

Applications of genomics in Plant Breeding

Ratemo .B. Omboki^{1*}, Weiren Wu¹, Xiaofang Xie¹, Gandeka Mamadou²

Fujian Provincial Key Laboratory of Crop Breeding by Design, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China.

Fujian Provincial Key Laboratory of Crop Molecular and Cell Biology, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China.

Corresponding author email: billenje@gmail.com

ABSTRACT: Genomics is a major discipline that contributes greatly to plant breeding research and food security at large. The goal of attaining food security would therefore be greatly boosted if we can employ this discipline in many of our projects. Its branches; comparative, functional and structural genomics generate vast amounts of data together with NGS technologies. Major cereal genomes have been sequenced which has greatly contributed to genome selection studies and their application in breeding. NGS has enabled assignment of markers to specific loci using gene mapping. In this paper we will review applications of genomics to plant breeding and contributions brought on board by other technologies and how they have been exploited.

Key Words: Collinear, Eco-TILLING, Metabolomics, Phenomics, Polonator, Transcriptomics.

INTRODUCTION

Recent plant breeding programs take into account available genomic information of cereal crops. Genomes that have been sequenced include rice, wheat, barley, maize among others. The population of the world has increased at a faster rate in the past decade as compared to the rates it increased decades ago. Plant genomic data is being utilized in genetic engineering to ensure better and fitter varieties of crops are available to ensure food security to the masses. According to estimates from the UN, world population is expected to increase by over 50% by the year 2050 (Ronald, 2008) Recent reports are suggesting that by the middle of this century, world population is expected to hit 9 billion. This will lead to increased demand for food by over 70% (Godfray, et al., 2010)

Cereals are very important crops in the lives of human beings. They provide over sixty percent of proteins and calories in our everyday meals (Rajeev, David , & Akhilesh, 2006). Vasanth & Zhangyue (2014) reported that countries with high populations such as India and China have invested heavily in research on cereals as a way to ensure that their people do not go hungry. Demand for food in both countries has been higher because of high economic growth rates that led to changed food consumption patterns.

A genome is defined as the entire genetic make up and it contains all the hereditary information of an individual organism which are known as genes. They control all traits exhibited by an organism. Some genes are known to be pleiotropic. Dissection of the mechanisms of activity of genes is one of the major goals of many research programs. Plant design aims to enable them resist biotic and abiotic stresses, resist attacks by pathogens, improve their yields and nutritive values. Inheritance of traits that control all these features is complex. This is where genomics comes into place to provide molecular insights to understand and improve on these traits. Plant genomics has vast potential in crop breeding since it provides very valuable information that can be exploited to improve on varieties that we currently rely on in our breeding programs (Siva, Ramesh, Ibrokhim, & Jafar, 2012)

Hence Genomics is the sequencing and analysing of an organism's entire genetic make up to understand their structure, functions and how they are passed down from one generation to the next.

Comparative, functional and structural genomics & Plant breeding

Comparative genomics aims to compare genomes which co-evolved to find regions of agronomic importance that can be exploited to improve on cereal varieties that are currently available. Polyploid nature of some of the triticeae genomes has slowed down the pace at which genomics can modify parameters to suit needs of consumers and processors of grain crops. (Feuillet & Eversole, 2007) Comparisons reveal gene re-

arrangements that have occurred over the years due to evolution. It has also helped in identification of genes that control observed phenotypes. Virulence genes in wheat and barley were identified through comparative analysis of wheat, barley and rice genomes (Yoshida, et al., 2010). Comparative genomics has also aided in the discovery of new regions of breeding interest. After this, targeted re-sequencing can be performed to identify new alleles in related wild varieties or cultivars (Saintenac, Jiang, & Akhunov, 2011). Identification of coding sequences in a genome is a rather complex adventure. However, a good thing is that they are generally conserved across closely related species. This aids in creation of gene models that help in design of algorithms that will help in their identification (Flicek, et al., 2013). On the other hand, identification of regulatory elements in a genome is difficult, but with the help of comparative genomics its easy to locate where they are found (Jessica & Kerstin, 2013)

Cereal genomes have been reported to be largely colinear in nature and thus comparative genomics can aid in their improvements (Devos, 2005). Apart from the major cereal crops i.e. wheat, rice and maize, other grain crops that can contribute to food security like sorghum, millet and rye have not received much attention. They are , in most cases referred to as orphan crops. Transfer of genomic information using bioinformatics and comparative genomics from the major cereal crops to the orphan crops can go a long way in improving them agronomically (Rebecca, Rosamond, & Molly, 2004).

