Genomics Researches and Their Applications in Plant Breeding at PAG XXIV Conference

Franz Marielle Nogoy1, Shadi Rahimi1, Kwon-Kyoo Kang2*, Yong-Gu Cho1*

1Department of Crop Science, Chungbuk National University, Cheongju 28644, Korea
2Department of Horticulture, Hankyong National University, Anseong 17579, Korea

ABSTRACT The rise of whole genome sequences of different plants provided more understanding about the gene regulation and genome evolution in further studying plants. More and more pathways and networks are identified by novel gene discoveries. Therefore, the Plant and Animal Genome Conference (PAG XXIV) provides a good venue to share the recent progress in the area of plant research genome sequencing technologies in various plants. However, this information can make a powerful system for developing improved crop varieties. By the way, the genome annotation and assembly is an essential key for breeding of stress-tolerant plants. PAG XXIV demonstrated different works about the extensive use of genomic databases accompanied by bioinformatics tools to accelerate breeding methods, discovery of new approaches to genomics, further increasing biomass of bioenergy crops, and explaining the genetic mechanisms in plant growth and defense. This review article summarizes some of the researches in various plants of rice, corn, wheat, cottonwood, switchgrasses, Thinopyrum, wheatgrass and Arabidopsis presented in PAG XXIV with the focus on genome technologies and their applications in plant breeding.

Keywords PAG conference, Plant breeding, Biotechnology, Genome, Database

INTRODUCTION

Over the past decades, there are many various research advancements that were made possible because of the rapid emergence of high-throughput technologies. In Plant Breeding and Biotechnology, it is very undeniable that the possibility of genome sequencing has generated new knowledge on crops' response to abiotic and biotic stresses, single nucleotide polymorphism (SNPs) and InDels variations, gene functions and the likes. In plants, since the completion of whole genome sequencing of Arabidopsis thaliana in 2000 (Arabidopsis Genome Initiative 2000), a lot more of crop sequencing projects have been done. One of the successful projects is in Oryza sativa, which was completed in 2005 (International Rice Genome Sequencing Project 2005).

Early this January, scientists from different parts of the world gathered in the Plant and Animal Genome XXIV Conference (PAG) held at the Town and Country Hotel, San Diego, California, USA. PAG is organized annually to gather researchers on plant and animal genome projects as well as product exhibitors to share recent developments and new ideas in their work. The program was full of plenary presentations, workshops, posters and genome database/software demonstrations. The event was an excellent venue to exchange ideas and applications on this globally important project (Intlpag.org 2016). PAG covers a very wide range of topics in plants alone, so this article is written to highlight only some of the recent studies in genome technologies and their applications in plant breeding based on the presentations given during the conference. In this case, PAG revealed identification of genomic locations of genes affecting a diversity of adaptive and productivity traits became possible through the
population and associated genome sequence data (Naithani et al. 2016) which could enable the genetic engineering of stress-tolerant, high-yielding varieties. Extensive comparative RNA-Sequencing among the different species could understand the evolution of gene regulation in the plant subgenomes, as well (Gordon et al. 2016).

In this review, we draw on some relative information gathered from PAG conference for a comprehensive understanding of the current courses in Plant Breeding and Biotechnology including applied breeding, approaches in genomics, bioenergy crops and its biomass productivity, stress responses and genomic databases and software. This review makes an overview of the current studies in genome technologies being done in different countries and helps them to improve the frameworks of their own studies as well as inspire those people working on national programs to alleviate hunger and poverty by connecting state-of-the-art science with the needs of the local farmers.

**APPLIED BREEDING**

Since the start of growing crops for food consumption, farmers already have instincts to select good traits based on physical appearance. They reserve the seeds of the good plants and plant them the next planting season. As the world increases its knowledge in science and technology, farmers and breeders become aware of genes with desirable phenotypes and thereby creating more varieties of crops available. But other than the genes that encode particular proteins that arise in a specific trait, it is now known that there are many other factors that transcribe the genome like non-coding RNAs. Here are some of the current studies in breeding and genetics from PAG conference that proves the advancement in this area.

First off, to study the complexity of molecular mechanism of hybrid vigor in maize, the first dynamic transcriptomes were presented to switch on heterosis in cultivated maize (Ji 2016). Ji (2016) analyzed the functions of NTs and DEGs with high significance between hybrids and parental inbred lines. These results can create a gene pool which is involved in switching on heterosis in maize for prediction of hybrid performance and improvement of the efficiency in maize hybrid breeding programs.

From maize to wheat, key flowering-time genes VRN1, VRN2, VRN3, PPDI, and TFL1 cloned and characterized in the forage grass species Thinopyrum elongatum and T. ponticum (tall wheatgrass) (Pahari et al. 2016). It was showed that the coding sequences for flowering genes in wheatgrass were highly conserved with annual cereals, however, non-coding regions were less conserved and can be important in the regulation of perennial trait. These results are helpful to develop perennial wheat breeding on a repeatable scale.

Diverging from traditional breeding, Cao et al. (2016) developed aneuploids in wheat using the established maize-pollen based doubled haploid (DH) technology. Eight stable DH lines were obtained with chromosome numbers ranging from 42 to 48, and these lines showed resistance to a particular disease (fusarium head blight [FHB], powdery mildew, leaf rust and stem rust), which can be used directly as new sources of resistance in wheat breeding programs.

