SCST Genetics Super Workshop

*SNP Testing*

////////

**Bryenna Warmann**
**Bayer Crop Science**
**St. Louis, MO**
Bryenna.Warmann@bayer.com
**February 2024**

# What is a SNP?

Single Nucleotide Polymorphism (SNP)

// SNP is a substitution of a single nucleotide at a specific position on the genome.

Variety 1: ATGCCGTAGTACA**T**TGTATGCTAATAGCG

Variety 2: ATGCCGTAGTACA**G**TGTATGCTAATAGCG

// Needs to be contained in at least 1% of the species population

// Usually, biallelic

// Can occur in non-coding or coding regions

// Caused by mutation and natural selection

// Are inherited

// 100Ks – 1Ms of SNPs can exist in whole genomes

// May or may not be trait-associated



https://www.socmucimm.org/resources/news-media/single-nucleotide-polymorphism-snp-allele-frequency-dna-pools/
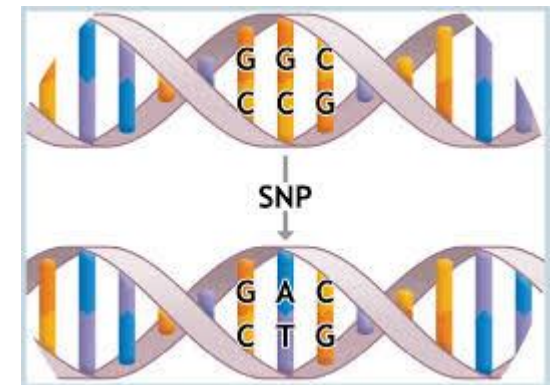
# Identifying SNPs

Genome Sequencing is used to find SNP locations

// Populations

// Breeding varieties

Reference -      ….ATGCCGTAGTACA**[T/G]**TGTATGCTAATAGCG….
Variety 1 –      ….ATGCCGTAGTACATTGTATGCTAATAGCG….
Variety 2 –      ….ATGCCGTAGTACAGTGTATGCTAATAGCG….
Variety 3 –      ….ATGCCGTAGTACATTGTATGCTAATAGCG….
Variety 4 –      ….ATGCCGTAGTACAGTGTATGCTAATAGCG….
Variety 5 –      ….ATGCCGTAGTACAGTGTATGCTAATAGCG….

// Genome-Wide Association Studies (GWAS)

// Aim to identify associations between SNPs and phenotypic traits

// Public Databases

// Research institutions

# Uses in Genetic Testing

SNPs are very useful genetic markers

// Marker Assisted Breeding Programs

// Quality Testing Programs

## Varietal Testing

// Hybrids

// X 2 different varieties give us a new variety

// Inbreds

// X Same variety gives us the same variety

## Hybridity Testing

// Hybrids

## Trait Testing

// Traits resulting from SNPs

// Co-linked traits



| | |
|---|---|
| Individual 1 | TC |
| Maternal | . . . C G A T A T T C C **T** A T C G A A T G T C . . . |
| Paternal | . . . C G A T A T T C C **C** A T C G A A T G T C . . . |
| Individual 2 | CC |
| Maternal | . . . C G A T A T T C C **C** A T C G A A T G T C . . . |
| Paternal | . . . C G A T A T T C C **C** A T C G A A T G T C . . . |
| Individual 3 | TT |
| Maternal | . . . C G A T A T T C C **T** A T C G A A T G T C . . . |
| Paternal | . . . C G A T A T T C C **T** A T C G A A T G T C . . . |
| Individual 4 | CT |
| Maternal | . . . C G A T A T T C C **C** A T C G A A T G T C . . . |
| Paternal | . . . C G A T A T T C C **T** A T C G A A T G T C . . . |

https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms

4

# Creating a SNP Panel

Informativity vs. Cost

Determine the Purpose: Line identification vs Line confirmation

// Breeding Programs
  // Early in the process
  // Ensure advancing correct lines
  // Small number of samples
  // More markers = high genome representation

// Quality Control Programs
  // Later in the process
  // Quality check to confirm production processes
  // Larger number of samples
  // Less markers = lower cost

