The Use of SNP Genotyping for QTL/ Candidate Gene Discovery in Plants

Salman F. Alamery
Biochemistry Department, College of Science, King Saud University,
Riyadh, P.O.Box 2455-11451, Saudi Arabia.

Abstract

Many plant species have been sequenced, which has enabled the identification of DNA sequence variations such as novel single-nucleotide polymorphisms (SNPs). SNP-based genotyping technologies have been developed over the last decade and play important roles in detecting SNPs in a cost and time effective manner, thereby facilitating plant genetic and genomic studies. The availability of genome and genetics resources and increase in sequencing data has led to SNPs discovery and array development using various approaches in many plant species. The development of SNP arrays for genotyping requires three steps: identification of a large number of SNPs, validation, and final selection. The discovery of SNPs and array development at various densities have been successfully developed and utilized in many diploid and polyploid plant species. SNP genotyping arrays have been useful for many applications such as cloning, association mapping, analyses of species diversity, evolution, genomic selection, and comparative genomics. The assaying of large segregating or mutant populations with an SNP array allows the accurate, efficient, and rapid determination of the genotypes of many individuals and genetic variations. Genotyping assays are an efficient way of generating a vast amount of genotypic data in crop species to produce highly informative SNPs for marker-assisted selection in breeding programs as well as functional genomic studies. This review highlights the SNP genotyping in the application of identification of SNPs in the discovery of quantitative trait locus (QTL) or candidate genes for important crop traits and dissection of other complex traits.

Keywords:
Single-Nucleotide Polymorphisms; Genotyping; Quantitative Trait Locus;
Candidate Genes; Crop Plants

1. Introduction

Genotyping is the process of determining the differences in DNA sequence between individuals. It involves the generation of allele-specific products and identification of the status of genotypes. Genotyping has become extremely important in association studies to determine genes or sequence variants linked to specific traits or diseases. It is also essential for gene mapping, analyses of species diversity and evolution, and marker-assisted selection [1]. Genotyping studies can be designed to identify DNA sequence differences at the single-nucleotide polymorphism (SNP) level. Hence, SNP genotyping is now widely applied in plant research such as candidate genes discovery, quantitative trait locus (QTL) analysis, linkage mapping and compara-

tive genomics [2].

A single nucleotide polymorphism (SNP) is defined as single base-pair change at specific position in the genome and the most common genetic variations in plant genomes. SNP allelic variations in a population are often associated with a particular trait phenotype [3]. Therefore, SNPs have potential to be used as markers for genetic studies in functional genomics and breeding [4].

It has been reported that the frequency of SNPs in plant species is about 1 SNP in every 100–300 base pairs [5]. For instance, 37,000 SNPs have been identified in Arabidopsis Columbia (Col) and Landsberg erecta (Ler) [6]. This provides powerful tool for genotyping assays enabling these SNPs to be associated with of economically important traits [6].

The development of genomic and bioinformatic databases in plant species has dramatically increased the identification of SNP polymorphisms and scoring of the variation in specific targets. More importantly, a large number of potential SNPs and their surrounding sequences would provide a foundation for highly multiplexed automated genotyping analysis [7]. The availability of genomic resources permits the identification of SNPs and their association with particular phenotype in which it can be employed for marker-assisted selection and candidate gene approach [8, 9].

2.SNP array development and selection

SNP arrays called “SNP chips” are a type of DNA microarray designed with a large number of SNPs (up to one million). This approach involves high-throughput genome scan providing a time and cost-efficient tool for genotyping individuals at millions of different positions across the genome [10]. These arrays have become commonly used in plant genetic studies such as gene mapping, genome-wide association, genomic selection, etc. [11].

The completion of human genome sequencing and genotyping arrays were initially developed in human before other organisms. The SNP Genotyping has been a powerful tool for comprehensive genome-wide association studies in human [1]. It also has significant impact on the genetic analysis of human disease and cancer. This facilitated understanding complex diseases and further characterized the human genome [2]. In addition, SNP genotyping was also played a role in animal genetic and breading studies such as in genome selection and identification of disease resistance genes [3]. However, the genotyping array is still limited to certain plant species and challenging due to the genome nature and size. SNP identification and calling is more difficult and complex. Even though the level of diversity and functionality in plant species compared to animal and human, there is still ongoing works on adopting this technology for functional genomics studies [4].

Various approaches have been used to

identify SNP genotyping arrays in many plant species depending on the availability of genome sequences and genomic/ genetic resources. The primary method is the direct sequencing of DNA fragments. This is the most reliable approach and has been proven to successfully detect SNPs in many organisms [12]. It is appropriate when sequence information is limited or when polymorphisms of interest in a specific genotype or candidate gene are being investigated [13]. Direct sequencing can be used as a validation tool to differentiate true polymorphisms from sequencing errors. This requires the sequencing of enough individuals to achieve sufficient the accuracy and coverage to distinguish real SNPs from sequencing artifacts [7, 14]. However, this is generally costly and time-consuming and is impractical for large-scale genetic studies.

The mining of expressed sequence tags (ESTs) have also been used for identifying SNPs. ESTs are short sequences of cDNA. The number of identified ESTs range from tens of thousands for species that have been little investigated to over a million in well studied plant species [15]. For instance, in Arabidopsis, 1.5 million ESTs have been identified, whereas in Brassica species; namely B. napus, B. rapa and B. oleracea 643,944, 213,605 and 179,213 have been identified, respectively. The EST sequences provide sufficient redundancy for screening of the presence of polymorphisms using bioinformatic analysis tools, allowing subsequent identification of SNPs polymorphism [16]. This approach mainly provides limited SNPs located only in transcribed regions (coding and UTR regions), resulting in a small number of SNPs, but it offers a low-cost source of informative and abundant SNPs and high quality sequence data.

RNA sequencing (RNA-seq) is another cost effective approach for SNPs discovery within coding regions of the genome[17]. It can be used to identify SNPs in the transcriptome and help to reduce genome complexity. Geraldes, Pang [18] were able to identify over 0.5 million putative SNPs in 26,595 genes in Populus trichocarpa. RNA-seq has been found to provide more accurate functional annotation due to the enrichment of expressed genes, but, it has not been widely used due to its cost and laboriousness [19].

The most comprehensive approach for discovering SNPs is next-generation sequencing (NGS). It provides cheap and reliable large-scale SNPs identification. SNPs discovery by NGS is not limited to protein-coding sequences leading to the identification of SNPs widely distributed across the genome and enabling determination of the correlation between phenotypes and SNPs in non-coding regions. This has enabled the discovery of thousands of SNPs in closely related species [20]. NGS-derived SNPs have been reported in many plant species [21-23]. SNPs identification by NGS requires a complete genome sequence as a reference, although it can be achieved without a full reference sequence using de novo read assembly mapping [23,24] . The genome sequences of a number of plant

species with relatively small genome are available as suitable reference genomes [25]. Genotyping-by-sequencing (GBS) has newly emerged as approach for sequence-based genotyping. The strength of this approach is that sequence polymorphism [mainly SNPs] and genotyping are completed at the same time. This approach enables the search for SNPs and presence/absence variations in diverse species with and without reference genomes [26, 27].

