

Genomics of plant genetic resources: an introduction

Roberto Tuberosa^{1*}, Andreas Graner² and Rajeev K. Varshney³

¹Department of Agroenvironmental Science and Technology, Viale Fanin 44, 40127 Bologna, Italy, ²Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstraße 3, D-06466 Gatersleben, Germany and ³International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad 502 324, Andhra Pradesh, India

This special issue of Plant Genetic Resources has assembled 52 short articles selected among the over 350 oral and poster communications presented during the 2nd International Symposium on Genomics of Plant Genetic Resources (GPGR2; www.GPGR2.com) held in Bologna, Italy, from April 24 to 27, 2010. The second edition of GPGR2, co-organized by Bioversity International, the Leibniz Institute of Plant Genetics and Crop Plant Research and the University of Bologna, followed the first edition organized in 2005 by the Chinese Academy of Agricultural Science in Beijing, China.

The overall objective of GPGR2 was to critically evaluate how the latest advances in genomics platforms and resources have enhanced our capacity to investigate plant genetic resources and harness their potential for improving crop productivity and quality. The unifying picture that emerges from the articles collected in this issue shows the increasingly pivotal role of genomics for characterizing germplasm collections, best managing genebanks, elucidating plant functions and identifying superior alleles at key loci for the selection of improved genotypes. In this brief introduction, we present an overview of the main topics covered and have included additional references to provide some further reading opportunities to the interested reader who wishes for a more comprehensive overview of the merits and limitations of genomics-based approaches.

The first group of articles (from page 155–184) offers a glimpse of the tools, platforms and resources currently available to investigate the structural and functional diversity present in both the coding and non-coding regions of the plant genome. This complexity, largely

inaccessible until recently, is now receiving increasing attention in view of its importance in the regulation of quantitative trait expression. Structural variability (e.g. copy number variations and/or indels) in what, as a reflection of our past ignorance, was often referred to as 'junk DNA' has instead been shown to be an important driver of phenotypic variability (Magalhaes *et al.*, 2007; Salvi *et al.*, 2007). Operationally, the new paradigm has been set by next generation sequencing (NGS) and bioinformatics, quickly adopted as the gold standard required to deliver an exhaustive, accurate characterization of DNA variation, the discovery of single nucleotide polymorphisms (SNPs) in massive number (Akhunov *et al.*, 2009; Varshney *et al.*, 2009) and the analysis of synteny (Bolot *et al.*, 2009). The cost of sequencing has already fallen dramatically and keeps dropping, thus allowing for the direct analysis of large sets of accessions at a fraction of the cost of such an operation just a few short years ago. NGS is also becoming the preferred choice for transcriptome profiling (Forrest and Carninci, 2009; Tamura and Yonemaru, 2010). Unlike microarray platforms (Gupta *et al.*, 2008; Pietsch *et al.*, 2009), NGS offers the distinct advantage of being able to report on changes across the entire transcriptome, including in rare transcripts. The power and benefits of NGS are particularly evident in species such as apple, potato or maize that suffer from low linkage disequilibrium while enjoying a high level of polymorphism, two features which require a highly detailed analysis at the DNA level to identify haplotype diversity in germplasm collections. A level of genetic resolution sufficient to validate candidate genes and, in some cases, even identify causal polymorphisms can be attained by association mapping, an approach increasingly adopted to dissect the genetic basis of target traits (Ersoz *et al.*, 2007; Rafalski, 2010). Genome-wide association mapping greatly benefits from the

*Corresponding author. E-mail: roberto.tuberosa@unibo.it

utilization of genome-wide SNP genotyping (Waugh *et al.*, 2009). Compared with other classes of molecular markers, SNPs are amenable to high-throughput automation at a relatively low cost (Edwards *et al.*, 2009). Although SNPs are already routinely utilized in a number of important crops (e.g. rice, barley and maize; McNally *et al.*, 2009; Hayden *et al.*, 2010; Yan *et al.*, 2010), considerable work is still required to establish suitable platforms in polyploid species such as wheat, in view of the additional difficulties caused by the presence of homoeologous loci (Ganal and Röder, 2007; Berard *et al.*, 2009; Trebbi *et al.*, 2011).

The second group of articles (from page 185–280) deals with one of the most difficult challenges faced by genebank managers, namely the characterization of germplasm collections. Over the past two decades, molecular profiling has greatly improved the accuracy of this characterization (Glaszmann *et al.*, 2010), particularly in the so-called ‘orphan’ species, which have been largely neglected also due to the lack of the means to properly investigate their biodiversity (Varshney *et al.*, 2010). An accurate characterization at the genomic level is a fundamental step required for (1) a more cost-efficient management of germplasm collections, both *in situ* and *ex situ*, (2) understanding phylogenetic relationships among species (Bolot *et al.*, 2009), (3) the assembly of core collections suitable for association mapping studies (Maccaferri *et al.*, 2011) and (4) assessing genetic similarity among accessions sharing common ancestors (Maccaferri *et al.*, 2007).

