NIGERIAN AGRICULTURAL JOURNAL
ISSN: 0300-368X
Volume 49 Number 2, October 2018. Pp. 161-174
Available online at: http://www.ajol.info/index.php/naj

A REVIEW ON PLANT GENOMIC DEVELOPMENT; ITS IMPORTANCE, CONSTRAINTS AND PROSPECTS

Njoku, D.N. and Ano, C.U.C.
National Root Crops Research Institute, Umudike, PMB 7006 Umuahia, Nigeria
Corresponding Author's email: njokudn2012@gmail.com

ABSTRACT
This review presents an overview of how genomics have developed over the years and how genomic tools have helped to change our approach to crop improvement. In the years preceding now, Plant breeding has been very successful in developing improved varieties using conventional tools and methodologies. Nowadays, the availability of genomic tools and resources has led to a new revolution of plant breeding, as they facilitate the study of the genotype and its relationship with the phenotype, in particular for complex traits. Recent technological advancements like Next Generation Sequencing (NGS), Genome-wide expression studies, Re-sequencing of genomes for the genome-wide discovery of markers amenable for high-throughput genotyping platforms, like SSRs and SNPs, or the construction of high density genetic maps, have substantially expanded our ability to analyze and understand plant genomes and to reduce the gap existing between genotype and phenotype. The fast evolving field of genomics allows scientists to analyze thousands of genes in parallel, to understand the genetic architecture of plant genomes and also to isolate the genes responsible for mutations. Conclusively, advances in genomics are providing breeders with new tools and methodologies that allow a great leap forward in plant breeding, including the 'superdomestication' of crops and the genetic dissection and breeding for complex traits.

Keywords: Genotyping, phenotyping, genomics, and sequencing

Introduction

The goals of agricultural plant science are to increase crop productivity, increase the quality of agricultural products, and maintain the environment. Each of these goals has significant economic value. Increased productivity has accounted for nearly all of the added value in germ plasm until recently. Quality is rapidly replacing productivity as the most valuable property of genomic improvement. The cassava market, for example, is moving from a homogenous commodities market to a segmented, specific-use market where the value of unique cassava root is preserved from the farm through to the end-user, the same could be said of sweetpotato. Maintaining productivity and quality without compromising environmental quality is growing in importance. The real cost of agriculture to the environment will be increasingly factored into production costs. These goals are interrelated. The greatest environmental impact of agriculture is the use of land. Increased productivity directly reduces the amount of land needed.

One of the ways by which these goals will be met include germ-plasm improvement. Germ-plasm improvement can be achieved through both conventional and molecular means. Germ-plasm improvement will continue to depend on non-transgenic methods that use sophisticated assays and molecular genetic markers. It is difficult to envisage a replacement for meiosis-based approaches to environmental adaptation. Nevertheless, gene technology will be the principal means by which value-added traits are created over the next several years. Genomics in particular will accelerate the discovery of genes that confer key traits, enabling their rapid improvement.

Application of conventional pre-genomics scientific breeding methodologies has led to the development of modern cultivars, which have contributed to the dramatic improvement of yield of most major crops since the middle of the 20th century. The success of plant breeding in the last century has relied in the utilization of natural and mutant induced genetic variation and in the efficient selection, by using suitable breeding methods, of the favorable genetic combinations. In this respect, the evaluation and identification of genetic variants of interest as well as the selection methodologies used have largely been

based in the phenotypic evaluation. Genomics provides breeders with a new set of tools and techniques that allow the study of the whole genome, and which represents a paradigm shift, by facilitating the direct study of the genotype and its relationship with the phenotype (Tester and Langridge, 2010). While classical genetics revolutionized plant breeding at the beginning of the 20th century, genomics is leading to a new revolution in plant breeding at the beginning of the 21th century.

The field of genomics and its application to plant breeding are developing very quickly. The combination of conventional breeding techniques with genomic tools and approaches is leading to a new genomics-based plant breeding. 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. 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 breeders (large collections of markers, high-throughput genotyping strategies, high density genetic maps, new experimental populations, etc.) that can be incorporated into existing breeding methods (Tester and Langridge, 2010; Lorenz et al., 2011; Varshney and Tuberosa, 2007). 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.) (Lorenz et al.,2011; Peleman and van der Voort, 2003; Collard and Markill, 2008).

Genomics approaches are particularly useful when dealing with complex traits, as these traits usually have a multi-genic nature and an important environmental influence. Thanks to these technological improvements. It is now feasible for a small laboratory to generate in a short time span (e.g., several months) enough molecular data to obtain a set of mapped quantitative trait loci (QTLs), even in a species lacking any previous genomic information (Varshney et al., 2010). Genomic tools are thus facilitating the detection of QTLs and the identification of existing favorable alleles of small effect, which have frequently remained unnoticed and have not been included in the gene pool used for breeding (Morgante and Salamini, 2003; Vaughan et al., 2007). Many plant genomes are large and complex due to an abundance of transposable elements and a long history of repeated genome duplication, making genome sequencing a major challenge (Schatz et al., 2012). The era of plant genomics began with release of the *Arabidopsis* genome sequence in 2000 (Nature, 2000). It was a milestone in plant biology and made *Arabidopsis* one of the most popular species for basic plant research. Rice, a staple food in most of the world, was the second available plant genome in 2002 (Goff et al., 2002; Yu et al., 2002). Rapid progress in the development of new sequencing technology and bioinformatic tools in recent years has allowed faster and more efficient sequencing, and assembly of genomes at lower cost. Genome sequences of economically important monocots, such as rice, maize, sorghum, and so on, have now been decoded (Goff et al., 2002; Yu et al., 2002; Banks et al., 2011; Rensing et al., 2008; Schnable et al., 2009; Paterson et al., 2009), including cassava and potato (Xu et al., 2011; Simon et al., 2012). These genomes will not only promote plant genomics and breeding studies for crop improvement programs, but also provide an unprecedented opportunity for basic plant biological research in the area of development and evolution.

Genomic Tools and Resources for Plant Breeding

Genomic Selection or Genomics-assisted Selection

The biggest driving force for genomics-assisted crop breeding in the plant genomics era has been the inexpensive sequencing and re-sequencing opportunity for population individuals of genetic crosses and breeding lines. This helps to precisely identify and link genetic variations to the phenotypic expressions, taking into account the rare and private allelic variations that are abundant in crop line population or germplasm resources (Poland, 2015; Abdurakhmonov and Abdukarimov, 2008; Kumpatla et al., 2012). The Sanger technology has been the predominant sequencing method for the past thirty years. It has been used to sequence several genomes as well as many transcriptomes. The first international collaborative project resulted in the whole genome sequence of the model plant *Arabidopsis thaliana* (Nature, 2010). After that, reference genomes of selected genotypes were completed in a limited number of crops such as rice (Nature, 2005), maize (Schnable et al, 2009), sorghum (Paterson et al., 2009), populous (Tuskan et al., 2006), grapevine (Jaillon et al., 2007), papaya (Ming et al., 2008), or soybean (Schmutz et al., 2010). The transcriptomes of most major crops, to a greater or lesser extent, were also sequenced. A global view of the genomes and transcriptomes sequenced can be obtained from the Gene Index Project (http://compbio.dfci.harvard.edu/tgi/plant.html) or in the NCBI Unigene database (http://www.ncbi.nlm.nih.gov/unigene). At present, the genomes for many agricultural plants including