This approach has been demonstrated to be robust across a range of species and producing high number of molecular markers [28]. However, it is limited in terms of the number of individuals for whom NGS data are available, requires bioinformatic analysis, and is technically challenging [4].

With the advancement of sequencing technologies and SNP discovery approaches, computational tools and databases for SNP markers have been developed such as AutoSNPdb [29], TreeSNPs [30], dbSNP [31], and NABIC SNP [32]. Such databases provide more detail about SNP sequences, names, gene definitions, locations, functions, and associations among others.

The development of SNP arrays requires three steps: identification of a large number of SNPs, validation, and final selection. The discovery of SNPs aims to identify as many SNPs as possible. This requires background of bioinformatics analyses, including sequence or reads mapping or alignment, reads trimming, SNPs calling, and SNPs filtering [11, 33]. SNP arrays can also include SNPs developed from specific gene sequences "candidate genes of interest" selected from databases. Various SNP analyses can yield different numbers of variants among species [34]. Therefore, SNP validation is essential to select true SNPs polymorphisms and to decrease the false positive SNPs.

Given that sequencing data remains prone to inaccuracies at rate as high as one error every 100 base pairs [35]. SNPs selections can be significantly improved by increase the stringency of SNPs detection and meeting the requirement of accurate SNP calling [36]. A number of important filtering criteria should be considered including the following:

(1) Information of the SNP flanking sequences is a key element. This refers to sequence length on either side of the SNP, for instance, 50 bp for Illumina. There should be no SNPs present within 50 bp on either side of the SNP.

(2) Sequence depth over the entire mapping assembly. A minimum sequence /read depth per individual/genotype represents less sequencing error at SNPs.

(3) The allele depth defined as a number of different sequences in which an allele appeared for both variant and reference alleles. SNPs were selected where the minor allele was present in more than one genotype. SNPs present in less than two genotypes are excluded.

(4) The general considerations for array SNP selection include the presence of repetitive or palindromic sequences, GC content, SNP depth, SNP types and SNP frequency.

(5) Location and distribution of SNPs

throughout the genome and their genetic effects.

(6) Minimum quality score measuring the probability that a base is called incorrectly.

(7) Minimum coverage of reads, supporting the presence of the allele in a given SNP.

(8) Maximum coverage of reads, which should be less than the average read depth of all SNPs.

(9) Average copy number of the SNP flanking region, which should be less than two.

(10) Sequences with more than 4 SNPs per 100 bp should be removed to avoid selecting any SNPs located in hypervariable regions.

Furthermore, during selection, SNPs can be classified according to their physical position on the chromosomes and best Blast hits. According to their relative location, SNPs are classified into different categories because they are part of coding sequences, within exons, introns, between exons or within promoter sequences. This would help to integrate the information with existing genetic mapping for further genetic studies.

The majority of SNPs are most likely to be false positive, leading to SNPs being undetermined or incorrectly genotyped [37]. filtering criteria for SNP selection and reduction of genomic complexity would increase the SNP genotyping efficiency by achieving a higher proportion of correctly genotyped SNPs relative to incorrectly genotyped ones [37].

Approaches for SNP identification involving the discovery and validation of predicted SNP polymorphisms have been optimized and developed for simple genomes in diploid plant species [13, 38]. However, SNP identification in polyploid crops with complex genomes has remained very challenging. The higher levels of genome complexity and presence of polymorphisms between subgenomes in polyploid crops constitute major additional challenges for SNP prediction [2].

Polyploidy is extremely common in the plant kingdom, particularly, in important crops such as potato
(Solanum tuberosum L.), alfalfa
(Medicago sativa L.), durum wheat
(Triticum durum Desf.), cotton
(Gossypium hirsutum L.) and canola
(Brassica napus). The major concern with plant species is the complexity of their genomes, due to highly repetitive sequence, homologous genome and intragenomic duplication. These factors influence the efficiency of SNPs identification and impede the accurate discrimination of candidate SNPs between homologous and paralogous sequences. In polyploid species, it is very common to find polymorphic SNP within a single genotype due to the presence of either homoeologous loci from individual subgenomes or paralogous loci from duplicated regions of the genome [12].

In Brassica napus, most polymorphisms between subgenomes are homoeologous making it difficult to assign them to the A or C subgenome [Kaur et al., 2012]. Hence, stringent mapping parameters are often essential to avoid false SNP calls [11, 39].

Polyploid species are characterized by a

large genome size and complex structure, consisting of subgenomes of two homologous genomes from progenitor diploid species. They also contain high level of repetitive DNA sequences. For example, Brassica napus is an allotetraploid that has undergone extensive duplication events over the course of evolution. High levels of intragenomic duplication lead to significant intragenomic sequence paralogy and a complex gene family. Owing to this genome duplication, a gene could have a second copy at another position in the genome [40]. Therefore, duplication events in Brassica influence the alignment comparison process of NGS sequencing assembly and SNPs calling. Thus, there is a need for rigorous discrimination between the origin of identified SNPs [2]. Maize is an ancient polyploid species with large regions of genomic duplication. It has been shown that a considerable number of SNP markers have a pattern of shifting of the clusters to one side or five clusters. This is indicative of detection of more than one locus that correspond to the duplicated regions identified in the maize genome [14].

The discovery of SNPs and array development at various densities have been successfully achieved and applied in many diploid and polyploid plant species [34]. Examples of these include wheat 20K, 820K and 660K SNP arrays [41-43], Maize 600K SNP array [44], Brassica napus 60K and 6K SNP arrays [45, 46], rice 700K SNP array [47], Apple 8K and 480K SNP array [48,49] , peanut 58K SNP arrays [50], strawberry 90K SNP array [51], sugarcane 345K, 76K and 84K SNP arrays [52-54], cotton 63K SNP Array[55] and oat 6K SNP array [56]. However, there are additional obstacles to SNP identification and utilization for genotyping in polyploid species, so this field of study has progressed slowly due to the complex nature of the genomes and polyploid inheritance [11, 57].

Against this background, to reduce the genome complexity, additional genetic information about progenitor relationships with allopolyploid specie is extremely useful to assist in discriminating SNPs polymorphism through the ability to compare polyploid-derived sequence genotypes to diploid counterparts in which strict alignment and assembly criteria should be applied [2]. SNP arrays have significantly accelerated the molecular studies in plant genomics. As a large number of SNP array platforms have been developed, there are still demands to address their limitations. The efficiency and accuracy of genotyping rate is still a challenging. SNP selection and validation requires careful consideration to increase the success rate and to produce high proportions of correctly genotyped SNPs. Another issue is the number of informative polymorphic SNP in a population. Only a small number of SNPs can be suitable for further analysis as the majority of SNPs cannot be detected, genotyped or being monomorphic. This is due to significant numbers of false positive SNPs. Moreover, the level of sequence diversity and structural genome variations including translocations and copy number variation may reduce both the efficiency of SNP dis-

covery and the ability of correctly identifying the allelic state of each individual [58, 59]. The main drawback of SNP arrays is that they still remain expensive. There is a huge cost associated with design and validation of the array. The array technique requires specific equipment and high labour cost, Thus resulting in the limit use by most researchers. Commercially available SNP arrays with variable number of fixed SNPs have a limited use and cannot be modified. They are often customized to be specific to certain species or populations [4].

