Automated, High Throughput SNP Genotyping of *Zea mays*

Lena Emadi[1], Mike Farrell[2], Jeff Roeder[2], Travis Butts[2], Honrado Lopez[1], Paul Streng[1]
[1]Fluidigm Corporation, [2]Omega Bio-tek, Inc.

Introduction

Genotyping is a critical tool for agricultural plant selection and breeding to achieve optimal physiological traits. Successful screening of plant genotypes require rapid, robust, high throughput and cost-effective methods. Here we use *Zea mays* (maize) to demonstrate the utility of joining Omega Bio-tek's automated plant DNA extraction technology with Fluidigm's Juno system and Integrated Fluidic Circuits (IFCs) to rapidly extract and genotype 48 samples for 47 SNPs in less than 6 hours.

Methods & Materials

Fresh samples of different maize products were purchased from various produce markets. Exact varieties of the samples are unknown. From these, 8 replicates of approxmiately 100 mg of maize seed material from each of the 6 varieties (n = 48) were extracted for DNA utilizing Omega Bio-tek's Mag-BIND® Plant DNA DS Kit (M1130). This extraction chemistry has previously been shown to be highly effective across a broad range of agricultural products. Each sample was homogenized using a Spex 2010 Geno/Grinder® and then lysed by incubation at 55°C in CSPL Lysis Buffer and Proteinase K solution. The remaining steps of the extraction procedure, which includes binding DNA to Omega Bio-tek's Mag-BIND® particles, washing to remove impurities, and eluting the DNA into 50 µL of Elution Buffer, were automated on the Hamilton Microlab® STAR™ robotic liquid handling system. The entire extraction procedure was completed in 1 hour, 55 minutes. Purified DNA was quantified via PicoGreen® dsDNA quantitation assay. The amount of plant genomic DNA recovered per mg of sample input was 30.6 ± 18.8 ng/mg.

SNP Type Assay for this project were designed by Fluidigm's Assay Design Group (Zea mays assembly build "RefGen_v2/ ZmDGB101 [18 Dec 2010]) targeting 47 common loci used in maize genotyping. Sample loading and PCR were performed automatically by the Juno system in less than 4 hours. Upon completion of PCR on the Juno system, the IFC was scanned on the EP1 system to collect genotyping data for later analysis.

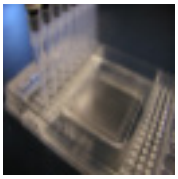
Maize Genotyping Results

Genotypes were determined using Fluidigm SNP Genotyping Analysis Software v4.1.3. Analysis setting include SNP Type Normalization and a confidence threshold of 90. The software determines the genotype based on a K-means clustering algorithm, where the sample is associated with the cluster based on the proximity of a sample to the center of the cluster. Genotype calls were labeled as XX (homozygous for Allele 1), YY (homozygous for Allele 2) and XY (heterozygous). Samples with a confidence score of less than 90 were labeled as "No call."

Overall Call Rate

Overall auto call rate is 94.37%. For some clusters, as shown in Figure 1 and Figure 2, the clusters are not as compact and the genotypes of these samples could not be scored with high confidence. This excludes assay for 2 SNPs whose genotypes could not be determined because all samples expressed the same genotype; because the software assumes 3 clusters will form, the algorithm could not determine the genotype call based on a single cluster with high confidence.

1. Prepare reagents



30 minutes

2. Load

3. Preamplify

Automated on Juno system < 3.5 hours

4. PCR

5. Endpoint reading

5 minutes

Figure 1.