**PHENOTYPING SYSTEM**

There are some advance works on field phenotyping that accelerates the genetic progress in plant breeding. Gouache et al. (2016) showed the first prototype based on reflectance measurements and Red Green Blue (RGB) cameras that were developed and tested to screen wheat genotypes for their response to nitrogen. This prototype was improved and automated, allowing higher throughput, and integrating light detection and ranging to access canopy architecture. This system has been tested and developed on wheat and maize response to drought via an autonomous vehicle, the Phenomobile, and a gurney mounted system on rails. Lastly, the reflectance and camera based approach has been developed into a commercial system, coined AirPhen, vectored by UAV, that is the first of a series of dedicated field phenotyping systems proposed by a spin-off company, HiPhen. The AirPhen system has been tested on maize, wheat and sugar beet to screen for nitrogen, drought and disease response. Along with cameras and sensors for field phenotyping, digital
phenotyping is also hype in crop breeding. KeyGene developed Phenoreport, an interactive data visualization tool that allows researchers to mine the results by creating tables and graphs and performing a comprehensive data analysis. It can handle large amounts of digital phenotype data and allows researchers to easily store and share the data within their own institute. It also contains a dashboard interface with all phenotypic data collected in a single research project, ranging from raw to outlier corrected data, from growth curves to scatterplots, and from results to reporting. A unique feature is that researchers can apply filters and sorting to easily draft their own tables and graphs (Biemond et al. 2016).

### APPROACHES IN GENOMICS

As said earlier, there are various factors now to look at if researchers would like to improve crops to be resistant from biotic stress, tolerant to abiotic stress, highly

#### Table 1. List of published plant genome sequences.

| Plant species (common name)                  | Family                  | Approach               | Reference                                           |
|----------------------------------------------|-------------------------|------------------------|-----------------------------------------------------|
| Arabidopsis lyrata                           | Brassicaceae            | WGS$^2$                | Hu et al. 2011                                      |
| Arabidopsis thaliana (mouse ear cress)       | Brassicaceae            | BAC-by-BAC             | Arabidopsis Genome Initiative 2000                  |
| Brachypodium distachyon                      | Poaceae                 | WGS                    | International Brachypodium Initiative 2010          |
| **Brassica rapa** (Chinese cabbage)          | Brassicaceae            | WGS                    | Brassica rapa Genome Sequencing Project Consortium 2011 |
| Cajanus cajan (pigeon pea)                   | Fabaceae                | WGS                    | Varshney et al. 2011                               |
| Carica papaya (papaya)                       | Caricaceae              | WGS                    | Ming et al. 2008                                   |
| Cucumis sativus (cucumber)                   | Cucurbitaceae           | WGS                    | Huang et al. 2009a, 2009b and Woycicki et al. 2011 |
| Fragaria vesca (woodland strawberry)         | Rosaceae                | WGS                    | Shulaev et al. 2011                                |
| Glycine max (soybean)                        | Fabaceae                | WGS                    | Schmutz et al. 2010                                |
| Glycine soja                                 | Fabaceae                | WGS                    | Kim et al. 2010                                    |
| **Medicago truncatula** (barrel medic)       | Fabaceae                | WGS and BAC-by-BAC     | Young et al. 2011                                  |
| Malus x domestica (apple)                    | Rosaceae                | WGS                    | Velasco et al. 2010                                |
| Oryza sativa (rice, indica)                  | Poaceae                 | WGS                    | Yu et al. 2002                                     |
| **Oryza sativa** (rice, japonica)            | Poaceae                 | WGS and BAC-by-BAC     | Goff et al. 2002; International Rice Genome Sequencing Project 2005 |
| Phoenix dactylifera (date palm)              | Arecaceae               | WGS                    | Al-Dous et al. 2011                                |
| Physcomitrella patens                        | Funariaceae             | WGS                    | Rensing et al. 2008                                |
| Populus trichocarpa (poplar)                 | Salicaceae              | WGS                    | Tuskan et al. 2006                                 |
| Ricinus communis (castor bean)               | Euphorbiaceae           | WGS                    | Chan et al. 2010                                   |
| Selaginella moellendorffii                   | Selaginellaceae         | WGS                    | Banks et al. 2011                                  |
| **Solanum tuberosum** (potato)               | Solanaceae              | WGS                    | Potato Genome Sequencing Consortium 2011            |
| Sorghum bicolor (sorghum)                    | Poaceae                 | WGS                    | Paterson et al. 2009                               |
| Thellungiella parvula                        | Brassicaceae            | WGS                    | Dassanayake et al. 2011                             |
| Theobroma cacao (cacao)                       | Malvaceae               | WGS                    | Argout et al. 2011                                 |
| Vitis vinifera (grapevine)                   | Vitaceae                | WGS                    | Jaillon et al. 2007                                |
| Zea mays (maize)                             | Poaceae                 | BAC-by-BAC             | Schnable et al. 2009                               |

Revised from the article of Hamilton and Buell (Plant J 2012;70:177-190).