With recent advancements of sequencing technologies, SNPs prediction and validation by sequencing could replace the genotyping arrays. The genotyping arrays are species-specific, expensive to design. It only covers all known SNP variants and not intended to discover rare or novel variants. In contrast, sequencing generates more coverage and the entire genome is scanned to find and genotype new SNP variants. The sequencing approach is less expensive, more power and computational process is becoming more routine. In addition, sequencing is ideal for uninvestigated species with limited genomic resources. Sequencing is powerful tool for genetic diversity and genomic selection with speed and efficiency. In the future, sequencing could become more widespread and ideal option for SNP discovery for more genetic and genomics applications [5, 6]

3. SNP genotyping and candidate gene association

SNP genotyping arrays are a powerful approach to detect SNPs in a cost- and time-effective manner and more importantly facilitate genetic and genomic studies such as QTL mapping and association studies. The discovery of a vast number of SNPs provides an ideal makers for the novel SNPs array to study the genetic mapping of QTL associated with candidate genes [26, 34, 60].

The candidate genes for particular traits in plants have been the particular focus in order to facilitate the reverse genetic approach and gene discovery or the identification of gene polymorphism that could be useful for marker-assisted selection [55,61, 62]. Thus, SNP discovery and genotyping in segregating populations with specific trait generally enable us to find a strong correlation between many more SNPs and genes controlling the trait in a population [2]. This approach can be successful if there is a known, strong functional SNP and phenotypic data and quantitative loci.

Cloning and characterization of candidate genes within QTL in plant genetics and breeding have been difficult and not feasible through genetic mapping and map based cloning due to large genetic and physical distances of QTL, low resolution of available genetic maps and complexity of the genome [63]. Conventionally, QTL is determined by linkage mapping approaches where two parents and their segregating population are screened for polymor-

phisms associated with the trait of interest. The linkage of a particular region to a given phenotype can be determined by the frequency of recombinants exhibiting phenotypic variations for a trait of interest [4]. In fact, SNP genotyping solves the problem of positional map cloning of a large number of candidate gene association studies in the case that many candidate genes are weak candidates with a low level of polymorphism. This is because candidate gene studies only examine one or a few SNPs which carry limited polymorphism information on the overall population variation within that QTL. All above, if a certain region corresponds to a QTL, the gene may be a candidate gene. This approach could greatly facilitate the map cloning approach[64].

Recently, the mapping of QTL of interest has been made possible by taking advantage of SNP genotyping technologies combined with NGS data. Candidate gene-based approaches have been used for identifying SNPs for specific candidate genes which might provide connection between allele function and specific trait variations. In addition, scanning may be performed to identify regions associated with a particular phenotype, while gene-wide scanning for the SNP distribution and association to a segregating population would significantly enhance large-scale efforts to identify all loci controlling multiple traits for crop improvement [7].

SNP genotyping plays an essential role in identifying a number of candidate genes / QTL linked to important crop traits and dissecting other complex traits [65]. For example, a candidate gene and QTL cluster associated with four fiber traits on chromosome A07 of cotton was identified [66]. In wheat, a candidate gene based SNP marker was developed to determine the locus that controls resistance to leaf rust, stripe rust, and powdery mildew diseases [67]. In addition, a study by Wu, Zhao [68] found 39 candidate genes linked to three loci DSRC4, DSRC6, and DSRC8 associated with resistance to sclerotinia stem rot in B. napus. Moreover, putative SNP markers were mapped to the marker flanking region linked to Fhb, a gene involved in resistance to fusarium head blight (FHB) disease [69]. This would enable fine mapping towards cloning of the Fhb1 gene.

Many studies have been conducted to identify SNP markers linked to a QTL region to produce high resolution QTL and narrow down the target locus for the analysis of candidate genes. For example, a high-density SNP map gave a better resolution RFLP/SSR-based QTL by condensing two QTL regions for grain weight in rice within 123 kb. Moreover, a few SNPs were functionally associated with the variation in yield [70]. In soybean, SNP genotyping narrowed down the QTL region for aphid resistance gene, Rag1 from 12 cM to 115 kb with two linked SNP markers and additional candidate genes being identified [71]. In another study, SNP markers of candidate genes for flowering time in Brassica were identified. [72]. Also, Three QTLs for flowering time were linked to three SNPs in the promoter of the BrFLC2 gene caus-

ing low expression. These markers facilitated comparative mapping between B. rapa and Arabidopsis.

SNP genotyping arrays with a large number of SNPs are useful for the analysis of many traits or individuals at high genetic resolution, leading to candidate gene mapping. The idea is that SNP markers at candidate genes selected for their association with a phenotype are physically mapped and the genotypic data are compared across contrasting plant genotypes. Hence, a cluster of SNPs situated in a small physical distance within QTL region are most likely the potential candidate genes. This strategy has a good advantage to the problems associated with map-based cloning and physical mapping [73, 74]. Taking advantage of SNP genotyping technologies combined with NGS data, candidate gene-based approaches have been used for the discovery of SNPs linked to candidate genes/QTL which might provide a direct correlation of allele function with specific trait variation [7].

4.Use of Illumina Infinium genotyping assay as an example

High throughput SNP genotyping platforms have been developed to assay up to 1 million SNPs which played a decisive role in the success of genomic studies. Many SNP genotyping platforms are currently available commercially such as Illumina GoldenGate and Infinium, TaqMan or GeneChip from Affymetrix. These technologies can be applied to genotype SNP markers and require a preliminary step of SNP discovery, technical expertise and an expensive laboratory set-up.

The Illumina Infinium assay is a whole-genome SNP genotyping assay that has been proven to be successful and efficient for many plant crops [23, 45, 75]. It is capable of multiplexing from about 6,000 up to 1 million SNPs. It enables genome-wide analysis offering an ideal method to identify candidate polymorphic SNPs for QTL mapping. The Infinium assay develops BeadChips with customized SNP content. Each SNP locus is assayed and analyzed independently for each sample yielding high intensity estimate and accurate genotype calls.

In this paper, we briefly describe the use of 6K B. napus custom SNP array with 5306 SNPs as implemented in the Infinium assay [Illumina Inc., San Diego, USA]. Details of the array and the Infinium assay are described previously

[Dalton-Morgan et al., 2014]. These SNPs were distributed over the 19 chromosomes of B. napus genome. This array was applied to genotype B. napus cultivars and map populations in order to study the SNP genotypes and identify the candidate SNPs located on QTL or chromosome regions associated with candidate genes. Samples were analyzed with the Infinium II assay protocol according to the manufacturer's instructions. Genotyping module in GenomeStudio Illumina with the default parameters was utilized for SNP data analysis clustering and genotype call. The clusters were reviewed and manually edited if

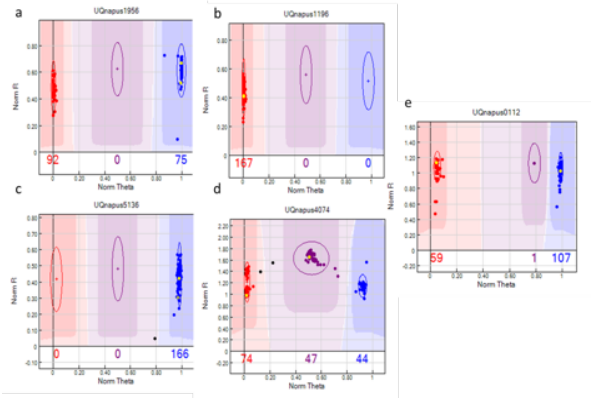
needed. SNPs that were difficult to score were either manually adjusted or eliminated from the analysis completely.

