AGRES – An International e-Journal , (2017) Vol. 6, Issue 1: 1-15 **ISSN 2277-9663**

GENOMIC TOOLS IN CROP BREEDING: A REVIEW

***[1]RAMESH KUMAR AND [2]MEHTA, D. R.**

DEPARTMENT OF GENETICS AND PLANT BREEDING
JUNAGADH AGRICULTURAL UNIVERSITY
JUNAGADH, GUJARAT– 362 001 (INDIA)

***E-MAIL: rameshr1005@gmail.com**

[1]Ph.D. Scholar, Department of Genetics and Plant Breeding, Junagadh Agricultural University, Junagadh, Gujarat– 362 001 (India)
[2]Assosiate Professor, Department of Genetics and Plant Breeding, Junagadh Agricultural University, Junagadh, Gujarat– 362 001 (India)

ABSTRACT

Plant breeding has been very successful in development of Climate Resilient and improved varieties using conventional tools and methodologies. Combination of various techniques which includes novel genetic tools and modern genetic and breeding approaches should increase efficiency and precision that have a great potential to impact crop breeding. Next Generation Sequencing (NGS) technologies are allowing the sequencing of genomes and transcriptomes, which is producing a vast array of genomic information with high precision. The analysis of NGS data by means of bioinformatics developments allows discovering new genes and regulatory sequences and their positions, and makes available large collections of molecular markers. TILLING and EcoTILLING make possible to screen mutant and germplasm collections for allelic variants in target genes. Advanced backcross QTL (AB-QTL) analysis is a potential solution by combining the discovery and transfer of valuable QTLs from wild germplasm into elite breeding lines into a single process. To comprehensively understand functional genomics regarding overall plant development, the advanced tools of metabolomics, together with QTL analysis, GWAS and knock-out/down technology have great importance. In conclusion, recent advances in genomics are providing plant breeders with new tools and methodologies that allow a great leap forward in plant breeding and the genetic dissection and breeding for complex traits.

KEY WORDS: AB-QTL, EcoTILLING, Genomics, Metabolomics, TILLING, Transcriptomes,

INTRODUCTION

Ever since the beginnings of the domestication of plants and 10,000 years ago, plant breeding has been extremely successful in developing crops and varieties that have contributed to the development of modern societies, and have successively beaten (neo-) Malthusian predictions (Fedoroff, 2010). It has been predicted for over two decades that molecular marker technology would reshape breeding programmes and facilitate rapid gains from selection (Stuber *et al.*, 1982; Tanksley *et al.* 1982). Currently, however, marker-assisted selection (MAS) has failed to significantly improve polygenic traits (Bernardo, 2008; Xu, Y. and Crouch, 2008). While MAS has been effective for the manipulation of large effect alleles with known association to a marker (Zhong *et al.*, 2006), it has been at an impasse when many alleles of small effect segregate and no substantial, reliable effects can be identified (Moreau *et al.*, 2004).

Nowadays, genomics provides plant breeders with a new set of tools and techniques that allow the study of the whole facilitating the direct study of the genotype and its relationship with the phenotype (Tester and Langridge, 2010). Genomes of plants have been subjected to structural as well as functional genomics research, which during the last two decades covered both basic and applied aspects of research. The combination of conventional breeding techniques along with genomic tools and approaches is leading to a new genomics-based plant breeding, also known as integrated plant breeding platform. In this new plant breeding context, genomics will be essential to develop more efficient plant cultivars, which are necessary, according to FAO, for the new 'greener revolution' needed to feed the world’s growing population while preserving natural resources (Perez-de-Castro *et al.*, 2012).

Advances in genomics can also contribute to crop improvement in two general ways. First, a better understanding of the biological mechanisms can lead to new or improved screening methods for selecting superior genotypes more efficiently. Second, new knowledge can improve the decision-making process for more efficient breeding strategies (Varshney *et al.*, 2005). One of the main pillars of genomic breeding is the development of high-throughput DNA sequencing technologies, collectively known as next generation sequencing (NGS) methods. These and other technical revolutions provide genome-wide molecular tools for plant breeders (large collections of markers, high-throughput genotyping strategies, high density genetic maps, new experimental populations) that can be incorporated into existing breeding methods (Varshney and Tuberosa, 2007a; Varshney and Tuberosa, 2007b; Tester and Langridge, 2010; Lorenz *et al.*, 2011).

Recent advances in genomics are producing new plant breeding methodologies, improving and accelerating the breeding process in many ways (e.g., association mapping, marker assisted selection, breeding by design, gene pyramiding, genomic selection, etc.) (Peleman and van der Voort, 2003; Collard and Mackill, 2008). Genomic tools are thus, facilitating the detection of QTLs and the identification of existing favourable alleles of small effect, which have frequently remained unnoticed and have not been included in the gene pool used for breeding purpose (Morgante and Salamini, 2003; Vaughan *et al.*, 2007).

In addition to above mentioned approaches, some novel technologies e.g. TILLING, ecotype TILLING (Eco TILLING), Genome Wide Association (GWA) and Sequencing of RNA Transcripts (RNA-seq) technologies have also emerged during last decade that are considered to have greater impact on plant genetics research and breeding programmes. This article provides an overview on some selected genomics technologies, their potential and limitations for accelerating crop improvement programmes.

