The Molecular Basis of Crop Domestication Studies

Paul Gepts
Plant Sciences
University of California, Davis

Outline

• Historical perspective
• Why molecular markers?
• Genealogical (phylogenetic) analysis
• Centers and patterns of domestication
• The inheritance of the domestication syndrome
• Isolation of domestication genes
• The fate of genetic diversity during and after domestication
• Search for genes involved in domestication
Historical Perspective

- A. de Candolle: 19th century
  - Four types of data: biological (distribution of wild relatives), archaeological (archaeobotany: macro-remains), historical, linguistic
- Vavilov: early 20th century
  - Biological data: Centers of diversity = center of origin
- Second half 20th century
  - Biological data:
    - Harlan: Cytogenetics, application of species concept
    - Various: Use of molecular markers
  - Archaeological data:
    - Accelerator Mass Spectrometry of $^{14}$C
    - Micro-remains

Why molecular markers?

- It was possible! Molecular capabilities:
  - Allozymes $\rightarrow$ Seed proteins $\rightarrow$ DNA-based markers (RFLP-RAPD-AFLP-SSR-Sequence)
- Large numbers $\rightarrow$ better coverage
- Devoid of environmental influence
- Neutrality $\rightarrow$ genetic distance
- Selectivity (directly, indirectly) $\rightarrow$ identification of genome regions under selection
Outline

- Historical perspective
- Why molecular markers?
- Genealogical (phylogenetic) analysis
- Centers and patterns of domestication
- The inheritance of the domestication syndrome
- Isolation of domestication genes
- The fate of genetic diversity during and after domestication
- Search for genes involved in domestication

Molecular Phylogeny of *Phaseolus*

- P. lunatus Group
- P. pauciflorus Group
- P. vulgaris Group
- P. leptostachyus Group
- P. filiformis Group
- P. polystachios Group
- P. tueckheimii Group
- P. pedicellatus Group

Combined nrDNA and cpDNA

4.1 Ma

Slide: A. Delgado-Salinas

Delgado-Salinas, Bibler & Lavin. 2006
Two principal *Phaseolus* clades

<table>
<thead>
<tr>
<th>Domesticated taxa</th>
<th>Clade A</th>
<th>Clade B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of species</td>
<td>24</td>
<td>32</td>
</tr>
<tr>
<td>Geographic distribution</td>
<td>Limited: AZ, NM, TX to Panama</td>
<td>Broader: SE Canada to Bolivia; islands</td>
</tr>
<tr>
<td>Elevations/Vegetations</td>
<td>Highlands; oak-pine forests</td>
<td>Lowlands to highlands: dry to wet forests</td>
</tr>
<tr>
<td>Elevation and latitude windows</td>
<td>690 +/- 545m; 4°26” +/- 5°06”</td>
<td>737 +/- 654 m; 6°22” +/- 10°21”</td>
</tr>
<tr>
<td>Flowering</td>
<td>Only rainy season</td>
<td>Dry or rainy season</td>
</tr>
<tr>
<td>Long frost periods</td>
<td>No tolerance</td>
<td>Tolerance</td>
</tr>
<tr>
<td>Habitat disturbance</td>
<td>Sensitive</td>
<td>Less sensitive</td>
</tr>
<tr>
<td>Infraspecific</td>
<td>Rare</td>
<td>More frequent</td>
</tr>
</tbody>
</table>

Multiple Domestications in *Phaseolus*

- Principal feature of domestication in *Phaseolus*
- Few other examples:
  - *Capsicum*
  - *Cucurbita*
  - *Oryza*
  - *Cicer*
  - *Gossypium*
- Questions:
  - Why non-random distribution of domestications?
  - Locations of domestication?
  - Same mutations/genes?
Wild beans in Ecuador and Northern Peru

Distribution of Wild Phaseolus vulgaris and Bean Domestications

Wild beans
- Mesoamerican
- Ancestral
- Andean

Domestications
- Mesoamerican
- Andean

Principal Component Analysis of Allozyme Diversity
Phaseolin as an evolutionary marker