$^2$WGS: whole genome sequencing; BAC-by-BAC: hierarchical shotgun sequencing or a clone-by-clone shotgun approach.
nutritious and high yielding. The quest for genes that would suffice the good characteristics of crops is aided by genomic technologies. Genomic technologies allow the discovery of genomes of model organisms. The data in Table 1 are from ‘Advances in plant genome sequencing’ (Hamilton and Buell 2012) which presents the list of published plant genome sequences and the approach of sequencing. Furthermore, the experimental characterization of the transcriptional units in a genome is a necessary work considering the inaccuracies and difficulties for ab initio gene prediction in eukaryotic genomes (Wellmer and Riechmann 2005). The earlier works for the characterization of transcriptional units were based on ESTs, large-scale full-length cDNA cloning, serial analysis of gene expression, massively parallel signature sequencing and rapid amplification of cDNA ends polymerase chain reaction (RACE PCR) (Wellmer and Riechmann 2005). More recently, aside from whole genome-based approaches, in PAG XXIV conference, there were researches on detecting and exploring the roles of small RNAs, DNA synthesis genes and gene promoters.

**SMALL RNAs**

Small RNAs are 21- to 24-nucleotide molecules, including microRNAs (miRNAs) and small interfering RNAs (siRNAs) (Huang et al. 2016). They play a variety of important regulatory roles in plants and animals. The direct detection of small RNAs could revolutionize analyses of small RNAs. Huang et al. (2016) described an RNA based, fluorescence sensor for small RNA detection that utilizes the fluorophore-binding sequence of ‘Spinach’, a green fluorescent protein-like RNA aptamer for which the RNA-fluorophore complex exhibits strong and consistent fluorescence under excitation wavelength when combined with the sxRNA switch for detection of miRNA-mRNA interactions.

**DNA SYNTHESIS GENES**

Aside from small RNAs, Wang (2016) has shown DNA synthesis genes with higher expression at meiocytes in Arabidopsis suggesting their potential roles in meiosis. Strong evidence was provided using molecular genetics, cell biological and biochemical approaches, that lagging strand DNA synthesis genes are required for the formation of meiotic interference-sensitive crossovers (COs), supporting the idea that meiotic recombination requires lagging strand DNA synthesis. It can suggest a novel role for DNA synthesis in the differentiation of meiotic recombination pathways.

It is indeed more complicated to analyze the interactions of genes, RNAs, proteins and other small compounds to each other despite the presence of large databases, so a graph-search empowered platform named HRGRN was developed to search known relationships and discover novel relationships among genes specific for Arabidopsis biological networks (Dai et al. 2016). The HRGRN includes over 51,000 ‘nodes’ that represent very large sets of genes, proteins, small RNAs, and compounds and approximately 150,000 ‘edges’ that are classified into nine types of interactions (interactions between proteins, compounds and proteins, transcription factors (TFs) and their downstream target genes, small RNAs and their target genes, kinases and downstream target genes, transporters and substrates, substrate/product compounds and enzymes, as well as gene pairs with similar expression patterns to provide deep insight into gene-gene relationships) to comprehensively model and represent the complex interactions between nodes.

**GENE PROMOTERS**

Promoters are important regulation factors/elements in gene expression. Thilmony (2016) used gene expression profiling data to identify candidate promoters with distinct leaf-, root-, pollen-, seed- or fruit-specific expression patterns. The corresponding upstream promoter sequences were then fused to a reporter gene and transformed into rice, Brachypodium, wheat, tobacco, tomato and/or Arabidopsis plants. The results showed two green tissue specific promoters predominantly expressed in rice leaves show light responsiveness and four additional rice pro-
motors exhibit cell type specificity within the roots of transgenic plants. Also, three rice promoters with pollen-specific expression have been characterized and used to successfully ablate transgenic pollen in Brachypodium. In addition, five fruit-specific promoter candidates from citrus have been isolated and characterized. These novel promoters can be harnessed for a more defined spatial expression of gene of interests.

Also, Zhu (2016) studied the use of heterologous promoters to regulate the expression of genes in transgenic receptors which can help prevent gene silencing. Heterologous maize promoters were engineered that regulate gene-specific expression in rice plant receptors. As a final result, a so called pZM-LEG6 with truncation (3′-deleted region of pZM-LEGF) was identified that can increase GUS activity. This can be used as a tool for transgenic breeding and genetic engineering works as well.

MARKERS

Nuclear and organelle markers were used to study the population differentiation of cottonwood (Populus deltoides). To analyze population structure and genetic diversity in eastern cottonwood, 495 unrelated individuals collected across the southeastern native range of the species were genotyped by resequencing 18,153 genes and 23,835 intergenic regions. Nuclear markers divided the population into two groups that follow largely a eastern and western distribution, with a pairwise FST value of 0.072. Organelle markers revealed genetic structure patterns that differed from the nuclear genome (Fahrenkrog et al. 2016). Their results highlight the advantages of including nuclear and organelle markers in population genetics studies to attain a higher resolution of a species’ genetic structure.