Things to Consider

// Informativity
  // The ability to differentiate between lines within a population

// Allele distribution by marker
  // PIC-value

// Number of markers needed

// Genome representation

// Inbred vs Hybrid informativity

# PIC Value

Polymorphism Information Content (PIC)

// Allele distribution by marker

// Ideally, 50:50 distribution of alleles

   // PIC value: ~0.5 = high diversity

   // PIC value: > 0.25 and <0.5 = intermediate diversity

   // PIC value: <0.25 = low diversity

// PIC = 1 - (allele $1^2$ + allele $2^2$)

| | Mk1 | Mk2 | Mk3 |
|---|---|---|---|
| Allele 1 | 50 | 90 | 10 |
| Allele 2 | 50 | 10 | 90 |
| Total lines | 100 | 100 | 100 |
| Allele 1 + Allele 2 | 100 | 100 | 100 |
| Allele 1% | 0.50 | 0.90 | 0.10 |
| Allele 2% | 0.50 | 0.10 | 0.90 |
| Allele 1%$^2$ | 0.25 | 0.81 | 0.01 |
| Allele 2%$^2$ | 0.25 | 0.01 | 0.81 |
| PIC : 1-(Allele $1^2$ + Allele $2^2$) | 0.50 | 0.18 | 0.18 |

# SNP Marker Panel Informativity

// Informativity
   // The ability to differentiate between lines within a population

// Example:
   // 10 varieties but only 5 unique genotypes
   // With this panel if varieties 1 and 3 were accidentally mixed together or swapped, the test wouldn't be able to tell them a part - maybe we care, maybe we don't

// Understanding testing scope will help determine correct panel informativity

|  | Mk1 | Mk2 | Mk3 | Mk4 | Mk5 |
|---|---|---|---|---|---|
| Var1 | Allele 1 | Allele 2 | Allele 1 | Allele 2 | Allele 1 |
| Var2 | Allele 1 | Allele 2 | Allele 1 | Allele 2 | Allele 2 |
| Var3 | Allele 1 | Allele 2 | Allele 1 | Allele 2 | Allele 1 |
| Var4 | Allele 1 | Allele 2 | Allele 1 | Allele 2 | Allele 2 |
| Var5 | Allele 1 | Allele 2 | Allele 1 | Allele 2 | Allele 1 |
| Var6 | Allele 1 | Allele 2 | Allele 2 | Allele 1 | Allele 2 |
| Var7 | Allele 1 | Allele 2 | Allele 2 | Allele 1 | Allele 1 |
| Var8 | Allele 1 | Allele 2 | Allele 2 | Allele 1 | Allele 2 |
| Var9 | Allele 1 | Allele 2 | Allele 2 | Allele 1 | Allele 1 |
| Var10 | Allele 2 | Allele 1 | Allele 2 | Allele 1 | Allele 2 |
| PIC : 1-(su | 0.18 | 0.18 | 0.50 | 0.50 | 0.50 |