- Phaseolin: major seed storage protein in beans; 35-50% of total seed N
- Small multigene family: 6-9 genes; single, complex locus
- Two major phaseolin types:
  - S: Mesoamerican
  - T: Andean
- Two classes of genes within S and T phaseolin gene families
  - alpha: have tandem direct repeats
  - beta: do not have tandem direct repeats

Polymerase chain reaction (PCR)
Assay of Phaseolin Diversity

[Diagram showing PCR primer binding sites, third exon, third intron, fourth exon, and direct repeat regions with 124(+21)bp and 15(+15)bp]
PCR amplification of the 15-bp region in phaseolin

Polyacrylamide gel showing PCR products from different haplotypes: 1) T phaseolin, 2) S phaseolin, 3) I phaseolin, M 123 bp ladder.

Survey of the distribution of tandem direct repeats in phaseolin

<table>
<thead>
<tr>
<th>Region</th>
<th>15 bp present</th>
<th>15 bp absent</th>
<th>21 bp present</th>
<th>21 bp absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mesoamerica: MEX, C. Am., COL</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Intermediate: ECD, N. PER</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>S. Andes: S. PER, BOL, ARG</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>
Interpretation of the distribution of repeats

- Absence of tandem direct repeats in wild populations from Ecuador and N. Peru
- Absence of repeats also in nearest relatives (*P. polyanthus, P. coccineus*)
- Probability of generating tandem direct repeat is higher than losing one:
  - suggests that absence of repeats is ancestral state
- Intermediate geographic distribution:
  - suggests dispersal northwards and southwards of wild populations (but how?) followed by domestications in Mesoamerica and southern Andes

Outline

- Historical perspective
- Why molecular markers?
- Genealogical (phylogenetic) analysis
- Centers and patterns of domestication
  - The inheritance of the domestication syndrome
  - Isolation of domestication genes
  - The fate of genetic diversity during and after domestication
- Search for genes involved in domestication
Microsatellite marker analysis of maize origin

Matsuoka et al. 2002

- **Plant material:**
  - Broad cross-section: 246 plants

- **Markers:**
  - 99 microsatellites
Phylogenetic analyses (I)

Phylogenetic analyses (II)
Age of domestication

- 33 microsatellite markers: stepwise variation
- estimate mutation rate in a known population: $F_{11}$
  - $4.28 \times 10^{-4}$
- result:
  - 9,188 B.P. (95% confidence limits of 5,689-13,093 B.P.)

Proposed Center of Domestication of Common Bean in Mesoamerica

Wild beans most closely related to domesticated beans

Other wild beans

M. Kwak & P. Gepts, in prep.
The Rio Lerma-Rio Grande de Santiago basin

- Climate: Cwa
  - Subtropical: t° coldest month: 5-18 °C
  - Subhumid: 4-6 months of humidity in summer
  - Semi-warm: average annual t°: 18-22 °C

- Vegetation:
  - Dry deciduous forest to drier thorn forest

Distribution of wild beans and teosinte in Mexico

Smith 1994
Site of Einkorn Wheat Domestication Identified by DNA Fingerprinting (Heun et al., 1997)

- Fingerprinting of 338 lines based on the presence vs. absence of 288 AFLP bands
- Lines were assigned into groups based on geographical origin.
- Cultivated einkorns are closely related among themselves and show a close phylogenetic similarity to group D

=> Are the D lines the closest relatives of the wild progenitors that gave rise to cultivated einkorn?
Multiple domestications in einkorn wheat?

Dispersed-specific model of domestication:
• Pre-domestication cultivation & dispersal of β race
• Multiple domestication in different localities

Kilian et al. 2007

The Origin of the Apple

Harris et al. 2002
**Outline**

- Historical perspective
- Why molecular markers?
- Genealogical (phylogenetic) analysis
- Centers and patterns of domestication
- The inheritance of the domestication syndrome
- Isolation of domestication genes
- The fate of genetic diversity during and after domestication
- Search for genes involved in domestication

**Domestication syndrome**

Wild → Early domesticate → Crop

- Suite of traits selected during domestication:
  - Initial domestication
  - Later selection: varietal differences
- Often deleterious in the wild → crops cannot survive without human intervention
- Seed dispersal and dormancy; reproductive system; growth habit; toxic compounds; consumer traits

Gepts 2004
Seed Dispersal

Aegilops sp.

