III International Symposium on Molecular Markers in Horticulture

Molecular Markers
in Horticulture Symposium
2013 ISHS

2013 September, 25-27

Riva del Garda (Trento)

Italy,

FONDAZIONE
EDMUND
MACH

Under the Auspices of the
Autonomous Province of Trento

POSTER - A pipeline for robust marker calling from Infinium SNP arrays for diploid crops.

Corresponding author: mario.diguardo@fmach.it

Di Guardo M.[1,2,3]., Koehorst-van Putten H.J.J.[1], Longhi S.[1], Noordijk Y.[1], Muranty H.[4], Banchi E.[2], Bianco L. [2], Troggio M.[2], Velasco R.[2], Costa F.[2], Visser R.G.F.[1], van de Weg E.[1]

[1] Wageningen UR Plant Breeding, Wageningen University & Research centre, Droevendaalsesteeg 1, P.O. Box 16, 6700 AA Wageningen, The Netherlands

[2] Department of Genetics and Biology of Fruit Crops, IASMA Research and Innovation Centre, Foundation Edmund Mach, Via Mach 1, 38010 Trento, Italy

3 Graduate school Experimental Plant Sciences, Wageningen University

[4] INRA, UMR1345 Institut de Recherche en Horticulture et Semences, SFR 4207 Quasav, Pres L'UNAM, F-49071 Beaucouzé, France

Abstract: Single-nucleotide polymorphisms (SNPs) have been recognized as the marker of choice in genetic studies on mapping populations due to their suitability for high throughput genome-wide genotyping. Several SNP genotyping platforms and related software for genotype calling are commercially available, including Illumina Infinium platform and the Genome Studio software from Illumina.
Although the SNP genotyping can benefit of a high automation rate, SNP-allele calling is far from being fully automatized and, equally important, a common approach for filtering of well performing markers has not been published yet. Moreover, the application of a 20K array on diploid apple learned that around 7% of the markers were incorrectly called due to limitations in the genotyping software. All this made the genotyping a time-consuming procedure with an unnecessary high error rate.
In this poster we present a semi-automatized, pipeline for SNP filtering and calling of progenies of mapping populations based on Excel using Genome Studio derived data as input. This pipeline efficiently distinguishes between robust informative markers and markers that are problematic or truly monomorphic. It also resolves miss-callings that are due to the presence of paralogous loci, null alleles, or of additional SNP at the primer site. The final set of SNP markers are grouped based on their segregation pattern and genotype calls are adjusted to JoinMap format. Completion of the pipeline from start to finish takes just 2-3 hours. The pipeline has been calibrated by the use of three mapping populations and has been applied on another ten populations. It showed to be highly efficient, as 99.7% of the filtered polymorphic markers mapped easily.
This work was performed in the framework of the EU-FruitBreedomics project, a central goal of which is to accelerate and increase the efficiency of Rosaceous cultivar release through the delivery of molecular markers that are of use for marker assisted breeding.