Indian J. Plant Genet. Resour. 30(1): 20-24 (2017)
DOI 10.5958/0976-1926.2017.00002.X

Application of Genomics to Enhance Utilization of Plant Genetic Resources

Robert J Henry
Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane QLD 4072, Australia

Genomics may greatly enhance utilization of plant genetic resources (Brozynska *et al*., 2016). This is especially critical with the prospect of major climate change (Aberton *et al.,* 2016). New cereal genotypes are needed to deliver desirable nutritional and functional characteristics in the environments of the future (Henry *et al*., 2016). Analysis of wild and domesticated genetic resources may identify new sources of genetic variation for breeding. Genomic analysis of cereal (especially rice and wheat) genetic resources may contribute significantly to global food security. Whole genome sequencing and transcriptome sequencing can contribute to discovery of valuable new genes and alleles and explain the molecular basis of economic traits. Utilizing plant genetic resources requires analysis of large amounts of genomic data (Rossetto and Henry, 2014)

Progress in Key Species

The following is a review of our recent progress in major plant species using DNA sequencing technology to discover the genes controlling key traits and use the increased understanding of their molecular basis to deliver new insights into strategies for accelerated plant improvement.

Rice

Re-sequencing of domesticated rice reveals parts of the genome that have little or no diversity in the domesticated gene pool (Krishnan *et al*., 2012). Sequencing of the genomes of wild rice (*Oryza*) populations has revealed new ancestral populations that widen the effective gene pool of rice for breeding. These newly identified genetic resources will provide a new source of diversity for use in breeding for disease resistance and tolerance to climate change in rice. These resources and genomics tools facilitate the development of rice genotypes to satisfy the quality requirements of increasingly discerning rice consumers (Anacleto *et al*., 2015). The breeding of hybrid rice may benefit from the use of genome re-sequencing to select parents (Waters *et al*., 2015). The diversity of repetitive sequences across the *Oryza* genus has been analysed (Copetti *et al*., 2015)

The AA Genome Primary Gene Pool of Rice

The primary gene pool of rice includes the wild *Oryza* species with AA genomes that are inter-fertile with domesticated rice. Analysis of the whole chloroplast genome of these species has been used to define their evolutionary relationships (Wambugu *et al*., 2015). Diverse grain quality attributes may be found in this gene pool (Wang *et al*., 2015). This gene pool also provides a source of biotic and abiotic stress tolerance genes for current and future environments.

New Sources of Genetic Diversity for Rice

Recent research has identified large poorly characterized populations of *Oryza* in northern Australia (Brozynska *et al*., 2014). Phylogenetic analysis of the AA genome taxa indicates that they are sister groups to the clade including domesticated taxa suggesting that this region may be an important centre of diversity for rice. Reference genome sequences for these taxa have been produced recently as a tool to facilitate analysis of diversity in these wild populations. This whole genome analysis (Brozynska *et al*., 2016) has suggested that northern Australia may be a centre of diversity and possibly of origin of the A genome clade that includes the domesticated rice in Asia and Africa. Sequencing of DNA from bulks of individuals is being used for association of polymorphisms with traits at the whole genome level in populations segregating for key traits.

Wheat

Sequencing of the transcriptome of developing seeds of diverse wheat germplasm has revealed genetic diversity that explains differences in carbon assimilation (Rangan *et al.*, 2016), flour yield in milling and bread quality on baking (Furtado *et al.*, 2015). Together these gene discoveries offer a significant opportunity to accelerate the rate of genetic gain in wheat breeding.

Carbon Assimilation

Photosynthesis in the pericarp of wheat proceeds by a C4 pathway. Genetic variation in grain photosynthesis

---

**Author for Correspondence: Email-Robert.henry@uq.edu.au* by this recently defined pathway (Rangan *et al*., 2016) may explain differences in yield due to re-fixation of carbon respired to support protein and starch biosynthesis during endosperm formation. Selection for enhanced photosynthesis in the grain may result in higher yields in environments limited by harsh conditions during late grain filling. The influence of heat stress on expression of this pathway and the apparent substantial variation in this trait within the wheat gene pool is being studied.

