INTRODUCTION / INTRODUCTION

Exploiting genome-wide association in oilseed *Brassica* species[1]

E. Balázs and W.A. Cowling

Plant and animal breeders measure the phenotype of individuals to select higher yielding, better quality, and more disease resistant types. This is not always an efficient or particularly rapid process. Animal breeders have improved estimations of breeding values and accelerated breeding cycles by including genetic relationships from pedigrees in the analysis, and more recently by implementing genomic selection (Goddard and Hayes 2009; Weller 2010). In plants, molecular markers for complex quantitative traits have been linked to genomic regions known as quantitative trait loci (QTL), based on linkage analysis in experimental populations, but with relatively poor genetic resolution. Despite the large number of reports of QTL in plants, relatively few have been used in applied plant breeding programs for selection purposes (Young 1999; Heffner et al. 2009). This has led to an interest in integrating quantitative and molecular genetics in plant breeding, in order to improve identification and manipulation of QTL (Gupta et al. 2010).

Human and animal geneticists have developed sophisticated molecular and bioinformatic tools to discover molecular genetic associations between phenotype and genotype, with genetic resolution in genome-wide single nucleotide polymorphism (SNP) panels high enough to identify genes related to health and disease (Manolio et al. 2007).

The principles of association genetics are now being applied to plants (Oraguzie et al. 2007). In *Arabidopsis*, linkage disequilibrium studies based on high density SNP maps have reached new heights, approaching the intensity of human and animal studies (Atwell et al. 2010). As in animal breeding, plant breeders are now considering genomic selection as a means of accelerating genetic improvement (Heffner et al. 2009), but major issues must be resolved before genomic selection is applied efficiently and profitably

to plant breeding. In plant breeding, genotypes are replicated and tested in multiple environments, and genotype by environment interactions have a major influence on estimations of the genetic or breeding value of genotypes.

We considered that it was important to ask the experts how they would apply association genetics to plant breeding, so we invited human, animal, and plant molecular geneticists; bioinformaticians; plant breeders; and biometricians to come together and explore the major questions in plant genetic improvement. Will genomic selection increase genetic gain and provide economical benefits to plant breeding programs? In animals, the advantage of genomic selection over nongenomic methods is related to the proportion of the genetic variance that is explained by the markers (Hayes and Goddard 2010). In plants, simulation studies have resulted in a correlation between "true" breeding value and genomic breeding value, up to 0.85 for polygenic traits with low heritability, making genomic selection a very attractive possibility for increasing genetic gain per unit of time (Heffner et al. 2009).

With the aid of a successful application for a conference grant from the OECD Co-operative Research Programme on Biological Resource Management for Sustainable Agricultural Systems, we invited several scientists from OECD countries to an international conference titled "Exploiting Genome-wide Association in Oilseed Brassicas: a model for genetic improvement of major OECD crops for sustainable future farming" (www.oecd-genomeassociation-oz09.com). The conference was held at The University of Western Australia, 9–12 November 2009, and was attended by 80 scientists representing more than 13 countries.

The conference was preceded by a workshop titled "Joint modelling of additive and non-additive (genetic line) effects in multi-environment trials", presented by Brian Cullis

E. Balázs. Department of Applied Genomics, H-2462 Martonvásár, Agricultural Research Institute, Brunszvik. u. 2, Hungary (previously Theme Coordinator Research Theme 3: The Food Chain [Plants], OECD Co-operative Research Programme: Biological Resource Management for Sustainable Agricultural Systems).

W. Cowling.[2] Deputy Director, International Centre for Plant Breeding Education and Research, The UWA Institute of Agriculture, c/o School of Plant Biology M084, The University of Western Australia, 35 Stirling Highway, Crawley, Western Australia, Australia 6009.

[1]This article is one of a selection of papers from the conference "Exploiting Genome-wide Association in Oilseed Brassicas: a model for genetic improvement of major OECD crops for sustainable farming".
[2]Corresponding author (e-mail: wallace.cowling@uwa.edu.au).

RIGHTSLINK®

(Australia) and Robin Thompson (UK), which was supported by the Grains Research and Development Corporation and The University of Western Australia. This workshop highlighted the importance of new biometrical techniques for analysing complex traits in field trials in multiple environments, which will form the basis of genomic selection in plant breeding.