Phaseolus vulgaris

Zea mays

Changes in Growth Habit

Phaseolus vulgaris

Zea mays

Pennisetum glaucum
Increase in Size of Inflorescence, Fruit and/or Grain

Zea mays

Pennisetum glaucum

Cucurbita sp.

Gigantism
Domestication Syndrome Traits

- Changes in reproductive system: Towards increased "reproducibility"
  - Outcrossing to selfing:
    - Tomato and peppers (Rick 19XX; Rice (Morishima 1984)
  - Outcrossing to sterility, parthenocarpy:
    - Banana-plantain (Simmonds 1959)
  - Outcrossing to vegetative:
    - Cassava (Elias and Mackey 2000)

Domestication Syndrome Traits

- Evolution of wheat yields in Mesopotamia (Araus et al. 2001):
  - c. 8000 BC: estimated grown yield was 1.56 Mg/ha
  - contemporary yields: roughly 1.0 Mg/ha
- Maize yields in U.S.A.
  - Flat until 20th century?
Summary of genetic studies on the domestication syndrome

- Linkage-map based analyses: QTL analyses
  - Outcrossers: Pearl millet, maize, sunflower, *Brassica oleracea*,
  - Selfers: Common bean, rice, tomato, barley, wheat, cotton, soybean, lentil, pea, tomato, eggplant
- Common features:
  - Few loci
  - Major phenotypic effect
  - Most of phenotypic variation accounted for in genetic terms = high heritability
  - Few regions of the genome = linked
- Exceptions:
  - Maize; sunflower

A comparison of linkages with other species

<table>
<thead>
<tr>
<th>Species</th>
<th>Chromosome</th>
<th>Growth habit, photoperiod sensitivity</th>
<th>Spikelet architecture and shedding, plant and spike morphologies and flowering</th>
<th>Plant and panicle morphology, shedding</th>
<th>Fruit traits, earliness</th>
<th>Growth habit, fruit traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common bean, 2n = 2x = 22</td>
<td>1</td>
<td>Ppd</td>
<td>sim, Tor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PD, NM, DF, DM, SW, HI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearl millet, 2n = 2x = 14</td>
<td>6,7</td>
<td>Pl, Al</td>
<td>C1, B1Pb</td>
<td>NNFT, Hmax, NS, WES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice, 2n = 2x = 24</td>
<td>1</td>
<td>TP, PNL, PHN, SBP, CC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tomato, 2n = 2x = 24</td>
<td>2</td>
<td>FS, FC, FW</td>
<td>D1, DR, FD, DTL, DFR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Consequences for domestication

- Genetic architecture observed could speed up domestication
- Domestication could have occurred within a few 100 generations, genetically speaking (>< archaeologically)
- Other factors affecting speed of domestication: genetic bottlenecks, strength of selection, effective population size, rates of gene flow, level of inbreeding
- Part of a potential for domestication?

Outline

- Historical perspective
- Why molecular markers?
- Genealogical (phylogenetic) analysis
- Centers and patterns of domestication
- The inheritance of the domestication syndrome
- Isolation of domestication genes
- The fate of genetic diversity during and after domestication
- Search for genes involved in domestication
Cloning of the first domestication gene

- Differences in growth habit (shape of the plant) between teosinte and maize, specifically with regard to lateral branches:

### Comparison of lateral branches of teosinte and maize

<table>
<thead>
<tr>
<th>Teosinte</th>
<th>Maize</th>
</tr>
</thead>
<tbody>
<tr>
<td>At most nodes</td>
<td>At 2-3 nodes</td>
</tr>
<tr>
<td>Elongated</td>
<td>Short</td>
</tr>
<tr>
<td>Tip of primary lateral branch: inflorescence = tassel</td>
<td>Tip of primary lateral branch: inflorescence = ear</td>
</tr>
<tr>
<td>Secondary lateral branch: carry ears</td>
<td>No secondary lateral branches</td>
</tr>
</tbody>
</table>