BIOENERGY CROPS AND BIOMASS PRODUCTIVITY

The main goal of breeders for crops with high potential source for biofuels is to increase biomass production and lower the gas emission produced by biofuels. In bioenergy crop researches the common trend is studying the genetics behind flowering time and biomass related genes by generating breeding population, generating SNP variation and quantitative trait locus (QTL) analysis to achieve high biomass yield. First on the list of bioenergy crops is switchgrass. The switchgrass fuel generated more than 19 million kW hours of electricity, which can meet the electrical energy needs of more than 1,800 homes for one year. Also, the sulfur emissions were reduced by 62 tons and CO₂ was reduced by over 50 thousand tons (Rinehart 2016). Rinehart (2016) studied more than 31 thousand bales of Iowa-grown switchgrass creating more than 15 thousand tons of biomass fuel. The switchgrass used in the test burn was able to replace over 12 thousand tons of Wyoming coal with a locally produced, renewable fuel. Nested association mapping (NAM) population of switchgrass (Panicum virgatum L.) was used to determine natural allelic variations for biomass yield and quality traits. 12.54 million SNPs were classified in the 18 chromosomes of a tetraploid switchgrass genome.

Miscanthus is a genus of C4 perennials grasses (Ali et al. 2016). The response of Miscanthus with high biomass potential was studied under drought stress (Malinowska et al. 2016) using RNA sequencing for gene expression profiling and compositional analysis of biomass. Sharma et al. (2016) created a mapping population between Miscanthus sacchariflorus ‘Robustus-Bluemel’ and Miscanthus sinensis var. condensatus ‘Cosmo Revert’. Chilling response RAD-seq (Restriction Associated Digestion-sequencing) data revealed significant SNP associations with Recovery Fv/Fm: percent retained after seventh day of chilling treatment minus percent retained after third day of chilling treatment. These results can show the photosynthesis efficiency and chilling treatment at maturity stage. Another mapping population was generated by Dong et al. (2016) from a cross between M. sacchariflorus ‘Robustus’ and M. sinensis var. condensatus ‘Cosmopolitan’ using restriction site-associated DNA sequencing. Comparative genomic analysis confirmed whole genome duplication in M. sacchariflorus relative to Sorghum bicolor. They identified 12 and 14 QTLs in female and male maps respectively.

In addition, manipulation of carbon flow positively have
an effect in plant biomass, De Paoli et al. (2016) used artificial microRNA (amiR) strategy to silence the NADP-malic enzyme family in *A. thaliana*. Their results revealed amiR plants with 16.4% increase in daily net CO₂ fixation rate with concomitant increase in dry biomass and 24% more sugar release from transgenic cell wall improving plant biomass for fuel production and making it as a possible solution for sustainable energy.

**FLOWERING TIME REGULATION**

The flowering time regulation in switchgrass can be helpful to improve switchgrass biomass yields. Exome-capture genotypes at 1,377,841 SNPs was used in 509 switchgrass individuals to identify associations with heading date measured over three years (Grabowski et al. 2016). Also, 372 switchgrass genotypes were studied for 11 genes involved in biomass production (Bahri et al. 2016). On average, 2/3 of SNPs in exons were non-synonymous and, of these, 50% resulted in non-conservative amino acid substitutions which one of those located in a PAD domain in Phytochrome B, a gene involved in photoperiod response.

High-throughput sequencing technologies like genome wide methylated DNA immunoprecipitation-sequencing (MeDIP-Seq) and whole genome bisulfite-sequencing (BS-Seq) were performed to compare methylation levels between two genotypes of switchgrass differed in flowering time, AP13 (lowland) and VS16 (upland) (Crampton et al. 2016). MeDIP-Seq surveys overall differences between the genomes and BS-Seq gives specific methylation status of individual cytosines. They found that AP13 had 4,667 more MeDIP-Seq peaks than VS16 (21,287 peaks) (Crampton et al. 2016). Furthermore, a comprehensive transcript profiling of the two ecotypes revealed that more than a hundred TFs were differentially expressed between the two ecotypes and several genes involved in stress responses related to drought, heat, flooding, salinity, and pathogen defense were found (Fiedler et al. 2016). However, the pathways related to biomass production including C4 photosynthesis, photorespiration, and phenylpropanoid pathways were identified via KEGG pathway analysis (Fiedler et al. 2016). Besides, RNA binding proteins, Mei2-like, pumilio, and RRM domain containing proteins and their implications were shown in posttranscriptional gene regulation. Tornqvist et al. (2016) studied a pseudo-F2 population generated from a cross between two F1 siblings that had been produced by a cross between the lowland variety ‘Ellsworth’ and the upland variety ‘Summer’ (switchgrass). These results suggested a genetic map and the results of QTL analysis for flowering time genes in switchgrass.

**MICROBIAL ENDOPHYTES**

Various plant functions and traits are co-dependent on the microbial communities (Mucyn 2016). The populations of bacterial endophytes identified within the rhizome, root, stem, leaf and seedlings grown from sterilized seeds of Miscanthus (Cope-Selby 2016). It was suggested that the vertical transmission through seed is an important process for sustaining bacterial endophytes in Miscanthus.

The well-studied root microbiome of *Arabidopsis thaliana* Col-0, was compared with two grass species model Brachypodium distachyon Bd21 and Setaria viridis A10-1 (Mucyn 2016). Microbial endophytes were studied in poplar and willow that can promote plant growth and resilience (Doty et al. 2016). These results help to promote the use of symbiosis with microbial endophytes in which substantial increase in plant biomass and yield could be reached with reduced inputs, thus, improving the economic and environmental sustainability of bioenergy production (Doty et al. 2016).

