Maize Genomics: Exploitation of Structural Variation (SV) and Opportunities for Climate Change Mitigation

NCGA Workshop
Chicago, IL
7-8 April 2010
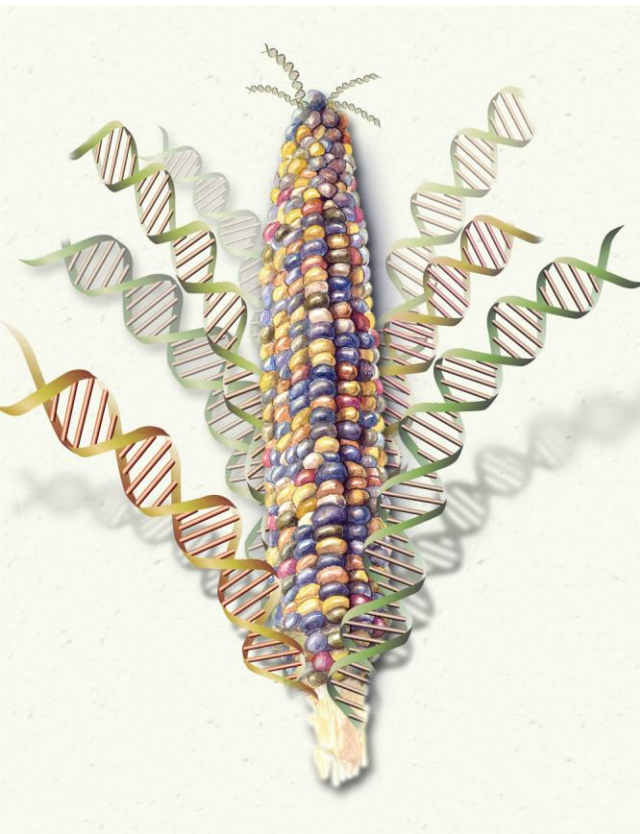
Patrick S. Schnable
Iowa State University

PLANT SCIENCES INSTITUTE

The $30M B73 Maize Genome Sequencing Project

- 2.3 GB maize genome has more complex structure than human genome
- WU Genome Sequencing Center (R. Wilson, PI); Arizona Genome Institute; Cold Spring Harbor Laboratory; Iowa State University
- Draft genome released 2/08

Nicolle Rager Fuller, NSF

Schnable et al., Science 2009

Genome Summary

- Centromeres positioned on genetic and physical maps (Dawe, Presting, J. Jiang, Bircher NSF project)

- Detailed TE annotation (Bennetzen, Wessler and N. Jiang NSF project): many new TEs discovered

- Automated gene calling, followed by extensive "filtering" for TEs, gene fragments etc. -> 32,540 genes in "filtered gene set"

- Manual inspection of subset of FGS-mostly "real" genes; most gene structures are mostly correct

- Sequencing error rate in *gene space* is estimated to be ~0.00025

Goal: Link Genes to Traits

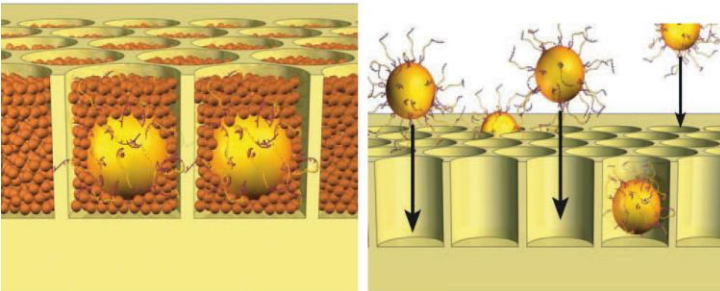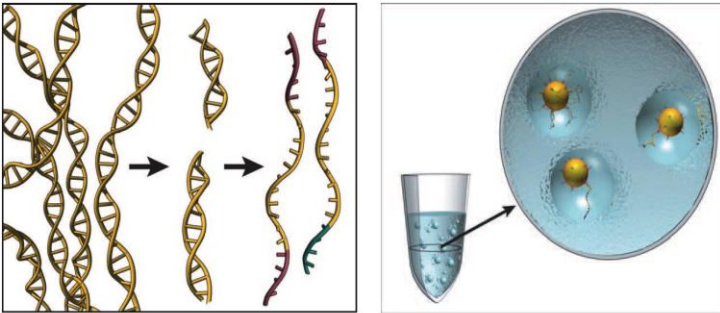Exploit genetic and phenotypic variation:
QTL mapping
Association studies
NAM and "next generation pop'ls"
Phenotyping resources are limiting