1.Sequencing

First-Generation Sequencing

The automated Sanger method is considered as a ‘first-generation’ technology. It has been used to sequence many genomes as well as several transcriptomes. The first international collaborative project resulted in the whole genome sequence of the model plant *Arabidopsis thaliana* (The Arabidopsis Genome Initiative. 2010) and then number of crops such as rice (International Rice Genome Sequencing Project. (2005), maize (Schnable *et al.*, 2009), sorghum (Paterson *et al.*, 2009) and soybean (Schmutz *et al.*, 2010). The transcriptomes of most major crops, to a greater or lesser extent, were also sequenced in recent past.

Next-Generation Sequencing (NGS)

Despite many technical improvements during this era, the limitations of automated Sanger sequencing showed a need for new and improved technologies for sequencing large numbers of human genomes. Recent efforts have been directed towards the development of new methods, leaving Sanger sequencing with fewer reported advances (Metzker, 2005; Hutchison, 2007) Sequencing technologies include a number of methods that are grouped broadly as template preparation, sequencing and imaging, and data analysis. The unique combination of specific protocols distinguishes one technology from another and determines the type of data produced from each platform.

Moreover, new "**third generation**" platforms are being developed and incorporated to sequencing projects, such as PacBio (Pacific Biosciences), Helicos or Ion Torrent. Many transcriptomes have also been sequenced, a number of them in several species such as sweet potato (Wang *et al.*, 2011) or buckwheat (Logacheva *et al.*, 2011) for which no previous sequence information was available. These assays are showing the great complexity of plant transcriptomes, allowing the identification of rare transcript variants that are being used to improve gene explanation and our knowledge with related to gene function and regulation.

Bioinformatics

NGS technologies are facilitating sequencing projects, but have brought new challenges, as millions of short DNA reads have to be analysed and assembled (Metzker, 2010). Therefore, it is necessary to develop new bioinformatics tools (algorithms and software), which allow the analyses of huge amounts of genome-wide data, but it is also necessary to change the approaches used to understand complex biological traits (Pop and Salzberg, 2008; Horner *et al.*, 2009). Two of the most common analyses carried out on these NGS reads are genome assembly and annotation and mapping. Genome assembly is a complex task requiring powerful computers and skilled bioinformaticians (Pop and Salzberg, 2008). Several bioinformatic tools and databases (Table 1) have been developed for DNA sequence analysis, marker discovery and querying and analyzing information.

2. TILLING and EcoTILLING

Progress in plant breeding in terms of development of superior and high yielding varieties of agricultural crops is possible by accumulation of beneficial alleles from vast plant genetic resources existing worldwide. But still, a significant portion of these superior alleles cannot be used, because those alleles are left behind during evolution and domestication (Reddy *et al.*, 2014). This untapped genetic variation existing in wild relatives and land races of crop plants could be exploited gainfully for development of agronomically superior cultivars. Introgressions of novel alleles from wild relatives of crop plants into cultivated varieties (McCouch *et al.*, 2007) have clearly demonstrated that certain alleles and their combinations potentially make dramatic changes in trait expression when moved to a suitable genetic background by overcoming the genetic bottlenecks which restricted their introgression to cultivars.

TILLING (Targeting induced local lesions in genomes), a newly developed general reverse genetic strategy, helps to locate an allelic series of induced point mutations in genes of interest. It allows the rapid and inexpensive detection of induced point mutations in populations of physically/chemically mutagenized individuals. In addition to allowing efficient detection of mutations by TILLING approach, EcoTILLING technology is also ideal for examining natural variation (Rashid *et al.*, 2011). Allele mining can also be used for screening and detection of plants with desired characters by knockdown and knockout mutations in specific genes, which makes TILLING and EcoTILLING as an attractive strategy for a wide range of applications from the basic functional genomic study to practical crop breeding approaches.

Success of the identification of present variation useful for breeding programmes will depend on the right identification of target genes. The availability of sequences coming from sequencing projects and the information provided by gene expression studies is significantly increasing the number and quality of candidates for TILLING and EcoTILLING studies. TILLING has been successfully used in Arabidopsis (Colbert *et al.*, 2001), Lotus (Perry *et al.*, 2003), barley (Caldwell *et al.*, 2004) and maize (Weil and Monde, 2007). EcoTILLING was first applied in Rice (Kadaru *et al.*, 2006) and subsequently, in barley (Mejlhede *et al.*, 2006) and wheat (Wang *et al.*, 2008)

3. Advancement in QTL Analysis

AB-QTL

Plant breeding involving wild species, there are two common weaknesses: first, the population segregates for a large percentage of genes from the wild parent, thus resulting in lower statistical power to detect QTLs with small effects (Chee *et al.*, 2005). Secondly, once potentially valuable QTLs are discovered, substantial backcrossing and intercrossing are likely to be required for the development of commercial cultivars. Separating QTL discovery and cultivar development into discrete and sequential steps not only increases the time required for new cultivar development, but also reduces the likelihood that the QTL information is used to create a superior crop cultivar (Tanksley and Nelson, 1996). Advanced backcross QTL (AB-QTL) analysis is a potential solution by combining the discovery and transfer of valuable QTLs from wild germplasm into elite breeding lines in a single process. In the advanced backcross (AB) design, QTL analysis will be delayed until a later generation like the BC_2 or BC_3 generation. The logic behind this approach is that the effect of individual QTLs can be more precisely measured because undesirable effects of wild species on the elite background are reduced since later generation progenies such as BC_2 or BC_3 carry a smaller number of genes from the donor parent (Baohua and Chee, 2010). AB-QTL has been successfully used in tomato (Stevens *et al.*, 2007), rice (Manosalva *et al.*, 2009), wheat (Naz *et al.*, 2008) and maize (Mano and Omori, 2008).

mQTL and mGWAS

Plants produce large numbers of metabolites of diversified structures and abundance that play important roles in plant growth, development and varied response to environments. These diverse small molecular weight metabolites, the chemical base of crop yield and quality, are also valuable nutrition and energy sources for human beings and live stocks (Hall *et al*., 2008). Although metabolomics is downstream of the other functional genomics (transcriptomics and proteomics), the practical size of the metabolome of a species, unlike transcriptome or proteome, cannot be speculated directly by known genomic information *via* central dogma. Therefore, metabolomics is used to obtain a large amount of valuable information for the discovery of genes and pathways through accurate and high throughput corollary peak annotation *via* snapshotting the plant metabolome (Tohge *et al*., 2014).