The "focus" crop of the conference was the genus *Brassica* (which is related to the model plant *Arabidopsis*), but papers spanned a wide range of crops and topics including statistical genetic issues in human bioinformatics (Weir 2010); animal genomic selection (Hayes and Goddard 2010); genome-wide association studies in *Brassica* (Honsdorf et al. 2010, Snowdon et al. 2010; Zhao et al. 2010; Zou et al. 2010), cereals (Raman et al. 2010; Waugh et al. 2010), and sugarcane (Wei et al. 2010); and issues surrounding accurate and reliable sequencing and mapping of the *Brassica* genome (Li et al. 2010; Parkin et al. 2010).

Several challenges must be overcome before genomic selection will be applied commercially in plant breeding. SNP or other marker technology must be developed to a level that is readily available at an economical price for genome-wide assessment of genotypes in the breeding program (Duran et al. 2010; Durstewitz et al. 2010). However, QTL derived from linkage mapping should enhance the efficiency of association mapping, especially in polyploid oilcrops (Zou et al. 2010). Overlying genetic factors, epigenetic variation may be harnessed or induced for genetic improvement of crops (King et al. 2010).

The breeding value of a plant genotype is directly influenced by the environment in which it is tested, and genotypes are usually tested in many environments across large geographical regions. Two papers specifically addressed this problem, based on the mixed model approach to the analysis of yield and oil data in a commercial canola breeding program, where the addition of pedigree relationships improved estimates of additive and nonadditive genetic values across environments (Beeck et al. 2010; Cullis et al. 2010). Assessment of the breeding value of plant genotypes across environments could be improved further by genomic selection, where genetic relationship is estimated by whole genome-wide markers.

Falk (2010) challenges plant breeders to consider a population breeding approach that aims to increase effective population size, the frequency of recombination, and genetic variability inside commercial breeding programs. Genomic selection is likely to be more successful in genetic populations with large effective population size and frequent genetic recombination. Canola (*B. napus*) is typical of many major crop breeding programs in its narrow genetic base (Cowling 2007). Breeding programs with a small effective population size will have large regions of linkage disequilibrium, and fewer alleles on which to base genomic selection, than those with a large effective population size.

It is our hope that this special issue of *Genome* will provide the stimulus for closer collaboration between plant and animal breeders and specialists in bioinformatics, molecular genetics, quantitative genetics, and biometrics to ensure that genomic selection results in more efficient and profitable plant breeding.

Acknowledgements

References

Atwell, S., Huang, Y.S., Vilhjálmsson, B.J., Willems, G., Horton, M., Li, Y., et al. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature, **465**(7298): 627–631. doi:10.1038/nature08800. PMID:20336072.

Beeck, C.P., Cowling, W.A., Smith, A.B., and Cullis, B.R. 2010. Analysis of yield and oil from a series of canola breeding trials. Part I. Fitting factor analytic mixed models with pedigree information. Genome, **53**(11): 992–1001. doi:10.1139/G10-051.

Cowling, W.A. 2007. Genetic diversity in Australian canola and implications for crop breeding for changing future environments. Field Crops Res. **104**(1–3): 103–111. doi:10.1016/j.fcr.2006.12.014.

Cullis, B.R., Smith, A.B., Beeck, C.P., and Cowling, W.A. 2010. Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. Genome, **53**(11): 1002–1016. doi:10.1139/G10-080.

Duran, C., Eales, D., Marshall, D., Imelfort, M., Stiller, J., Berkman, P., et al. 2010. Future tools for association mapping in crop plants. Genome, **53**(11): 1017–1023. doi:10.1139/G10-057.

Durstewitz, G., Polley, A., Plieske, J., Luerssen, H., Graner, E.M., Wieseke, R., and Ganal, M.W. 2010. SNP discovery by amplicon sequencing and multiplex SNP genotyping in the allopolyploid species *Brassica napus*. Genome, **53**(11): 948–956. doi:10.1139/G10-079.

Falk, D.E. 2010. Generating and maintaining diversity at the elite level in crop breeding. Genome, **53**(11): 982–991. doi:10.1139/G10-081.

Goddard, M.E., and Hayes, B.J. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet. **10**(6): 381–391. doi:10.1038/nrg2575. PMID:19448663.

Gupta, P.K., Langridge, P., and Mir, R.R. 2010. Marker-assisted wheat breeding: present status and future possibilities. Mol. Breed. **26**(2): 145–161. doi:10.1007/s11032-009-9359-7.