**STRESS RESPONSES**

Stress is always associated to bad things, but stress response helps us navigate an ever-changing environment and it is the reason why we are able to react to new situations to figure out if they are safe or dangerous (Borel 2015). Without it, we wouldn’t survive. The same goes for plants, which are stressed by environmental changes that threaten their health, such as flood or drought, or by critters...
that would harm or kill them, like pathogens and plant-eating insects (Borel 2015). Plants have a sophisticated manner of responding to stresses.

Now, to positively increase the response of plants to stress, a study on doubling chromosomes was done. After doubling the chromosomes of various diploid *Paulownia fortunei* plants, the obtained tetraploid plants exhibited considerable differences in phenotypic traits and tolerance to both biotic and abiotic stresses from the original diploids (Fan 2016). A large number of miRNAs can be differentially expressed between diploid and tetraploid *P. fortunei* plants and that both conserved and novel miRNAs can cleave their target mRNA sequences (Fan 2016).

**ABIOTIC STRESS TOLERANCE**

For plants’ response to abiotic stress like cold and salt, Sagehashi *et al.* (2016) identified a HSF gene, *HsfA2c*, that was induced after a mild cold treatment in rice. They inhibited the native suppressor of HSF activity, HSP90 by designing dominant negative HSP90 mutants (HSP90-D80N). As a result, they have generated transgenic rice plants by using *HsfA2c* and *Hsp90-D80N*, which showed extremely high gene expression of HSPs under cold conditions and exhibited improved cold tolerance during various growth stages. When trifoliate orange (*Poncirus trifoliata* (L.) Raf.) was exposed to cold stress, 60 TFs were responsive to the cold stress signals (Liu 2016).

A number of genes involved in the catabolism and signaling of hormones, such as abscisic acid and ethylene, were affected by the cold stress as well (Liu 2016). To better understand the transcriptional regulation rice undergoes during salt stress, genome-wide transcriptomes of two *O. sativa* Indica group varieties including the natural salt-tolerant, Pokkali and salt-sensitive breeding line, IR29 were constructed under salt stress (Geniza *et al.* 2016). More than 105,000 transcripts represent the de novo transcriptome assemblies of IR29 and Pokkali. IR29 had 6,595 unique transcripts differentially expressed whereas Pokkali had 7,209 in response to salt stress. After aligning the sequence reads with the reference Japonica genome, they identified ~190,000 SNPs and ~40,000 simple sequence repeat (SSR) sites that have potential as genetic markers.

Advances in biotechnology allow to study root development and respond to water deficit in two model monocot species, *S. viridis* and *B. distachyon* (Yee *et al.* 2016). Transgenic luciferase-expressing plants were created that can be imaged in thin sheets of soil termed GLO-Roots (Growth and Luminescence Observatory for Roots). These results can allow to analyze changes in root system architecture during growth in soil over a period of time and study water deficit responses under physiologically relevant conditions.

**BIOTIC STRESS TOLERANCE**

In stress caused by pathogens, researchers are unveiling the genes responsible for plants’ resistance. The dissection of QTL-*Fhb2* using flanking marker sequences identified 4-Coumarate-CoA Ligase (*4CL*), basic Helix Loop Helix (*bHLH*) TF, Glutathione S-transferase (*GST*) and ABC transporter as putative candidate genes localized within QTL-*Fhb2* locus of resistance gene from wheat, imparting resistance to FHB (Dhokane 2016). Soybean aphid is an important pest causing significant soybean yield losses. The *Rag2* locus confers resistance to soybean aphid biotypes 1 and 2 (Stacey 2016). RNA-seq transcriptomic analysis identified 2,361 genes significantly regulated between the resistant and susceptible lines. Two genes (unknown and mitochondrial protease) located within the defined *Rag2* locus, were expressed significantly higher in the resistant genotype. The expression of a putative nucleotide binding site-leucine-rich repeat resistant gene within the *Rag2* locus was not different between resistant and susceptible lines, but a second neuroblastoma-leucine-rich repeat gene located just at the border of the defined *Rag2* locus was. Therefore, this gene may be a candidate R gene controlling aphid resistance (Stacey 2016).

A strategy to increase tolerance to stress is by scanning available plant genome sequences for the full spectrum of nod-like receptor (NLR) fusion events to evaluate the diversity of such integrations of potential sensor domains across flowering plants, including 19 crop species (Sarris *et
NLR fusions that occur in multiple plant families were examined and it was identified that some domains show re-occurring integration across lineages. Proteins fused to NLRs overlap with previously identified pathogen targets confirming that they act as ‘baits’ for the pathogen. While some fusions have been previously implicated in disease resistance, others provide new targets for engineering durable resistance to plant pathogens (Sarris et al. 2016).