Tassel: male

Ear: female
Previous results from genetic analyses

- Analysis of domestication syndrome in maize (see Lecture 16)
  - LBIL: Lateral branch internode length; gene with largest effect for this trait, located on chromosome 1, in same region as \(tb1\) (teosinte branched)
- \(Tb1\), when introduced from maize into teosinte by hybridization, converts teosinte to maize plant type
- Effects of \(Tb1\):
  - Loss of apical dominance --> marked growth of axillary buds
  - Bottom of the plant: tillers; top of the plant: long branches ending in tassels (<> normal maize has short branches ending in ears)
  - \(F_2\) generation: teosinte: recessive allele; maize: dominant allele with increased mRNA --> repression of organ growth

---

**Fig. 3.** Diagram of the 10 teosinte–maize chromosomes showing distribution of MMLs used in this study. Thickness of cross-lines for each MML indicates any departures from Mendelian segregation ratios (see scale). Distances between the MMLs are shown as \(r\), the recombination fraction (see scale). Stippled blocks highlight six regions with major effects on the morphological differences between...
**Tb1 sequencing**  
(Wang et al. 1999, 2001)

- Sample of plants:
  - Maize (13): mostly Mex (6), also ECU, GTM, VEN; AZ, ND, WI  
  - Teosinte parviglumis (9): Gue (5), Mex, Jal (2), Mich  
  - Teosinte mexicana (8): Mex (4), Jal, Dur, Mich, Chi  
  - Teosinte diploperennis (1)

- Sample of sequence:
  - 2.9 kb of transcriptional unit (TU) and 1.1 kb of 5' non transcribed region

---

**Nucleotide polymorphism**

(from Wang et al. 1999, 2001)

- **Tb1**: TU  
  - Maize: \( \pi = 1.74 \); parviglumis: \( \pi = 4.62 \)

- **Tb1**: NTR  
  - Maize: \( \pi = 0.47 \); parviglumis: \( \pi = 28.68 \)

- **Adh1**:  
  - Maize: \( \pi = 15.72 \); parviglumis: \( \pi = 17.38 \)
Further analyses

- HKA test: subject to selective sweep?
  - Compare ratio (polymorphism w/in species/divergence to outgroup) for \( tb1 \) to same ratio for neutral gene
  - Test not significant for TU but well for NTR
  - Test also significant when TU used as neutral control; very short region of hitchhiking
- Regulatory sequence is key
  - But: no fixed differences in NTR between maize and teosinte --> further upstream?
- Selection coefficient: \( s = 0.04 \) to 0.08
- Time to fixation: 315 to 1,000 years
Domestication syndrome genes

- Some 30 genes isolated:
  - Often recessive, but exceptions: e.g., *tb-1*
  - Often transcription factors, but exceptions: e.g., starch biosynthesis enzymes
  - Often changes in regulatory, rather than structural regions
  - Often loss of function, but exceptions

Gepts 2004; Doebley et al. 2006; Burger et al. 2008

The Domestication Syndrome of Common Bean
Common bean determinacy: \textit{PvTFL1y}

Outline

- Historical perspective
- Why molecular markers?
- Genealogical (phylogenetic) analysis
- Centers and patterns of domestication
- The inheritance of the domestication syndrome
- Isolation of domestication genes
- The fate of genetic diversity during and after domestication
- Search for genes involved in domestication
Fate of Genetic Diversity

- General tendency:
  - Reduction of genetic diversity:
    - Depends on species, gene
  - Causes:
    - Demographic events: affect entire genome
    - Selection: affects specific gene
- Distinction between domestication and improvement (varietal differences) genes
- Exceptions:
  - Crops: e.g., einkorn wheat
  - Cropping systems: e.g., common bean

Experiments to measure evolution of genetic diversity at the molecular level

Chloroplast DNA (Doebley 1992)
Table 1. Ploidy, nucleotide diversity, and yield differences between domesticated crops and wild progenitors (from Buckler et al. 2001)