The benefits that accrue from transferring of genomic information from one species to another trickle down in different ways. They can lead to improved crop biodiversity analysis and the subsequent identification of better varieties, alleles and allele combinations through marker assisted selections and cloning of genes and their transfer between various taxa (Rebecca, Rosamond, & Molly, 2004). Sometimes, transfer of better alleles is not only from major crops to orphan crops, it can be reciprocal. For example, better alleles that enable superior resistance to stresses can be found in orphan varieties which if transferred to the major cereals can greatly boost efforts put in to ensure man is food secure. Therefore, it is wise to invest in research that encompasses all types of grain crops to improve nutrition and incomes of people in poor regions in the world.

Structural genomics aims to identify the shapes of proteins and other biomolecules coded for by genes via experimental and computational methods. Most of the time, they are rational and driven by targets in a region of interest in a genome. It works by linking structure to function (Structural genomics, 2006). Few cereal genomic sequences are available at the moment i.e. rice, maize etc. To fully exploit the genomic information from a genome, it is necessary to analyze the sequences. This is normally done using Expressed Sequence Tags (ESTs) to reveal gene contents of organisms to enable high density mapping. However, in most cases they are only estimates that need to be ascertained further (Nils, 2007). Since there is an increasing number of genomes that have been sequenced, proteomics technology is needed to study the proteins that are coded for by the genes. 2-D PAGE and various Mass Spectrometric techniques are the best methods which are widely used to study the expressed proteins. This will go along way in increasing our current knowledge on plant physiology and chemistry of cereal and crop plants (Helene & Jacque, 2004). Research on protein expression using mass spectrometry has advanced from observation and quantitation to interactions with each other, phosphorylation, quantitative profiling and sub-cellular localization (Silke, Helmut, & Bettina, 2009).

The study of molecular responses of plants using gene expression data as a result of perturbations is not enough to comprehensively reveal their initiation and regulation. Structural genomics can help in revealing how transcription and translation occurs since they also control the responses of plants due to stresses (A, et al., 2011). Proteomics studies have been applied in studying plants' responses to abiotic, biotic stresses, and plant development (Herve, Ezequiel, Ima, & Wilhelm, 2013). There is possibility of developing markers using crop proteomics by reference sequences such as of rice and arabidopsis or other cereal crops that have been sequenced which will aid in the use of proteomic biomarkers in plant breeding and genetics research; however this approach has not been explored comprehensively (M, et al., 2013). With advances in technology, plant proteomics can be able to give insights into flower development, cell division and differentiation, plant organs, development of seeds and their germination that can enable us improve crop cultivars by understanding more on plant development (Tomas, Tibor, & Jozef, 2011).

Functional genomics is a field that uses molecular biology techniques to study data that is generated from a genome. It studies functions of genes, their interaction with each other, their expression profiles, how they are regulated and the responses they exhibit phenotypically. Functional genomics also involves metabolomics and phenomics (Nachithu & Ponnusamy, 2011) One of the major challenges in the past few years in relation to plant functional genomics is the transfer of genomic models into crops. Such data needs to be integrated to realize improvements in breeding programs i.e. genomic and crop phenotypic data ought to be linked (K, K, M, & S, 2004) Development of technical and other resource rich platforms has contributed to the advances currently available for functional genomics. Microarrays and expression profiling, mutant libraries and full length c-DNAs are such