**GENOMIC DATABASES AND SOFTWARE**

Nowadays, researchers can work more efficiently with the help of different bioinformation available like the list of databases in Table 2 which are from ‘Translational genomics for plant breeding with the genome sequence explosion’ (Kang et al. 2015). The databases enlisted contain information on gene expression, genetic markers, QTLs, SNPs, genome, genomic resources or small RNA of various plant species. Researchers are not just simply doing statistics with their data now, but researchers especially those in breeding extend their skills in using codes, numbers, and symbols to further predict gene relations, protein interactions and the networking among the regulatory molecules. Large data sequences need bioinformatic tool to handle smaller sequences. Several recent bioinformatic tools for plant breeding analysis were revealed in PAG XXIV. Plant responses to abiotic and biotic stress involve large-scale transcriptional

| DB name    | Content                           | Plant species                                                                 | URL                               | Reference             |
|------------|-----------------------------------|-------------------------------------------------------------------------------|-----------------------------------|-----------------------|
| PLEXdb     | Gene expression                   | Arabidopsis, barley, Brachypodium, Citrus, cotton, grape, maize, medicago, poplar, rice, soya bean, wheat sugarcane, tomato | http://www.plexdb.org/            | Dash et al. 2012      |
| RiceXPro   | Gene expression                   | Rice                                                                          | http://ricexpro.dna.affrc.go.jp/  | Sato et al. 2013      |
| CerealsDB<sup>3</sup> | Genetic markers | Rice, Wheat                                                                 | http://wheat.pw.usda.gov/GG3/node/148 | Wilkinson et al. 2012 |
| SoyKb      | Genetic markers, genomic resources | Soya bean                                                                    | http://soyk.org/                  | Joshi et al. 2014     |
| PGDBj      | Genetic markers, QTLs, genomic resources | 80 plant species                                                             | http://pgdbj.jp/                  | Asamizu et al. 2014   |
| SoyBase    | Genetic markers, QTLs, genomic resources | Soya bean                                                                    | http://soybase.org/               | Grant et al. 2010     |
| SNP-Seek   | SNP                               | Rice                                                                          | http://www.oryzasnp.org/iric-portal/ | Alexandrov et al. 2015 |
| Phytozone  | Genome                            | 48 plant genomes                                                              | http://phytozone.jgi.doe.gov/pz/portal.html# | Goodstein et al. 2012 |
| PlantGDB   | Genome                            | 27 plant genomes                                                              | http://www.plantgdb.org/          | Duvick et al. 2008    |
| GrainGenes | Genome, Genetic markers, QTLs, genomic resources | Wheat, barley                                                                | http://wheat.pw.usda.gov/GG3/     | Carollo et al. 2005   |
| ASRP       | Small RNA                         | Arabidopsis                                                                   | http://asrp.danforthcenter.org/   | Backman et al. 2008   |
| CSRDB      | Small RNA                         | Maize, rice                                                                   | http://sundarlab.ucdavis.edu/smras/ | Johnson et al. 2007   |

Revised from the article of Kang et al. (Plant Biotech J 2015. doi: 10.1111/pbi.12449).

<sup>3</sup>URL updated.
reprogramming (Denby 2016). The gene regulatory networks underlying these transcriptional responses to the environment were elucidated using a combination of experimental and computational/mathematical tools (Denby 2016). High-resolution time series expression data was generated from Arabidopsis leaves following pathogen infection (bacterial and fungal), drought, and high light, which allowed to arrange the chronology of these stress responses and identify transient changes in gene expression. Also, transcriptional network models were generated to predict regulatory relationships between differentially expressed TFs and used network features and identify key regulators of Arabidopsis stress responses. Simulations of the network models can be used to predict how to enhance beneficial transcriptional responses to stress and as such increase the stress tolerance of Arabidopsis (Denby 2016). KBase provides a variety of plant resources and analytical services that are integrated with gene expression profiles, regulatory interactions, and metabolic networks. KBase also enables the construction and curation of metabolic models of specific microbes, fungal, and plant genomes and their interactions, to understand plant microbial and mycorrhizal interactions, and generate new hypotheses (Kumari et al. 2016).

In rice, various sequence databases enables breeders to improve agricultural production. The recent re-sequencing effort of 3,000 rice accessions by International Rice Research Institute enabled the group of Sanciangco et al. (2016) to use genome wide association mapping to examine novel alleles across the 3KRG to identify regions in the genome responsible for certain traits. They tested the robustness of historical phenotype data measured from a diverse panel from the 3K accessions to identify marker-trait associations. The historical data was collected over time in unreplicated trials as genebank characterization data for the source accession of the sequenced genetic stock. Their results showed significant associations for markers associated with several agronomic and yield traits, including culm number, grain length, grain width, 100-grain weight, and days to flowering, supporting findings from previous studies of rediscovered genes.

On the other hand, it was reported the development of a series of whole-genome SNP arrays, Rice6K, Rice60K and Rice90K (Yu et al. 2016). Rice6K SNP array contains about 4,500 high-quality SNP markers selected from re-sequencing data of more than 500 rice varieties. Rice60K SNP array (also called RiceSNP50) contains about 43,000 high-quality SNP markers selected from re-sequencing data of more than 800 rice varieties. The SNP markers are evenly distributed on the 12 chromosomes of rice, and can provide information for polymorphisms between indica and japonica subspecies as well as varieties within indica and japonica groups. Also, rice stress response interactome was created consisting of 500 rice (O. sativa) genes using WikiPathways based on published studies and analysis of transcriptomic data sets available in the public domain (Naithani et al. 2016). Set of up- and down-regulated network hubs were identified that are shared between rice biotic and abiotic stress-response subnetworks.