Flour Yield

The yield of flour obtained when wheat is milled is controlled by genes that have been identified by analysis of levels of expression in wheat genotypes differing in flour yield. This discovery will avoid the need to have large amounts of grain to assess flour milling performance and provide tools for early generation selection in wheat breeding.

Grain Hardness

The genetic variation in the hardness of wheat grain has been associated with the *pin* genes. Transcriptome analysis has demonstrated that the expression levels of these genes and not just their protein structures is the key to wheat hardness (Nirmal *et al*., 2016).

Breadmaking Quality

A highly differentially expressed gene encoding a small sulphur rich protein in wheat endosperm may explain differences in the breadmaking qualities of wheat (Furtado *et al*., 2015). However this gene may not be important for products such as chapatti. The gene has an open reading frame encoding a small sulphur rich protein with potential for involvement in cross linking of larger proteins in the gluten. Knowledge of this gene may allow the development of wheat genotypes with acceptable end use quality at much lower grain protein contents improving the nitrogen use efficiency of the crop. The combination of selection for flour yield and bread quality will ensure wheat genotypes can be selected with acceptable grain quality. A modest number of genetic loci may require selection to ensure wheat breeding delivers acceptable end use quality allowing more selection pressure for grain yield.

Sources of Diversity for Wheat

New hexaploids (synthetics) derived from the diploid progenitors may expand the range of variation found in wheat. The improved genomic tools identified by transcriptome analysis may also allow much more effective screening and utilization of variation within the existing hexaploid gene pool. Tools to support the analysis of sub-genome specific gene differences (at the genomic and transcriptomic levels) will greatly enhance progress in wheat genetic improvement.

Peanut

Peanut (groundnut) is an important pulse crop. The genomes of relatives of the diploid progenitors have been characterized. Analysis of the tetraploid domesticated peanut remains challenging. Sequencing of gene enriched genomic DNA from bulks of genotypes with extremes of the traits is being evaluated for gene and marker discovery in peanut.

Coffee

Coffee is a crop of great economic importance. The genome of coffee is being characterized (Tran *et al*., 2016). The diploid robusta coffee (Denoeud *et al*., 2014) and more complex tetraploid arabica genomes have been sequenced. Coffee bean quality results from an interaction between genotype and environment (Cheng *et al*., 2016). Introgression of genes from other species in the genus may be required to expand the gene pool of domesticated coffee to cope with the impact of climate change on current production areas. Re-sequencing of these diverse species is in progress. Sequencing of bulks is being used in association of traits with markers at the whole genome level in Arabica coffee. A transcriptome of the arabica coffee bean has been produced by long read sequencing.

Mango

Mango is a tree crop with substantial diversity in both cultivated and wild gene pools. A genome sequence will provide a platform for exploration of genetic variation in this genepool. Breeding of species with a long generation time such as trees is greatly advantaged by the availability of genome resources facilitating molecular selection at early growth stages. Tree architecture and fruit quality are among the key traits to be advanced in this and other fruit trees crops.

Macadamia

Macadamia is a good example of a recently domesticated crop. This species is genetically very distant from other species (a member of the Proteaceae) with characterized genomes. A genome sequence has the potential to make a big contribution to the understanding of such phylogenetically isolated taxa. A draft sequence of the genome has been produced (Nock *et al*., 20I6). Sequencing of the wild relatives will rapidly help suggest strategies for use in breeding improved varieties with desirable traits such as smaller tree size. Re-sequencing of this recent domesticate and the four related wild species in the genus will allow targeted expansion of the gene pool.

Sugarcane

Sugarcane is the leading industrial crop globally. This extremely complex genome is the subject of considerable sequencing effort (Souza *et al*., 2011; Hoang *et al*., 2015). The extreme polyploid of this species is a key constraint to application of genomic tools. Due to the challenge of assembly of this difficult genome, a first step towards a reference is being generated by sequencing sugarcane BAC clones that cover the sorghum genome. This 'monoploid genome' will support the analysis of the complete polyploid genome that is being generated with long read sequencing. This technology has also been used to generate a more complete set of transcripts for this complex species (Hoang *et al.*, 2016). Alternate transcripts from the same gene and transcripts from homologous genes make a very complex system in this 12X species. The genomes of the progenitor species have also been analysed. Hybrids with new combinations of these genomes and other related species from with the group may expand the sugarcane gene pool to support production of a more diverse range of products from sugarcane crops and enable production in more diverse environments.

