# Axiom<sup>®</sup> Genotyping Arrays: Automated analysis of complex plant genomes



Bridget Moore<sup>1</sup>, Nahla Bassil<sup>2</sup>, Michael Mittmann<sup>1</sup>, Ali Pirani<sup>1</sup>, Thomas Davis<sup>3</sup>, Laurent Bellon<sup>1</sup>, Amy Iezzoni<sup>4</sup>, Lisa Mahoney<sup>3</sup>, Daniel Sargent<sup>5</sup>, Eric Van de Weg<sup>6</sup>, Teresa Webster<sup>1</sup>

### ABSTRACT

Axiom<sup>®</sup> Genotyping Solution for custom design of high-density SNP microarrays has enabled large-scale high-throughput SNP genotyping for a diverse spectrum of non-human genomes. Processing SNPs on this scale requires automated workflows; however, non-human genomes, especially plants, present challenges, including polyploid genome structures and high levels of nucleotide and structural diversity between subpopulations of a species. High-throughput automated genotyping of plants requires a technology that is robust in the face of complex plant genomes, and cost-effective for processing large numbers of samples.

This analysis workflow, algorithms, and results are presented for allo-octoploid, cultivated strawberry samples, both *F.* ×*ananassa* and "non-ananassa" cultivars and selections, that were genotyped with the Axiom Strawberry Genotyping Array (IStraw90) designed by the International RosBREED Strawberry Consortium. This workflow includes three main components: (1) an adaptive genotyping algorithm, GT1, (2) *off-target variant (OTV)* calling, and (3) classification of markers into quality categories

#### Analysis workflow, continued

**3a)** Markers are sorted into relevant quality categories. The potentially most accurate and polymorphic markers are sorted into the *Polymorphic High Resolution* class.



#### Analysis workflow

There are three components to the Axiom analysis workflow:

1) Usage of the Axiom genotyping algorithm (AxiomGT1), which dynamically adapts to shifted and compressed locations produced by differences in the fixed sub-genomes of allo-octoploid strawberry SNPs. AxiomGT1 is a clustering algorithm that adapts pre-positioned genotype cluster locations called priors to the sample data in a Bayesian step and computes three posterior (solid ovals) cluster locations.



**Figure 3: Examples from allo-octoploid strawberry.** *Poly High Resolution* SNPs are recommended for downstream analysis, *No Minor Homozygotes* appear useful in strawberry, and *OTVs* are recommended after *OTV* genotyping.

**3b)** Markers with complex genotypes causing cluster splits are further filtered into additional classes using "Variance" filters.







Figure 1: Cluster Plots for allo-octoploid strawberry SNPs. X= Contrast

=Log2(A\_signal/B\_signal). Y= Size=[Log2(A\_signal) + Log2(B\_signal)]/2. Each sample is colored by the AxiomGT1 genotype call (blue, gold, red) for the sub-genome with the segregating allele. Each genotype cluster is labeled by the likely allo-octoploid genotype using the following notation: the genotypes of 4 sub-genomes are separated by dashes, the genotype of the sub-genome with the segregating allele is noted first (blue, gold, red). **A**. SNP with BB fixed sub-genomes. **B**. SNP with AA fixed sub-genomes

**2)** Novel genotyping algorithms are used for markers in regions of high genomic complexity. The Axiom *OTV* Caller detects and genotypes *OTV* SNPs. *OTVs* are genomic markers sites at which there are significant subpopulation sequence differences, resulting in a 4th genotype cluster (see arrow).





Figure 3: Examples from *F.* ×*ananassa* and "non-ananassa" samples

## Results

Genotypes for SNPs in the *Poly High Resolution* class have greater than 99.8% reproducibility. The following table shows reproducibility of technical replicates over SNPs in the *Poly High Resolution* class for 357 octoploid strawberry samples, genotyped on the Axiom® Strawberry Genotyping Array (IStraw90).

#### Next steps

Cost effective genotyping is enabled by the Axiom® 384-array layout, where 384 samples are simultaneously processed on a 384 microplate for analyzing up to 35,000 polyploid or 50,0000 diploid SNPs per sample.



**Figure 2: Effect of OTV calling on OTV cluster (arrow) genotypes. A**. Before OTV genotyping the OTV cluster is mis-called as AB (gold). **B.** After OTV genotyping the OTV cluster has been identified and re-labeled as a 4<sup>th</sup> OTV genotype cluster (cyan).

1) Affymetrix, Inc., CA. 2) USDA/ARS, NCGR, Corvallis, OR 3) University of New Hampshire, NH 4) Michigan State University, MI 5) Fondazione E. Mach, San Michele all'Adige, Italy 6) Plant Breeding Wageningen University & Research centre, Wageningen, Netherlands

The high-value markers, such as those in the *Poly High Resolution* class, identified by the Axiom® Genotyping Analysis on Axiom 96array format can be transferred to Axiom 384-format arrays with 100% fidelity. The multi-species capability of the Axiom 384-format enables the design of a single array for markers from all rosaceous species that can be used in marker-assisted breeding.

©2014 Affymetrix, Inc. All rights reserved. Affymetrix<sup>®</sup> and Axiom<sup>®</sup> are registered trademarks of Affymetrix, Inc. All other trademarks are the property of their respective owners. Unless indicated otherwise, Affymetrix products are For Research Use Only. Not for use in diagnostic procedures.