SNP and IDP discovery using 454 transcriptome sequencing

- Emrich et al., 2007 Genome Research*
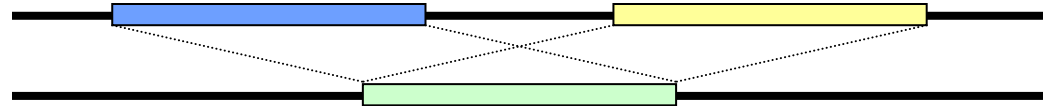- Barbazuk et al., 2007, Plant Journal*
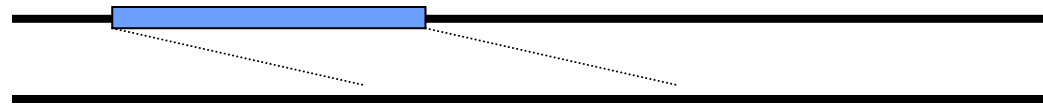
  open access article

Scott Emrich
Notre Dame

Brad Barbazuk
Univ of FL

Structural Variation (CNV & PAV)

CNV

PAV

- In humans SV can be associated with disease ("traits")
- What is overall level of (genic) SV in maize?
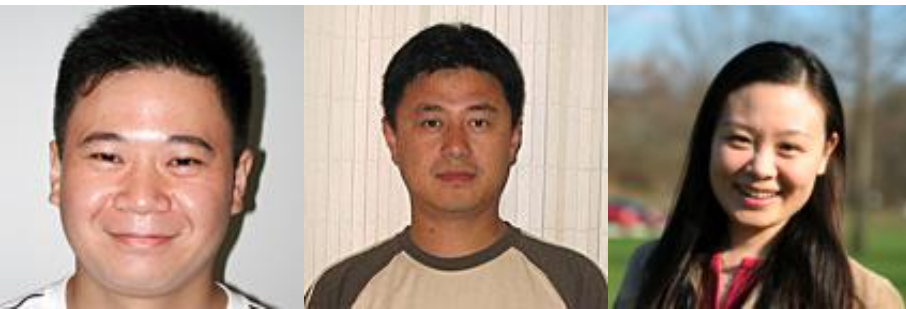- Does SV contribute to phenotypic diversity of maize?