Eucalypts

Eucalypts are some of the most widely planted trees in the world with potential as a biomass or wood crop in many environments (Healy *et al*., 2016). New reference genome sequences are being generated for diverse species within the large group of Eucalypts. Hybrids between species show significant promise as tree crops. Sequencing of bulks of genotypes is being used to associate polymorphisms with traits. These relatively small genomes (< 400Mbp) are relatively easy targets for whole genome analysis and genotyping by whole genome sequencing. Comparative genomic analysis has explored variation across the Eucalypts from *Corymbia* to *Eucalyptus*.

Domestication of New Species

Crop Wild Relatives

Wild crop relatives may be source of novel genes for crop improvement but may also become targets for domestication in their own right (Shapter *et al*., 2013). New diversified crops will contribute to the resilience of the food production system and may deliver the genotypes necessary for production in new altered environments of the future. Re-sequencing is an important tool for the characterization of lines generated in the introgression of genes from crop wild relatives into the domesticated gene pool.

Totally New Crops

Analysis of the food needs of human populations and available wild plants may identify new opportunities for totally new crop species. Most suitable plants found in agricultural societies and able to be consumed by conventional processing and food preparation technologies have probably already been domesticated. However, new regions and new technologies may reveal new possibilities. The domestication of distant relatives of rice as new food crops for production in cool climates is an example of this strategy (Shapter *et al*., 2013). In this case, the mutation of the shattering gene in a temperate grass may rapidly generate a potential new cereal crop for these environments.

Climate Adaptation

Adapting agriculture to new and variable climates (Kole *et al*., 2015) will require a strong response in crops genetics (Abberton *et al*, 2016). Knowledge of the genome and variation in the genome in response to environmental selection in wild populations can be a critical guide to strategies for agricultural selection (Henry and Nevo, 2014). These analyses has emphasised the importance of adapting crops to the new array of pests and diseases found in new environments and that successful adaptation to higher temperatures may be largely associated with adaptation to new pathogens (Cronin *et al*., 2007; Fitzgerald *et al*., 2011).

Future Prospects

Advances in Technology

Advances in DNA sequencing technology will continue to accelerate the rate of application of genomics in crop improvement. Whole genome assembly had been made simplified by recent advances in long read sequencing. Annotation of genomes and more complete characterization of transcriptomes has also been facilitated by these developments. Continued developments will be required to enable widespread whole genome analysis of plant genetic resources and association with phenotype to achieve routine use of whole genome sequence data in crop improvement. Advances in computing and bioinformatics will be necessary to fully realize this opportunity.

International Collaboration

Coordinated efforts to capture and utilize genomic and other data on plant genetic resources are essential to ensure this technology is able to impact effectively crop performance and food security. The DIvseek (www.divseek.org) initiative is a program that aims to capture the progress in genomics and in phenomics to support global crop improvement for food security. These cooperative approaches will be essential to ensure optimal use of global plant genetic resources for agriculture and food production. Plant breeders need to be able to find within global genetic resources the alleles that they need to adapt their local varieties to the challenges of future environments.

References

Abberton M, A Abbott, J Batley, A Bentley, M Blakeney, J Bryant, H Cai, J Cockram, A Costa de Oliveira, LJ Cseke, H Dempewolf, C De Pace, D Edwards, P Gepts, A Greenland, AE Hall, R Henry, K Hori, GT Howe, S Hughes, M Humphreys, AM Ismail, D Lightfoot, A Marshall, S Mayes, HT Nguyen, FC Ogbonnaya, R Ortiz, AH Paterson, PW Simon, J Tohme, R Tuberosa, B Valliyodan, R Varshney, SD Wullschleger and M Yano (2016) Global agricultural intensification during climate change: a role for genomics. *Plant Biotechnol. J.* **14:** 1095-1098.