The SNP genotype results show consistent and clear clustering for most of the 5306 SNPs across the mapping population with 91% of SNPs successfully genotyped and clustered, taking into account both monomorphic and polymorphic SNPs (Fig 1 and Table 1). Successful SNP markers primarily produced distinct clustering patterns with three possible genotypes (AA, AB, BB), detecting polymorphic or monomorphic loci. A good-quality SNP genotyping cluster shows three clearly defined and tight clusters with the homozygotes and heterozygotes. The cluster definition is relatively simple in diploid species. However, it is more complex in crop plants that are polyploid such as B. napus because SNP clusters can often detect more than three overlapping clusters. This is sometimes seen in polyploid species as five clusters (AAAA, AAAB, AABB, ABBB, and BBBB) instead of the typical three clusters (AA, AB and BB). Therefore, SNP clusters in polyploid cases often make the distinction of the allelic status more challenging since the clusters need to be defined more carefully and mostly in a manual fashion.

Table 1. Classification of SNP marker results from B. napus Infinium SNP genotyping assay on the population in this study

SNP position	No. of total SNPs	Proportion of SNPs		
		Failed	Monomorphic	Polymorphic
Genome-wide	5306	487 (9%)	3535 (67%)	1283 (24%)
Chromosome A7	285	32	178	74
Scaffold 3 on A7	109	12	44	57

Fig 1. Examples of successfully clustered SNPs based on the GenomeStudio software. Clusters display the area where the three different genotypes with homozygous allele A (red), heterozygous AB (purple) and homozygous allele B (blue) are called. A) Two clusters with AA and BB genotypes and a monomorphic state for the parents (both in yellow). B) One cluster with AA genotype and a monomorphic state for both parents and all individuals. C) One cluster with BB genotype and a monomorphic state for both parents and all individuals. D) Three clusters represent the genotypes AA, AB and BB, and a polymorphic state for parents. E) Two clusters with AA and BB genotypes and a polymorphic state for both parents. The black dot outside of the cluster represents failed samples and thus is scored as "no call".



Two different indexes were initially used to evaluate the quality of the raw data before the optimization of SNP clustering. The call rate index showed that more than 90% of SNPs were successfully genotyped for the mapping population. In addition, the GenTrain score index gives an indication of the quality of the SNP clustering. According to Illumina, for an SNP to be retained, a minimum GenTrain score of 0.15 is advisable. In the present study, an SNP had to get a minimum GenTrain score of 0.15 and had to be segregating in the related mapping population to be declared successful. In the DH mapping population, the majority of clustering SNPs had a GenTrain score of more than 0.4, indicating that SNP genotyping was highly reliable

with a low rate of missing data. Low GenTrain score may indicate ambiguous cluster separation.

Of the 5306 SNPs assayed, 1284 (24%) were polymorphic between the two parental lines and thus could be used for clustering (Table 1). This provides a large number of novel markers that will be used in genetic mapping. Another 3535(67%) SNPs from this assay were monomorphic between the two parental lines, meaning that they clustered together, having the same base at that specific locus. These SNPs are uninformative in genetic studies. Only 487(9%) of the 5306 SNPs failed to give a genotype or could not be easily clustered and were not investigated further in this study. This probably indicates false SNPs, resulting from possible sequencing mistakes.

The Infinium assay is sensitive to the number of allelic copies [target locus + duplicated or paralogous alleles] being assayed (Fig 2 and 3). This can be a problem, particularly in polyploids such as B. napus, where two homozygous clusters or more than three clusters are most likely due to the existence of non-unique genomic regions. This is because B. napus is a highly duplicated genome. It is anticipated that the amphidiploid nature of the B. napus genome will usually result in homeologous pairs of genes, originating from the A and C genomes. This sometimes results in polymorphisms detected at a given position, one corresponding to the A genome and the other to the C genome. The Infinium assay cannot distinguish the status of SNP alleles from homologues (76). This makes automated SNP genotyping more challenging in a complex genome. The absence of cluster separation can be due to a non-allele-specific match of the primers, for example (Fig 3).

Fig 2. Vertically separated clusters generally polymorphic for a different locus than the source of the targeted SNP. This may indicate the presence of a third polymorphic allele.

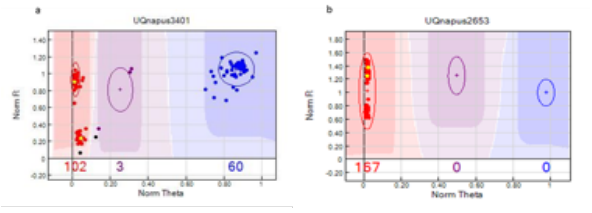
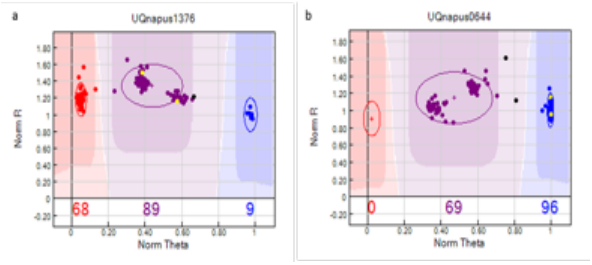


Fig 3. An example of more than two genotype clusters observed due to the polyploid nature of the B. napus genome



The majority of SNPs on chromosome A7 (252 out of 285) gave a clear genotype (Table 1). Of these, 74 were successfully polymorphic. Fifty seven of these SNPs were on scaffold 3, scaffold with the largest length in chromosome A7, representing the QTL region underlying a disease resistance candidate gene [77]. The SNPs located the within genes in scaffold 3 were generally more polymorphic than the overall set of SNPs in chromosome A7 suggesting their potential as candidate genes.

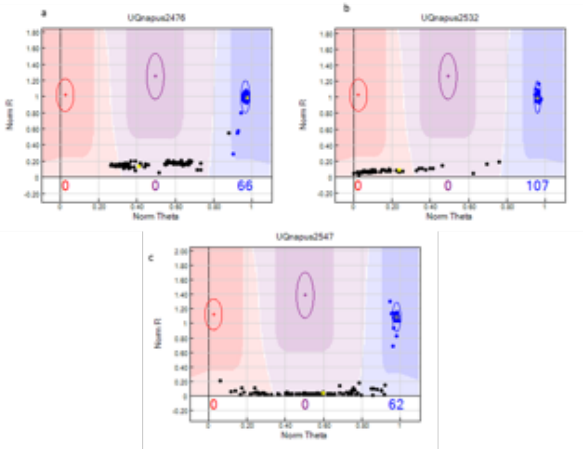
The results also showed that the 57 SNPs markers were distorted, with a 2:1 segregation ratio, and the direction of distortion

was mainly toward the maternal parent. In addition, the number of genotypes corresponding to the resistant parent was almost twice that corresponding to the susceptible parent, with averages of 102 and 58, respectively. This suggests that segregation of the resistance might be controlled by multiple genes. This finding is consistent with a previous study [78], which found a distorted segregation ratio for phenotype. This distortion might be the result of chromosomal abnormalities biasing the estimation of the recombination fraction between markers on the linkage group. This may subsequently lead to the loss of informative data in QTL mapping, if these markers are not handled carefully.