With the advance of sequencing technology, dozens of plant species have been sequenced. To comprehensively understand the functional genomics regarding plant development, importance of advanced tools of metabolomics, together with QTL (quantitative trait locus) analysis, GWAS (genome-wide association study), and knock-out/down technology, has been increasingly recognized within the plant science community (Hong *et al*., 2016). In *Arabidopsis*, the analysis of 369 recombinant inbred lines and 41 introgression lines indicated that the metabolite heterosis is primarily contributed by epistasis (Schauer *et al*., 2006). In tomato, metabolite profiling in seeds of 76 introgression lines in two consecutive harvest seasons revealed the presence of 30 metabolite quantitative trait loci (mQTLs) and dissected partial mechanisms, underlying the variational contents of main primary metabolites (Lisec *et al*., 2009). Similar mQTL analyses have been performed in other plant species, such as wheat (Hill *et al*., 2015), rice (Matsuda *et al*., 2012) and rape (Feng *et al*., 2012).

Marker- Assisted Backcrossing

Once the markers associated with a trait of interest is identified through linkage mapping, association mapping, AB-QTL or transcriptomics approach, the next step is to use these markers in actual breeding programme (Utomo and Linscombe, 2009). In this context, the selection of one or a few genes (QTLs) through molecular markers using backcrossing is a highly efficient technique (Collard and Mackill, 2008). There are three levels of MABC (i) **foreground selection (**Hospital, F. and Charcosset, 1997), which includes screening of target gene or QTL using molecular markers, this step can also be used for selection of recessive allele for backcrossing as recessive alleles require one generation of selfing for its expression, (ii) **recombinant selection** involves selection of the BC progeny containing the target gene and recombination events (between the target locus and linked flanking markers). The purpose of this selection step is to minimize the '**linkage drag'** by using markers that flank the target gene. This linkage drag poses a big problem during selection through conventional breeding methods. Furthermore this recombination selection event is usually carried out using two BC generations (Frisch et al., 1999), and (iii**) background selection** involves use of markers that are unlinked to the target locus for the selection of BC progeny containing the highest proportion of recurrent parent (RP).

4. Genomic selection

The weaknesses of traditional MAS come from the way MAS splits the task into two components, first identifying QTL and then estimating their effects (Jannink *et al.,* 2010). QTL identification methods can make MAS poorly suited to crop improvement: (i) biparental populations may be used that are not representative and in any event do not have the same level of allelic diversity and phase as the breeding programme as a whole (Jannink *et al.*, 2001; Sneller *et al.*, 2009); (ii) the necessity of generating such populations is costly such that the populations may be small and therefore, underpowered; (iii) validation of discoveries is then warranted, requiring additional effort; (iv) the separation of QTL identification from estimation means that estimated effects will be biased (Beavis, 1994; Melchinger *et al.*, 1998; Schon *et al.*, 2004), and small-effect QTL will be missed entirely (Lande and Thompson, 1990) as a result of using stringent significance thresholds.

To minimize the limitations for successful MAS, Lande and Thompson (1990) proposed a visionary two steps approach: (i) select significant markers from large marker sets, and (ii) combine phenotypic information with significant markers in a selection index that would explain a significant proportion of additive genetic variance. In the first step, they were unable to estimate all marker effects simultaneously with simple regression due to the lack of degrees of freedom. Therefore, they proposed selecting the most significant markers from the previous generation *via* multiple linear regressions and then re-estimating effects of the selected markers in the current generation with independent multiple regressions. Genomic selection is a form of MAS that simultaneously estimates all locus, haplotype, or marker effects across the entire genome to calculate genomic estimated breeding values (GEBVs) (Meuwissen *et al.*, 2001). This approach contrasts greatly with traditional MAS because there is not a defined subset of significant markers used for selection. Instead, GS analyzes jointly all markers on a population attempting to explain the total genetic variance with dense genome wide marker coverage through summing marker effects to predict breeding value of individuals (Meuwissen *et al.*, 2001). The central process of GS is the calculation GEBVs for individuals having only genotypic data using a model that was “trained” from individuals having both phenotypic and genotypic data (Figure 1) (Meuwissen *et al.*, 2001). The population of individuals with both phenotypic and genotypic data is known as the “training population” as it is used to estimate model parameters that will subsequently be used to calculate GEBVs of selection candidates (e.g., breeding lines) having only genotypic data (Figure 1). These GEBVs are then used to select the individuals for advancement in the breeding cycle. Therefore, selection of an individual without phenotypic data can be performed by using a model to predict the individual’s breeding value (Meuwissen *et al.*, 2001). To maximize GEBV accuracy, the training population must be representative of selection candidates in the breeding program to which GS will be applied.