Anacleto R, RP Cuevas, R Jimenez, C Llorente, E Nissila, R Henry and N Sreenivasulu (2015) Prospects of breeding high-quality rice in the post-genomic era. *Theor. Appl. Genetics* **128(8):** 1449-1466. doi:10.1007/s00122-015-2537-6

Brozynska M, ES Omar, A Furtado, D Crayn, B Simon, R Ishikawa and RJ Henry (2014) Chloroplast genome of novel rice germplasm identified in northern Australia. *Trop. Plant Biol.* **7:** 111-120.

Brozynska M, A Agnelo Furtado and RJ Henry (2016) Genomics of crop wild relatives: Expanding the genepool for crop improvement. *Plant Biotech. J.* **14:** 1070-1085.

Brozynska M, D Copetti, A Furtado, R Wing, D Crayn, G Fox, R Ishikawa and RJ Henry (2016) Sequencing of Australian wild rice genomes reveals ancestral relationships with domesticated rice. *Plant Biotech. J.* accepted 24/11/2016.

Cheng B, A Furtado, HE Smyth and RJ Henry (2016) Influence of genotype and environment on coffee quality. *Trends Food Sci. Technol.* doi: 10.1016/j.tifs.2016.09.003.

Copetti D, J Zhang, M El Baidouri, D Gao, J Wang, E Barghini, RM Cossu, A Angelova, CE Maldonado, S Roffler, H Ohyanagi, T Wicker, C Fan, A Zuccolo, M Chen, AC de Oliveira, B Han, RJ Henry, Y Hsing, N Kurata, W Wen Wang, SA Jackson, O Panaud and RA Wing (2015) *Oryza* Repeat Database: a resource for Genus-wide rice genomics and evolutionary biology. *BMC Genomics* **16:** 538.

Cronin, JK, PC Bundock, RJ Henry and E Nevo (2007) Adaptive Climatic molecular evolution in wild barley at the Isa Defense Locus. *Proc. Natl. Acad. Sci. USA* **104:** 2773-2778.

Denoeud F, L Paulet Carretero, A Dereeper, G Droc, R Guyot, M Pietrella, C Zheng, A Alberti, F Anthony, G Aprea, M J Aury, P Bento, M Bernard, S Bocs, C Campa, A Cenci, CM Combes, D Crouzillat, C Da Silva, L Daddiego, F De Bellis, S Dussert, O Garsmeur, T Gayraud, V Guignon, K Jahn, V Jamilloux, T Joët, K Labadie, T Lan, J Leclerc, M Lepelley, T Leroy, Li L-T, P Librado, L Lopez, A Muñoz, B Noel, A Pallavicini, G Perrotta, V Poncet, D Pot, Priyono, M Rigoreau, M Rouard, J Rozas, C Tranchant-Dubreuil, R VanBuren, Q Zhang, AC Andrade, X Argout, B Bertrand, A de Kochko, G Graziosi, RJ Henry, Jayarama, R Ming, C Nagai, S Rounsley, D Sankoff, G Giuliano, VA Albert, P Wincker and P Lashermes (2014) The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. *Science* **345:** 1181-1184.

Fitzgerald TL, FM Shapter, SMC Donald, DLE Waters, IH Chivers, A Drenth, E Nevo and RJ Henry (2011) Genome diversity in wild grasses under environmental stress. *Proc. Natl. Acad. Sci. USA* **108:** 21139-21144.

Furtado, A, P Bundock, P Banks, G Fox, X Yin and RJ Henry (2015) A novel highly differentially expressed gene in wheat endosperm associated with bread quality. *Sci. Rep.*5, 10446; doi: 10.1038/srep10446.

Healey A, DJ Lee, A Furtado, BA Simmons and RJ Henry (2015) Efficient eucalypt cell wall deconstruction and conversion for sustainable lignocellulosic biofuels. *Frontiers Bioeng. Biotechnol.* **3:** 190.

Henry RJ (2014) Genomics strategies for germplasm characterization and the development of climate resilient crops. *Front. Plant. Sci.* doi: 10.3389/fpls.2014.00068

Henry RJ and Nevo E (2014) Exploring natural selection to guide breeding for agriculture. *Plant Biotechnol. J.* **12:** 655-662.

Henry RJ, P Rangan and A Furtado (2016) Functional cereals for production in new and variable climates. *Curr. Opin. Plant Biol.* **30:** 11–18.