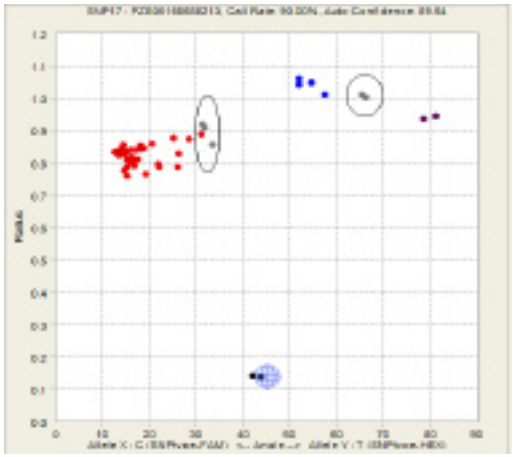
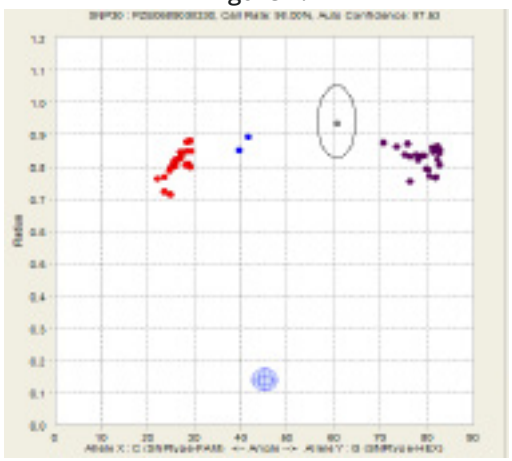


Figure 2.



Figures 1 & 2. Examples of cluster plots with genotype for homozygous X (red), homozygous Y (purple), and heterozygous (blue). Samples that failed due to a low confidence score are labeled as "No call" and are circled in both plots.

Concordance Between Technical Replicates

Concordance between the 2 technical replicates of each sample on the IFC is 96.64%. Discordant calls were traced to one of the two replicates labeled as a "No call" by the software because of the slightly lower confidence score.

Homogeneous Genotypes

Of the 47 SNPs evaluated in this panel, 30 generated the same genotype call in all 6 varieties of maize. Gene IDs for these SNPs are included in Table 1. A call map of each genotype is included in Figure 3.

Diverse Genotypes

14 SNPs resulted in a diverse range of genotypes across the 6 varieties of maize. Gene IDs for these SNPs are included in Table 2. Call maps of each genotype is included in Figure 4.

Table 1. Gene IDs from 30 homozygous SNPs across the 6 maize varieties.

Gene IDs	
GRMZM2G154290	GRMZM2G164854
GRMZM2G114667	GRMZM2G140901
GRMZM2G172794	GRMZM2G364068
GRMZM2G082916	GRMZM2G438895
AC202989.3_FG007	GRMZM2G089400
GRMZM2G176595	GRMZM2G390641
GRMZM2G131793	GRMZM2G093789
GRMZM2G131793	GRMZM2G322328
GRMZM2G101867	GRMZM2G023051
GRMZM2G152512	GRMZM2G044398
GRMZM2G173035	GRMZM2G082322
GRMZM2G074351	GRMZM2G155015
GRMZM2G171600	GRMZM2G016250
GRMZM5G878561	GRMZM2G136311
GRMZM2G436742	GRMZM5G862101

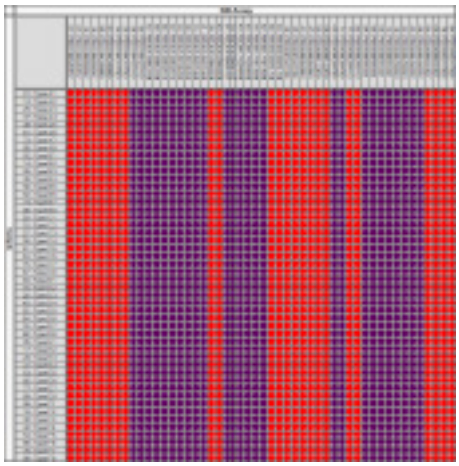


Figure 3. Call map displaying genotypes of 47 maize samples with 303 homologous SNPs. Homozygous allele X shown in red and homozygous allele Y shown in purple.