specialty crops have been sequenced, as reviewed by Michael and VanBuren (Michael and VanBuren, 2015), which created a new paradigm for modern crop breeding. Crop breeding, which is powered and enriched by molecular markers, genetic linkage maps, QTL mapping, association mapping, and marker-assisted selection methods in the past century (Morrell et al., 2011; Abdurakhmonov and Abdukarimov, 2008), has now greatly accelerated and become ever productive and efficient in the plant genomics era (Poland, 2015). This is due to the (1) availability of large-scale transcriptome and whole-genome reference sequences (Michael and VanBuren, 2015); (2) high-throughput SNP marker collection and cost-effective, automated, and high-throughput genotyping platforms (HTP) and technologies (e.g., genotyping by sequencing or GBS), allowing breeders to screen multiple genotypes within a short time (Jimenez-Goimez, 2011; Poland, 2015); (3) identification and use of expression QTLs (genetical genomics) in breeding (Joosen et al., 2009); and (4) opportunity to perform genome-wide selection (i.e., genomic selection) (Poland, 2015). Also, when whole-genome sequences are not available and SNP markers are present in a limited number, the breeders using GBS and HTS platforms can readily genotype their mapping population and can provide genomic selections for the targeted crops of interest (Jimenez-Goimez, 2011; Poland, 2015; Kumpatla et al., 2012). Although it was first applied for animal breeding (Meuwissen et al., 2001), recently genomic selection has been successfully applied to a number of plant species (Cros et al., 2015; Longin et al., 2015; Iwata et al., 2015; Spindel et al., 2015; Cros et al., 2015; Beaulieu et al., 2014; Lipka et al., 2014), including studies using GBS in the context of genomic selection (Poland, 2015). Most importantly, the application of available genomics tools and a large number of high-throughput DNA markers and new-generation genotyping platforms have made the "breeding by design" (Peleman, 2003) possible and have developed "virtual breeding" approaches (Andersen, 2012) for efficient crop improvement. Advances have been made toward plant resistance genomics and molecular breeding against bacterial diseases (Takahashi et al., 2014) as well as biotic/ abiotic stress tolerance in agriculture crops (Onaga and Wydra, 2016). The determination of the functions of all the genes in a plant genome is the most challenging task in the postgenomic era of plant biology. However, several techniques or platforms, like serial analysis of gene expression (SAGE), massively parallel signature sequencing (MPSS), and micro- and macroarrays, have been used in several crops for the estimation of mRNA abundance for large number of genes simultaneously. The microarrays have also been successfully used in wheat for understanding alterations in the transcriptome of hexaploid wheat during grain development, germination and plant development under abiotic stresses (Wilson et al., 2004; Wilson et al., 2005). Comparison was made between Affymetrix GeneChip Wheat Genome Array (an in-house custom-spotted complementary DNA array) and quantitative reverse transcription-polymerase chain reaction (RT-PCR) for the study of gene expression in hexaploid wheat (Poole et al., 2007). Also, functional genomics approach in combination with "expression genetics" or "genetical genomics" provides a set of candidate genes that can be used for understanding the biology of a trait and for the development of perfect or diagnostic marker(s) to be used in map-based cloning of genes and MAS (Jordan et al., 2007). A similar example was provided by Jordan et al. (Jordan et al., 2009), when they identified regions of wheat genome controlling seed development by mapping 542 eQTLs, using a DH mapping population that was earlier used for mapping of SSRs and QTL analysis of agronomic and seed quality traits (McCartney et al., 2005).

Marker Assisted Backcross Selection

Marker assisted selection (MAS) is an indirect process where selection is carried out on the basis of a marker instead of the trait itself. The successful application of MAS relies on the tight association between the marker and the major gene or QTL responsible for the trait. As we have described before, the new genomic tools accelerate the identification of markers tightly linked to target genomic regions. On the other hand, the new dense genotyping platforms available today accelerate the genotyping of large amounts of samples during the MAS process in a rapid and economically feasible manner. MAS can take benefit from these technologies, speeding up the release of new varieties. MAS is also frequently applied to perform background selection in the context of backcrossing programmes. Background selection consists in the identification of plants with lower contents in donor genome to continue the breeding scheme, in order to achieve the recovery of the recipient genome. The use of background markers facilitates the quick recovery of the recurrent parent genome (Hospital et al., 1992). Background selection is being used extensively in rice breeding. High-density genome maps are being effectively used in background analysis. For example, background selection integrated with foreground selection of bacterial blight resistance (*xa13* and *Xa21* genes), amylose content (*waxy* gene) and fertility restorer gene has been performed in order to identify superior lines with maximum recovery of Basmati rice genome along with the quality traits and minimum non-targeted genomic introgressions of the donor chromosomes (Gopalakrishnan et al., 2008). Frequently, current breeding programmes involve the

introgression of more than one gene or QTL controlling traits of interest into one genetic background, in a process that is called pyramiding. The most useful application of MAS in the process of pyramiding is related to the introgression of different genes or QTLs whose effect on the phenotype is undistinguishable. The accumulation of genes from different sources which confer resistance against the same disease is an example, and is indeed one of the most widespread applications of gene pyramiding (Huang et al., 1997).

Identification of Molecular Markers Linked to Single Genes and QTLs

NGS and high-resolution maps have led to a significant improvement in the identification of molecular markers linked to specific genes and to QTLs. The most important advantage comes from the dense genome coverage, which allows the identification of markers closely linked to any target genomic region, with the advantages that this tight linkage provides. There are increasing reports describing accurate QTLs mapping with different NGS or high-throughput genotyping strategies. For example, a high density rice map constructed by whole-genome re-sequencing of a RILs population, was used to identify four QTLs controlling plant height (Garg et al., 2011). On a different study (Yu et al., 2011) an ultra-high density genetic map based on SNPs, obtained with Illumina GA, was compared with a linkage map based on RFLPs/SSRs in rice. The positions of several cloned genes, two major QTLs for grain length and grain width, and a QTL for pigmentation were evaluated in a RIL population, arising the expected result that the SNPs map detected more QTLs and more accurately than a RFLPs/SSRs based linkage map. Association mapping is just rising in model species and major crops. Maize is the most widely studied crop regarding association analysis. Many candidate genes have been successfully associated to morphological or quality traits. As an example, candidate genes *Dwarf8*, *Vgt1* and *ZmRap2.7* were successfully associated to flowering time (Buckler et al., 2009). Other candidate genes have been associated, among others, to forage quality, carotenoid content, oil content and kernel quality (Andersen et al., 2008; Harjes et al., 2008; Zheng et al., 2008; Manicacci., 2009). GWA studies have been more limited, probably due to the large genome of maize (2300 Mbp) and the great number of markers needed to cover it. The first study identified a fatty acid desaturase gene (*fad2*) associated with increased oleic acid levels (Belo et al., 2008). Examples of association mapping approaches in other crops are more limited. Studies based on the candidate gene approach have been reported in some crops, like grape, or conifers (Emannueli et al., 2010; Beaulieu et al., 2011). However, GWA studies have only been developed either in the model species *A. thaliana* (Atwell et al., 2010) or in major crops such as rice (Huang et al., 2010), barley (Massman et al., 2011), or wheat (Neuman et al., 2011). Although genetic association mapping is in its early steps, it is a promising tool for the dissection of complex traits in crop plants.