# Technologies for SNP Marker Detection

// TaqMan

    // Qualitative end-point

    // Quantitative real-time

// KASP® – Kompetitive Allele-Specific PCR – LGC/Biosearch Technologies

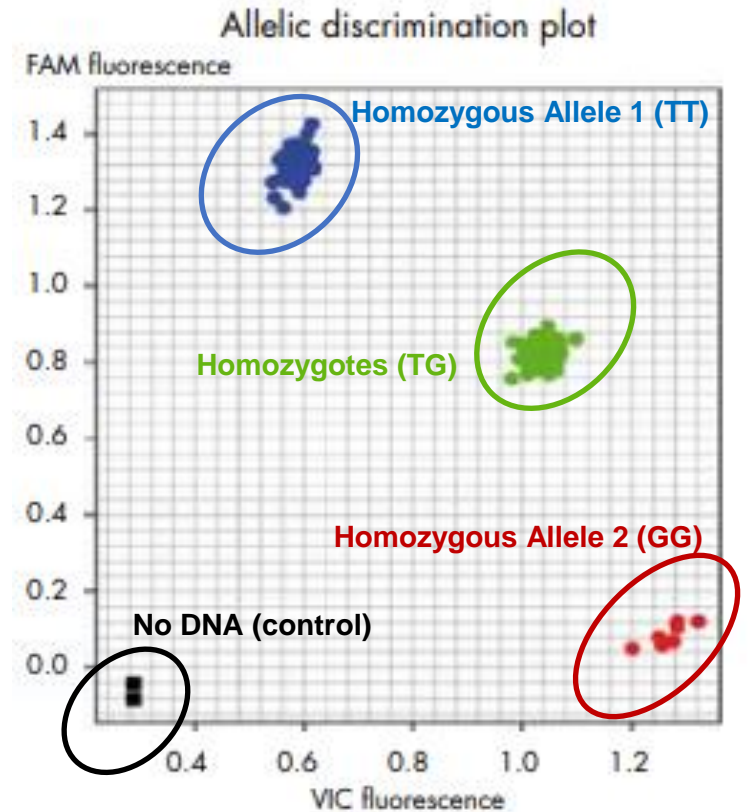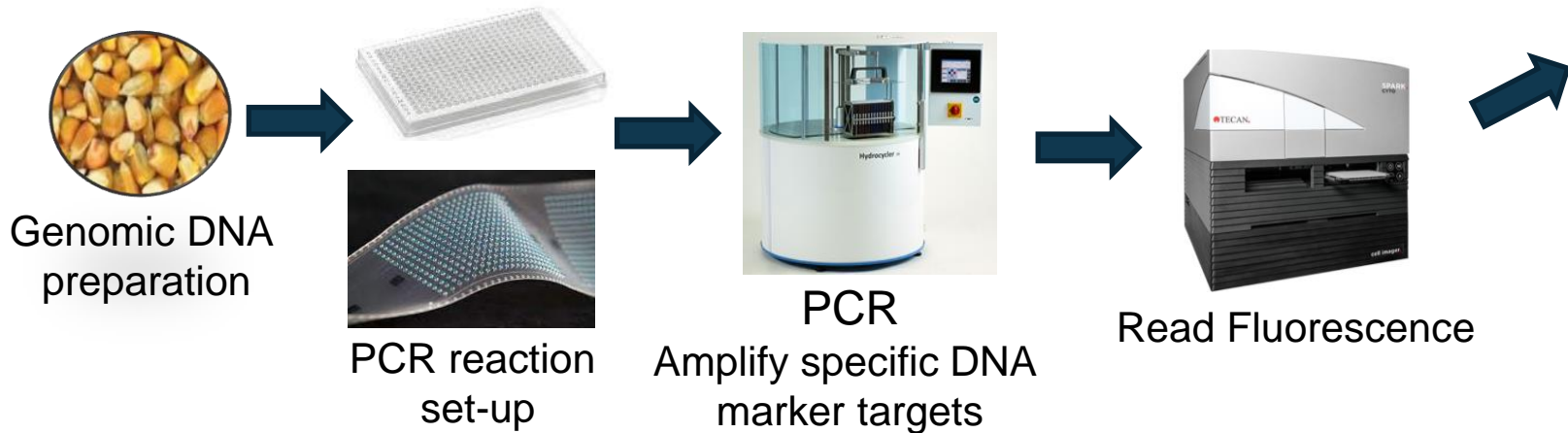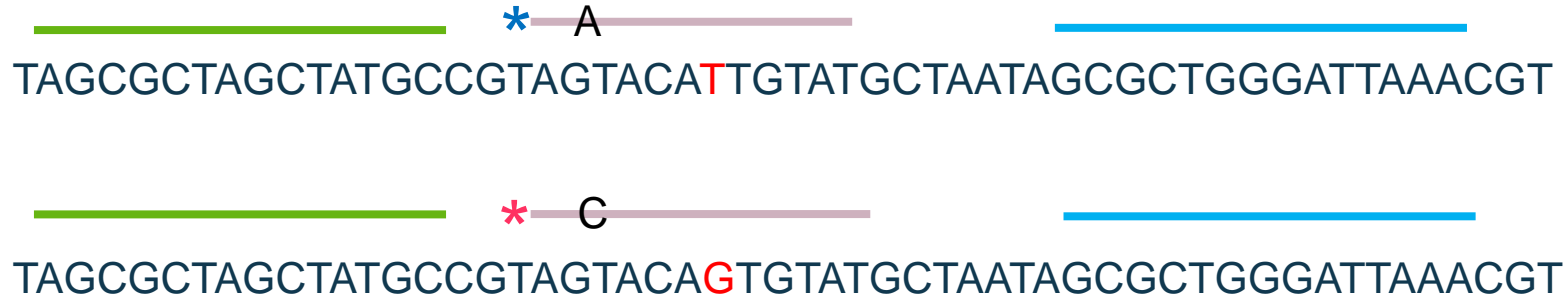// rhAMP® - Integrated DNA Technologies