Hoang NV, A Furtado, FC Botha, BA Simmons and RJ Henry (2015) Potential for genetic improvement of sugarcane as a source of biomass for bioenergy. *Frontiers Bioeng. Biotechnol.* **3:** 182.

Hoang NV, A Furtado, PJ Mason, A Marquardt, L Kasirajan, PP Thirugnanasambandam, FC Botha and RJ Henry (2016) A survey of the complex transcriptome from the highly polyploid sugarcane genome using PacBio full-length isoform sequencing compared with *de novo* assembly from Illumina RNA sequencing. *BMC Genomics* accepted 5/12/16.

Kole C, M Muthamilarasan, R Henry, D Edwards, R Sharma, M Abberton, J Batley, A Bentley, M Blakeney, J Bryant, H Cai, M Cakir, LJ Cseke, J Cockram, AC Oliveira, CD Pace, H Dempewolf, S Ellison, P Gepts, A Greenland, A Hall, K Hori, GT Howe, S Hughes, MW Humphreys, M Iorizzo, AM Ismail, A Marshall, S Mayes, HT Nguyen, FC Ogbonnaya, R Ortiz, AH Paterson, PW Simon, J Tohme, R Tuberosa, B Valliyodan, RK Varshney, SD Wullschleger, M Yano and M Prasad (2015) Application of genomics-assisted breeding for generation of climate resilient crops: Progress and prospects. *Front. Plant Sci.* 6: 563. doi:10.3389/fpls.2015.00563.

Krishnan SG, DLE Waters, SK Katiyar, AR Sadananda, V Satyadev and R Henry (2012) Genome-wide DNA polymorphisms in elite indica rice inbreds discovered by whole-genome sequencing. *Plant Biotechnol. J.* **10:** 623-634.

Nock CJ, Baten A, Barkla BJ, Furtado A, Henry RJ and King G (2016) Genome and transcriptome sequencing characterises the gene space of *Macadamia integrifolia* (Proteaceae) *BMC Genomics* **17:** 937.

Nirmal RC, A Furtado, C Wrigley and RJ Henry (2016) Influence of gene expression on hardness in wheat. *PloS One* 11(10): e0164746. doi:10.1371/journal.pone.0164746.

Rangan P, A Furtado, and RJ Henry (2016) New evidence for grain specific C4 photosynthesis in wheat. *Sci. Rep.* 6, 31721; doi:10.1038/srep31721.

Rossetto M and RJ Henry (2014) Escape from the laboratory: new horizons for plant genetics. *Trends Plant Sci*. **19:** 554-555.

Shapter FM, M Cross, G Ablett, S Malory, IH Chivers, GJ King and RJ Henry (2013) High-throughput sequencing and mutagenesis to accelerate the domestication of *Microlaena stipoides* as a new food crop. *PloS one* 8(12) e82641. doi:10.1371/journal.pone.0082641.

Souza GM, H Berges, S Bocs, R Casu, AD Hont, JE Ferreira, R Henry, R Ming, B Potier, MA Van Sluys, M Vincentz and AH Paterson (2011) The sugarcane genome challenges: Strategies for sequencing a highly complex genome. *Trop. Plant Biol.* **4:** 145-156.

Tran H, SL Lee, H Smyth, A Furtado and RJ Henry (2016) Advances in genomics for the improvement of quality in Coffee. *J. Sci. Food Agr.* **96:** 3310-3312.

Wambugu PW, M Brozynska, A Furtado, DL Waters and RJ Henry (2015) Relationships of wild and domesticated rice species based on whole chloroplast genome sequences. *Sci. Rep.* **5:** 13957 doi: 10.1038/srep13957.

Wang K, PW Wambugu, B Zhang, AC Wu, RJ Henry and RG Gilbert (2015) The biosynthesis, structure and gelatinization properties of starches from wild and cultivated African rice species (*Oryza barthii* and *Oryza glaberrima*). *Carbohydrate Polymers* **129:** 92-100. doi:10.1016/j.carbpol.2015.04.035.

Waters D, GS Krishnan, E Mani, S Singh, S Vaddadi, R Henry and A Baten (2015) Genome wide polymorphisms and yield heterosis in rice (*Oryza sativa* subsp. *indica*) *Trop. Plant Biol.* **8:** 117-125.