Microarrays and RNA sequencing (Expression Studies)

New genomic tools are also of interest to expand and accelerate gene expression studies. The analysis of gene expression has provided a rich source of biological information, which allows breeders to understand the molecular basis of complex plant processes, leading to the identification of new targets for manipulating these processes. Gene expression studies were at first based on the classical Northern blot method that only allowed the quantification of tens of genes simultaneously. The QRT-PCR is a more affordable and quantitative technique; but the number of genes analyzed by experiment is also limited (VanGuilder et al., 2008). Other approaches allowing the study of thousands of genes were differential display and cDNA amplified fragment length polymorphisms (cDNA-AFLPs) (Bachem et al., 1996). However, these methods are not really quantitative and are limited by the ability of the developed libraries to capture low-abundance transcripts. Other methods that overcome part of these problems are the serial analysis of gene expression (SAGE) (Anisimov, 2008) and massively parallel signature sequencing (MPSS) (Reinartz et al., 2002). Nevertheless, the most employed methods at present to analyze transcript profiling are the hybridization-based platforms or microarrays (Schena et al., 1995). Expression arrays have several advantages when compared with other methods. They can measure tens of thousands of different transcripts in the same reaction, they are semi-quantitative and sensitive to low-abundance transcripts if those are represented in a given array. The most extensive data are from the model species *A. thaliana* (Schmid et al., 2005), but an increasing number of studies in crops like maize, wheat, rice, barley, or soybean are already available. Microarrays make use of the existing EST collections and genome sequence data. The vast increase provided by NGS in the number of sequences opens the possibilities of expression studies in a large number of species lacking previous sequence information. Also, deep NGS sequencing of RNA transcripts (RNA-seq) is emerging as an alternative to microarray studies to quantify gene expression (Marioni et al., 2008; Stiglic et al., 2010). RNA-seq does not depend on genome annotation or on the probes contained in the array platform. This technology is also very useful to

improve genome annotation, improving the detection of rare transcripts and splicing variants and the mapping of exon/intron boundaries. Moreover, RNA-seq avoids bias introduced during hybridization of microarrays and saturation level problems, has a greater sensibility, and shows high reproducibility (Marioni et al., 2008; Cloonan et al., 2008). This approach has been already used in different crops with different breeding objectives, leading to the identification of genes involved in several metabolic pathways, disease response, fruit development, etc. (Alagna et al., 2009; Zenoni et al., 2010; Wang et al., 2011). All these studies show the potential of RNA-seq for complex traits breeding.

Breeding by Design

This is simply the possibility of predicting the outcome of a set of crosses on the basis of molecular markers information (Peleman and Van der Voort, 2003). It involves 3 steps: mapping *loci* involved in all agronomically relevant traits, assessment of the allelic variation at those *loci*, and, finally, breeding by design. In the method as initially described by Peleman and van der Voort (Peleman and Van der Voort, 2003), the first step was proposed to be completed by either using mapping populations segregating for the trait of interest or based on a candidate gene approach (mainly exploiting information from model plant species and increasing understanding of gene function). Also linkage disequilibrium (LD) mapping was suggested, focused on the region previously identified as related to the trait ('targeted LD mapping'). Currently, other possibilities such as GWA studies allow a more efficient way to accomplish this first step, avoiding limitations of biparental populations. The second step of the process consists in the identification of allelic variation for the *locus* of interest and the assignation of the phenotypic value to each of them. This step cannot be based on biparental populations, given that only two alleles per *locus* are segregating in this case. The analysis should then include plant materials representing the variability of the species. Genotypic and phenotypic data for each plant are required. Application of this breeding strategy has been used for different crops and with different objectives, such as breeding for heading date in rice (Wei et al., 2010) or seed length in soybean (Lu et al., 2011). This procedure has also been used in patent applications; as an example, 'breeding by design' has been reported as part of the development of higher quality maize varieties. However, the most effective application of the 'breeding by design' approach will come from the incorporation of the most advanced genomic tools into the process, which will allow the improvement of the predictions.

Benefits and Challenges

With respect to the recent advances in the plant sciences, as the sequences of many plant genomes become known, the power of genomics for applied breeding has to be one of the most exciting advances of recent years. Extremely valuable to breeders in the national agricultural research systems is the ability to genotype their collections to get a clear picture of their diversity and how such diversity might be enhanced through sharing and access to global collections. The use of marker-assisted selection in cases where phenotyping presents a challenge or to trace introgression of known genes or important regions from wild relatives should also become part of every serious national breeding program (Deborah, 2005). Also, many plant genomes are large and complex due to an abundance of transposable elements and a long history of repeated genome duplication, making genome sequencing a major challenge (Schatz et al., 2012). Complete sequence information, maps, and a huge array of molecular markers exist for rice; with more sequence information for other crops, new techniques for assessing allelic diversity, and a better understanding of synteny (Delseny, 2004), these are now being adapted for the breeding of other crops. Yet, for orphan crops like cowpea, common bean, the millets, tef, and cassava, we still have insufficient numbers of ESTs, bacterial artificial chromosome libraries, molecular maps, and markers (Nelson et al., 2004). Programs such as the Generation Challenge Program and crop-specific initiatives such as Phaseomics are beginning to address these limitations, but a glance at the number of ESTs available for different organisms (www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html) indicates that more funds and efforts are clearly warranted. Good value can also be had through sequencing of the genomes of major plant pathogens. In addition, there are many challenges in creating the needed infrastructure, including high-throughput analysis systems and critical high-speed Internet access to the tools of bioinformatics; development of a pool of breeders well-versed in the use of these tools also still limits progress on this front. Networks in Asia that brought together rice (the Asian Rice Biotechnology Network, ARBN) and maize breeders (the Asian Maize Biotechnology Network, AMBIONET) to build capacity and better interactions among molecular breeders have been most successful; a similar network called AMMANET (African Molecular Marker Applications Network), which holds promise for African breeders, is another welcome development. The regional center in Nairobi called Biosciences for East and Central Africa (BECA) is serving as a center of excellence for agricultural biotechnology by interacting with, and serving the various universities and national