Our understanding of gene expression has changed dramatically over the past decade, largely catalysed by technological developments. High-throughput methods of microarrays and next-generation sequencing have generated large amounts of genome-wide gene expression data that are collected in public archives. Added-value databases process analyse and annotate these data further to make them accessible to every biologist. Here are just some of the topics covered in PAG conference during the computer demo sessions. 1) how to use public resources in The Arabidopsis Information Resource (TAIR) including registration, data searches, stock ordering, data submission and downloading large datasets (Reiser 2016), 2) presentation of Transcriptome ENcyclopedia Of Rice (TENOR) (http://tenor.dna.affrc.go.jp), a database that encompasses large-scale mRNASeq data obtained from rice under a wide variety of conditions (Kawahara et al. 2016), 3) TreeGenes, a web-based database that provides interfaces and tools to collect, curate, visualize, and analyze data from over 1,700 forest tree species. The map-based CartograTree interface connects geo referenced trees with curated phenotypic and genetic data as well as environmental metrics from sources such as Ameriflux and WorldClim (Demurjian 2016), 4) Apollo, composed of a web-based client, an annotation-editing engine, and a
CONCLUSION

Recent advances in sequencing technologies made it possible to sequence whole genome of many plants. Researchers and breeders can even harness these whole sequences using different bioinformatic tools for analysis. The presence of publicly available databases and analysis softwares can help us increase the discovery of useful traits, present novel genes and improve genetic gain. Reuse of these data can be very helpful, but there are many obstacles in its preparation and analysis and in the interpretation of the results. The gains of the future of plant breeding lies on bioinformatics groups, researchers and breeders who will work hand in hand to develop and improve plants that can adapt to the ever-changing environment.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Next-Generation BioGreen 21 Program (National Center for GM Crops No. PJ01191601), Rural Development Administration, Republic of Korea.

REFERENCES

Ali S, Chang J, Serba DD, Bhandari HS, Bartley L, Saha MC. 2016. Cataloging natural allelic variations in switchgrass (Panicum virgatum L.) NAM parental genomes. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0012.

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408: 796-815.

Bahri B, Daverdin G, Cheng JF, Barry KW, Devos KM. 2016. Single nucleotide polymorphism (SNP) variation in switchgrass (Panicum virgatum) for genes involved in biomass production. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0013.

Biemond C, Speckmann GJ, Hekkelman M, Huvenaars K, Doeswijk T, Schrieg Mv, et al. 2016. Data management and visualization in digital phenotyping. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0209.

Borel B. 2015. Do plants get stressed? [Internet]. Popular Science, Harlan, IA: [cited 2016 Feb 13]. Available from: http://www.popsci.com/do-plants-get-stressed.

Cao W, Humphreys G, Fedak G, Wolfe D, McCallum B, Fetch T, et al. 2016. Development of aneuploids in wheat by means of doubled haploid technology-corn pollination method. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0257.

Cope-Selby N, Smith D, Cookson A, Donnison I, Flavell R, Farrar K. 2016. Seed transmission of bacterial endophytes in miscanthus. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0011.

Crampton MC, Sripathi V, Ayyappan V, Saha MC, Kalavacharla V. 2016. Comparison of AP13 and VS16 switchgrass (Panicum virgatum) methylomes using high-throughput sequencing. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0014.

Dai X, Li J, Liu T, Zhao PX. 2016. HRGRN: enabling graph search and integrative analysis of Arabidopsis signaling transduction, metabolism and gene regulation networks. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W057.

De Paoli HC, Engle N, Tschaplinski TJ, Abraham PE, Hettich RL, Mewalal R, et al. 2016. Rechanneling carbon flow with potential to increase valuable biomass feedstocks. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0005.

Demurjian S, Grau E, Vasquez-Gross H, Gessler D, Neale D, Wegrzyn J. 2016. TreeGenes and CartograTree:
community resources for forest tree genomics. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0383.

Denby K, Polanski K, Penfold C, Gherman L, Consortium P. 2016. Re-designing regulatory networks underlying plant stress responses. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W932.

Dhokane D, Kushalappa AC, McCartney CA. 2016. Identification of fusarium head blight resistance genes from wheat QTL-Fhb2, through forward and reverse genetics. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0138.

Dong H, Liu S, Clark LV, Gifford J, Juvik JA, Sacks EJ. 2016. QTL mapping for biomass yield and yield components in a Miscanthus sacchariflorus x M. sinensis population. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0010.

Doty SL, Sher AW, Khan Z, Fleck ND, Firrincieli A, DeLuca TH, et al. 2016. Sustainable biomass production using plant growth promoting microbial endophytes. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0026.

Fahrenkrog AM, Neves LG, Resende M, Dervinis C, Barbazuk B, Kirst M. 2016. Nuclear and organelle population structure in eastern cottonwood (Populus deltoids). Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0437.

Fan G. 2016. The increased tolerance to biotic and abiotic stresses and the changes in transcriptome and MicroRNA activity after chromosome doubling in paulownia trees. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W848.

Fiedler E, Ayyappan V, Sripathi VR, Saha MC, Thimmapuram J, Bhide K, et al. 2016. Transcriptome analysis of two switchgrass ecotypes: AP13 (lowland) and VS16 (upland). Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0017.

Geniza M, Fox S, Fonseca N, Petryszak R, Jaiswal P. 2016. Transcriptome characterization and expression analyses in salt-tolerant and salt-sensitive cultivars of Oryza sativa after exposure to salt stress. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0128.

Gordon S, Session A, Jenkins J, Shu S, Schmutz J, Barry KW, et al. 2016. The genus Brachypodium as a tool to study polyploidy. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0006.