Hayes, B., and Goddard, M. 2010. Genome-wide association and genomic selection in animal breeding. Genome, **53**(11): 876–883. doi:10.1139/G10-076.

Heffner, E.L., Sorrells, M.E., and Jannink, J.-L. 2009. Genomic selection for crop improvement. Crop Sci. **49**(1): 1–12. doi:10.2135/cropsci2008.08.0512.

Honsdorf, N., Becker, H.C., and Ecke, W. 2010. Association mapping for phenological, morphological, and quality traits in canola quality winter rapeseed (*Brassica napus* L.). Genome, **53**(11): 899–907. doi:10.1139/G10-049.

King, G.J., Amoah, S., and Kurup, S. 2010. Exploring and exploiting epigenetic variation in crop. Genome, **53**(11): 856–868. doi:10.1139/G10-059.

Li, X., Ramchiary, N., Choi, S.R., Nguyen, D.V., Hossain, M.J., Yang, H.K., and Lim, Y.P. 2010. Development of a high density integrated reference genetic linkage map for the multinational *Brassica rapa* Genome Sequencing Project. Genome, **53**(11): 939–947. doi:10.1139/G10-054.

Manolio, T.A., Rodriguez, L.L., Brooks, L., Abecasis, G., Ballinger, D., Daly, M., et al.. 2007. New models of collaboration in genome-wide association studies: the Genetic Association Information Network. Nat. Genet. **39**(9): 1045–1051. doi:10.1038/ng2127. PMID:17728769.

Oraguzie, N.C., Rikkerink, E.H.A., Gardiner, S.E., and de Silva, N.H. 2007. Association Mapping in Plants. Springer, New York.

Parkin, I.A.P., Clarke, W.E., Sidebottom, C., Zhang, W., Robinson, S.J., Links, M.G., et al. 2010. Towards unambiguous transcript mapping in the allotetraploid *Brassica napus.* Genome, **53**(11): 929–938. doi:10.1139/G10-053.

Raman, H., Stodart, B., Ryan, P.R., Delhaize, E., Emebiri, L., Raman, R., et al. 2010. Genome-wide association analyses of common wheat (*Triticum aestivum* L.) germplasm identifies multiple loci for aluminium resistance. Genome, **53**(11): 957–966. doi:10.1139/G10-058.

Snowdon, R.J., Wittkop, B., Rezaidad, A., Hasan, M., Lipsa, F., Stein, A., and Friedt, W. 2010. Regional association analysis delineates a sequenced chromosome region influencing antinutritive seed meal compounds in oilseed rape. Genome, **53**(11): 917–928. doi:10.1139/G10-052.

Waugh, R., Marshall, D., Thomas, B., Comadran, J., Russell, J., Close, T., AGOUEB, BarleyCAP, and Ramsay, L. 2010. Whole-genome association mapping in elite inbred crop varieties. Genome, **53**(11): 967–972. doi:10.1139/G10-078.

Wei, X., Jackson, P.A., Hermann, S., Kilian, A., Heller-Uszynska, K., and Deomano, E. 2010. Simultaneously accounting for population structure, genotype by environment interaction, and spatial variation in marker–trait associations in sugarcane. Genome, **53**(11): 973–981. doi:10.1139/G10-050.

Weir, B.S. 2010. Statistical genetic issues for genome-wide association studies. Genome, **53**(11): 869–875. doi:10.1139/G10-062.

Weller, J.I. 2010. Whole genome marker-assisted selection. CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources 5, No. 023

Young, N.D. 1999. A cautiously optimistic vision for marker-assisted breeding. Mol. Breed. **5**(6): 505–510. doi:10.1023/A:1009684409326.

Zhao, J., Artemyeva, A., Del Carpio, D.P., Basnet, R.K., Zhang, N., Gao, J., et al. 2010. Design of a *Brassica rapa* core collection for association mapping studies. Genome, **53**(11): 884–898. doi:10.1139/G10-082.

Zou, J., Jiang, C., Cao, Z., Li, R., Long, Y., Chen, S., and Meng, J. 2010. Association mapping of seed oil content in *Brassica napus* and comparison with quantitative trait loci identified from linkage mapping. Genome, **53**(11): 908–916. doi:10.1139/G10-075.