agricultural research systems of the region. At BECA, the modern tools of genomics can be shared with breeding programs through training, provision of markers, high-throughput analysis coupled with a sophisticated bioinformatics platform, and joint efforts to genotype key crops and identify projects suitable for marker-assisted selection (Deborah, 2005). The use of molecular markers has helped highlight the importance of genes from wild relatives for use in crop improvement (Tanskey and McCouch, 1997; Koornneef et al., 2004) and, as evidenced by recent work on tomato improvement, the results can sometimes be spectacular (Frydman et al., 2004). African farmers are showing real enthusiasm for new interspecific hybrids that combine the best of both Asian and African rices (Jones et al., 1997). For complex traits, the identification of quantitative trait loci (QTL) has advanced to a considerable degree, to the point where it is now becoming somewhat more feasible to identify specific genes that control the traits underlying the QTL (Ashikari et al, 2005). Advances in genomics should also be able to contribute new insights to our currently vague understanding of that most important of traits, heterosis (hybrid vigor). Can the recent work showing how inbred lines of maize differ strikingly in gene sequences (Brunner, 2005) and gene expression patterns (Gue et al., 2004) provide some clues? Can such understanding help us determine whether there is good value in promoting the development of hybrid sorghum and millets for Africa and to explore further the potential of heterosis in many crops beyond maize? Certainly, development of hybrid seed is one way to promote viable seed markets for crops. But do we understand well enough the cost–benefit equations for small farmers with respect to purchase of high-quality seed (hybrid or not) vs. the saving of seed, and is the development of a strong private-sector seed business a necessary part of moving such farmers beyond the subsistence level? Such questions go beyond the realm of science into that of sociology and economics, but good answers clearly require input from the scientific community. Other practical benefits of the new knowledge and understanding that can come from Plant genome research when applied through the new tools of plant breeding are: Accelerated improvements in the safety, quality, and diversity of food and other products of plants; Greater assurance of food security worldwide in the face of a doubling of the world population over the next 30–35 years and declining agricultural land and quality water for irrigation; Cleaner, healthier, environment and greater energy efficiency through improvements in fertilizer-use efficiency, thereby reducing production costs and concerns for groundwater contamination, and because of more sustainable disease and pest control through defenses delivered with seeds rather than with pesticides; Expanded use of plant products, including higher-quality animal feeds, industrial feedstocks, and other value-added applications;

The development and marketing of new improved seeds. The seed industry has become a major growth industry worldwide, raking up to $5 billion dollars annually, for USA alone (James, 1998).

Future Directions

The revolutionizing advances made in the past three decades in plant genomics and its subdisciplines provided a mass of novel opportunities with easy-solution applications and highthroughput, cost-effective, and time-effective technologies. Plant genomics era increased our understanding of the basis of complex life processes/traits in plants and crop species, and it paved a way for effective improvement of plants to fulfill our diet and other needs. However, it also piled up challenging grand tasks ahead for current genomics and post-genomics era (Ibrokhim, 2016). Due to tireless effort, tremendous achievements have been made toward sequencing more than hundreds of plant genomes including major crop species and specialty, model/non model, wild, vascular, flowering, and polypoid plants (Micheal and Jackson, 2013; Michael and Van Buren, 2015). However, the first current and future task ahead is to extend such large-scale, multiple accession genome sequencing initiatives for each priority agricultural and specialty crop species including their wild relatives and ancestor-like genome representatives. Take for instance, Germplasm from hundreds of African cassava cultivars are characterized in this approach, allowing marker-assisted breeding schemes to be developed for improving nutrient content as well as tolerance of both drought and viral cassava mosaic disease (CMD) and CBSD (Simon et al., 2012). Although it sounds largely ambitious, this task will be mandatory and important for the next plant genome sequencing phase. This is to effectively use all variations existing among plant/crop germplasm resources and its ecotypic populations and to design efficient GWAS analysis and consequent genomic selections as well as tools/software programs for better analyzing plant genomes and improving genome assembly issues (Weigel and Mott, 2009; Leebens-Mack, 2015). This is especially needed for polyploidy crops (Song and Chen, 2015; Michael and VaBuren, 2015; Morrell et al., 2011) because the sequencing of many polyploids and their subgenomes would increase our understanding of the complexity of polypoidy, gene silencing, epigenetics, and biased retention and expression of genes after polyploidization (Song and Chen, 2015; Chaudhary et al., 2009; Renny-Byfield et al., 2014; Yoo et al., 2013). Furthermore, it also helps to discover all natural variations and lost genes during crop

domestication that should be useful to restore the key agriculturally important traits in the future. A consequent grand task and challenge with the completion of the above-highlighted tasks is the handling, organizing, systematizing, and visualizing a huge amount of plant genome sequencing ("Big Data") data that require urgent attention, effort, collaborative work, and investment. There is an urgent need to develop more efficient bioinformatics platforms to handle plant genome data due to challenges, specificities, complexities, and sizes of currently available and future sequenced plant genomes mentioned herein (Schatz et al., 2012; Sinha, 2011). Funding this aspect of plant genomics and bioinformatics research is a necessary key step [1] for future advances on this task ahead. Furthermore, there is a need to make sequenced genomes "functional" (Michael and Jackson, 2013) and biologically meaningful (Fernie, 2012; Morrell et al., 2011). This can be done by linking the sequence variation(s) with phenotype(s), trait expression, and epigenetic and adaptive features of plants to their living environment and extreme conditions. The successful completion of this task will require the combined approaches of genomics with bioinformatics, proteomics, metabolomics, phenomics, genomic selections, genetical genomics, reverse genomics, system biology, etc. (Prohens, 2011; Stokes and McCourt, 2014; Fernie, 2012; Andersen, 2012; Ricroch and Henard-Damave, 2015; Sinha, 2011). This also requires the integration of all available genomic and phenotypic data to identify key networks that also require downstream effort of integration of specific networks to networks of other systems in order to connect heterogeneous data (Fernie, 2012). There are suggested thoughts and tasks for plant genomics that should target to develop plant genome-specific "Encyclopedia of DNA Elements (ENCODE)" (Michael and Jackson, 2013; Michael and VanBuren, 2015), which will be an important achievement in the next phases of development. There is a need to use molecular phenotyping (i.e., using molecular process such as protein-RNA interactions, translation rates, etc.) in QTL mapping (Jimenez-Goimez, 2011) that would help to precisely link the sequence variation(s) to its phenotype(s). These particular grand tasks further highlight a need for extended effort and work on the development of inexpensive highthroughput plant phenotyping (Fahlgren et al., 2015; Poland, 2015) and plant proteome and metabolome profiling tools and instrumentation (Deal, 2011; Heazlewood, 2011) by utilizing small amount single-cell-derived samples (Deal, 2011; Heazlewood, 2011; Fernie, 2012). Another task is to optimize and better design novel transgenomics and genome editing technologies for the key priority crops and plant by-product production. In addition, there are needs to identify the appropriate choice of plant tissues for genome editing, reduce or eliminate side effects and off-target toxicity and mutagenesis of application of novel genome modification technologies, and develop reliable screens for the detection of edited genome samples (Puchta and Hohn, 2010). The revolutionizing effects of these novel genome-editing/manipulation. Of all these, the biggest task ahead will be in the preparation of well-qualified next-generation scientists capable of continuing plant genomics tasks highlighted, well versed in conventional plant biology, ecology, plant breeding, evolution, taxonomy, modern "omics" disciplines, and cross-related scientific disciplines (e.g., mathematics, computing, and modeling) (Schatz, 2012; Sinha,2011). Importantly, they are required to have a capability to utilize modern computing and instrumentation platforms and bioinformatics knowledge (Fernie, 2012). For instance, there is a huge need for a new generation of molecular breeders (Moose and Mumm, 2008) with full knowledge and appreciation of conventional plant breeding aspects including the understanding of agrotechnology methodologies, genetic diversity of crop germplasm, and randomized multi-environmental field trails. These breeders also need to have abilities to handle, work, and utilize the sequenced genomes, high-throughput genotyping, and phenotyping platforms. This is a bottleneck for plant genomics at present, which requires urgent awareness, attention, and investment.