As the population is double haploid (DH), most of the SNPs were expected to show two clear main clusters, representing the two homozygous genotypes. A small additional cluster in the middle of the graph corresponding to heterozygous genotypes may represent a third allele or an additional null allele. Doubled haploids are genetically homozygous lines. However, DH samples sometimes behave as heterozygotes. This might be because the SNP was initially identified as polymorphic in sequence comparisons but behaved monomorphically in the analyzed sample or was genotyped as heterozygous reflecting structural chromosomal rearrangements: duplication, translocation or transpositions. Taken the findings together, our analyses of the DH mapping population showed distinguishable segregating allelic variants at a single locus underlying the QTL.

Three SNPs were null and showed the presence of an insertion or deletion surrounding the SNP locus (Fig 4), presumably due to a mutation or the presence of a third allele. Further sequencing and validation of these SNPs might give an indication of whether the SNP allele status has an association with a candidate gene.

Fig 4. Individuals with homozygous deletion cluster at the bottom of the graph; their genotype calls are missing/null. This may indicate the presence of chromosomal deletion or a third allele. The BB cluster remains intact.



5. Conclusion

This work demonstrates how NGS technologies associated with SNP genotyping could be a feasible strategy for the genotyping of thousands of SNPs and correlating them with QTL or candidate genes. It also provides an alternative approach for identification of candidate gene. The results can be further analyzed for genetic mapping and association in the B. napus segregating population. Hence, the genotyping quality and physical position data obtained for the SNPs can be useful for consensus genetic maps, positional cloning, or association mapping. Despite geno-

typing assay are being efficient to generate a vast amount of genotypic data in polyploid crops, analysis of SNP calls is somewhat challenging in polyploids due to the multiallele combinations in the genotypes. However, high density SNP arrays would be efficient to produce highly informative SNPs for marker-assisted selection in breeding programs as well as functional genomic studies.

References

[1] Syvanen, A.-C., 2005. Toward genome-wide SNP genotyping, Nature Genetics 378,

[2] Kaur, S., Francki, M. G., and Forster, J. W., 2012. Identification, characterization and interpretation of single-nucleotide sequence variation in allopolyploid crop species, Plant Biotechnology Journal. 10, no. 2, pp. 125-138,

[3] Wang, D. G., Fan, J.-B., Siao, C.-J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L., Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E., Mittmann, M., Morris, M. S., Shen, N., Kilburn, D., Rioux, J., Nusbaum, C., Rozen, S., Hudson, T. J., Lipshutz, R., Chee, M., and Lander, E. S., 1998. Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome, Science. 280, no. 5366, pp. 1077-1082

[4] Deschamps, S., Llaca, V., and May, G. D., 2012. Genotyping-by-Sequencing in Plants, Biology. 1, no. 3, pp. 460-483,

[5] Gupta, P. K., Roy, J. K., and Prasad, M., 2001. Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants, Current Science. 80, no. 4, pp. 524-535

[6] Jander, G., Norris, S. R., Rounsley, S. D., Bush, D. F., Levin, I. M., and Last, R. L., 2002. Arabidopsis Map-Based Cloning in the Post-Genome Era, Plant Physiology. 129, no. 2, pp. 440-450

[7] Rafalski, A., 2002. Applications of single nucleotide polymorphisms in crop genetics, Current Opinion in Plant Biology. 5, pp. 94 - 100,

[8] Westermeier, P., Wenzel, G., and Mohler, V., 2009. Development and evaluation of single-nucleotide polymorphism markers in allotetraploid rapeseed (Brassica napus L.), Theoretical and Applied Genetics. 119, no. 7, pp. 1301-1311,

[9] Edwards, D., and Batley, J., 2010. Plant genome sequencing: applications for crop improvement, Plant Biotechnology Journal. 8, no. 1, pp. 2-9,

[10] von Thaden, A., Nowak, C., Tiesmeyer, A., Reiners, T. E., Alves, P. C., Lyons, L. A., Mattucci, F., Randi, E., Cragnolini, M., Galián, J., Hegyeli, Z., Kitchener, A.

C., Lambinet, C., Lucas, J. M., Mölich, T., Ramos, L., Schockert, V., and Cocchiararo, B., 2020. Applying genomic data in wildlife monitoring: Development guidelines for genotyping degraded samples with reduced single nucleotide polymorphism panels, Molecular Ecology Resources. 20, no. 3,

[11] You, Q., Yang, X., Peng, Z., Xu, L., and Wang, J., 2018. Development and Applications of a High Throughput Genotyping Tool for Polyploid Crops: Single Nucleotide Polymorphism (SNP) Array, Frontiers in Plant Science. 9, pp. 104-104,

[12] Ganal, M. W., Altmann, T., and Röder, M. S., 2009. SNP identification in crop plants, Current Opinion in Plant Biology. 12, no. 2, pp. 211-217,

[13] Edwards, D., Forster, J., Cogan, N. I., Batley, J., and Chagné, D., "Single Nucleotide Polymorphism Discovery," Association Mapping in Plants, N. Oraguzie, E. A. Rikkerink, S. Gardiner et al., eds., pp. 53-76: Springer New York, 2007.

[14] Durstewitz, G., Polley, A., Plieske, J., Luerssen, H., Graner, E. M., Wieseke, R., and Ganal, M. W., 2010. SNP discovery by amplicon sequencing and multiplex SNP genotyping in the allopolyploid species Brassica napus, Genome. 53, no. 11, pp. 948-956

[15] Schmid, K. J., Sörensen, T. R., Stracke, R., Törjék, O., Altmann, T., Mitchell-Olds, T., and Weisshaar, B., 2003. Large-Scale Identification and Analysis of Genome-Wide Single-Nucleotide Polymorphisms for Mapping in Arabidopsis thaliana, Genome Research. 13, no. 6a, pp. 1250-1257

[16] Batley, J., Barker, G., O'Sullivan, H., Edwards, K. J., and Edwards, D., 2003. Mining for Single Nucleotide Polymorphisms and Insertions/Deletions in Maize Expressed Sequence Tag Data, Plant Physiology. 132, no. 1, pp. 84-91,

[17] Barbazuk, W. B., Emrich, S. J., Chen, H. D., Li, L., and Schnable, P. S., 2007. SNP discovery via 454 transcriptome sequencing, The Plant journal : for cell and molecular biology. 51, no. 5, pp. 910-918,

[18] Geraldes, A., Pang, J., Thiessen, N., Cezard, T., Moore, R., Zhao, Y., Tam, A., Wang, S., Friedmann, M., Birol, I., Jones, S. J. M., Cronk, Q. C. B., and Douglas, C. J., 2011. SNP discovery in black cottonwood (Populus trichocarpa) by population transcriptome resequencing, Molecular Ecology Resources. 11 Suppl 1, pp. 81-92