Note that a few samples amplified, but the final genotype call was ambiguous and could not be determined accurately by the software (Figure 5). We believe that these SNPs are particularly difficult to genotype because of the complexity of the maize genome, which is replete with chromosome duplications and repetitive regions and may result in a second copy at another position in the maize genome (2). Because of this, it is possible to observe a more broad cluster phenotype where the final call cannot be determined with high confidence. With a larger sample cohort and a more sophisticated confidence genotype call algorithm, it may be possible to determine the genotypes of these failed samples. Here, we report only the genotype calls with the highest confidence.

Table 2. Gene IDs for 30 homologous SNPs across 6 maize varieties.

Gene IDs
GRMZM2G021427
GRMZM2G021621
GRMZM2G038034
GRMZM2G062458
GRMZM2G342246
GRMZM2G405622
GRMZM2G414639
GRMZM2G131820
GRMZM2G030038
GRMZM2G171622
GRMZM2G878561
GRMZM2G027302
GRMZM2G062084
GRMZM2G027955

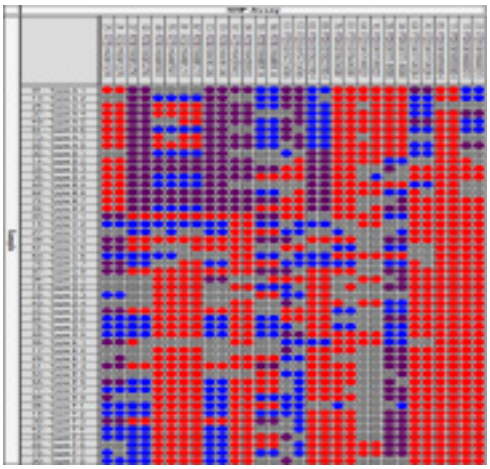


Figure 4. Call map displaying genotypes of 47 maize samples with 14 highly polymorphic SNPs. Calls are defined as a homozygous allele X (red), homozygous allele Y (purple), and heterozygous XY (blue). Samples with a high confidence score lower than 90 are labeled as "No call" and highlighted in gray.

Conclusions

Omega Bio-tek's Mag-BIND® Plant DNA DS Kit (M1130) extracts high quality genomic DNA from maize in less than 2 hours. The amount of plant gDNA recovered per mg of sample input was 30.6 ± 18.8 ng/mg. With Fluidigm's Juno system and Integrated Fluidic Circuits (IFCs), it is possible to genotype 96 samples for 96 SNPs in less than 6 hours.

Here, we've evaluated 47 common loci across 6 varieties of maize using SNP Type Assays. 30 SNPs were homogeneous for all samples, 14 SNPs were found to be highly polymorphic.

Overall autocall rate is 94.37% with 96.64% concordance between genotype calls of technical replicates.

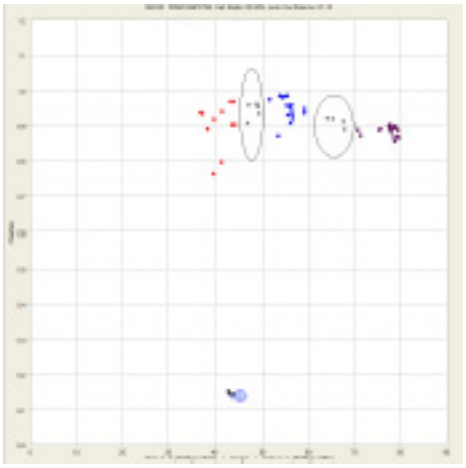


Figure 5. Example of a broad cluster plot with some samples labeled as "No call."

Product Information

Description	Product No.	Preps
Mag-BIND® Plant DNA DS Kit	M1130-00	1 x 96
	M1130-01	4 x 96

References

[1]. Omega Bio-tek (2015) Rapid, high performance and cost-effective plant DNA extractions.

[2]. Ganal, MW, *et al*. (2011). A large maize SNP genotyping array: Development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. PLoS ONE 6(12):e17224. doi: 10.1371/journal.pone.0028334.

400 Pinnacle Way, Suite 450
Norcross, GA 30071

www.omegabiotek.com
info@omegabiotek.com
(800) 832-8896