// Microarray-based Detection

// Targeted genotyping by sequence (tGBS)

# Endpoint TaqMan PCR

Technologies cont.



* A

TAGCGCTAGCTATGCCGTAGTACA**T**TGTATGCTAATAGCGCTGGGATTAAACGT

* C

TAGCGCTAGCTATGCCGTAGTACA**G**TGTATGCTAATAGCGCTGGGATTAAACGT

Genomic DNA
preparation

PCR reaction
set-up

PCR
Amplify specific DNA
marker targets

Read Fluorescence

Allelic discrimination plot

FAM fluorescence

Homozygous Allele 1 (TT)

Homozygotes (TG)

Homozygous Allele 2 (GG)

No DNA (control)

VIC fluorescence

Type-it® Fast SNP Probe PCR Kit,
product insert (ver.12/2015). Qiagen

# Multiplexing

Technologies cont.

// Targeting multiple SNP locations in a single reaction or tube/well

  // Thousands of SNPs detected in a single reaction

// Specific primers and probes for each allele
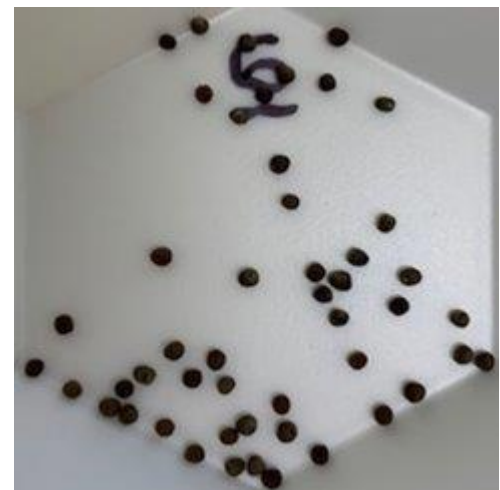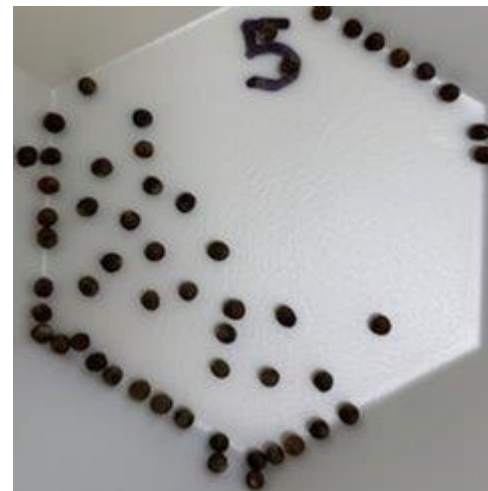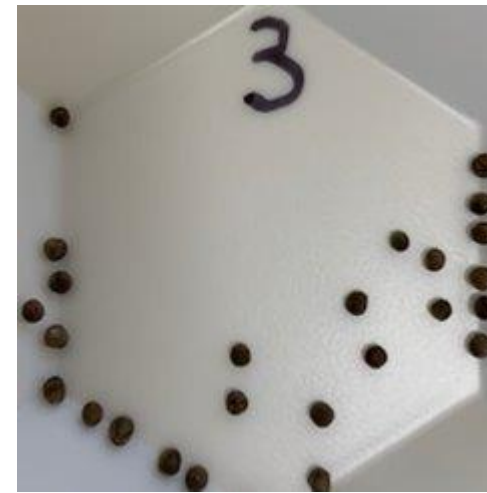
// GeneChip – 900,000+ SNP detection
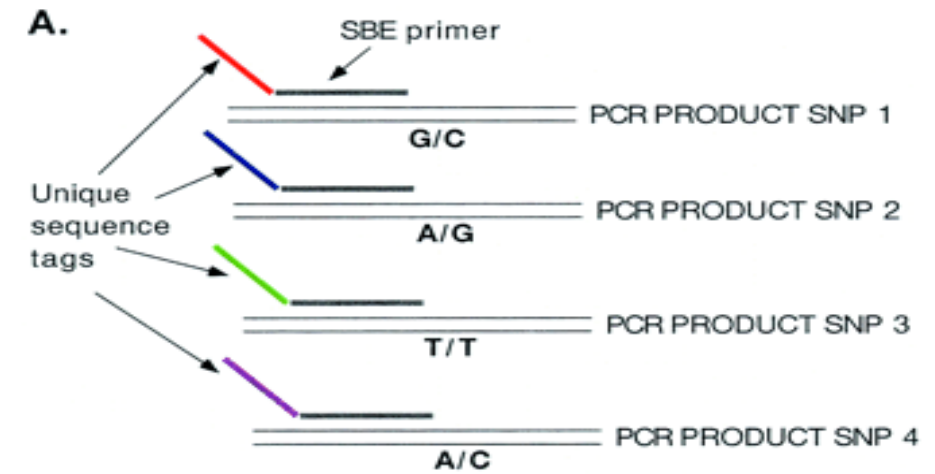