[19] Rogier, O., Chateigner, A., Amanzougarene, S., Lesage-Descauses, M.-C., Balzergue, S., Brunaud, V., Caius, J., Soubigou-Taconnat, L., Jorge, V., and Segura, V., 2018. Accuracy of RNAseq based SNP discovery and genotyping in Populusnigra, BMC Genomics. 19, no. 1, pp. 909

[20] Varshney, R. K., Nayak, S. N., May, G. D., and Jackson, S. A., 2009. Next-gen-

eration sequencing technologies and their implications for crop genetics and breeding, Trends in biotechnology. 27, no. 9, pp. 522-530,

[21] Yamamoto, T., Nagasaki, H., Yonemaru, J.-i., Ebana, K., Nakajima, M., Shibaya, T., and Yano, M., 2010. Fine definition of the pedigree haplotypes of closely related rice cultivars by means of genome-wide discovery of single-nucleotide polymorphisms, BMC Genomics. 11, no. 1, pp. 267,

[22] Ossowski, S., Schneeberger, K., Clark, R. M., Lanz, C., Warthmann, N., and Weigel, D., 2008. Sequencing of natural strains of Arabidopsis thaliana with short reads, Genome Research. 18, no. 12, pp. 2024-2033

[23] Kumar, S., Banks, T. W., and Cloutier, S., 2012. SNP Discovery through Next-Generation Sequencing and Its Applications, International Journal of Plant Genomics. 2012, pp. 15,

[24] You, F., Huo, N., Deal, K., Gu, Y., Luo, M.-C., McGuire, P., Dvorak, J., and Anderson, O., 2011. Annotation-based genome-wide SNP discovery in the large and complex Aegilops tauschii genome using next-generation sequencing without a reference genome sequence, BMC Genomics. 12, no. 1, pp. 59,

[25] Feuillet, C., Leach, J. E., Rogers, J., Schnable, P. S., and Eversole, K., 2011. Crop genome sequencing: lessons and rationales, Trends in Plant Science. 16, no. 2, pp. 77-88,

[26] Hirakawa, H., Shirasawa, K., Ohyama, A., Fukuoka, H., Aoki, K., Rothan, C., Sato, S., Isobe, S., and Tabata, S., 2013. Genome-Wide SNP Genotyping to Infer the Effects on Gene Functions in Tomato, DNA Research,

[27] Poland, J. A., and Rife, T. W., 2012. Genotyping-by-Sequencing for Plant Breeding and Genetics, Plant Gen. 5, no. 3, pp. 92-102, 2012/11,

[28] He, J., Zhao, X., Laroche, A., Lu, Z.-X., Liu, H., and Li, Z., 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding, Frontiers in Plant Science. 5, pp. 484-484,

[29] Duran, C., Appleby, N., Clark, T., Wood, D., Imelfort, M., Batley, J., and Edwards, D., 2009. AutoSNPdb: an annotated single nucleotide polymorphism database for crop plants, Nucleic Acids Research. 37, no. Database issue, pp. D951-D953,

[30] Clément, S., Fillon, J., Bousquet, J., and Beaulieu, J., 2010. TreeSNPs: a laboratory information management system (LIMS) dedicated to SNP discovery in trees, Tree Genetics & Genomes. 6, no. 3, pp. 435-438

[31] Smigielski, E. M., Sirotkin, K., Ward, M., and Sherry, S. T., 2000. dbSNP: a database of single nucleotide polymorphisms, Nucleic Acids Research. 28, no. 1, pp. 352-355,

[32] Kim, C.-K., Won, S.-Y., Seol, Y.-J., and Lee, T.-H., 2015. NABIC SNP: an integrated database for SNP markers, Bioinformation. 11, no. 7, pp. 369-372,

[33] Clevenger, J., Chu, Y., Chavarro, C., Agarwal, G., Bertioli, D. J., Leal-Bertioli, S. C. M., Pandey, M. K., Vaughn, J., Abernathy, B., Barkley, N. A., Hovav, R., Burow, M., Nayak, S. N., Chitikineni, A., Isleib, T. G., Holbrook, C. C., Jackson, S. A., Varshney, R. K., and Ozias-Akins, P., 2017. Genome-wide SNP Genotyping Resolves Signatures of Selection and Tetrasomic Recombination in Peanut, Molecular plant. 10, no. 2, pp. 309-322,

[34] Montanari, S., Bianco, L., Allen, B. J., Martínez-García, P. J., Bassil, N. V., Postman, J., Knäbel, M., Kitson, B., Deng, C. H., Chagné, D., Crepeau, M. W., Langley, C. H., Evans, K., Dhingra, A., Troggio, M., and Neale, D. B., 2019. Development of a highly efficient Axiom™ 70 K SNP array for Pyrus and evaluation for high-density mapping and germplasm characterization, BMC Genomics. 20, no. 1, pp. 331-331,

[35] Appleby, N., Edwards, D., and Batley, J., "New Technologies for Ultra-High Throughput Genotyping in Plants," Plant Genomics, Methods in Molecular Biology™ J. P. Gustafson, P. Langridge and D. J. Somers, eds., pp. 19-39: Humana Press, 2009.

[36] Verde, I., Bassil, N., Scalabrin, S., Gilmore, B., Lawley, C. T., Gasic, K., Micheletti, D., Rosyara, U. R., Cattonaro, F., Vendramin, E., Main, D., Aramini, V., Blas, A. L., Mockler, T. C., Bryant, D. W., Wilhelm, L., Troggio, M., Sosinski, B., Aranzana, M. J., Arús, P., Iezzoni, A., Morgante, M., and Peace, C., 2012. Development and Evaluation of a 9K SNP Array for Peach by Internationally Coordinated SNP Detection and Validation in Breeding Germplasm, PLoS ONE. 7, no. 4, pp. e35668,

[37] Trebbi, D., Maccaferri, M., Heer, P., Sørensen, A., Giuliani, S., Salvi, S., Sanguineti, M., Massi, A., Vossen, E., and Tuberosa, R., 2011. High-throughput SNP discovery and genotyping in durum wheat (Triticum durum Desf.), Theoretical and Applied Genetics. 123, no. 4, pp. 555-569

[38] Chagné, D., Batley, J., Edwards, D., and Forster, J., "Single Nucleotide Polymorphism Genotyping in Plants," Association Mapping in Plants, N. Oraguzie, E. A. Rikkerink, S. Gardiner et al., eds., pp. 77-94: Springer New York, 2007.

