Large scale SNP genotyping with optimized molecular marker sets for cost-efficient plant breeding in the Brassica species (B. napus, B. oleracea and B. rapa)

Jörg Plieske, Jana Lemm, Andreas Polley and Martin W. Ganal

TraitGenetics GmbH, Am Schwabeplan 1b, D-06466 Stadt Seeland OT Gatersleben, Germany

Web: www.traitgenetics.com e-mail: contact@traitgenetics.com

Introduction

Through the development of the large Brassica 60K Illumina Infinium array, it has now become routine to generate a wealth of genotyping data for individual Brassica lines and varieties. However, with costs in the area of approximately 100 € per sample, this array is quite expensive. Furthermore, many of the genotype data generated in this way constitute of redundant information since many markers are in perfect linkage disequilibrium in breeding material and varieties.

Using the Brassica 60K genotyping array, large sets of varieties have been analyzed at TraitGenetics for Brassica napus, B. oleracea and B. rapa during our own internal research activities. Based on these data, we have generated marker databases that include mapping data, allele frequency data and haplotype information for the markers on the 60K array. Such information is crucial for the development of optimized marker sets for plant breeding that can be analyzed in a very cost-effective way. We present here our current activities in the Brassica species towards the development of an optimized, high quality genotyping array that can be used in Brassica napus and its diploid ancestors at much reduced cost.

Illumina SNP calling for Brassica napus

Genetic BIN map of B. napus derived from various populations

Number of markers on the array: 13,714 (100%)
Number of functional markers: 13,465 (98%)
Number of polymorphic markers in 984 B. napus samples: 12,792 (93%)
Number of polymorphic markers with MAF >0.03: 11,917 (87%)

This 15K array can now be used at much reduced costs compared to other arrays and without much loss of information compared to larger and more expensive arrays.

In parallel, we have generated an optimized marker collection based on individual SNP markers (KASP) for variety identification, variety purity analysis, marker-assisted backcrossing and other purposes.

Summary

With genotyping data from large sets of lines and varieties we have now selected specific marker sets based on polymorphism information content, haplotype structure, marker quality, genome-specificity in allopolyploid species, genetic map position and other quality features that are relevant for plant breeding. In this way, we can now offer very cost-effective molecular marker analyses based on an optimized array with around 13,000 markers (approx. 29-49 € including array, sample processing and the generation of the genotype table) and also with individual marker sets for various breeding applications including marker-assisted selection, cost-effective genome selection, marker-assisted backcrossing, variety identification and other purposes.

If you are interested in using these arrays or marker sets through our services, please do not hesitate to contact us.