PNAS October 24, 2000 97 (22) 12164-12169; https://doi.org/10.1073/pnas.210394597



https://www.bumc.bu.edu/microarray/services/affymetrix-genechip/#:~:text=Affymetrix%20GeneChip%C2%AE%20System%20for%20Genotyping&text=This%20new%20approach%20increases%20the,markers%20in%20the%20replication%20phase.

10

# Can You Guess the Varieties (or Crop)?

# Cheddar  1

**PRODUCT DETAILS**

| Days to Maturity | Head Color | Head Curd |
|---|---|---|
| **80-84 Days** | **Medium-orange** | **Dense** |
| Head Shape | Plant Size | |
| **Domed** | **Medium-large** | |

SNP Profile - H  1  2  H  1  1  2  H  2  1

# Freedom  6

**PRODUCT DETAILS**

| Days to Maturity | Head Color | Head Curd |
|---|---|---|
| **72-77 Days** | **Very white** | **Dense** |
| Head Shape | Plant Size | |
| **Domed** | **Medium** | |

SNP Profile - 1  1  2  1  H  1  2  1  H  1

# Omphalos  4

**Product Details**

| | |
|---|---|
| **Fruit Colour** | Dark green |
| **Plant Type** | |
| **Related Maturity / Picking** | 75 - 80 days |
| **Fruit Size** | 1.5 to 1.8kgs |
| **Fruit Shape/Skin** | Round |

SNP Profile - 2  H  H  1  1  H  1  2  H  2

# Green Voyager  2

**Product Details**

| | |
|---|---|
| **Fruit Colour** | Bluie green |
| **Plant Type** | |
| **Related Maturity / Picking** | 15 - 20 days |
| **Fruit Size** | 1.4 to 1.6kgs |
| **Fruit Shape/Skin** | Round |

SNP Profile - H  H  1  H  H  H  H  2  H  2

# Castle Dome  3

**PRODUCT DETAILS**

| Relative Maturity | Bead Size | Color |
|---|---|---|
| **Early** | **Fine** | **Green** |
| Head Type | | |
| **Tight dome** | | |

SNP Profile - H  2  1  2  H  H  1  H  H  H

# Ironman  5

**PRODUCT DETAILS**

| Relative Maturity | Season | Bead Size |
|---|---|---|
| **Mid-late** | **Warm** | **Medium fine** |
| Color | Head Type | Desert or Coastal |
| **Blue green** | **High dome** | **Desert** |

SNP Profile - 1  H  1  H  2  1  2  1  H  H

12

*Thank you!*

**Questions?**