[39] Byers, R., Harker, D., Yourstone, S., Maughan, P., and Udall, J., 2012. Development and mapping of SNP assays in allotetraploid cotton, Theoretical and Applied Genetics. 124, no. 7, pp. 1201-1214

[40] Lysak, M. A., Koch, M. A., Pecinka, A., and Schubert, I., 2005. Chromosome triplication found across the tribe Brassiceae, Genome Research. 15, no. 4, pp. 516-525

[41] Sun, C., Dong, Z., Zhao, L., Ren, Y., Zhang, N., and Chen, F., 2020. The Wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat, Plant Biotechnology Journal. 18, no. 6, pp. 1354-1360,

[42] Winfield, M. O., Allen, A. M., Burridge, A. J., Barker, G. L. A., Benbow, H. R., Wilkinson, P. A., Coghill, J., Waterfall, C., Davassi, A., Scopes, G., Pirani, A., Webster, T., Brew, F., Bloor, C., King, J., West, C., Griffiths, S., King, I., Bentley, A. R., and Edwards, K. J., 2016. High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool, Plant Biotechnology Journal. 14, no. 5, pp. 1195-1206,

[43] Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L., 2012. Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Ap proach, PLoS ONE. 7, no. 2, pp. e32253,

[44] Unterseer, S., Bauer, E., Haberer, G., Seidel, M., Knaak, C., Ouzunova, M., Meitinger, T., Strom, T. M., Fries, R., Pausch, H., Bertani, C., Davassi, A., Mayer, K. F. X., and Schön, C.-C., 2014. A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array, BMC Genomics. 15, no. 1, pp. 823

[45] Mason, A. S., Higgins, E. E., Snowdon, R. J., Batley, J., Stein, A., Werner, C., and Parkin, I. A. P., 2017. A user guide to the Brassica 60K Illumina Infinium™ SNP genotyping array, Theoretical and Applied Genetics. 130, no. 4, pp. 621-633

[46] Raman, H., Dalton-Morgan, J., Diffey, S., Raman, R., Alamery, S., Edwards, D., and Batley, J., 2014. SNP markers-based map construction and genome-wide linkage analysis in Brassica napus, Plant Biotechnology Journal. 12, no. 7, pp. 851-860

[47] McCouch, S. R., Wright, M. H., Tung, C.-W., Maron, L. G., McNally, K. L., Fitzgerald, M., Singh, N., DeClerck, G., Agosto-Perez, F., Korniliev, P., Greenberg, A. J., Naredo, M. E. B., Mercado, S. M. Q., Harrington, S. E., Shi, Y., Branchini, D. A., Kuser-Falcão, P. R., Leung, H., Ebana, K., Yano, M., Eizenga, G., McClung, A., and Mezey, J., 2016. Open access resources for genome-wide association mapping in rice, Nature Communications. 7, no. 1, pp. 10532

[48] Chagné, D., Crowhurst, R. N., Troggio, M., Davey, M. W., Gilmore, B., Lawley, C., Vanderzande, S., Hellens, R. P., Kumar, S., Cestaro, A., Velasco, R., Main, D., Rees, J. D., Iezzoni, A., Mockler, T., Wilhelm, L., Van de Weg, E., Gardiner, S. E., Bassil, N., and Peace, C., 2012. Ge-

nome-Wide SNP Detection, Validation, and Development of an 8K SNP Array for Apple, PLoS ONE. 7, no. 2, pp. e31745,

[49] Bianco, L., Cestaro, A., Linsmith, G., Muranty, H., Denancé, C., Théron, A., Poncet, C., Micheletti, D., Kerschbamer, E., Di Pierro, E. A., Larger, S., Pindo, M., Van de Weg, E., Davassi, A., Laurens, F., Velasco, R., Durel, C.-E., and Troggio, M., 2016. Development and validation of the Axiom®Apple480K SNP genotyping array, The Plant Journal. 86, no. 1, pp. 62-74

[50] Pandey, M. K., Agarwal, G., Kale, S. M., Clevenger, J., Nayak, S. N., Sriswathi, M., Chitikineni, A., Chavarro, C., Chen, X., Upadhyaya, H. D., Vishwakarma, M. K., Leal-Bertioli, S., Liang, X., Bertioli, D. J., Guo, B., Jackson, S. A., Ozias-Akins, P., and Varshney, R. K., 2017. Development and Evaluation of a High Density Genotyping 'Axiom_Arachis' Array with 58 K SNPs for Accelerating Genetics and Breeding in Groundnut, Scientific Reports. 7, no. 1, pp. 40577

[51] Bassil, N. V., Davis, T. M., Zhang, H., Ficklin, S., Mittmann, M., Webster, T., Mahoney, L., Wood, D., Alperin, E. S., Rosyara, U. R., Koehorst-vanc Putten, H., Monfort, A., Sargent, D. J., Amaya, I., Denoyes, B., Bianco, L., van Dijk, T., Pirani, A., Iezzoni, A., Main, D., Peace, C., Yang, Y., Whitaker, V., Verma, S., Bellon, L., Brew, F., Herrera, R., and van de Weg, E., 2015. Development and preliminary evaluation of a 90 K Axiom® SNP array for the allo-octoploid cultivated strawberry Fragaria× ananassa, BMC Genomics. 16, no. 1, pp. 155

[52] Aitken, K. S., Farmer, A., Berkman, P. J., Muller, C., Wei, X., Demano, E., Jackson, P., Magwire, M., Dietrich, B., and Kota, R., "Generation of a 345K sugarcane SNP chip." pp. 1165-1172.

[53] Yang, X., Song, J., You, Q., Paudel, D. R., Zhang, J., and Wang, J., 2017. Mining sequence variations in representative polyploid sugarcane germplasm accessions, BMC Genomics. 18, no. 1, pp. 594

[54] Balsalobre, T. W. A., da Silva Pereira, G., Margarido, G. R. A., Gazaffi, R., Barreto, F. Z., Anoni, C. O., Cardoso-Silva, C. B., Costa, E. A., Mancini, M. C., Hoffmann, H. P., de Souza, A. P., Garcia, A. A. F., and Carneiro, M. S., 2017. GBS-based single dosage markers for linkage and QTL mapping allow gene mining for yield-related traits in sugarcane, BMC Genomics. 18, no. 1, pp. 72-72,

[55] Hulse-Kemp, A. M., Lemm, J., Plieske, J., Ashrafi, H., Buyyarapu, R., Fang, D. D., Frelichowski, J., Giband, M., Hague, S., Hinze, L. L., Kochan, K. J., Riggs, P. K., Scheffler, J. A., Udall, J. A., Ulloa, M., Wang, S. S., Zhu, Q.-H., Bag, S. K., Bhardwaj, A., Burke, J. J., Byers, R. L., Claverie, M., Gore, M. A., Harker, D. B., Islam, M. S., Jenkins, J. N., Jones, D. C., Lacape, J.-M., Llewellyn, D. J., Percy, R. G., Pepper, A. E., Poland, J. A., Mohan Rai,

K., Sawant, S. V., Singh, S. K., Spriggs, A., Taylor, J. M., Wang, F., Yourstone, S. M., Zheng, X., Lawley, C. T., Ganal, M. W., Van Deynze, A., Wilson, I. W., and Stelly, D. M., 2015. Development of a 63K SNP Array for Cotton and High-Density Mapping of Intraspecific and Interspecific Populations of Gossypium spp, G3: Genes|Genomes|Genetics. 5, no. 6, pp. 1187,

[56] Tinker, N. A., Chao, S., Lazo, G. R., Oliver, R. E., Huang, Y.-F., Poland, J. A., Jellen, E. N., Maughan, P. J., Kilian, A., and Jackson, E. W., 2014. A SNP Genotyping Array for Hexaploid Oat, The Plant Genome. 7, no. 3, pp. plantgenome2014.03.0010,

[57] Clevenger, J., Chavarro, C., Pearl, Stephanie A., Ozias-Akins, P., and Jackson, Scott A., 2015. Single Nucleotide Polymorphism Identification in Polyploids: A Review, Example, and Recommendations, Molecular plant. 8, no. 6, pp. 831-846