A selection index integrates and weights multiple traits to achieve greater gains than if traits with independent thresholds are individually or collectively selected (Hazel and Lush, 1942; Hazel, 1943). Selection indices can incorporate marker data as indirect selection traits ((Meuwissen *et* *al.*, 2001; Neimann-Sorensen and Robertson, 1961; Smith, 1967). However, current MAS applied to loci selected by SR violates the selection index assumptions of multivariate normality and small changes in allele frequencies because selection is based on only few large effect loci ((Meuwissen *et al.*, 2001; Dekkers, 2007). Because GS is based on many markers distributed throughout the genome, index selection assumptions are met, providing an opportunity to use index selection theory to predict response to GS (Dekkers, 2007).

Traditional marker-assisted selection has been ineffective for complex traits. The introduction of genomic selection (GS), however, has shifted that paradigm. Rather than seeking to identify individual loci significantly associated with a trait, GS uses all marker data as predictors of performance and consequently delivers more accurate predictions. Selection can be based on GS predictions, potentially leading to more rapid and lower cost gains from breeding (Jannink *et al.*, 2010).

Integrated Plant Breeding

Genomics research is generating new tools, such as functional molecular markers and bioinformatics, as well as new knowledge about statistics and inheritance phenomena that could increase the efficiency and precision of identification of QTL. Sequencing can allow the identification of rare transcript variants that should be used for improving gene explanation. TILLING and EcoTILLING an attractive strategy with genome sequencing can generate and identify new alleles, which become source of variation in breeding population. Combination of various techniques showing in Figure 2 (Varshney *et al.*, 2005), which includes novel genetic tools and modern genetic and breeding approaches increase efficiency and precision that have a great potential to impact crop breeding. Due to reduced costs on sequencing and genotyping combined with advances in biometrics and bioinformatics, we envisage a bright future on applications of these novel approaches in plant breeding.

ABBREVIATIONS

AB-QTL = Advanced-backcross QTL
EcoTILLING = Ecotype TILLING
GEBV = Genomic estimated breeding value
GS = Genomic selection
GWA = Genome-wide association
MABC = Marker assisted back crossing
MAS = Marker assisted selection
mGWAS = Metabolomics genome-wide association study
mQTL = Metabolomics QTL
NGS = Next generation sequencing
QTL = Quantitative Trait Locus
TILLING = Targeting induced local lesions on genomes

CONCLUSION

Recent advances in genomics are providing plant breeders with new tools and methodologies that allow a great leap forward in plant breeding and the genetic dissection and breeding for complex traits.

REFERENCES

Baohua, W. W. and Chee, P. W. (2010). Application of advanced backcross quantitative trait locus (QTL) analysis in crop improvement. *J. Plant Breed. Crop Sci.*, **2**(8): 221-232.

Beavis, W. D. (1994). The power and deceit of QTL experiments: lessons from comparative QTL studies. In: Wilkinson DB (ed). *Proceedings of the 49th Annual Corn and Sorghum Research Conference*. Washington, DC:

American Seed Trade Association, pp. 250-265.

Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci.*, **48**: 1649-1664.

Caldwell, D. G.; McCallum, N.; Shaw, P.; Muehlbauer, G. J.; Marshall, D.F. and Waugh, R. (2004). A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L). *Plant J.*, **40**: 143-150.

Chee, P.; Draye, X.; Jiang, C. X.; Decanini, L.; Delmonte, T. A.; Bredhauer, R.; Smith, C. W. and Paterson, A. H. (2005). Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcross-self approach: I. Fiber elongation. *Theor. Appl. Genet.*, **111**: 757-763.

Colbert, T.; Till, B. J.; Tompa, R.; Reynolds, S.; Steine, M. N.; Yeung, A. T.; McCallum, C. M.; Comai, L. and Henikoff, S. (2001). High throughput screening for induced point mutations. *Plant Physiol.*, **126**: 480-484.

Collard, B. C. Y. and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Phil. Transact. Royal Soc. B.*, **363**: 557-572.

Dekkers, J. C. M. (2007). Prediction of response to marker-assisted and genomic selection using selection index theory. *J. Anim. Breed. Genet.*, **124**: 331-341.

Fedoroff, N. V. (2010). The past, present and future of crop genetic modification. *New Biotechnol.*, **27**: 461-465.

Feng, J.; Long, Y.; Shi, L.; Shi, J.; Barker, G. and Meng, J. (2012). Characterization of metabolite quantitative trait loci and metabolic networks that control glucosinolate concentration in the seeds and leaves of *Brassica napus*. *New Phytol.*, **193**: 96-108.

Frisch, M.; Bohn, M. and Melchinger, A. E. (1999). Comparison of selection strategies for marker assisted backcrossing of a gene. *Crop Sci.*, 39: 1296-1301.

Hall, R. D.; Brouwer, I. D. and Fitzgerald, M. A. (2008). Plant metabolomics and its potential application for human nutrition. *Physiol. Plant*., **132**: 162-175.

Hazel, L. (1943). The genetic basis for constructing selection indexes. *Genetics*, **28**: 476-490.

Hazel, L. and Lush, J. L. (1942). The efficiency of three methods of selection. *J. Hered.*, **33**: 393-399.

Hill, C. B.; Taylor, J. D.; Edwards, J.; Mather, D.; Langridge, P.; Bacic, A. and Roessner, U. (2015). Detection of QTL for metabolic and agronomic traits in wheat with adjustments for variation at genetic loci that affect plant phenology. *Plant Sci.* **233**: 143-154.

Hong, J.; Yang, L.; Zhang, D. and Shi, J. (2016). Plant Metabolomics: An Indispensable System Biology Tool for Plant Science. *Int. J. Mol. Sci.*, **17**: 767.