examples (Yunhe, et al., 2012). The behavior of genes when they have been knocked-out; knocked-down, or ectopically expressed is contained in the mutant libraries by use of mobile genetic elements or T-DNAs. They are a few of the many ways in which gene functions are investigated (L.J, et al., 2005). Microarrays investigate expression of genes which determines if they are tissue or specific to some cells. This information can also give insights into the way pathways are regulated in a genome. RiceXpro database contains genetic expression information of japonica rice cultivar in different natural environments (Y, et al., 2011). Collection of full lengths of c-DNAs can be very helpful for breeding programs because they enable the identification of introns, exons, regulatory elements and other regions of agronomic interest by researchers. Over 20,000 of Indica rice's c-DNAs have been collected from its related cultivars formed gave excellent data for comparative genomics (X, et al., 2007) Plant functional genomics has led to advancement of other related fields. One such area is the area of QTLs. QTL 'ome' is the science of identifying their locations, major and minor alleles, investigating their effects in crops via cloning and how they can be exploited for crop improvement purposes. QTL 'omics' has come of age since hundreds of genes have been mapped, cloned and their functions ascertained in crops (Silvio & Roberto, 2015)

Next Generation Sequencing technologies and plant breeding

Next Generation Sequencing (NGS) has revolutionized plant genomics. It has increased the rate at which a genome sequenced and its reads assembled to make sense. Genome sequencing technology has progressed from the Sanger or rather the chain termination methods, to more recent ones like 454 from Roche, Solexa from Illumina, SOLiD, a creation of Applied Biosystems, Heliscope from Helicos Biosciences and the Danaher/Azco polonator (Olena & Marco, 2008, N, 2007)

Roche's 454 sequencer uses pyrosequencing process that applies the Sequencing By Synthesis (SBS) approach. DNA coated beads and Emulsion PCR are other key components of this technology. Calling of raw bases using this technology is over 90% and it is considered to be very good. However it has a problem of INDELS that makes it error-prone and which can make further processing of reads labor intensive (S & A, 2010, Jun, Rod, Ahmed, & Genfa, 2011)

Solexa's Illumina technology also uses the SBS approach. It is the most commonly used platform. It has the advantage of generating superior reads. Illumina has MiSeq and HiSeq platforms with the latter generating over 2000Gb per run. Their raw base accuracy is over 99%. (S B. , 2004) Quality filtering was reported to reduce rates of errors in reads generated by the HiSeq platform (Minoche, Dohm, & Himmelbauer, 2011).

SOLiD platform adopts the Sequencing By Ligation (SBL) approach which is massively parallel. Its able to process two slides at a time, and provides very good quality of data. However its library preparation which begins with an emulsion PCR as in the case of 454 prior to sequencing, is very intensive and time consuming. The Azco polonator technology uses SBL technique and it aims to rapidly reduce the costs of sequencing genomes. It is bead based and uses emulsion PCR to amplify fragments. Reads generated are short and are easily mappable. (Olena & Marco, 2008). Heliscope is a technology which at times is referred to as Single Molecule Real Time (SMRT) sequencing that is massively parallel. It detects each nucleotide as it is being synthesized using fluorescence technology. It has the ability to read a single sequence at a very fast speed and hence could contribute to the reduction of costs of sequencing (K, S, & J, 2009, E, J, & A, 2009). Next-NGS will be much more involved with Single Molecule Sequencing (SMS) technology because it can read a sequence at very fast rates and thus lower down costs. A variety of platforms are being developed to promote this technology. Pacific Biosciences has developed a single fluorescence based single molecule sequencing technology that identifies nucleotides which have been linked with distinct colors. Nanotechnology has been employed in SMS to monitor activity of polymerase enzyme on a DNA strand in SBS technology. LightSpeed Genomics is developing a technology that will employ the use of electron microscopy in detection of signals. They all aim to generate accurate data and high quality reads (Pushpendra, 2008, Xu, Fujita, & Hagamata, 2009)

Bioinformatics is a discipline that manages vast biological data that is generated from comparative, functional and structural genomics and other related omics technologies. It has evolved over the years due to emergence of powerful algorithms that can be used to manage and manipulate the existing data in order to be biologically meaningful. Bioinformatics data is stored in public and private databases. Databases that are of research interest include DNA Databank of Japan, NCBI, Gramene database among others. Data needs to be modelled to help in creating desired varieties of crop cultivars.