<table>
<thead>
<tr>
<th></th>
<th>Zea mays</th>
<th>Sorghum bicolor</th>
<th>Oryza sativa</th>
<th>Avena sativa</th>
<th>Hordeum vulgare</th>
<th>Triticum aestivum</th>
<th>Pennisetum glaucum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maize</td>
<td>Sorghum</td>
<td>Rice</td>
<td>Oats</td>
<td>Barley</td>
<td>Wheat</td>
<td>Pearl millet</td>
</tr>
<tr>
<td>Ploidy</td>
<td>Tetraploid</td>
<td>Diploid</td>
<td>Diploid</td>
<td>Hexaploid</td>
<td>Diploid</td>
<td>Hexaploid</td>
<td>Diploid</td>
</tr>
<tr>
<td>Progenitor diversity</td>
<td>0.0210</td>
<td>0.0035</td>
<td>0.0035′</td>
<td></td>
<td>0.0014′</td>
<td>0.0070′</td>
<td>0.0036′</td>
</tr>
<tr>
<td>Cultivar diversity</td>
<td>0.0163</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0016′</td>
<td>0.0050</td>
</tr>
<tr>
<td>Cult/prop diversity</td>
<td>78%</td>
<td>60%</td>
<td>71%</td>
<td>65%</td>
<td>117%</td>
<td>71%</td>
<td>67%</td>
</tr>
<tr>
<td>Wild yield</td>
<td>0.16″</td>
<td>&lt;0.60′</td>
<td>1.12″</td>
<td>2.93″</td>
<td>0.65″</td>
<td>0.65″</td>
<td>&lt;0.55″</td>
</tr>
<tr>
<td>Yield 1961</td>
<td>2.5</td>
<td>2.4</td>
<td>4.1</td>
<td>1.8</td>
<td>2.4</td>
<td>2.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Yield 2004</td>
<td>9.1</td>
<td>5.9</td>
<td>5.6</td>
<td>4.5</td>
<td>6.4</td>
<td>7.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Recent yield gain</td>
<td>3.6</td>
<td>2.5</td>
<td>1.4</td>
<td>2.5</td>
<td>2.7</td>
<td>3.0</td>
<td>1.4</td>
</tr>
<tr>
<td>Gain over progenitor</td>
<td>55.9</td>
<td>&gt; 9.9</td>
<td>5.0</td>
<td>1.5</td>
<td>9.8</td>
<td>11.0</td>
<td>&gt; 2.7</td>
</tr>
</tbody>
</table>

Evolution of Genetic Diversity

[Graph showing the decline in genetic diversity from wild progenitor to primitive domesticate to improved cultivar, with labels for neutral and selected populations.]

Burger et al. 2008
Analysis of "old DNA" in maize

• What is old DNA?
  – technical challenges:
    • degraded DNA: small size --> small fragments for PCR that distinguish maize and teosinte
    • contamination: controls, location of analysis: here, two locations: Leipzig and Cambridge!

• Application to crop evolution studies:
  – Cloning of domestication genes:
    • in maize:
      – $tb1$
      – *prolamin box binding factor* (pbf: seed storage protein expression)
      – *sugary1* (*su1*: starch composition --> texture of tortillas)

• Availability of archaeological samples

Ocampo Caves (Valenzuela cave), dated to 3890 ± 60 years before the present

Choice of DNA sequences

$tb1$: 56 bp fragment: distinguishes maize and teosinte; 1 allele: 100% in maize, 36% in teosinte (total of 6 alleles in teosinte)

$pbf$: 25 bp fragment: two alleles: 97% and 3% in maize; 17% and 83% in teosinte

$su1$: 60 bp fragment: two alleles: 30% and 62% in maize; both 7% in teosinte
Geographic and chronological differentiation

- Alleles of modern maize were already present 4,500 years ago
- Possible exception: *su1*, where 2,000 yr old cobs still carried alleles known now only from teosinte
- Southwest U.S.: possible origin of Northern Flint, one of the two parents of Corn Belt maize.

