Corn breeding and genetic analysis of agronomic traits relies on phenotypic variation and genetic variation in the genes controlling these traits. My research program focuses on understanding genetic diversity in maize so we can mine beneficial alleles from the appropriate germplasm sources for continued corn improvement.

The process of domestication that began 9000 years ago has had profound consequences on maize, where modern corn has moderately reduced genetic diversity across nearly all genes in the genome relative to teosinte, and severely reduced levels of diversity for key genes targeted by domestication. The question that remains is whether these reductions in genetic diversity have impacted our ability to make progress in corn breeding today.

As an outcrossing species, maize has tremendous genetic variation compared to most other crops. The complementary combination of genome-wide association mapping (GWAS) approaches, large HapMap datasets, and germplasm resources are leading to important discoveries of the relationship between genetic diversity and phenotypic variation in inbred lines. However, among the traits targeted during domestication and breeding are many yield component traits, including number of ears, kernel row number, seed size, and kernel composition. Therefore, we must reintroduce variation from landraces and/or teosinte if we hope to learn how domestication has impacted these yield component traits, and yield itself.
Using Genetics, Genomics, and Breeding to Understand Diverse Maize Germplasm

Sherry Flint-Garcia
USDA-ARS Columbia, MO
Outline

Introduction to Maize Domestication & Diversity

Inbred Lines

Teosinte, the wild ancestor

Breeding with Zea
Maize Domestication

Domesticated from *Zea mays* ssp. *parviglumis*

Single domestication event ~9,000 years ago in Mexico

Intermediate form of landraces; populations adapted across the Americas to specific microclimates and/or human uses
Consequences of Artificial Selection

Artificial Selection

Domestication

Plant Breeding

Teosintes

Maize Landraces

Maize Inbred Lines

Unselected (Neutral) Gene

Domestication Gene

Improvement Gene

Zea Synthetic

Teosinte NILs

Teosinte Synthetic

98% (~49,000) maize genes

2% (~1,000) maize genes

Nested Association Mapping (NAM)
Maize ATLAS Project

Germplasm Enhancement of Maize (GEM)
Inbred Lines
Linkage-Based QTL Mapping

“Genome Scan”

Identify genomic regions that contribute to variation and estimate QTL effects

Parent 1       Parent 2

F_1            F_2 population

Genotype
Phenotype

Statistics for Mapping

Position (cM)

LOD Score

0  10  20  30  40  50  60  70  80  90  100  110  120  130  140
u1622  u1552  b2277  m231  b248  b1225  b2077  b2070  b1520
## Association Analysis

Utilize natural populations

Exploit extensive historical recombination

### Candidate gene approach:

<table>
<thead>
<tr>
<th>Line</th>
<th>Gene X</th>
<th>Ancestral chromosomes</th>
<th>Subject Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line 1</td>
<td>A C G A</td>
<td><img src="image1" alt="Ancestral chromosomes" /></td>
<td><img src="image2" alt="Subject Population" /></td>
</tr>
<tr>
<td>Line 2</td>
<td>A C G A T</td>
<td><img src="image1" alt="Ancestral chromosomes" /></td>
<td><img src="image2" alt="Subject Population" /></td>
</tr>
<tr>
<td>Line 3</td>
<td>A T A A G</td>
<td><img src="image1" alt="Ancestral chromosomes" /></td>
<td><img src="image2" alt="Subject Population" /></td>
</tr>
<tr>
<td>Line 4</td>
<td>C T A G T</td>
<td><img src="image1" alt="Ancestral chromosomes" /></td>
<td><img src="image2" alt="Subject Population" /></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td><img src="image1" alt="Ancestral chromosomes" /></td>
<td><img src="image2" alt="Subject Population" /></td>
</tr>
<tr>
<td>Line N</td>
<td>A C G G</td>
<td><img src="image1" alt="Ancestral chromosomes" /></td>
<td><img src="image2" alt="Subject Population" /></td>
</tr>
</tbody>
</table>

1.3m
2.0m

1.4m
2.5m

1.5m
1.8m
# Linkage vs. Association Analysis

## Linkage (QTL) Mapping

- **Structured population**
  - BC, F2, RIL, etc
- **Analysis of 2 alleles**
- **High power**
- **Low resolution**
  - 10-20 cM (10-20 Mb in maize)
- **Genome scan**
  - Don’t know anything about the genetics underlying the trait

## Association Mapping

- **Unstructured population**
  - Unrelated or distantly-related
- **Analysis of many alleles**
- **Low power**
- **High resolution**
  - 1000-5000 bp in maize *
- **Candidate gene testing**
  - Pathway or candidates previously identified. Used as a validation method.
- **Genome Scan**
  - Don’t know anything about the genetics underlying the trait. Used as a discovery method

* Depends on the species/population
Based on 89 SSR loci

NAM Development

Linkage Mapping

Association Mapping

Corn Tracker
NAM Kernel Composition

Starch

Amylose

Amylopectin

Fiber

NIR conducted by Syngenta Seeds, Inc.