Access to data is a critical step in research. Systems need to be adopted and developed which can increase the rates at which bioinformatics skills can be applied analyze, compare and describe crop traits that can be of agricultural interest (Kasprzyk, et al., 2004). Data from NGS can be meaningless unless the reads generated are aligned and assembled. Raw data generated can be problematic to analyze (Pop & Saltzberg, 2008).

Sequencing technologies generate different types of reads due to their usage of different platforms, which further compounds the problem of data analysis. Several tools and kits have been developed which use Perl and R to statistically analyze the data for mostly 454 and Illumina platforms (Tongwu, et al., 2011, A, et al., 2009). Alignment of genomes to reference genome, commonly referred to as *De Novo* genome assembly, is the most important step in making sense out of NGS data (Flicek & Birney, 2009). Short reads is an important characteristic of advanced sequencing technologies. A mapping and assembly tool has been developed which produces high quality using a reference genome and has been used in several projects (S, et al., 2010). After assembly the data need be annotated and functionally predicted to identify functional variants like SNPs etc and the changes they bring to sequences (J & A, 2009). Many companies have developed software packages that can be easily used by researchers and other people with considerably less informatics knowledge to input, analyze like CLCBio, SoftGenetics, Roche, DNA STAR and many more. Due to vast data that is generated from NGS projects, cloud computing has been adopted to provide high level technological support for high level processing (Editorial, 2010).

Other Omics Technologies and plant breeding

Metabolomics, transcriptomics, phenomics, epigenomics and phylogenetics are other technologies employed in plant breeding research. A metabolome is an entire set of molecules found in an organism which is due to gene expression and interaction of expressed proteins among themselves and with other environmental factors (J, et al., 2011). Metabolite profiling is necessary to reveal compositional quality of plant products. It has been done using a variety of techniques i.e. chromatographic and mass spectrometric techniques, and Nuclear Magnetic Resonance. They have served to reveal metabolomes and provided insights into metabolic networks involved in their regulation which have in turn assisted in breeding of crops with desired compositional qualities (Alistair & Nicolas, 2009, Bekzod, Soren, & Soren, 2014). Transcriptomics studies the entire set of RNA sequences expressed during the process of transcription in a genome. It has helped in identification of transcripts which control yields, tolerance to biotic and abiotic stresses and respective mechanisms that are involved which has contributed more to food security (Shin & Makoto, 2008). RNA-sequencing has revealed genes and networks involved in phenotypic and genotypic plant responses due to changing environmental factors. RNA-Seq data from many plants has been made available which has enabled comparative transcriptomics to better identify genes and networks involved in plant responses (Pinky, et al., 2014). Phenomics is a discipline that aims to bridge gaps between plant biology, genomics and traits of agronomic value. Due to the increasing number of available genomic sequences, there is need to have a database that links certain gene sequences to plant structure, their development and performance in a given environment. Thus phenomics aims to measure physiobiochemical responses of plants due varying genetic and environmental pressures (Robert & Mark, 2011) and subsequent identification of QTLs involved among others (M, Y, A, & P, 2009). Epigenomics is concerned with the study of inheritance of traits that is not fully explained by variation observed in DNA sequences. Nathan (2013) reported that information generated has revealed how plant epigenomes respond to changing stresses in the environment, and how epigenetic information has been inherited down generations; which has contributed to discovery of patterns and mechanisms involved in plant evolution (Concepcion, Kyria, & Brandon, 2014). Epigenomics thus has contributed to designing of better plant varieties and better breeding programs. Phylogenetics aims to investigate and discover evolutionary relationships that exist amongst organisms. Due to the advances in next generation sequencing, more genomic data is readily available that can help to infer relationships and identify, if any, signatures of common ancestry which can promote breeding technologies. Earlier on, ribosomal and Chloroplast DNA were used to study relationships in the process of evolution. Presently, SNPs, MADS-box, Homeo box genes, and enzymatic genes like Amylases among others are being used as nuclear markers to study plant evolution (Elizabeth & Jun, 2013) and how they are distributed all over the world in phylogeography (John, Sarah, Amanda, Bryan, & Robb, 2013).