Horner, D. S.; Pavesi, G.; Castrignanò, T.; Meo, P.D.O.D.; Liuni, S.; Sammeth, M.; Picardi, E. and Presole, G. (2009). Bioinformatics approaches for genomics and post genomics applications of next generation sequencing. *Brief. Bioinform*., **2**: 181-197.

Hospital, F. and Charcosset. (1997). Marker assisted introgression of quantitative trait loci. *Genetics*, **147**: 1469-1485.

Hutchison, C. A. III. 2007. DNA sequencing: bench to bedside and beyond. *Nucleic Acids Res*., **35**: 6227-6237.

International Rice Genome Sequencing Project. (2005). The map-based sequence of the rice genome. *Nature*, **436**: 793-800.

Jannink J. L.; Bink, M. C. A. M. and Jansen, R. C. (2001). Using complex plant pedigrees to map valuable genes. *Trends Plant Sci*., **6**: 337-342.

Jannink, J. L.; Lorenz, A. J. and Iwata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Brief. Funct. Genomics*, **9**: 166-177.

Kadaru, S. B.; Yadav, A. S.; Fjellstrom, R. G. and Oard, J. H. (2006). Alternative EcoTILLING protocol for rapid, cost-effective single-nucleotide polymorphism discovery and genotyping in rice (*Oryza sativa* L.). *Plant Mol. Biol. Reporter*, **24**: 3-22.

Lande, R. and Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, **124**: 743-756.

Lisec, J.; Steinfath, M.; Meyer, R. C.; Selbig, J.; Melchinger, A. E.; Willmitzer, L. and Altmann, T. (2009). Identification of heterotic metabolite QTL in *Arabidopsis thaliana* RIL and IL populations. *Plant J*., **59**: 777-788.

Logacheva, M. D.; Kasianov, A. S.; Vinogradov, D. V.; Samigullin, T. H.; Gelfand, M.S.; Makeev, V. J. and Penin, A.A. (2011). De novo sequencing and characterization of floral transcriptome in two species of buckwheat (Fagopyrum). *BMC Genomics*, 12: 30.

Lorenz, A. J.; Chao, S.; Asoro, F. G.; Heffner, E. L.; Hayashi, T.; Iwata, H.; Smith, K. P.; Sorrells, M. K. and Jannink, J. L. (2011). Genomic selection in plant breeding: knowledge and prospects. *Adv. Agron*., **110:** 77-123.

Mano, Y. and Omori, F. (2008). Verification of QTL controlling root aerenchyma formation in a maize × teosinte "*Zea nicaraguensis*" advanced backcross population. *Breed. Sci*., **58**: 217-223.

Manosalva, P. M.; Davidson, R. M.; Liu, B.; Zhu, X.; Hulbert, S. H.; Leung, H. and Leach, J. E. (2009). A germin-like protein gene family functions as a complex quantitative trait locus conferring broad-spectrum disease resistance in rice. *Plant Physiol*., **149**: 286-296.

Matsuda, F.; Okazaki, Y.; Oikawa, A.; Kusano, M.; Nakabayashi, R.; Kikuchi, J.; Yonemaru, J. I.; Ebana, K.; Yano, M. and Saito, K. (2012). Dissection of genotype-phenotype associations in rice grains using metabolome quantitative trait loci analysis. *Plant J*. **70**: 624-636.

McCouch, S. R.; Sweeney, M.; Li, J.; Jiang, H.; Thomson, M. and Septiningsih, E. (2007). Through the genetic bottleneck: O. rufipogon as a source of trait-enhancing alleles for O. sativa. *Euphytica*, **154**: 317-339.

Mejlhede, N.; Kyjovska, Z.; Backes, G.; Burhenne, K.; Rasmussen, S. K. and Jahoor, A. (2006). EcoTILLING for the identification of allelic variation in the powdery mildew resistance genes mlo and Mla of barley. *Plant Breed*., **125**: 461-467.

Melchinger, A. E.; Utz, H. F. and Schon, C. C. (1998). Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics*, **149**: 383-403.

Metzker, M. (2010). Sequencing technologies-the next generation. *Nat. Rev. Genet*., **11**: 31.

Metzker, M. L. 2005. Emerging technologies in DNA sequencing. *Genome Res*. **15**: 1767-1776.

Meuwissen, T. H. E.; Hayes, B. J. and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, **157**: 1819–1829.

Moreau, L.; Charcosset, A. and Gallais, A. (2004). Experimental evaluation of several cycles of marker-assisted selection in maize. *Euphytica*, **137**: 111-118.

Morgante, M. and Salamini, F. (2003). From plant genomics to breeding practice. *Curr. Opin. Biotechnol*., **14**: 214-219.

Naz, A. A.; Kunert, A.; Lind, V.; Pillen, K. and Leon, J. (2008). AB-QTL analysis in winter wheat: II. Genetic analysis of seedling and field resistance against leaf rust in a wheat advanced backcross population. *Theor. Appl. Genet*., **116**: 1095-1104.

Neimann-Sorensen, A. and Robertson. A. (1961). The association between blood groups and several production characteristics in three Danish cattle breeds. *Acta Agric. Scand*., **11**: 163-196.