The third group of articles (from page 281–360) presents examples on how genomics-based approaches can provide information useful for crop improvement programmes by providing the breeder with effective indirect selection schemes. In particular, a number of articles (from page 324–351) deal with the improvement of crop quality and nutritional value, a topic of great interest in countries where malnutrition – the so-called ‘hidden hunger’ – is rampant as a consequence of an unbalanced diet. To address this problem, programmes such as HarvestPlus (www.harvestplus.org) are exploiting both natural and artificially induced variation to increase the iron, zinc and provitamin A content in crops. For the genomics-based improvement of crops for resistance to abiotic and biotic stresses, the generation challenge programme (GCP; www.generationcp.org) has developed a valuable molecular marker toolkit which provides easy access to existing information on molecular markers used in breeding programmes. Additionally, the GCP’s genotyping support service offers cost-efficient genotyping services, both for fingerprinting and the analysis of genetic diversity, as well as for molecular breeding. The most significant results achieved by the GCP through the application of genomics approaches have

been summarized in two recent articles (Glaszmann *et al.*, 2010; Varshney *et al.*, 2010) and some examples are also presented in this special issue. The adoption of genomics-assisted breeding has considerably enhanced the effectiveness of breeding programmes and the response to selection (Varshney *et al.*, 2005; Varshney and Tuberosa, 2007; Xu, 2010). Marker-assisted selection (MAS) is now routinely included in many breeding programmes, particularly for traits controlled by major loci (Ejeta and Knoll, 2007; Gupta and Langridge, 2010). This notwithstanding, MAS for complex quantitative traits (e.g. drought tolerance) remains a highly challenging undertaking, mainly because so much of the variation for these traits is under the control of many genes, where the contribution of each is too small to allow their ready identification and so justify the implementation of an MAS breeding strategy (Collins *et al.*, 2008). Nonetheless, some major loci which affect yield *per se* (i.e. not linked to phenology) across a broad range of environments have been described (Maccaferri *et al.*, 2008; Yadav *et al.*, 2011). When major loci that affect organ growth (e.g. leaf size) and yield components (e.g. seed number and seed weight) are identified, modelling the effects of the relevant quantitative trait loci (QTLs) in response to environmental cues should provide a highly effective approach for predicting yield performance across different environmental conditions (Parent *et al.*, 2010; Tardieu and Tuberosa, 2010). From a breeding standpoint, an interesting alternative to phenotypic selection for improving yield *per se* is provided by genome-wide selection, which is increasingly being adopted for the improvement of those major crops where SNP platforms allow for a cost-effective, high-throughput profiling of large populations (Bernardo, 2009).

Over the past decade, genomics has ushered in novel approaches which have produced a quantum leap in our ability to characterize and utilize plant genetic resources (Tuberosa *et al.*, 2002; Varshney *et al.*, 2005; Ribaut *et al.*, 2010; Yano and Tuberosa, 2009; Langridge and Fleury, 2011). Overall, molecular profiling suggests that the diversity stored in genebanks has been only marginally tapped into so far. Cloning of the key genes which underlie agronomically valuable traits will allow breeders to mine genebank collections much more effectively for novel alleles (Salvi and Tuberosa, 2007). Cloning will also provide perfect MAS markers and the opportunity to identify novel alleles *via* TILLING in mutant collections (Talamè *et al.*, 2008; Sestili *et al.*, 2010). Importantly, when access to genes is hampered by low recombination, genetic engineering will facilitate the transfer of these cloned genes, especially when sourced from wild relatives. In this regard, recent progress regarding site-specific recombination may open the door to further improve plant performance through the replacement of distinct alleles.

Meeting the challenges posed by climate change and the fast increasing demand for food, feed, fibre and fuel will require an acceleration of the rate of crop improvement, which in some key crops (e.g. wheat) has worriedly started to slow down (Tester and Langridge, 2010). Achieving higher gains from selection will require enlarging the pool of genetic resources and exploiting wild relatives of crops (Feuillet *et al.*, 2008; Kovach and McCouch, 2008) to identify superior alleles not yet utilized in the cultivated gene pool. This brief introduction and the articles of this special issue clearly show that genomics of plant genetic resources is having a tangible impact on the way genebanks are being managed and how germplasm collections are being exploited to improve crop performance.

Acknowledgements

The GPGR2 organizers thank the sponsors (for a complete list, see the congress website at www.GPGR2.com/sponsors.html) for their generous financial support and all participants for their contributions to the scientific programme of the Congress.