Marker Assisted Selection and Plant Genomics

Molecular markers are unique molecules that are able to distinguish an organism from the rest, reveals special features in a sample that contains genetic information. Although markers have been used for the past two decades in the form of cDNA-RFLPs, functions could not be assigned to them. Markers are used to assess genetic diversity which is important for optimization of germplasms for utilization in agriculture (R, B, & C, 2014). SSRs and SNPs are the most commonly used molecular markers for plant breeding research. Advances in NGS has led to rapid development of DNA markers for molecular research (Stephane, Victor, & Gregory, 2012) and have also substantially reduced time spent in identifying and locating DNA markers linked to traits of interest (H, et al., 2012). Predictive DNA markers associated with genomic regions are essential in genomic and integrated crop breeding

projects. QTLs and Bulk Segregant Analysis, BSAs, are predictive and hence are perfectly fit for this adventure. They have been used to find out genomic region(s) that control a phenotype of interest or identify likeliset DNA markers that could be precise for indirect selection or of agronomic interest (Yunbi, 2010). Genomic selection has been used in plant breeding. It uses high density markers and is thought to revolutionize crop breeding and contribute massively to food security in the near future. It has the advantage of predicting effects of all markers on all loci and even captures effects of small QTLs with minimal contributions (Zeratsion & Rodomino, 2014). Since plant breeding is a discipline that is highly dependent on prediction, accuracy is a key factor for its success. Gains from selection are directly proportional to the intensity of selection, accurate predictions, and the square root of additive genetic variance. Selection is also indirectly proportional to breeding cycle of a crop (Hans, Mario, Ricardo, Gustavo, & John, 2013).

Dominant and co-dominant markers are commonly used for molecular research. Studies have revealed types of markers that are best for genotyping, clustering and assigning of varieties into populations. DArT (Diversity Array Technology), SNPs and SSRs were compared and contrasted. SSRs were found to be the best because they gave good results after using only few loci followed by SNPs and DArT that need considerably many loci to assign best values (Ivan, Imad, & Theo, 2012). Construction of genomic, transcriptomic & association maps and gene have proved powerful in dissection of traits that control complex plant responses to various abiotic and abiotic stresses in the environment (Alice, Maneesha, Deepak, Laxmi, & Swarup, 2013). For orphan crops whose genomes have not been sequenced, de novo methods are employed to identify microsatellites e.g. fractionated genomic & cDNA and microsatellite rich libraries which were created for sanger sequencing are of paramount importance before designing of traditional satellite markers (Parida, et al., 2009). TILLING (Targetting Induced Local Lesions IN Genomes) has been promoted by availability of crop genomic sequences. This technique employs reverse genetics to identify traits that control phenotypes, a variant known as Eco-TILLING that can reveal polymorphisms that exist between various alleles (Mejlhede, et al., 2006)

CONCLUSIONS

Genomics and its related omics technologies has, to a fair extent, revolutionized the science of plant breeding and it is expected to contribute more due to rapid advances in sequencing of plant genomes using high throughput methods. Development of end user softwares to handle and analyze vast amounts of data from these experiments is one area that much effort should be put on so as to make sense out of such information and apply it in breeding programs. Marker technology has rapidly exerted influence in plant breeding due to advances in genomics. Functional markers have contributed to diagnostics technology that enable identification of molecules or sequences that contribute or participate during plant response to various stresses. This has enabled designing of better crop/plant varieties. The future of food security depends on availability of funding to improve agricultural practises for millions of people in developing countries who depend on it as a source of income and to ensure the worlds poor are foods secure. This can only be realized with the application of the above stated techniques in plant breeding.