Translating Basic Genome Research to Benefit Subsistence Farmers

Despite the considerable and continuing breakthroughs in plant genetic and genomic technologies, there has been relatively little global government investment into funding basic plant science and in translating these discoveries into food crops beneficial to farmers in less developed countries. To fill the gap, some foundations and public–private partnerships have launched programs. For example, the Bill and Melinda Gates Foundation is supporting a large program, called *Stress-Tolerant Rice for Africa and South Asia* (IRRI, 2007), which is assisting with the development and dissemination of the Sub1 rice variety, which resulted from a ten-year basic research collaboration funded primarily by the US Department of Agriculture. With the help of the Gates Foundation, last year more than 4 million farmers grew Sub1 rice (Xu and Ronald, 2013). The Rockefeller Foundation was instrumental in funding the development of Golden Rice (GRHB, 2005), a genetically engineered rice enriched for provitamin A that is expected to be released soon (Harmon, 2013). Worldwide, over 124 million children are vitamin A-deficient; many go blind or become ill from diarrhea, and nearly 8 million preschool-age children die each

year as a result of this deficiency. A public–private partnership advanced the development of second generation Golden Rice (Paine et al., 2005; Tang et al., 2009). One report estimates that improved vitamin A nutritional status obtained from eating vitamin A rice could prevent the deaths of thousands of young children each year (Stein et al., 2006). The positive effects of Golden Rice are predicted to be most pronounced in the lowest income groups at a fraction of the cost of the current supplementation programs (Stein et al., 2006; AATF, 2012), which are not only costly to run but also not always continued (GRHB, 2005). The Water Efficient Maize for Africa (WEMA) project is another important public–private partnership, which aims to develop drought-tolerant and insect-protected maize using conventional breeding, MAS, and biotechnology. The goal is to make these varieties available royalty free to small-hold farmers in sub-Saharan Africa through African seed companies (AATF, 2012). The introduction of drought-tolerant maize to Africa, where three-quarters of the world's severe droughts have occurred over the past ten years, is predicted to dramatically increase yields of this staple food crop for local farmers (AATF, 2012; AATF, 2010). Another exciting development is the US Agency for International Development (USAID) "Feed the Future" program, which partners with diverse countries to enhance local food security (USAID, 2012). For example the Maharashtra Hybrid Seed Company and Cornell University have jointly developed Bt eggplant that is resistant to fruit and shoot borers (USAID, 2004). Bt eggplant was recently made available on a royalty-free basis to smallholder farmers in Bangladesh. Researchers estimate that farmers growing the new Bt eggplant varieties could obtain yield increases of 30%–45% while reducing insecticide use. The USAID has also funded projects to enhance the productivity of banana, a staple food crop for more than 100 million people in East Africa, and which is susceptible to several serious diseases. Many strategies to control this disease rely on genetic engineering because most bananas don't produce seed and are propagated clonally (Peed, 2011; Studholme et al., 2010; Tripathi et al., 2009; FAO, 2004). Bananas with resistance to banana *Xanthomonas* wilt disease (BXW), have recently been genetically engineered with the rice XA21 resistance gene (Tripathi et al., 2014). These examples demonstrate the success of non-profit and public–private partnerships in translating basic research discoveries into benefits at the farm. Well-funded, long-term, multinational, multidisciplinary collaborations are vital if we are to continue making significant progress in developing new crop varieties to enhance food security in the developing world. In a recent report, leading scientists highlighted the need for significant investment in plant breeding and estimated that US$200 million annually is needed to carry out such a systematic, concerted, collaborative global effort (McCouch et al., 2013).

Conclusion

In the past three decades, plant genomics has evolved from the enrichment and advances made in conventional genetics and breeding, molecular biology, molecular genetics, molecular breeding, and molecular biotechnology in the land of high-throughput DNA sequencing technologies powering the plant research community to sequence and understand the genetic compositions, structures, architectures, and functions of full plant genomes. The technological and instrumentation advancements as well as the desire and need to feed the increasing human population, overcome biosecurity issues, and sustain agricultural production in the era of global climate change, the societal globalization, and technological advancements have been the main driving forces for plant genomics development. These led to sequence and assemble entire plant genomes including very complex polyploid plants, annotate gene functions, link the sequence variation(s) to the phenotype(s), and exploit sequence variation(s) in plant/crop improvement in genome-wide scale or through targeted native modification of plant genomes in a highly sequence-specific manner. Therefore, while conventional pre-genomics plant breeding has been, is, and will be successful at improving our crops, the application of genomic tools and resources to practical plant breeding will push forward the genetic gains obtained by breeding programmes. New genomic advances, many of which are already being developed, will make easier for breeders to obtain new cultivars with improved characteristics, either by facilitating selection or by improving the variation available for breeders by using precision breeding approaches. In particular, the present and new genomics tools are of great value for the genetic dissection and breeding of complex traits.

References

Abdurakhmonov IY, Abdukarimov A (2008). Application of association mapping to understanding the genetic diversity of plant germplasm resources. International Journal of Plant Genomics. 2008:574927. DOI:10.1155/2008/574927.

African Agricultural Technology Foundation (2012) Water efficient maize for Africa (WEMA). Available: http://wema.aatf-africa.org/aatf-projects/water-efficient-maize-africa-wema). African Agricultural Technology Foundation (AATF-Africa).

Alagna F, D'Agostino N, Torchia L, Servili M, Rao R, Pietrella M, Giuliano G, Chiusano ML, Baldoni L, Perrotta G. (2009). Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development.BMC Genomics. 2009;10:399.

Andersen JR, Zein I, Wenzel G, Darnhofer B, Eder J, Ouzunova M, Lubberstedt T (2008). Characterization of phenylpropanoid pathway genes within European maize (*Zea mays* L.) inbreds. BMC Plant Biol. 2008;8:2.

Andersen SB (2012). Virtual breeding. In: Abdurakhmonov IY, editor. Plant Breeding. Rijeka: InTech; 2012. p. 107-128. DOI: 10.5772/29802

Anisimov SV (2008). Serial Analysis of Gene Expression (SAGE): 13 years of application in research. Curr. Pharm. Biotechnol. 2008;9:338–350.

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles R R, Qian Q, Kitano H, Matsuoka M (2005). Science **309 ,** 741-745.

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng DZ, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JDG, Michael T, Nemri A, Roux F, Salt DE, Tang CL, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M. (2010). Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature. 2010;465:627–631.

Bachem CW, van der Hoeven RS, de Bruijn SM, Vreugdenhil D, Zabeau M, Visser RG (1996). Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. Plant J. 1996;9:745–753.

Beaulieu J, Doerksen T, Boyle B, Clément S, Deslauriers M, Beauseigle S, Poulin P, Lenz P, Caron S, Rigault P, Bicho P, Bousquet J, Mackay J. (2011). Association genetics of wood physical traits in the conifer white spruce and relationships with gene expre ssion. Genetics. 2011;188:197–214.