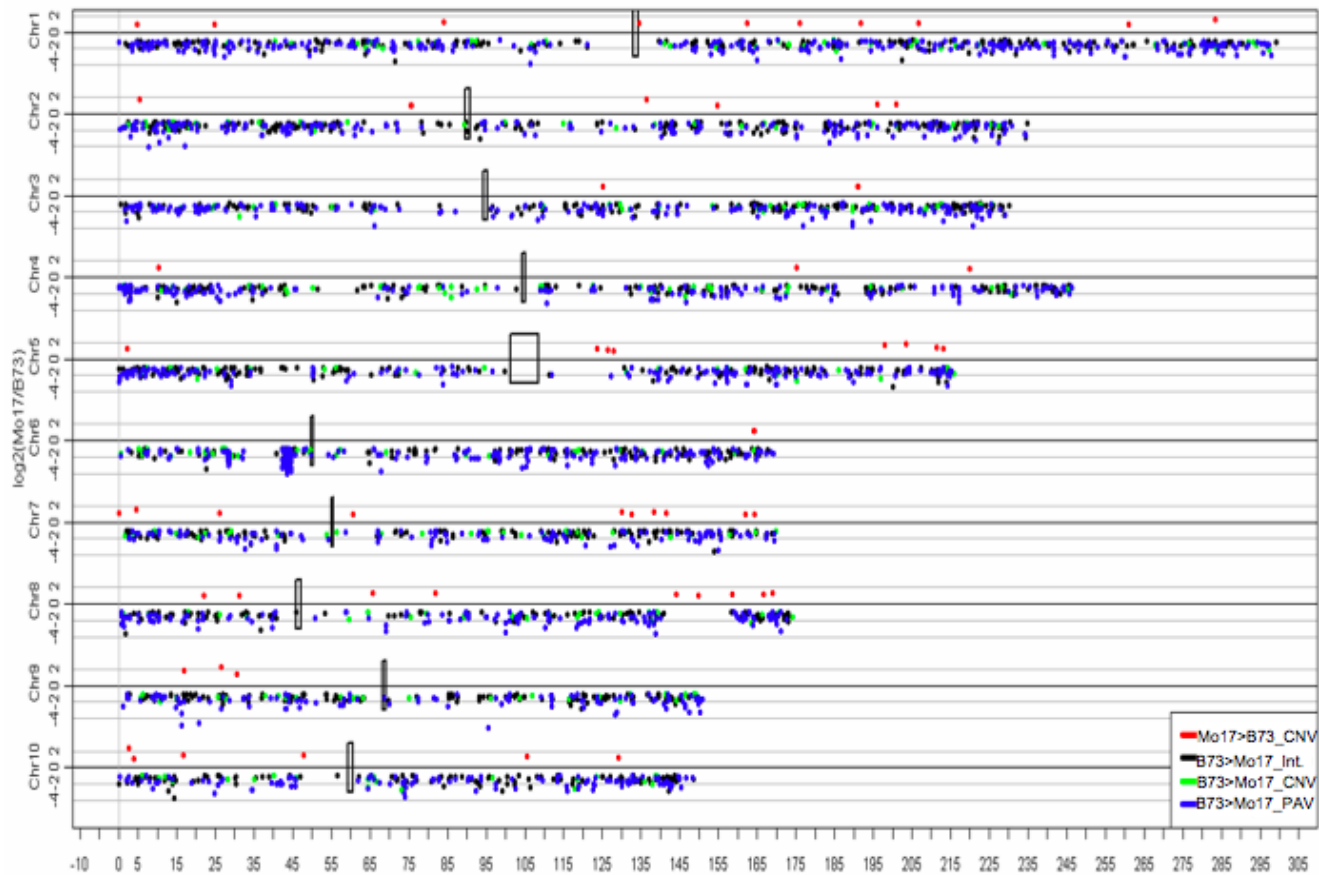Kai Ying   Yan Fu   Wei Wu
(应开)   **(傅延)**   **(吴薇)**
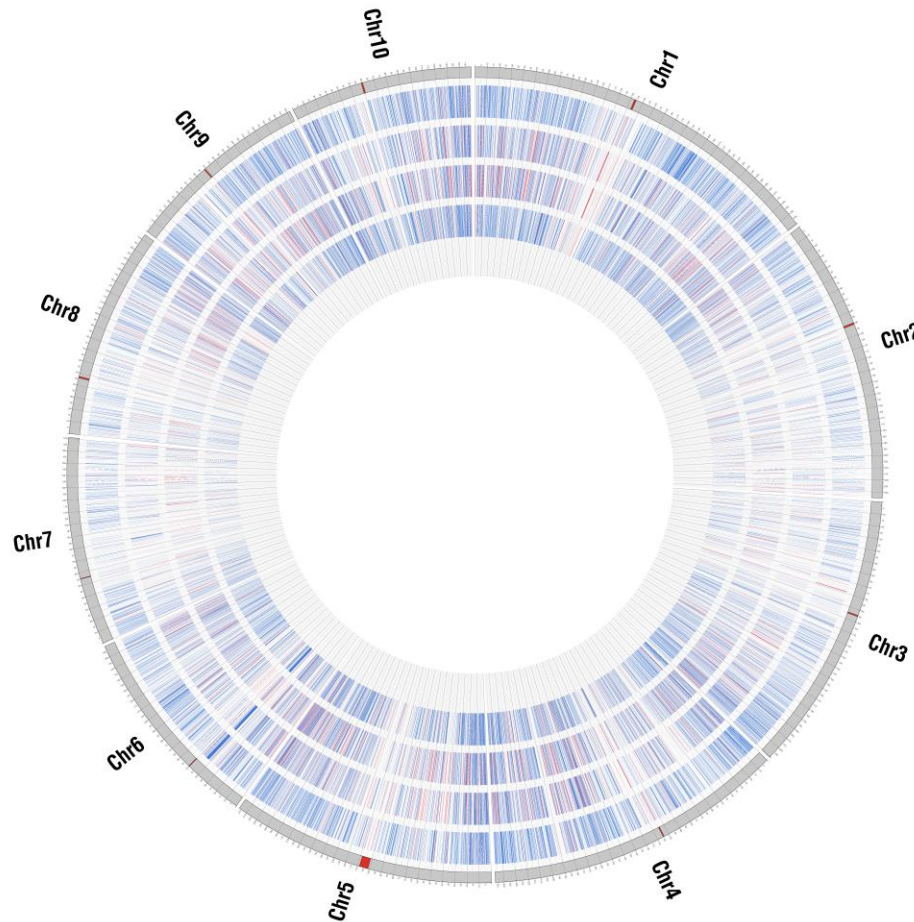
Nathan Springer   Jeff Jeddeloh   Brad Barbazuk