ACKNOWLEDGEMENT

Many thanks to The China Scholarship Council for financial support and Prof. Weiren Wu and Dr.Xiaofang Xie for proof reading the manuscript.

REFERENCES

- A, G, B, B, V, AP, L, O, R, H, I, NM, et al.2011. Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genetics*, e1001393.
- A, MA, E, B, C, F, M, R, H, K, V, YF, Y, F.2009. PIQA: pipeline for Illumina G1 genome analyzer data quality assessment. *Bionformatics*, 25(18), 2438-2439.
- Alice K, Maneesha SS, Deepak B, Laxmi, Swarup KP.2013. Integrated genomics and molecular breeding approaches for dissecting the complex quantitative traits in crop plants. *Journal of Biosciences*, 38(5), 971-987.
- Alistair RF, Nicolas S. 2009. Metabolomics-assisted breeding: a viable option for crop improvement. *Trends in Genetics*, 25(1), 39-48.
- Bekzod K, Soren B, Soren BE. 2014. High-throughput cereal metabolomics: Current analytical technologies, challenges and perspectives. *Journal of Cereal Science*, 59(3), 393-418.
- Concepcion MD, Kyria R, Brandon SG.2014. Epigenetics and plant genome evolution. *Current Opinion in Plant Biology*, 18, 1-8.
- Devos KM. 2005. Updating the 'crop circle'. *Current Opinion in Plant Biology*, 8, 155-162.
- E, P, J, L, A, A.2009. Generations of sequencing technologies. *Genomics*, 93, 105-111.