Beaulieu J, Doerksen TK, MacKay J, Rainville A, Bousquet J. (2014). Genomic selection accuracies within and between environments and small breeding groups in white spruce. BMC Genomics. 2014;15:1048. DOI: 10.1186/1471-2164-15-1048

Belo A, Zheng PZ, Luck S, Shen B, Meyer DJ, Li BL, Tingey S, Rafalski A (2008). Whole genome scan detects an allelic variant of fad2 associated with increased oleic acid levels in maize. Mol. Genet. Genomics. 2008;279:1–10.

Brunner S, Fengler K, Morgante M, Tingey S, Rafalski A (2005). Plant Cell **17 ,** 343-36.

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li HH, Mitchell SE, Pressoir G, Peiffer JA, Rosas MO, Rocheford TR, Romay MC, Romero S, Salvo S, Villeda HS, da Silva HS, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu JM, Zhang ZW, Kresovich S, McMullen MD (2009). The genetic architecture of maize flowering time. Science. 2009;325:714–718.

Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Bethel G, Robertson AJ, Perkins AC, Bruce SJ, Lee CC, Ranade SS, Peckham HE, Manning JM, McKernan KJ, Grimmond SM (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat. Methods.2008;5:613–619.

Collard BCY, Mackill DJ (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil. Transact. Royal Soc. B. 2008;363:557–572.

Cros D, Denis M, Bouvet JM, Sánchez L (2015). Long-term genomic selection for heterosis without dominance in multiplicative traits: case study of bunch production in oil palm. BMC Genomics. 2015;16:651. DOI: 10.1186/s12864-015-1866-9

Cros D, Denis M, Sánchez L, Cochard B, Flori A, Durand-Gasselin T, Nouy B, Omoré A, Pomiès V, Riou V, Suryana E, Bouvet JM (2015). Genomic selection prediction accuracy in a perennial crop: case study of oil palm (Elaeis guineensis Jacq.). Theor Appl Gen- et. 2015;128:397 410. DOI: 10.1007/s00122-014-2439-z

Deal RB (2011). Grand Challenge: Accelerating Discovery through Technology Develop- ment. Front Plant Sci. 2011;2:41. DOI: 10.3389/fpls.2011.00041

Deborah P D (2012). Agriculture in the developing world: Connecting innovations in plant research to downstream applications. vol. 102 no. 44.15739–15746, doi: 10.1073/pnas.0505895102

Delseny M (2004) Curr. Opin. Plant Biol. **7 ,** 126-131

Emanuelli F, Battilana J, Costantini L, Le Cunff L, Boursiquot JM (2010). This P, Grando MS. A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.) BMC Plant Biol. 2010;10:241.

Fahlgren N, Gehan MA, Baxter I (2015). Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. Curr Opin Plant Biol. 2015;24:93-9. DOI:10.1016/ j.pbi.2015.02.006

FAO (2004). Banana improvement. FAO Corporate Document Repository, editor. Available: http://www.fao.org/docrep/007/ae216e /ae216e01.html. Agriculture and Consumer Protection.

Fernie AR (2012). Grand Challenges in Plant Systems Biology: Closing the Circle(s). Front Plant Sci. 2012;3:35. DOI:10.3389/fpls.2012.00035

Frydman E, Carrari F, Liu Y, Fernie AR, Zamir D (2004). Science **305 ,** 1786-1789.

Garg R, Patel RK, Tyagi AK, Jain M. (2011). De novo assembly of chickpea transcriptome using short reads for gene discovery and marker identification. DNA Res. 2011;18:53–63.

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science. 2002, 296: 92-100. 10.1126/science.1068275.

Golden Rice Humanitarian Board (2005). Golden Rice Project. Available:http://www.goldenrice.org/index.php. Rockefeller Foundation.

Gopalakrishnan S, Sharma RK, Anand Rajkumar KA, Joseph M, Singh VP, Bhat KV, Singh NK, Mohapatra T (2008). Integrating marker assisted background analysis with foreground selection for identification of superior bacterial blight resistant recombinants in Basmati rice. Plant Breed. 2008;127:131–139.

Gue M, Rupe MA, Zinselmeier C, Habben J, Bowen BA, Smith OS (2004). Plant Cell **16,** 1701-1716

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan JB, Buckle ES (2008). Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science. 2008;319:330–333.

Harmon A (2013). Golden rice: lifesaver? The New York Times Sunday Review.

Heazlewood JL (2011). The green proteome: challenges in plant proteomics. Front Plant Sci. 2011;2:6. DOI:10.3389/fpls.2011.00006

Hospital F, Chevalet C, Mulsant P (1992). Using m,02arkers in gene introgression breeding programs. Genetics.1992;132:1199–1210.

Huang N, Angeles ER, Domingo J, Magpantay G, Singh S, Zhang G, Kumaravadivel N, Bennett J, Khush GS (1997). Pyramiding of bacterial blight resistance genes in rice: marker-assisted selection using RFLP and PCR. Theor. Appl. Genet. 1997;95:313–320.

IbrokhimYA (2016). Plant Genomics: Genomics era for plants and crop species- advances made and needs tasks ahead. ISBN 978-953-51-2456-6, print ISBN 978-953-51-2455-9, published: july 14, 2016 under CC BY 3.0 license.

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

IRRI (2007). Stress-tolerant rice for Africa and South Asia. Available:https://sites.google.com/a/irri.org/strasa /. International Rice Research Institute and AfricaRice**.**

Iwata H, Ebana K, Uga Y, Hayashi T (2015). Genomic prediction of biological shape: elliptic Fourier analysis and kernel partial least squares (PLS) regression applied to grain shape prediction in rice (Oryza sativa L.). PLoS One. 2015;10:e0120610. DOI: 10.1371/ journal.pone.0120610

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pè ME, Valle G, Morgante M, Caboche M, Adam-Blondon AF, Weissenbach J, Quétier F, Wincker P. (2007). French-Italian Public Consortium for Grapevine Genome Characterization. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449:463–467.

James RC (1998). Toward a successful multinational crop plant genome initiative. Proc Natl Acad Sci USA. 1998 Mar 3; 95(5): 1993–1995.

Jimenez-Goimez JM (2011). Next generation quantitative genetics in plants. Front Plant Sci. 2011;2:77. DOI:10.3389/fpls.2011.00077

Jones M P, Dingkuhn M, Aluko G K, Semon M (2009). Euphytica **92 ,** 237-246

Joosen RV, Ligterink W, Hilhorst HW, Keurentjes JJ. Advances in genetical genomics of plants. Curr Genomics. 2009;10:540-9. DOI: 10.2174/138920209789503914.

Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004). Annu. Rev. Plant Biol. **55 ,** 141-172.