Paterson, A. H.; Bowers, J. E.; Bruggmann, R.; Dubchak, I.; Grimwood, J.; Gundlach, H.; Haberer, G.; Hellsten, U.; Mitros, T.; Poliakov, A.; Schmutz, J.; Spannagl, M.; Tang, H.; Wang, X.; Wicker, T.; Bharti, A. K.; Chapman, J.; Feltus, F. A.; Gowik, U.; Grigoriev, I. V.; Lyons, E.; Maher, C. A.; Martis, M.; Narechania, A.; Otillar, R. P.; Penning, B. W.; Salamov, A. A.; Wang, Y.; Zhang, L.; Carpita, N. C.; Freeling, M.; Gingle, A. R.; Hash, C. T.; Keller, B.; Klein, P.; Kresovich, S.; McCann, M. C.; Ming, R.; Peterson, D. G.; Rahman, M.; Ware, D.; Westhoff, P.; Mayer, K. F. X.; Messing, J. and Rokhsar, D. S. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature*, **457**: 551-556.

Peleman, J. D. and van der Voort, J. R. (2003). Breeding by design. *Trends Plant Sci*., **8**: 330-334.

Perez-de-Castro, A. M.; Vilanova, S.; Canizares, J.; Pascual, L.; Blanca, J. M.; Diez, M. J.; Prohens, J. and Pico, B. (2012). Application of Genomic Tools in Plant Breeding. *Curr. Genomics*, **13**: 179-195.

Perry, J. A.; Wang, T. L.; Welham, T. J.; Gardner, S.; Pike, J. M.; Yoshida, S. and Parniske, M. (2003). A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume Lotus japonicus. *Plant Physiol.*, **131**: 866-871.

Pop, M. and Salzberg, S. L. (2008). Bioinformatics challenges of new sequencing technology. *Trends Genet.*, **24**: 142-149.

Rashid,, M.; He, H.; Guanxiao, Y. and Khurram, Z. (2011). Relevance of tilling in plant genomics. *AJCS.*, **5**(4): 411-420.

Reddy, A. M.; Gambhire, V. B. and Reddy, R. T. (2014). Allele mining in crop improvement. *Int. J. Develop. Res.*, **4**(2): 300-305.

Schauer, N.; Semel, Y.; Roessner, U.; Gur, A.; Balbo, I.; Carrari, F.; Pleban, T.; Perez-Melis, A.; Bruedigam, C. and Kopka, J. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.*, **24**: 447-454.

Schmutz, J.; Cannon, S. B.; Schlueter, J.; Ma, J.; Mitros, T.; Nelson, W.; Hyten, D. L.; Song, Q.; Thelen, J. J.; Cheng, J.; Xu, D.; Hellsten, U.; May, G. D.; Yu, Y.; Sakurai, T.; Umezawa, T.; Bhattacharyya, M. K.; Sandhu, D.; Valliyodan, B.; Lindquist, E.; Peto, M.; Grant, D.; Shu, S.; Goodstein, D.; Barry, K.; Futrell- Griggs, M.; Abernathy, B.; Du, J.; Tian, Z.; Zhu, L.; Gill, N.; Joshi, T.; Libault, M.; Sethuraman, A.; Zhang, X. C.; Shinozaki, K.; Nguyen, H. T.; Wing, R. A.; Cregan, P.; Specht, J.; Grimwood, J.; Rokhsar, D.; Stacey, G.; Shoemaker, R. C. and Jackson, S. A. (2010). Genome sequence of the palaeopolyploid soybean. *Nature*, **463**: 178- 83.

Schnable, P. S.; Ware, D.; Fulton, R. S.; Stein, J. C.; Wei, F.; Pasternak, S.; Liang, C.; Zhang, J.; Fulton, L.; Graves, T. A.; Minx, P.; Reily, A. D.; Courtney, L.; Kruchowski, S. S.; Tomlinson, C.; Strong, C.; Delehaunty, K.; Fronick, C.; Courtney, B.; Rock, S. M.; Belter, E.; Du, F.; Kim, K.; Abbott, R. M.; Cotton, M.; Levy, A.; Marchetto, P.; Ochoa, K.; Jackson, S. M.; Gillam, B.; Chen, W.; Yan, L.; Higginbotham, J.; Cardenas, M.; Waligorski, J.; Applebaum, E.; Phelps, L.; Falcone, J.; Kanchi, K.; Thane, T.; Scimone, A.; Thane, N.; Henke, J.; Wang, T.; Ruppert, J.; Shah, N.; Rotter, K.; Hodges, J.; Ingenthron, E.; Cordes, M.; Kohlberg, S.; Sgro, J.; Delgado, B.; Mead, K.; Chinwalla, A.; Leonard, S.; Crouse, K.; Collura, K.; Kudrna, D.; Currie, J.; He, R.; Angelova, A.; Rajasekar, S.; Mueller, T.; Lomeli, R.; Scara, G.; Ko, A.; Delaney, K.; Wissotski, M.; Lopez, G.; Campos, D.; Braidotti, M.; Ashley, E.; Golser, W.; Kim, H.; Lee, S.; Lin, J.; Dujmic, Z.; Kim, W.; Talag, J.; Zuccolo, A.; Fan, C.; Sebastian, A.; Kramer, M.; Spiegel, L.; Nascimento, L.; Zutavern, T.; Miller, B.; Ambroise, C.; Muller, S.; Spooner, W.; Narechania, A.; Ren, L.; Wei, S.; Kumari, S.; Faga, B.; Levy, M. J.; McMahan, L.; Van Buren, P.; Vaughn, M. W.; Ying, K.; Yeh, C. T.; Emrich, S. J.; Jia, Y.; Kalyanaraman, A.; Hsia, A. P.; Barbazuk, W. B.; Baucom, R. S.; Brutnell, T. P.; Carpita, N. C.; Chaparro, C.; Chia, J. M.; Deragon, J. M.; Estill, J. C.; Fu, Y.; Jeddeloh, J. A.; Han, Y.; Lee, H.; Li, P.; Lisch, D. R.; Liu, S.; Liu, Z.; Nagel, D. H.; McCann, M. C.; SanMiguel, P.; Myers, A. M.; Nettleton, D.; Nguyen, J.; Penning, B. W.; Ponnala, L.; Schneider, K. L.; Schwartz, D. C.; Sharma, A.; Soderlund, C.; Springer, N. M.; Sun, Q.; Wang, H.; Waterman, M.; Westerman, R.; Wolfgruber, T. K.; Yang, L.; Yu, Y.; Zhang, L.; Zhou, S.; Zhu, Q.; Bennetzen, J. L.; Dawe, R. K.; Jiang, J.; Jiang, N.; Presting, G. G.; Wessler, S. R.; Aluru, S.; Martienssen, R. A.; Clifton, S. W.; McCombie, W. R.; Wing, R. A. and Wilson, R. K. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Sci.*,**326**: 1112-1115.