Zeins

Protein

Amino Acid Profiles

Oil

Fatty Acid Profiles

Jason Cook
Oil Composition QTL in NAM

Joint Linkage Mapping - Oil

Genetic Distance (cM)

Physical Distance (bp)

Chr. 6 Oil Candidate: DGAT1-2

Encodes acyl-CoA:diacylglycerol acyltransferase

Fine mapped by Pioneer-Dupont

High parent = 19% oil
High allele = 0.29% additive effect
High allele has Phenylalanine insertion in C-terminus

NAM Genome Wide Association (GWAS)

1.6 Million HapMap.v1 SNPs projected onto NAM
Bootstrap (80%) sampling to test robustness of models

**DGAT 1-2 (Chr 6: 105,013,351-105,020,258)**

NAM Population: 24 HapMap.v1 SNPs in DGAT

281 Association Panel: 2 55K SNPs in DGAT (plus the 3 bp indel)

<table>
<thead>
<tr>
<th>Marker</th>
<th>Trait</th>
<th>Population</th>
<th>Analysis Method</th>
<th>BPP</th>
<th>P-Value</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Oil</td>
<td>282 Assn.</td>
<td>MLM (Q+K)</td>
<td>-</td>
<td>1.2E-04</td>
<td>0.18</td>
</tr>
<tr>
<td>M2</td>
<td>Oil</td>
<td>282 Assn.</td>
<td>MLM (Q+K)</td>
<td>-</td>
<td>9.9E-04</td>
<td>0.16</td>
</tr>
<tr>
<td>M3</td>
<td>Oil</td>
<td>NAM</td>
<td>GWAS - Bootstrap</td>
<td>31</td>
<td>-</td>
<td>0.18</td>
</tr>
<tr>
<td>M4</td>
<td>Oil</td>
<td>282 Assn.</td>
<td>MLM (Q+K)</td>
<td>-</td>
<td>4.3E-05</td>
<td>0.19</td>
</tr>
<tr>
<td>M4</td>
<td>Starch</td>
<td>NAM</td>
<td>GWAS - Bootstrap</td>
<td>51</td>
<td>-</td>
<td>-0.38</td>
</tr>
<tr>
<td>M5</td>
<td>Oil</td>
<td>NAM</td>
<td>GWAS - Bootstrap</td>
<td>67</td>
<td>-</td>
<td>0.13</td>
</tr>
<tr>
<td>M5</td>
<td>Starch</td>
<td>NAM</td>
<td>GWAS - Bootstrap</td>
<td>11</td>
<td>-</td>
<td>-0.31</td>
</tr>
</tbody>
</table>

DGAT 1-2 (Chr6: 105,013,351-105,020,258)

Rare allele?

= B73 Genotype
= Non-B73 Genotype

Summary – NAM Kernel Composition

Genetic Architecture of Kernel Composition Traits

Governed by many QTL ($N = 21-26$) with small to moderate effects

GWAS results confirm many QTL

DGAT is our favorite gene, but we still don’t have the complete story!

NAM In General

We can identify the common alleles in maize, but still have problems with rare alleles.

Ames Plant Introduction Station Inbreds

2,815 inbred lines from the Ames PI Station

Genotyping-by-sequencing (GBS) - 681,257 SNPs

Romay, et al. (2013)
Genome Biology
Ames Plant Introduction Station Inbreds

More than half of the SNPs in collection are rare!

302 Association panel = 75%, NAM founders = 57%.

Romay, et al. (2013) Genome Biology
Teosinte - The Wild Ancestor
Development of Teo Introgression Libraries

B73 × teosinte
(parviglumis)

10 accessions

B73 × teosinte

F1

BC1

BC2

BC3

BC4
Library Development

10 libraries of 887 BC4S2/DH Near Isogenic Lines (NILs)

BC4S4 NILs: GBS & RAD sequencing = 33,000–600,000 SNPs

Lines to be released in 2013
Library Coverage

10 maize-teosinte libraries
804 BC4S2 NILs and 83 BC4DH NILs
Each line:
  2.3 chromosomal segments
  4.1% of the teosinte genome
3.3X genome coverage
BC4S2 vs BC4DH from PI384071 donor
Applications of Teosinte NILs
Application 1. Empirical Genetics

1000 Selected Genes

What do these selected genes do?

What traits were targeted by artificial selection during domestication/breeding?

Are selected genes important?