Gouache D, de Solan B, Baret F, Comar A. 2016. From PhénoBlé to AirPhen: field phenotyping technologies take off. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0208.

Grabowski P, Evans J, Ramstein GP, Crisovan E, Vaillancourt B, Costich D, et al. 2016. Genome-wide associations with flowering time in switchgrass. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0001.

Hamilton JP, Buell CR. 2012. Advances in plant genome sequencing. Plant J. 70: 177-190.

Huang K, Doyle F, Wurz ZE, Tenenbaum SA, Caplan JL, Meyers BC. 2016. FastMiR: an RNA-based sensor for in vitro detection of small RNA. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0027.

Humann JL, Jung S, Zheng P, Cheng CH, Lee T, Frank M, et al. 2016. Cool season food legume genome database: an up-to-date resource enabling genetics, genomics and breeding research in pea, lentil, faba bean and chickpea. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W223.

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

Ji J. 2016. Investigation of novel transcripts and differential expression genes for the heterosis metabolic function in maize transcriptomes. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0037.

Kang YJ, Lee T, Lee J, Shim S, Jeong H, Satyawan D, et al. 2015. Translational genomics for plant breeding with the genome sequence explosion. Plant Biotechnol. J. doi: 10.1111/pbi.12449. [Epub ahead of print]

Kawahara Y, Oono Y, Wakimoto H, Ogata J, Kanamori H, Sasaki H, et al. 2016. TENOR: Database for comprehensive miRNA-Seq experiments in rice. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0320.

Kumari S, Ware D, Weston D, Kumar V, Ramakrishnan S, Ranjan P, et al. 2016. The DOE systems biology knowledgebase (KBase): progress towards advanced modeling of sustainable bioenergy crop development.
Liu JH. 2016. Transcriptome-based elucidation of molecular mechanism underlying cold tolerance of trifoliate orange (Poncirus trifoliata (L.) Raf.) and exploration of cold-responsive genes. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W168.

Malinowska M, Vendramin V, Scalabrin S, Schnabel S, Donnison I, Robson PRH. 2016. Physiological and molecular analysis of drought stress responses in miscanthus: identifying important traits for biofuel production from analysis of networks. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W086.

Mucyn TS. 2016. Harvesting the root microbiome of grasses toward sustainable increase crop production. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0007.

Munoz-Torres MC, Dunn NA, Unni D, Diesh C, Elsik CG, Holmes I, et al. 2016. Apollo: improving collaborative genome curation. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0387.

Naithani S, Partipilo C, Mohanty B, Lee DY, Jaiswal P. 2016. Rice stress-response gene-network: an example of community curation in WikiPathways for plants. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W926.

Pahari S, Cradduck M, Eudes F, Laroche A, Larsen J. 2016. A step toward the development of perennial wheat: characterization of flowering genes in perennial wheatgrass and their relationship to wheat. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0042.

Reiser L. 2016. Publicly available resources for plant genomics research at the arabidopsis information resource (TAIR). Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. C04.

Rinehart L. 2006. Switchgrass as a bioenergy crop. IP302 (Slot 297). ATTRA, Butte, MT.

Sagehashi Y, Yasuda H, Sato Y. 2016. Improvement of cold tolerance in rice by cold induced expression of heat shock genes. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0132.

Sanciangco MD, Alexandrov N, Capilit GLS, Chebotarov D, Hamilton NRS, Juanillas VM, et al. 2016. New discoveries from old data: exploiting the world's largest collection of sequenced germplasm through GWAS. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0713.

Sarris P, Cevik V, Dagdas G, Jones J, Krasileva KV. 2016. Comparative analysis of plant immune receptor architectures uncovers host proteins targeted by pathogens. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0028.

Sharma S, Glowacka K, Clark LV, Hoi CG, Salgado M, Sharma S, et al. 2016. Mapping chilling tolerant photosynthesis in an interspecific miscanthus population. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0009.

Stacey G. 2016. Identification of soybean proteins and genes differentially regulated in near isogenic lines differing in resistance to aphid infestation. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W780.

Thilmoney R. 2016. A suite of crop promoters with precise organ-specific expression patterns. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W728.

Tornqvist CE, Ramstein G, Evans J, Vaillancourt B, Crisovan E, Buell CR, et al. 2016. Genetic Analysis of flowering time in switchgrass using a Pseudo-F2 population from an upland x lowland cross. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0015.

Wang Y. 2016. New insights into meiotic recombination from analyses of DNA synthesis gene functions. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. W678.

Wellmer F, Riechmann JL. 2005. Gene network analysis in plant development by genomic technologies. Int. J. Dev. Biol. 49: 745-759.

Yee MC, Sebastian J, Lindner H, Priest HD, Mockler TC, Dinneny J. 2016. Regulation of crown root growth in monocots. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0018.

Yu H, Chen H, Xie W, He H, Tang X, Zhou F, et al. 2016. Rice whole-genome SNP arrays (Rice6K, Rice60K, Rice90K) for genomic research and breeding. Proc. of the Plant and Animal Genome Conference (PAGXXIV). San
Zhu S. 2016. Cloning and characterization of a multifunctional promoter from maize (Zea mays L.). Proc. of the Plant and Animal Genome Conference (PAGXXIV). San Diego, CA, USA. P0120.