Schon, C. C.; Utz H. F. and Groh, S. (2004). Quantitative trait locus mapping based on re-sampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. *Genetics*, **167**: 485-498.

Smith, C. (1967). Improvement of metric traits through specific genetic loci. *Anim. Prod.*, **9**: 349-358.

Sneller, C. H.; Mather, D. E. and, Crepieux, S. (2009). Analytical approaches and population types for finding and utilizing QTL in complex plant populations. *Crop Sci.*, **49**:363-380.

Stevens, R.; Buret, M.; Duffé, P.; Garchery, C.; Baldet, P.; Rothan, C. and Causse, M. (2007). Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three tomato populations. *Plant Physiol.*, **143**: 1943-1953.

Stuber C. W.; Goodman M. M. and Moll R. H. (1982). Improvement of yield and ear number resulting from selection at allozyme loci in a maize population. *Crop Sci.*, **22**: 737-740.

Tanksley, S. D. and Nelson, J. C. (1996). Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.*, **92**: 191-203.

Tanksley, S. D.; Young N. D. and Paterson A. H. (1989). RFLP mapping in plant breeding: new tools for an old science. *Biotechnol.*, **7**:257-264.

Tester, M. and Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Sci.*, ***327***: 818-822.

The Arabidopsis Genome Initiative. (2010). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana. Nature,* **408**: 796-815.

Tohge, T.; de Souza, L. P. and Fernie, A. R. (2014). Genome-enabled plant metabolomics. *J. Chromatogr.*, **966**: 7-20.

Utomo, H. S. and Linscombe, S. D. (2009). Current patents and future developments underlying marker assisted breeding in major grain crops, *Recent Pat DNA Gene Seq*, **3(1)**: 53-62.

Varshney, R. K. and Tuberosa, R. (2007a). *Genomics-Assisted Crop Improvement. Vol. 1: Genomics Approaches and Platforms.* Springer*;* New York.

Varshney, R. K. and Tuberosa, R. (2007b). *Genomics-Assisted Crop Improvement. Vol. 2: Genomics Approaches and Platforms.* Springer, New York.

Varshney, R. K.; Graner, A. and Sorrells, M. E. (2005). Genomics-assisted breeding for crop improvement. *Trends Plant Sci.*, **10**(12): 621-630.

Vaughan, D. A.; Balasz, E. and Heslop-Harrison, J. S. (2007). From crop domestication to super-domestication. *Ann. Bot.*, **100**: 893- 901.

Wang, J.; Sun, J. Z.; Liu, D. C.; Yang, W. L.; Wang, D. W.; Tong, Y. P. and Zhang, A. M. (2008). Analysis of Pina and Pinb alleles in the microcore collections of Chinese wheat germplasm by EcoTILLING and identification of a novel Pinb allele. *J. Cereal Sci.*, **48**: 836-842.

Wang, Z.; Fang, B.; Jingyi, C.; Zhang,X.; Luo, X.; Huang,L.; Chen, X.; Li, Y. 2011. De novo assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweetpotato (*Ipomoea batatas*). *BMC Genomics*, **11**: 726.

Weil, C. F. and Monde, R. A. (2007). Getting to the point - mutations in maize. *Crop Sci.*, **47**: S60-S67.

Xu, Y. and Crouch, J. H. (2008). Marker-assisted selection in plant breeding: from publications to practice. *Crop Sci.*, **48**: 391-407.

Zhong, S.; Toubia-Rahme, H. and Steffenson, B. J. (2006). Molecular mapping and marker-assisted selection of genes for septoria speckled leaf blotch resistance in barley. *Phytopathol.*, **96**: 993-999.

Table 1: List of some important bioinformatics tools and databases

Database	Description	URL
CotthonDB	Cotton information resource	http://cottondb.org/
CropNet	Genomic plant database	http://ukcrop.net/
Gramene	Grass information resource	http://www.gramene.org/
PlantMarkers	A database of predicted plant molecular markers	http://markers.btk.fi/
NCBI	Public databases and software tools	http://www.ncbi.nih.gov/
TASSEL	software package	http://www.maizegenetics.net/bioinformatics/tasselindex.htm