Kumpatla SP, Buyyarapu R, Abdurakhmonov IY, Mammadov JA. Genomics-assisted plant breeding in the 21st century: technological advances and progress. In: Abdur- akhmonov IY, editor. Plant Breeding. Rijeka: InTech; 2012. p 134-184. DOI: 10.5772/37458

Leebens-Mack J. OneKp Capstone Wiki [Internet]. (2015). Available from: https:// pods.iplantcollaborative.org/wiki/display/iptol/OneKP±Capstone±Wiki [Accessed: 2015-11-10]

Lipka AE, Lu F, Cherney JH, Buckler ES, Casler MD, Costich DE (2014). Accelerating the switchgrass (Panicum virgatum L.) breeding cycle using genomic selection ap- proaches. PLoS One. 2014;9:e112227. DOI: 10.1371/journal.pone.0112227

Longin CF, Mi X, Würschum T (2015). Genomic selection in wheat: optimum allocation of test resources and comparison of breeding strategies for line and hybrid breeding. Theor Appl Genet. 2015;128:1297-306. DOI:10.1007/s00122-015-2505-1

Lorenz AJ, Chao S, Asoro FG, Heffner EL, Hayashi T, Iwata H, Smith KP, Sorrells MK, Jannink JL. (2011). Genomic selection in plant breeding: knowledge and prospects. Adv. Agron. 2011;110:77–123

Lü HY, Liu XF, Wei SP, Zhang YM (2011). Epistatic association mapping in homozygous crop cultivars. PLoS One . 2011;6(3):e17773.

Manicacci D, Camus-Kulandaivelu L, Fourmann M, Arar C, Barrault S, Rousselet A, Feminias N, Consoli L, Frances L, Mechin V, Murigneux A, Prioul JL, Charcosset A, Damerval C (2009). Epistatic interactions between Opaque2 transcriptional activator and its target gene CyPPDK1 control kernel trait variation in maize. Plant Physiol. 2009;150:506–520.

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res. 2008;18:1509–1517.

Massman J, Cooper B, Horsley R, Neate S, Dill-Macky R, Chao S, Dong Y, Schwarz P, Muehlbauer GJ, Smith KP (2011). Genome-wide association mapping of Fusarium head blight resistance in contemporary barley breeding germplasm. Mol. Breed. 2011;27:439–454.

McCartney CA, Somers DJ, Humphreys DG (2005). Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452 × 'AC Domain' *Genome*. 2005;48(5):870–883.

McCouch S, Baute GJ, Bradeen J, Bramel P, Bretting PK (2013) Feeding the future. Nature 499: 23–24.

Meuwissen TH, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using ge- nome-wide dense marker maps. Genetics. 2001;157:1819-29

Michael TP, Jackson S (2013). The First 50 Plant Genomes. The Plant Genome. 2013;6:1-7. DOI:10.3835/plantgenome2013.03.0001.

Michael TP, VanBuren R (2015). Progress, challenges and the future of crop genomes. Curr Opin Plant Biol. 2015;24:71-81. DOI: 10.1016/j.pbi.2015.02.002

Ming R, Hou S, Feng Y, Yu O, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KLT, Salzberg SL, Feng L, Jones MR, Skelton RL, Murray JE, Chen C, Qian W, Shen J, Du P, Moriah Eustice ME, Tong E, Tang H, Lyons E, Paull RE, Michael TP, Wall K, Rice DW, Albert H, Wang ML, Zhu YJ, Schatz M, Nagarajan N, Acob RA, Guan P, Blas A, Wai CM, Ackerman CM, Ren Y, Liu C, Wang J, Wang J, Na JK, Shakirov EV, Brian Haas B, Thimmapuram J, Nelson D, Wang X, Bowers JE, Gschwend AR, Delcher AL, Singh R, Suzuki JY, Tripathi S, Neupane K, Wei H, Irikura B, Paidi M, Jiang N, Zhang W, Presting G, Windsor A, Navajas-Pérez R, Torres MJ, Feltus FA, Porter B, Li Y, Burroughs AM, Luo MC, Liu L, Christopher DA, Mount SM, Moore PH, Sugimura T, Jiang J, Schuler MA, Friedman V, Mitchell-Olds T, Shippen DE, dePamphilis CW, Palmer JD, Freeling M, Paterson AH, Gonsalves D, Wang L, Alam M (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus) Nature . 2008;452:991–996.

Moose SP, Mumm RH (2008). Molecular plant breeding as the foundation for 21st century- crop improvement. Plant Physiol. 2008;147:969-77. DOI:10.1104/pp.108.118232.

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

Morrell PL, Buckler ES, Ross-Ibarra J. (2011). Crop genomics: advances and applications. Nat Rev Genet. 2011;13:85-96. DOI: 10.1038/nrg3097

Nature (2000). The Arabidopsis Genome Initiative: Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*.2000. 408: 796-815. 10.1038/35048692.

Nelson RJ, Naylor RL, Jahn MM (2004). Crop Sci. **44 ,** 1901-1904.

Neumann K, Kobiljski B, Dencic S, Varshney RK, Borner A (2011). Genome-wide association mapping: a case study in bread wheat (*Triticum aestivum* L.) Mol. Breed. 2011;27:37–58.

Onaga G, Wydra K (2016). "Advances in plant tolerance to abiotic stress," *In Plant Genomics*, ed. Abdurakhmonov IY, editor. (Rijeka: InTech;). 10.5772/64350

Paine JA, Shipton CA, Chaggar S, Howells RM, Kennedy MJ (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. Nat Biotechnol 23: 482–487.

Paterson AH, Bowers JE, 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 AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Rahman M, Ware D, Westhoff P, Mayer KFX, Messing J, Rokhsar DS (2009). The *Sorghum bicolor* genome and the diversification of grasses. Nature. 2009;457:551–556.

Peed M (2011). We have no bananas: can scientists defeat a devastating blight? The New Yorker.

Peleman JD, van der Voort JR (2003). Breeding by design. Trends Plant Sci. 2003;8:330-4. DOI:10.1016/S1360-1385(03)00134-1

Poland J (2015). Breeding-assisted genomics. Curr Opin Plant Biol. 2015;24:119-24. DOI: 10.1016/j.pbi.2015.02.009.

Poole R, Barker G, Wilson ID, Coghill JA, Edwards KJ (2006). Measuring global gene expression in polyploidy; a cautionary note from allohexaploid wheat.

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, Orjeda G, Guzman F, Torres M, Lozano R, Ponce O, Martinez D, De la Cruz G, Chakrabarti SK, Patil VU, Skryabin KG, Kuznetsov BB, Ravin NV, Kolganova TV, Beletsky AV, Mardanov AV, Di Genova A, Bolser DM, Martin DM, Li G (2011). Genome sequence and analysis of the tuber crop potato. Nature. 2011, 475: 189-195. 10.1038/nature10158.

Prohens J (2011). Plant Breeding: A Success Story to be Continued Thanks to the Advances in Genomics. Front Plant Sci. 2011;2:51. DOI:10.3389/fpls.2011.00051

Reinartz J, Bruyns E, Lin JZ, Burcham T, Brenner S, Bowen B, Kramer M, Woychik R (2002). Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms. Brief. Funct. Genomics Proteomics. 2002;1:95–104.

Renny-Byfield S, Gallagher JP, Grover CE, Szadkowski E, Page JT, Udall JA, Wang X, Paterson AH, Wendel JF (2014). Ancient gene duplicates in Gossypium (cotton) exhibit near-complete expression divergence. Genome Biol Evol. 2014;6:559-71. DOI: 10.1093/gbe/evu037

Ricroch AE, Henard-Damave MC (2015). Next biotech plants: new traits, crops, developers and technologies for addressing global challenges. Crit Rev Biotechnol. 2015: Early online:1-16. DOI:10.3109/07388551.2015.1004521

Schatz MC, Witkowski J, McCombie WR(2012). Current challenges in de novo plant genome sequencing and assembly. Genome Biol. 2012, 13: 243

Schena M, Shalon D, Davis RW, Brown PO (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995;270:467–470.