References

- Akhunov E, Nicolet C and Dvorak J (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theoretical and Applied Genetics* 119: 507–517.
- Berard A, Le Paslier MC, Dardevet M, Exbrayat-Vinson F, Bonnin I, Cenci A, Haudry A, Brunel D and Ravel C (2009) High-throughput single nucleotide polymorphism genotyping in wheat. (*Triticum spp.*). *Plant Biotechnology Journal* 7: 364–374.
- Bernardo R (2009) Genomewide selection for rapid introgression of exotic germplasm in maize. *Crop Science* 49: 419–425.
- Bolot S, Abrouk M, Masood-Quraishi U, Stein N, Messing J, Feuillet C and Salse J (2009) The 'inner circle' of the cereal genomes. *Current Opinion in Plant Biology* 12: 119–125.
- Collins NC, Tardieu F and Tuberosa R (2008) Quantitative trait loci and crop performance under abiotic stress: where do we stand? *Plant Physiology* 147: 469–486.
- Edwards KJ, Reid AL, Coghill JA, Berry ST and Barker GLA (2009) Multiplex single nucleotide polymorphism (SNP)-based genotyping in allohexaploid wheat using padlock probes. *Plant Biotechnology Journal* 7: 375–390.
- Ejeta G and Knoll JE (2007) Marker-assisted selection in sorghum. In: Varshney RK and Tuberosa R (eds) *Genomics-assisted Crop Improvement*, vol. 2. Houten: Springer, pp. 187–206.
- Ersoz ES, Yu J and Buckler ES (2007) Applications of linkage disequilibrium and association mapping. In: Varshney RK and Tuberosa R (eds) *Genomics-assisted Crop Improvement*, Vol. 1: *Genomics Approaches and Platforms*. Houten: Springer, pp. 97–120.
- Feuillet C, Langridge P and Waugh R (2008) Cereal breeding takes a walk on the wild side. *Trends in Genetics* 24: 24–23.
- Forrest ARR and Carninci P (2009) Whole genome transcriptome analysis. *RNA Biology* 6: 107–112.
- Ganal MW and Röder MS (2007) Microsatellite and SNP markers in wheat breeding. In: Varshney RK and Tuberosa R (eds) *Genomics-assisted Crop Improvement*, Vol. 2: *Genomics Applications in Crops*. Houten: Springer, pp. 1–24.
- Glasmann JC, Kilian B, Upadhyaya HD and Varshney RK (2010) Accessing genetic diversity for crop improvement. *Current Opinion in Plant Biology* 13: 167–173.
- Gupta PK and Langridge P (2010) Marker-assisted wheat breeding: present status and future possibilities. *Molecular Breeding* 26: 145–161.
- Gupta PK, Rustgi S and Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. *Heredity* 101: 5–18.
- Hayden MJ, Tabone TL, Nguyen TM, Coventry S, Keiper FJ, Fox RL, Chalmers KJ, Mather DE and Eglinton JK (2010) An informative set of SNP markers for molecular characterization of Australian barley germplasm. *Crop and Pasture Science* 61: 70–83.
- Kovach MJ and McCouch SR (2008) Leveraging natural diversity: back through the bottleneck. *Current Opinion in Plant Biology* 11: 193–200.
- Langridge P and Fleury D (2011) Making the most of 'omics' for crop breeding. *Trends in Biotechnology* 29: 33–40.
- Maccaferri M, Sanguineti MC, Xie C, Smith JSC and Tuberosa R (2007) Relationships among durum wheat accessions. II. A comparison of molecular and pedigree information. *Genome* 50: 385–399.
- Maccaferri M, Sanguineti MC, Corneti S, Ortega JLA, Ben Salem M, Bort J, DeAmbrogio E, del Moral LFG, Demontis A, El-Ahmed A, *et al.* (2008) Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics* 178: 489–511.
- Maccaferri M, Sanguineti MC, Demontis A, El-Ahmed A, Garcia del Moral L, Maalouf F, Nachit M, Nserallah N, Ouabbou H, Rhouma S, Royo C, Villegas D and Tuberosa R (2011) Association mapping in durum wheat grown across a broad range of water regimes. *Journal of Experimental Botany* 62: 409–438.
- Magalhaes JV, Liu J, Guimaraes CT, Lana UGP, Alves VMC, Wang YH, Schaffert RE, Hoekenga OA, Pineros MA, Shaff JE, *et al.* (2007) A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nature Genetics* 39: 1156–1161.
- McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, Hoen DR, Bureau TE, Stokowski R, Ballinger DG, Frazer KA, Cox DR, Padhukasahasram B, Bustamante CD, Detlef Weigel D, Mackill DJ, Bruskiewich RM, Ratsch G, Buell CR, Leung H and Leach JE (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proceedings of the National Academy of Sciences of the United States of America* 106: 12273–12278.
- Parent B, Suard B, Serraj R and Tardieu F (2010) Rice leaf growth and water potential are resilient to evaporative demand and soil water deficit once the effects of root system are neutralized. *Plant Cell and Environment* 33: 1256–1267.
- Pietsch C, Sreenivasulu N, Wobus U and Roder MS (2009) Linkage mapping of putative regulator genes of barley grain