Conclusion: In conclusion, by 4400 years ago, early farmers had already had a substantial homogenizing effect on allelic diversity at three genes associated with maize morphology and biochemical properties of the corn cob. Thus, selection by farmers had profound genomic effects relatively early in the history of this crop.

Farmer-based Maintenance of Genetic Diversity

- Gene flow between wild and domesticated
- Exchange of varieties with other farmers

(Zizimbo-Villarreal et al. 2005)
Evolution of Genetic Diversity

Outline

- Historical perspective
- Why molecular markers?
- Genealogical (phylogenetic) analysis
- Centers and patterns of domestication
- The inheritance of the domestication syndrome
- The fate of genetic diversity during and after domestication
- Further search for genes involved in domestication
Top-down vs. bottom-up approaches

- **Top-down:**
  - QTL & LD analyses
  - Requires phenotypic segregation

- **Bottom-up:**
  - Population genetic analysis of genome scan
  - Does not require *a priori* knowledge of phenotypic segregation

Ross-Ibarra et al. 2007

Expectations in genome scan

- Sharp variation in polymorphism
- Detection of selection:
  - Strength & history of selection
  - Rates of mutation & recombination
  - Demography of population

- Assumptions:
  - Panmictic population
  - Constant population size

- Departures: inaccurate estimates

Ross-Ibarra et al. 2007; Walsh 2008
Dealing with demographic problems

- Develop demographic model → cumbersome model testing
  - Only in maize (see Wright et al. 2007)
- Empirical ranking by:
  - Parameters of diversity: no. of SNPs in gene, Tajima’s D
  - But false positive rate can be high

Examples of genome scans

- Maize:
  - Vigouroux et al. (2002)
    - 501 SSRs invariant in US maize
    - Neutrality tests, coalescent tests
    --> SSRs in 15 genes with evidence of selection
    - Also in sorghum and sunflower: clustered near domestication QTLs
  - Wright et al. (2005, 2007)
    - 774 genes: sequence information
    - 2-4% of genes with selection signature in this sample → in entire genome: 1,200 genes expected in maize genome with selection signature
    - Clustered near domestication QTLs; transcription factors, plant growth, amino acid biosynthesis
Examples of genome scans (II)

- Maize (contd.):
  - Yamasaki et al. (2005)
    - 1200 maize genes: nucleotide polymorphism data
    - 35 genes with no sequence diversity in maize inbreds = selection
    - Sequence information in teosinte: \( \rightarrow \) 8/35 showed signs of selection
- Sorghum:
  - Hamblin et al. (2004, 2006)
    - 371 loci
    - None with signature of selection (sample size?) \( \geq \) SSRs

Examples of genome scans (III)

- Common bean:
  - Papa et al. (2007):
    - \( \sim \) 2500 AFLP markers (1/250 kb) in 7 wild and 7 domesticated bulks
    - Significantly stronger reduction in genetic diversity for selected markers
    - Large fraction of the genome (16%) subject to effects of selection
    - Most markers under selection located near genes and QTLs of domestication \( \sim \) enrichment procedure

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>D markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putatively neutral</td>
<td>47</td>
<td>15 (32%)</td>
</tr>
<tr>
<td>Putatively selected</td>
<td>23</td>
<td>15 (65%)</td>
</tr>
</tbody>
</table>
Difficulties with Bottom-up approaches

- When the approach fails to uncover pattern of selection in nucleotide polymorphism:
  - No target of selection: not in history of crop, not in genetic background
  - Assayed polymorphism in wrong region
  - Low underlying levels of diversity; distribution of genetic diversity; inadequate sampling
  - History of domestication allele: pre- or post-domestication appearance
  - Issues of polyploidy, population structure
- No association with phenotype

Doebley et al. 2006
Applications to Germplasm Conservation & Breeding

• Conservation:
  – What and where to conserve?
  – Association with ecological data
• Role of wild relatives
• Limited genetic differences between wild and domesticated \( \rightarrow \) makes introgression easier
• In some species, extensive hitchhiking: positive or negative?
• Associated microorganisms: pathogens, beneficials