Auxin response factor, ARF1

Diversity ($\pi$)

Inbreds
Teosinte

Tillering/branching?
Application 2. QTL Mapping

Phenotyped at up to 18 reps:
Maturity, plant & ear height,
kernel row number, kernel weight,
kernel shape, leaf length-width,

Grain composition
(protein, starch, oil)
Days to Anthesis

NAM RILs

Teosinte NILs
### Kernel Row Number

<table>
<thead>
<tr>
<th>Pop</th>
<th>Add. Eff.(rows)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z029</td>
<td>-3.4</td>
</tr>
<tr>
<td>Z030</td>
<td>-1.8</td>
</tr>
<tr>
<td>Z031</td>
<td>-0.6</td>
</tr>
<tr>
<td>Z032</td>
<td>-1.4</td>
</tr>
<tr>
<td>Z033</td>
<td>-1.0</td>
</tr>
<tr>
<td>Z034</td>
<td>-1.1</td>
</tr>
<tr>
<td>Z035</td>
<td>-1.1</td>
</tr>
<tr>
<td>Z036</td>
<td>-0.3</td>
</tr>
<tr>
<td>Z037</td>
<td>0.2</td>
</tr>
<tr>
<td>Z038</td>
<td>-0.9</td>
</tr>
</tbody>
</table>

- **Z029**
  - Add. Eff.(rows): -3.4
- **Z030**
  - Add. Eff.(rows): -1.8
- **Z031**
  - Add. Eff.(rows): -0.6
- **Z032**
  - Add. Eff.(rows): -1.4
- **Z033**
  - Add. Eff.(rows): -1.0
- **Z034**
  - Add. Eff.(rows): -1.1
- **Z035**
  - Add. Eff.(rows): -1.1
- **Z036**
  - Add. Eff.(rows): -0.3
- **Z037**
  - Add. Eff.(rows): 0.2
- **Z038**
  - Add. Eff.(rows): -0.9

### Additional Data

<table>
<thead>
<tr>
<th>Pop</th>
<th>Add. Eff.(rows)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z029</td>
<td>---</td>
</tr>
<tr>
<td>Z030</td>
<td>-1.4</td>
</tr>
<tr>
<td>Z031</td>
<td>-1.4</td>
</tr>
<tr>
<td>Z032</td>
<td>-1.1</td>
</tr>
<tr>
<td>Z033</td>
<td>-2.1</td>
</tr>
<tr>
<td>Z034</td>
<td>-2.3</td>
</tr>
<tr>
<td>Z035</td>
<td>-1.7</td>
</tr>
<tr>
<td>Z036</td>
<td>-0.6</td>
</tr>
<tr>
<td>Z037</td>
<td>-0.7</td>
</tr>
<tr>
<td>Z038</td>
<td>-1.4</td>
</tr>
<tr>
<td>Pop</td>
<td>Add.Eff.(g)</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>Z029</td>
<td>-1.2</td>
</tr>
<tr>
<td>Z030</td>
<td>-0.9</td>
</tr>
<tr>
<td>Z031</td>
<td>-0.6</td>
</tr>
<tr>
<td>Z032</td>
<td>-1.2</td>
</tr>
<tr>
<td>Z033</td>
<td>-0.4</td>
</tr>
<tr>
<td>Z034</td>
<td>-1.3</td>
</tr>
<tr>
<td>Z035</td>
<td>-2.3</td>
</tr>
<tr>
<td>Z036</td>
<td>-1.2</td>
</tr>
</tbody>
</table>

![Graph showing Seed Weight (50 kernels)](image)
Application 3. Reintroduce Variation

Biological hypothesis: A loss of genetic variation results in a loss of phenotypic variation.

Breeding hypothesis: We can improve modern traits.

Flint-Garcia et al. (2009) TAG
Breeding With Zea
Synthetic Populations

\[ \text{NAM Synthetic} \times \text{Teosinte Synthetic} \rightarrow \text{Zea Synthetic} \]

B73 & NAM parents are inbreds, teosinte has NEVER been inbred
Zea Synthetic Inbreeding Depression

38% B73, 2% each NAM parent, 12% teosinte

Selfed (S1) \( F \approx 0.5 \)

Full Sib (FS) \( F \approx 0 \)

924 pairs
Zeaa Synthetic Doubled Haploids

38% B73, 2% each NAM parent, 12% teosinte

Goal: 2000 DH
First 800+ in 2013
Another 700 in Puerto Rico
AgReliant producing more
2014 trial: MO, NC, NY, IA
“Crazy?” Idea

7000 BC

Agronomics
Fertilizer
Density
Mechanization

Select only on yield

2014

ideotype

?
Teosinte Synthetic

75% B73 (SS), 25% teosinte
Acknowledgements

NSF Maize Diversity Project

Syngenta Seeds AgReliant Genetics

Susan Melia-Hancock, Kate Guill, Jason Cook, Zhengbin Liu, Ginnie Morrison, Christopher Bottoms, Avi Karn, Anna Selby

www.panzea.org