- development characterized by expression profiling. *BMC Plant Biology* 9: 4.
- Rafalski JA (2010) Association genetics in crop improvement. *Current Opinion in Plant Biology* 13: 174–180.
- Ribaut JM, de Vicente MC and Delannay X (2010) Molecular breeding in developing countries: challenges and perspectives. *Current Opinion in Plant Biology* 13: 213–218.
- Salvi S and Tuberosa R (2007) Cloning QTLs in plants. In: Varshney RK and Tuberosa R (eds) *Genomics-assisted Crop Improvement, Vol. 1: Genomics Approaches and Platforms*. Houten: Springer, pp. 207–226.
- Salvi S, Sponza G, Morgante M, Tomes D, Niu X, Fengler KA, Meeley R, Ananiev EV, Svitashv S, Bruggemann E, Li B, Hainey CF, Radovic S, Zaina G, Rafalski JA, Tingey SV, Miao GH, Phillips RL and Tuberosa R (2007) Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proceedings of the National Academy of Sciences of the United States of America* 104: 11376–11381.
- Sestili F, Botticella E, Bedo Z, Phillips A and Lafiandra D (2010) Production of novel allelic variation for genes involved in starch biosynthesis through mutagenesis. *Molecular Breeding* 25: 145–154.
- Talamè V, Bovina R, Sanguineti MC, Tuberosa R, Lundqvist U and Salvi S (2008) TILLMore, a resource for the discovery of chemically induced mutants in barley. *Plant Biotechnology Journal* 6: 477–485.
- Tamura K and Yonemaru J (2010) Next-generation sequencing for comparative transcriptomics of perennial ryegrass (*Lolium perenne* L.) and meadow fescue (*Festuca pratensis* Huds.) during cold acclimation. *Grassland Science* 56: 230–239.
- Tardieu F and Tuberosa R (2010) Dissection and modelling of abiotic stress tolerance in plants. *Current Opinion in Plant Biology* 13: 206–212.
- Tester M and Langridge P (2010) Breeding technologies to increase crop production in a changing world. *Science* 327: 818–822.
- Trebbi D, Maccaferri M, de Heer P, Sørensen A, van der Vossen E, Andries G, Giuliani S, Salvi S, Sanguineti MC, Massi A and Tuberosa R (2011) High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). *Theoretical Applied Genetics*. DOI:10.1007/s00122-011-1607-7.
- Tuberosa R, Gill BS and Quarrie SA (2002) Cereal genomics: ushering in a brave new world. *Plant Molecular Biology* 48: 445–449.
- Varshney RK and Tuberosa R (2007) *Genomics-assisted Crop Improvement, Vol. 1: Genomics Approaches and Platforms*, p. 386 and *Vol. 2: Genomics Applications in Crops*, p. 516. Houten: Springer.
- Varshney RK, Glaszmann J-C, Leung H and Ribaut JM (2010) More genomic resources for less-studied crops. *Trends Biotechnology* 28: 452–460.
- Varshney RK, Graner A and Sorrells ME (2005) Genomics-assisted breeding for crop improvement. *Trends in Plant Science* 10: 621–630.
- Varshney RK, Nayak SN, May GD and Jackson SA (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in Biotechnology* 27: 522–530.
- Waugh R, Jannink JL, Muehlbauer GJ and Ramsay L (2009) The emergence of whole genome association scans in barley. *Current Opinion in Plant Biology* 12: 218–222.
- Xu Y (2010) *Molecular Plant Breeding*. Wallingford: CABI.
- Yadav RS, Sehgal D and Vadez V (2011) Using genetic mapping and genomics approaches in understanding and improving drought tolerance in pearl millet. *Journal of Experimental Botany* 62: 397–408.
- Yan JB, Yang XH, Shah T, Sanchez-Villeda H, Li JS, Warburton M, Zhou Y, Crouch JH and Xu Y (2010) High-throughput SNP genotyping with the GoldenGate assay in maize. *Molecular Breeding* 25: 441–451.
- Yano M and Tuberosa R (2009) Genome studies and molecular genetics-from sequence to crops: genomics comes of age. *Current Opinion in Plant Biology* 12: 103–106.