- Editorial. 2010. Gathering clouds and a sequencing storm: why cloud computing could broaden community access to next-generation sequencing. *Nature Biotechnology*, 28, 1.
- Elizabeth AZ, Jun W. 2013. Reprint of: Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molecular Phylogenetics and Evolution*, 66(2), 539-550.
- Feuillet C, Eversole K. 2007. Physical mapping of the wheat genome: a coordinated effort to lay the foundation for genome sequencing and develop tools for breeders. *Journal of plant science*, 5, 307-313.
- Flicek P, Ahmed I, Amode M, Borell D, Beal K, Brent S, et al. 2013. Ensemble 2013. *Nucleic Acids Research*, 41, D48-D55.
- Flicek P, Birney E. 2009. Sense from sequence reads: methods for alignment and assembly. *Nature Methods*, 6, S6-S12.
- Godfray HC, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Toulmin C. 2010. Food Security: The challenge of feeding 9 billion people. *Science*, 327, 812-818.
- H, Y, Y, T, Z, Z, C, L, M, WS, J, GH. 2012. Application of next-generation sequencing for rapid marker development in molecular plant breeding: a case study on anthracnose disease resistance in *Lupinus angustifolius* L. *BMC Genomics*, 13, 318.
- Hans DD, Mario PC, Ricardo PW, Gustavo DC, John MH. 2013. Genomic Prediction in Animals and Plants: Simulation of Data, Validation, Reporting, and Benchmarking. *Genetics*, 193(2), 347-365.
- Helene BB, Jacque J. 2004. Focus on plant proteomics. *Plant Physiology and Biochemistry*, 42(12), 913-917.
- Herve V, Ezequiel L, Ima Z, Wilhelm G. 2013. Proteomics of model and crop plant species: Status, current limitations and strategic advances for crop improvement. *Journal of proteomics*, 93, 5-19.
- Ivan S, Imad E, Theo Jv. 2012. Empirical evaluation of DaRT, SNP, and SSR marker-systems for genotyping, clustering, and assigning sugar beet hybrid varieties into populations. *Plant Science*, 184, 54-62.
- J, vO, A, J. 2009. Mining SNPs from DNA sequence data; computational approaches to SNP discovery and analysis. *Methods in Molecular Biology*, 578, 73-91.
- J, WA, Ric DD, Annick M, Catherine D, Alexander E, Joachim K, Robert DH. 2011. Chapter sixteen – Plant Metabolomics and Its Potential for Systems Biology Research: Background Concepts, Technology, and Methodology. In *Methods in Enzymology* (Vol. 500, pp. 299-336).
- Jessica A, Kerstin LT. 2013. Comparative genomics as a tool to understand evolution and disease. *Genome Research*, 23, 1063-1068.
- John EM, Sarah MH, Amanda JZ, Bryan CC, Robb TB. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, 66(2), 526-538.
- Jun Z, Rod C, Ahmed B, Genfa Z. 2011. The impact of next-generation sequencing on genomics. *Journal of Genetic and Genomics*, 38(3), 95-109.
- K, J, K, P, M, S, S, T. 2004. A cross-genomic approach for systematic mapping of phenotypic traits to genes. *Genome Research*, 14, 109-115.
- K, VV, S, AD, J, DD. 2009. Next-generation sequencing: from basic research to diagnostics. *Clinical Chemistry*, 55, 641-658.
- Kasprzyk A, Keefe D, Smedley D, London D, Spooner W, Melsopp C, et al. 2004. EnsMart: a generic system for fast and flexible access to biological data. *Genome Research*, 160-169.
- LJ, v E, G, D, P, P, R, G, C, G, W, et al. 2005. EU-OSTID: a collection of transposon insertional mutants for functional genomics in rice. *Plant Molecular Biology*, 59, 99-110.
- M, H, G, F, K, N, X, Q, P, DO, W, M, et al. 2013. Construction of plastid reference proteomes for maize and Arabidopsis and evaluation of their orthologous relationships; the concept of orthoproteomics. *Journal of proteome research*, 491-504.
- M, R, Y, M, A, I, P, L. 2009. Phenotyping approaches for physiological breeding and gene discovery in wheat. *Annals of Applied Biology*, 155(3), 309-320.
- Mejlhede N, Backes G, Rasmussen S, Kyjovska Z, Burhenne K, Jahoor A. 2006. Eco-TILLING for the Identification of allelic variation in the powdery mildew resistance genes *mlo* and *m1a* in Barley. *Plant Breeding*, 125, 461-467.
- Minoche A, Dohm J, Himmelbauer H. 2011. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biology*, 12(11), R112.
- N, H. 2007. Advanced sequencing technologies and their wider impact in microbiology. *Journal of experimental biology*, 210, 1518-1525.
- Nachithu S, Ponnusamy R. 2011. 9 – Functional genomics. In *Concepts and Techniques in Genomics and Proteomics* (pp. 123-145). WoodHead Publishing.
- Nathan MS. 2013. Epigenetics and crop improvement. *Trends in Genetics*, 29(4), 241-247.
- Nils S. 2007. Triticeae genomics: advances in sequence analysis of large genome cereal crops. *Chromosome Research*, 15, 21-31.
- Olena M, Marco AM. 2008. Applications of next-generation sequencing technologies in functional genomics. *Genomics*, 92(5), 255-264.
- Parida S, Kalia S, Kaul S, Dalal V, Hemaprabha G, Selvi A, et al. 2009. Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theoretical Applied Genetics*, 118, 327-338.
- Pinky A, Swarup KP, Arunima M, Sweta D, Iny EM, Naveen M, Akhilesh KT. 2014. Expanding frontiers in plant transcriptomics in aid of functional genomics and molecular breeding. *Biotechnology Journal*, 9(12), 1480-1492.
- Pop M, Saltzberg SL. 2008. Bioinformatics challenges of new sequencing technology. *Trends Genetics*, 24, 142-149.
- Pushpendra KG. 2008. Single-molecule DNA sequencing technologies for future genomics research. *Trends in Biotechnology*, 26, 602-611.
- R, KS, B, BG, C, NS. 2014. From genomics to functional markers in the era of next-generation sequencing. *Biotechnology Letters*, 36, 417-426.
- Rajeev KV, David AH, Akhilesh KT. 2006. Advances in cereal genomics and applications in crop breeding. *Trends in Biotechnology*, 24(1), 490-499.
- Rebecca JN, Rosamond LN, Molly MJ. 2004. The role of genomics research in improvement of 'orphan' crops. *Crop Science*, 44, 1901-1904.
- Robert TF, Mark T. 2011. Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, 16(12), 635-644.
- Ronald DL. 2008. *Population* (2nd ed.).
- S, B. 2004. Solexa Ltd. *Pharmacogenomics*, 5, 433-438.
- S, BS, K, JB, C, L, A, WB, H, KT, K, MD, M, JB. 2010. Exome sequencing identifies the cause of a mendelian disorder. *Nature Genetics*, 42, 30-35.
- Saintenac C, Jiang D, Akhunov E. 2011. Targeted analysis of nucleotide and copy number variation by exon capture in allotetraploid wheat genome. *Genome Biology*, 12(9), R88.
- Shin T, Makoto M. 2008. Genetic approaches to crop improvement: responding to environmental and population changes. *Nature Reviews Genetics*, 9, 444-457.
- Silke O, Helmut EM, Bettina W. 2009. Advancements in plant proteomics using quantitative mass spectrometry. *Journal of Proteomics*, 72(3), 545-554.