[58] Zhang, J., Yang, J., Zhang, L., Luo, J., Zhao, H., Zhang, J., and Wen, C., 2020. A new SNP genotyping technology Target SNP-seq and its application in genetic analysis of cucumber varieties, Scientific Reports. 10, no. 1, pp. 5623

[59] Vanderzande, S., Zheng, P., Cai, L., Barac, G., Gasic, K., Main, D., Iezzoni, A., and Peace, C., 2020. The cherry 6+9K SNP array: a cost-effective improvement to the cherry 6K SNP array for genetic studies, Scientific Reports. 10, no. 1, pp. 7613

[60] Chong, X., Zhang, F., Wu, Y., Yang, X., Zhao, N., Wang, H., Guan, Z., Fang, W., and Chen, F., 2016. A SNP-Enabled Assessment of Genetic Diversity, Evolutionary Relationships and the Identification of Candidate Genes in Chrysanthemum, Genome biology and evolution. 8, no. 12, pp. 3661-3671,

[61] Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., Maccaferri, M., Salvi, S., Milner, S. G., Cattivelli, L., Mastrangelo, A. M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., International Wheat Genome Sequencing, C., Lillemo, M., Mather, D., Appels, R., Dolferus, R., Brown-Guedira, G., Korol, A., Akhunova, A. R., Feuillet, C., Salse, J., Morgante, M., Pozniak, C., Luo, M.-C., Dvorak, J., Morell, M., Dubcovsky, J., Ganal, M., Tuberosa, R., Lawley, C., Mikoulitch, I., Cavanagh, C., Edwards, K. J., Hayden, M., and Akhunov, E., 2014. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array, Plant Biotechnology Journal. 12, no. 6, pp. 787-796,

[62] Dalton-Morgan, J., Hayward, A., Alamery, S., Tollenaere, R., Mason, A. S., Campbell, E., Patel, D., Lorenc, M. T., Yi, B., Long, Y., Meng, J., Raman, R., Raman, H., Lawley, C., Edwards, D., and Batley, J., 2014. A high-throughput SNP array in the amphidiploid species Brassica napus shows diversity in resistance genes, Functional & Integrative Genomics. 14, no. 4,

pp. 643-655, 2014/12/01,

[63] Mauricio, R., 2001. Mapping quantitative trait loci in plants: Uses and caveats for evolutionary biology, Nature Reviews Genetics. 2, no. 5, pp. 370-381,

[64] 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, X. W., and Zhang, Q., 2014. A High-Density SNP Genotyping Array for Rice Biology and Molecular Breeding, Molecular plant. 7, no. 3, pp. 541-553

[65] Majeed, S., Rana, I. A., Atif, R. M., Ali, Z., Hinze, L., and Azhar, M. T., 2019. Role of SNPs in determining QTLs for major traits in cotton, Journal of Cotton Research. 2, no. 1, pp. 5

[66] Islam, M. S., Thyssen, G. N., Jenkins, J. N., Zeng, L., Delhom, C. D., McCarty, J. C., Deng, D. D., Hinchliffe, D. J., Jones, D. C., and Fang, D. D., 2016. A MAGIC population-based genome-wide association study reveals functional association of GhRBB1_A07 gene with superior fiber quality in cotton, BMC Genomics. 17, no. 1, pp. 903-903,

[67] Lagudah, E., Krattinger, S., Herrera-Foessel, S., Singh, R., Huerta-Espino, J., Spielmeyer, W., Brown-Guedira, G., Selter, L., and Keller, B., 2009. Gene-specific markers for the wheat gene Lr34/Yr18/Pm38 which confers resistance to multiple fungal pathogens, Theoretical and Applied Genetics. 119, no. 5, pp. 889-898

[68] Wu, J., Zhao, Q., Liu, S., Shahid, M., Lan, L., Cai, G., Zhang, C., Fan, C., Wang, Y., and Zhou, Y., 2016. Genome-wide Association Study Identifies New Loci for Resistance to Sclerotinia Stem Rot in Brassica napus, Frontiers in Plant Science. 7, pp. 1418-1418,

[69] Bernardo, A., Ma, H., Zhang, D., and Bai, G., 2012. Single nucleotide polymorphism in wheat chromosome region harboring Fhb1 for Fusarium head blight resistance, Molecular Breeding. 29, no. 2, pp. 477-488

[70] Yu, H., Xie, W., Wang, J., Xing, Y., Xu, C., Li, X., Xiao, J., and Zhang, Q., 2011. Gains in QTL Detection Using an Ultra-High Density SNP Map Based on Population Sequencing Relative to Traditional RFLP/SSR Markers, PLoS ONE. 6, no. 3, pp. e17595,

[71] Kim, K.-S., Hill, C., Hartman, G., Hyten, D., Hudson, M., and Diers, B., 2010. Fine mapping of the soybean aphid-resistance gene Rag2 in soybean PI 200538, Theoretical and Applied Genetics. 121, no. 3, pp. 599-610

[72] Li, F., Kitashiba, H., Inaba, K., and 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 Research. 16, no. 6, pp. 311-

323,

[73] McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. A., and Hirschhorn, J. N., 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges, Nature Reviews Genetics. 9, no. 5, pp. 356-369

[74] Matukumalli, L. K., Lawley, C. T., Schnabel, R. D., Taylor, J. F., Allan, M. F., Heaton, M. P., O'Connell, J., Moore, S. S., Smith, T. P. L., Sonstegard, T. S., and Van Tassell, C. P., 2009. Development and Characterization of a High Density SNP Genotyping Assay for Cattle, PLoS ONE. 4, no. 4, pp. e5350,

[75] Rasheed, A., Hao, Y., Xia, X., Khan, A., Xu, Y., Varshney, R. K., and He, Z., 2017. Crop Breeding Chips and Genotyping Platforms: Progress, Challenges, and Perspectives, Molecular plant. 10, no. 8, pp. 1047-1064

[76] Hu, Z., Huang, S., Sun, M., Wang, H., and Hua, W., 2012. Development and application of single nucleotide polymorphism markers in the polyploid Brassica napus by 454 sequencing of expressed sequence tags, Plant Breeding. 131, no. 2, pp. 293-299,

[77] Tollenaere, R., Hayward, A., Dalton-Morgan, J., Campbell, E., Lee, J. R. M., Lorenc, M. T., Manoli, S., Stiller, J., Raman, R., Raman, H., Edwards, D., and Batley, J., 2012. Identification and characterization of candidate Rlm4 blackleg resistance genes in Brassica napus using next-generation sequencing, Plant Biotechnology Journal. 10, no. 6, pp. 709-715,

[78] Raman, R., Taylor, B., Marcroft, S., Stiller, J., Eckermann, P., Coombes, N., Rehman, A., Lindbeck, K., Luckett, D., Wratten, N., Batley, J., Edwards, D., Wang, X., and Raman, H., 2012. Molecular mapping of qualitative and quantitative loci for resistance to Leptosphaeria maculans causing blackleg disease in canola (Brassica napus L.), Theoretical and Applied Genetics. 125, no. 2, pp. 405-418