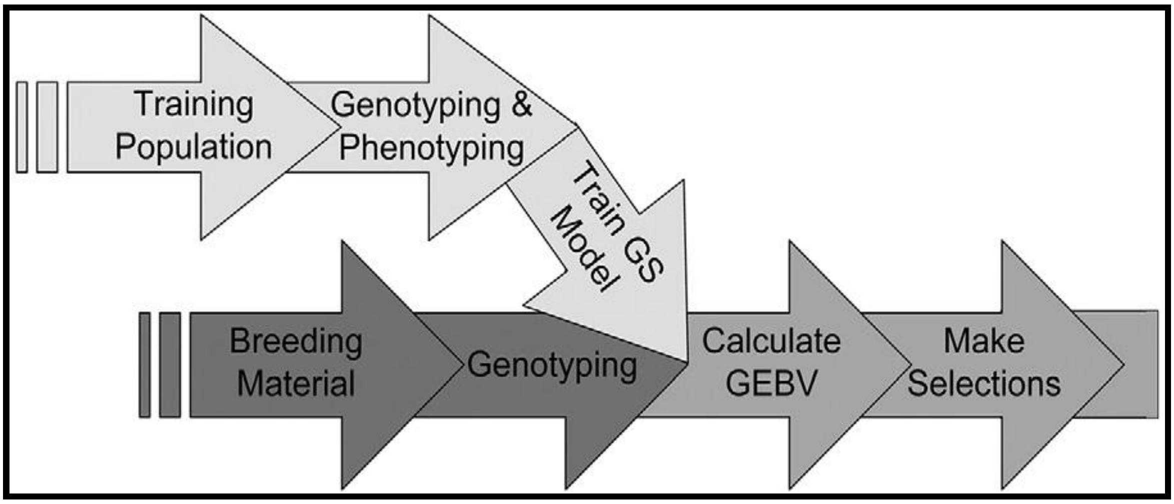

Figure 1: Genomic selection (GS) processes starting from the training population and selection candidates continuing through to genomic estimated breeding value (GEBV)–based selection

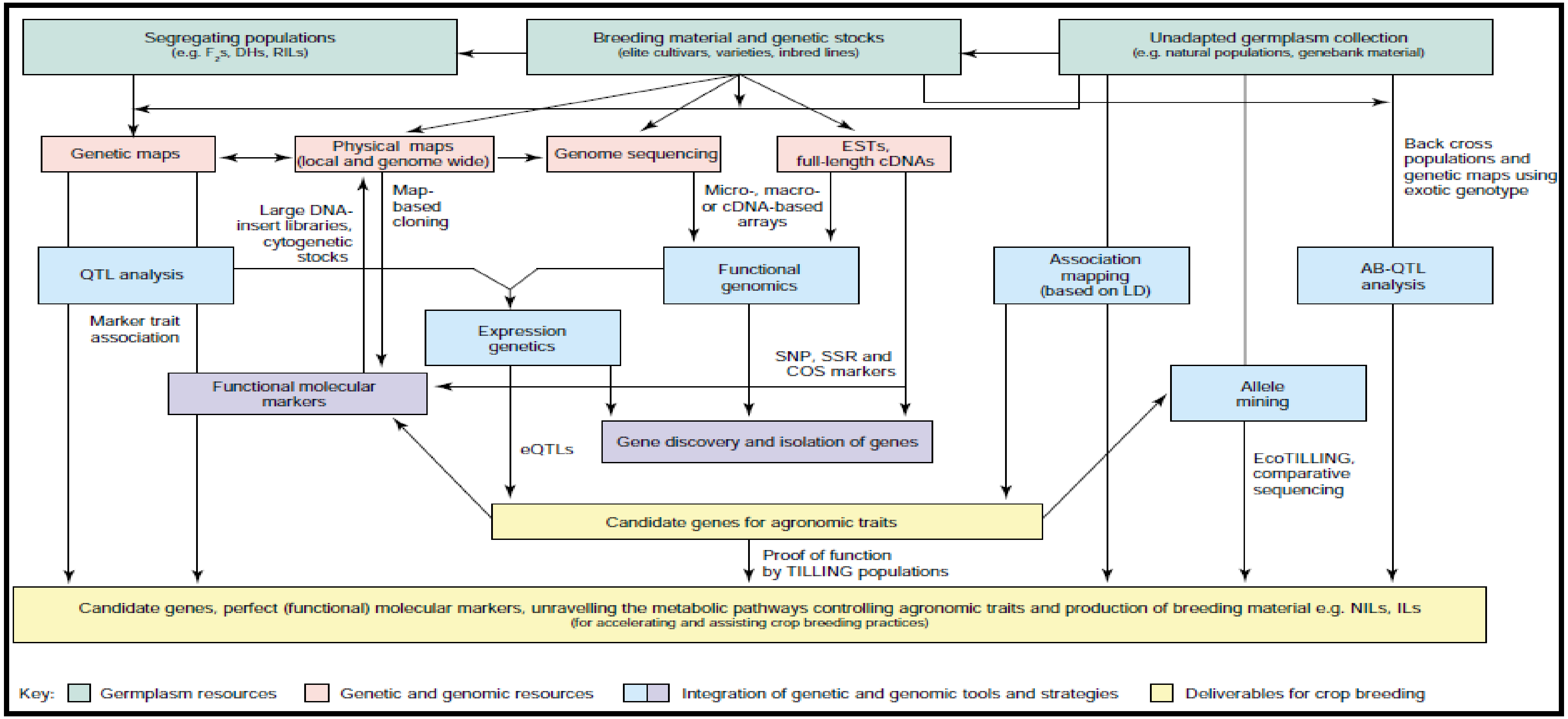

Figure 2: An integrated view of exploitation of genomic resources for crop improvement via different genetic and genomic strategies.

[MS received: January 23, 2017] *[MS accepted: February 26, 2017]*