- Silvio S, Roberto T. 2015. The crop QTLome comes of age. *Current Opinion in Biotechnology*, 32, 179-185.
- Siva PK, Ramesh B, Ibrokchim YA, Jafar AM. 2012. *Genomics-Assisted Plant breeding in the 21st century: Technological advances and progress, Plant Breeding*. InTech.
- SM, AB.2010. Exploring research frontiers in microbiology: the challenge of metagenomics in soil microbiology. *Research Microbiology*, 161, 497-505.
- Stephane D, Victor L, Gregory DM. 2012. Genotyping-by-sequencing in plants. *Biology*, 1, 460-483.
- Structural genomics.2006. In Ganten, Detlev, Ruckpaul, & Klaus (Eds.), *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine*. Springer Berlin Heidelberg.
- Tomas T, Tibor P, Jozef S.2011. Differential proteomics of plant development. *Journal of Proteomics*, 74(5), 577-588.
- Tongwu Z, Yingfeng L, Kan L, Linlin P, Bing Z, Jun Y, Songnian H.2011. BIGpre: A Quality Assessment Package for Next-Generation Sequencing Data. *Genomics, Proteomics & Bioinformatics*, 9(6), 238-244.
- Vasanth PG, Zhangyue Z. 2014. Food demand and the food security challenge with rapid economic growth in the emerging economies of India and China. *Food Research International*, 63(Part A), 108-124.
- X, L, T, L, S, Y, Y, L, Y, H, T, H, et al. 2007. A collection of 10,096 indica rice full-length cDNAs reveals highly expressed sequence divergence between *Oryza sativa* indica and japonica subspecies. *Plant Molecular Biology*, 65, 403-415.
- Xu M, Fujita D, Hagamata N. 2009. Perspectives and challenges of emerging single-molecule DNA sequencing technologies. *Small*, 5(23), 2638-2649.
- Y, S, B, AA, N, N, H, T, H, M, K, K, et al. 2011. RiceXpro: A platform for monitoring gene expression in japonica rice grown under natural field conditions. *Nucleic Acids Research*, 39, D1141-D1148.
- Yoshida T, Nishida H, Zhu J, Nitcher R, Distefeld A, Akashi Y, Dubcovsky J. 2010. Vrn-D4 is a vernalization gene located on the centromeric region of chromosome 5D in hexaploid wheat. *Theoretical Applied Genetics*, 120, 543-552.
- Yunbi X.2010. *Molecular Plant Breeding*. CAB International.
- Yunhe J, Zhaozia C, Weibo X, Tuan L, Huihui Y, Qifa Z.2012. Rice functional genomics research: Progress and implications for crop genetic improvement. *Biotechnology Advances*, 30(5), 1059-1070.
- Zeratsion AD, Rodomino O. 2014. Genomic selection: genome-wide prediction in plant improvement. *Trends in Plant Science*, 19(9), 592-601.