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU (2005). A gene expression map of Arabidopsis thaliana development. Nat. Genet. 2005;37:501–506.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, 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 MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddeloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK. (2009). The B73 maize genome: complexity, diversity, and dynamics. Science. 2009;326:1112–1115.

Simon Prochnik, Pradeep Reddy Marri, Brian Desany, Pablo D. Rabinowicz, Chinnappa Kodira, Mohammed Mohiuddin, Fausto Rodriguez, Claude Fauquet, Joseph Tohme, Timothy Harkins, Daniel S. Rokhsar, Steve Rounsley (2012). The Cassava Genome: Current Progress, Future Directions. Trop Plant Biol. 2012 Mar; 5(1): 88–94. Published online 2012 Jan 5. doi: 10.1007/s12042-011-9088

Sinha NR (2011). Plant developmental biology in the post-genomic era. Front Plant Sci. 2011;2:11. DOI:10.3389/fpls.2011.00011

Song Q, Chen ZJ (2015). Epigenetic and developmental regulation in plant polyploids. Curr Opin Plant Biol. 2015;24:101-9. DOI: 10.1016/j.pbi.2015.02.007.

Spindel J, Begum H, Akdemir D, Virk P, Collard B, Redoña E, Atlin G, Jannink JL, McCouch SR (2015). Genomic selection and association mapping in rice (Oryza sativa): effect of trait genetic architecture, training population composition, marker number and statistical model on accuracy of rice genomic selection in elite, tropical rice breeding lines. PLoS Genet. 2015;1:e1004982. DOI: 10.1371/journal.pgen.1004982.

Stein AJ, Sachdev HPS, Qaim M (2006). Potential impact and cost-effectiveness of Golden Rice. Nat Biotechnol 24: 1200–1201.

Stiglic G, Bajgot M, Kokol P (2010). Gene set enrichment meta-learning analysis: next-generation sequencing versus microarrays. BMC Bioinformatics. 2010;11:176.

Stokes ME, McCourt P (2014). Towards personalized agriculture: what chemical genomics can bring to plant biotechnology. Front Plant Sci. 2014;5:344. DOI: 10.3389/fpls. 2014.00344

Studholme DJ, Kemen E, MacLean D, Schornack S, Aritua V (2010). Genome-wide sequencing data reveals virulence factors implicated in banana *Xanthomonas* wilt. FEMS Microbiol Lett 310: 182–192.

Takahashi W, Yuichi M, Tohru S, Tadashi T (2014). Identification of novel major locus for gray leaf spot resistance in Italian ryegrass (*Lolium multiforum* Lam.) BMC Plant Biol. 2014;14: 303. Doi: 10.1186/s12870-014-0303-6

Tang G, Qin J, Dolnikowski GG, Russell RM, Grusak MA (2009). Golden Rice is an effective source of Vitamin A. Am J Clin Nutr 89: 1776–1783.

Tanksley, S. D. & McCouch, S. R. (1997). Science **277,** 1063-1066.

Tester M, Langridge P (2010). Breeding technologies to increase crop production in a changing world. *Science. 2010 Feb 12; 327(5967):818-22.*

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

Tripathi JN, Lorenzen J, Bahar O, Ronald PC, Tripathi L (2014). Transgenic expression of the rice *Xa21* pattern recognition receptor in banana (Musa sp.) confers resistance to *Xanthomonas campestris* pv. *musacearum*. Plant Biotechnol J In press.

Tripathi L, Mwangi M, Abele S, Aritua V, Tushemereirwe WK (2009). *Xanthomonas* Wilt: a threat to banana production in East and Central Africa. Plant Disease 93: 440–451.

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Déjardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006). The genome of black cottonwood *Populus trichocarpa* (Torr & Gray) Science. 2006;313:1596–1604.

USAID (2004). Pest resistant eggplant: India, Bengladesh and Phillipines Available:http://pdf.usaid.gov/pdf_docs/pdacp867.pdf). Washington (D.C.): USAID.

USAID (2012). Feed the future: the U.S. government's global hunger and food security initiative. Available: http://feedthefuture.gov/about. Washington (D.C.): USAID.

VanGuilder HD, Vrana KE, Freeman WM (2008). Twenty-five years of quantitative PCR for gene expression analysis. Biotechniques. 2008;44:619–626.

Varshney RK, Glaszmann JC, Leung H, Ribaut JM (2010). More genomic resources for less-studied crops. Trends Biotechnol. 2010;28:452–460.

Varshney RK, Tuberosa R (2007). Genomics-assisted crop improvement vol. 1: Genomics approaches and platforms.New York: Springer; 2007

Vaughan DA, Balász E, Heslop-Harrison JS (2007). From crop domestication to super-domestication. Ann. Bot.2007;100:893–901.

Wang Y, Zhang H, Li H, Miao X (2011). Second-generation sequencing supply an effective way to

screen RNAi targets in large scale for potential application in pest insect control. PLoS One. 2011;6(4):e18644.

Wei X, Liu L, Xu J, Jiang L, Zhang W, Wang J, Zhai H, Wan J (2010). Breeding strategies for optimum heading date using genotypic information in rice. Mol. Breed. 2010;25:287–298.

Weigel D, Mott R (2009). The 1001 genomes project for Arabidopsis thaliana. Genome Biol. 2009;10:107. DOI: 10.1186/gb-2009-10-5-107

Wilson ID, Barker GL, Lu C (2005). Alteration of the embryo transcriptome of hexaploid winter wheat (*Triticum aestivum* cv. Mercia) during maturation and germination. *Functional and Integrative Genomics*.2005;5(3):144–154.

Wilson ID, Barker GLA, Beswick RW (2004). A transcriptomics resource for wheat functional genomics.*Plant Biotechnology Journal*. 2004;2(6):495–506.

Xu K, Ismail A, Ronald P (2013). Flood tolerance mediated by the rice Sub1A transcription factor 1. Jenks M, Hasegawa P, editors. Plant abiotic stress, 2nd edition. Ames (Iowa): Wiley-Blackwell.

Yoo MJ, Szadkowski E, Wendel JF (2013). Homoeolog expression bias and expression level dominance in allopolyploid cotton. Heredity (Edinb). 2013;110:171-80. DOI: 10.1038/ hdy.2012.94

Yu HH, Xie WB, Wang J, Xing YZ, Xu CG, Li XH, Xiao JH, Zhang QF (2011). Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. PLoS One, 2011;6(3):e17595.

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science. 2002, 296: 79-92. 10.1126/science.1068037.

Zenoni S, Ferrarini A, Giacomelli E, Xumerle E, Fasoli M, Malerba G, Bellin D, Pezzotti M, Delledonne M (2010). Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. Plant Phisiol . 2010;152:1787–1795.

Zheng P, Allen WB, Roesler K, Williams ME, Zhang S, Li J, Glassman K, Ranch J, Nubel D, Solawetz W, Bhattramakki D, Llaca V, Deschamps S, Zhong GY, Tarczynski MC, Shen B. (2008). A phenylalanine in DGAT is a key determinant of oil content and composition in maize. Nat. Genet. 2008;40:367–372