimento LC, Mondego JMC, Pereira GAG, Carazzolle MF (2012) Identification of SNPs in RNA-seq data of two cultivars of *Glycine max* (soybean) differing in drought resistance. *Genet Mol Biol* 35:331-334
- Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Sel Evol* 34:275-305
- Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L (2014) Characterization of polyploid wheat genomic diversity using a high-density 90000 single nucleotide polymorphism array. *Plant Biotechnol J* 12:787-796
- Waugh R, Jannink JL, Muehlbauer GJ, Ramsay L (2009) The emergence of whole genome association scans in barley. *Curr Opin Plant Biol* 12:218-222
- Wei O, Peng Z, Zhou Y, Yang Z, Wu K, Ouyang Z (2011) Nucleotide diversity and molecular evolution of the WAG-2 gene in common wheat (*Triticum aestivum* L.) and its relatives. *Genet Mol Biol* 34:606-615
- Wei Q, Wang Y, Qin X, Zhang Y, Zhang Z, Wang J, Li J, Lou Q, Chen J (2014) An SNP-based saturated genetic map and QTL analysis of fruit-related traits in cucumber using specific-length amplified fragment (SLAF) sequencing. *BMC Genomics* 15:1158
- Xia Y, Li R, Ning Z, Bai G, Siddique KH, Yan G, Baum M, Varshney RK, Guo P (2013) Single nucleotide polymorphisms in HSP17.8 and their association with agronomic traits in barley. *PLoS One* 8:e56816
- Xie W, Feng Q, Yu H, Huang X, Zhao Q, Xing Y, Yu S, Han B, Zhang Q (2010) Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing. *Proc Natl Acad Sci USA* 107:10578-10583
- Xu J, Yuan Y, Xu Y, Zhang G, Guo X, Wu F, Wang Q, Rong T, Pan G, Cao M, Tang Q, Gao S, Liu Y, Wang J, Lan H, Lu Y (2014) Identification of candidate genes for drought tolerance by whole-genome resequencing in maize. *BMC Plant Biology* 14:83
- Yu H, Xie W, Li J, Zhou F, Zhang Q (2014) A whole-genome SNP array (RICE6K) for genomic breeding in rice. *Plant Biotechnol J* 12:28-37
- Yu J, Hu S, Wang J, Wong GKS, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science* 296:79-92
- Zhang X, Borevitz JO (2009) Global analysis of allele-specific expression in *Arabidopsis thaliana*. *Genetics* 182:943-954
- Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung AM, Bustamante CD, McCouch SR (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat Commun* 2:467
- Zhou Z, Jiang Y, Wang Z, Gou Z, Lyu J, Li W, Yu Y, Shu L, Zhao Y, Ma Y (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nature Biotechnology* 33:408-414