Several hundred *intact,* expressed, phylogenetically conserved genes exhibit CNVs and PAVs

Segmentation Results

CNV and PAV Loss (blue) & CNV Gain (red) Intervals relative to B73

Outer to Inner rings:
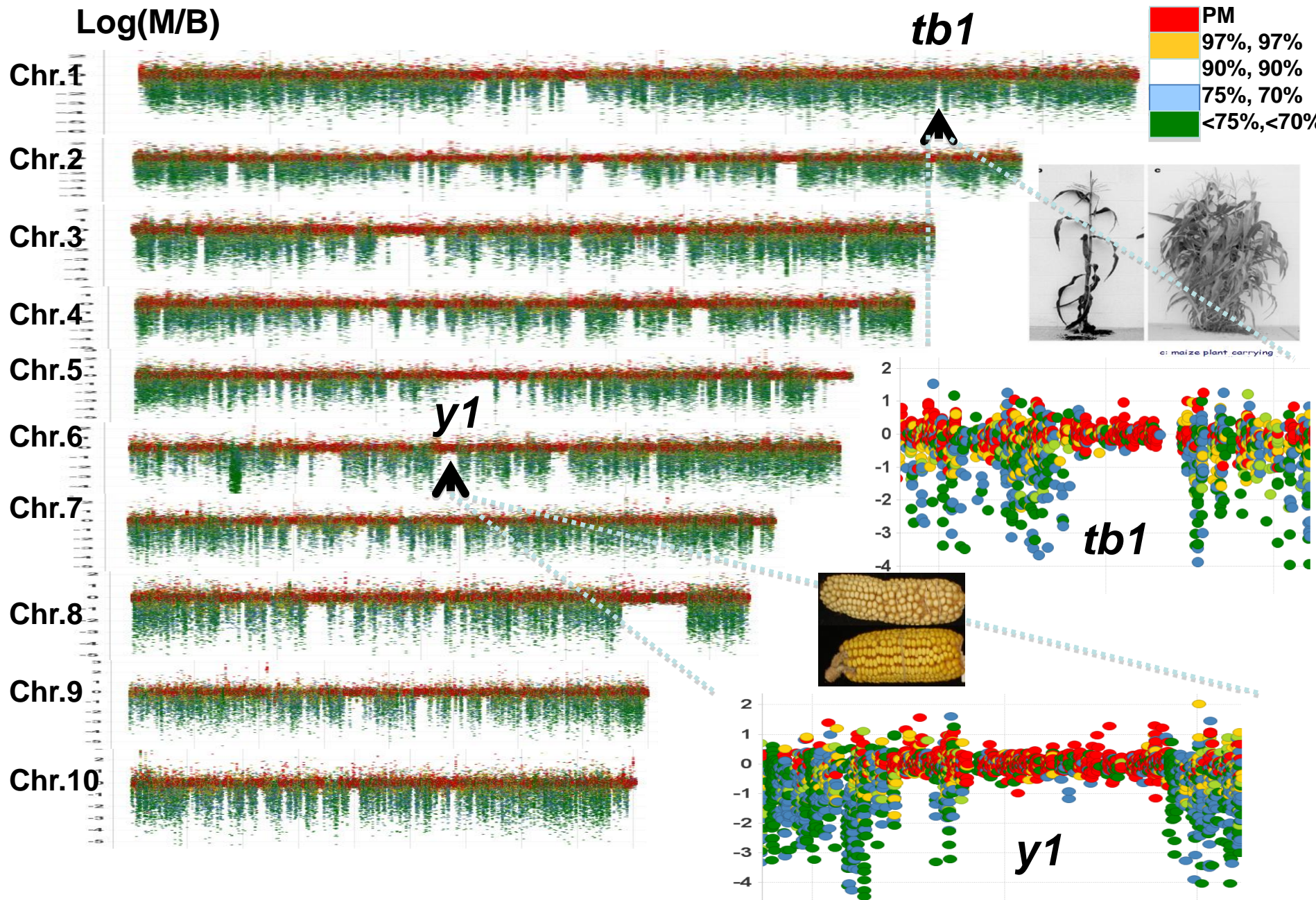Teosinte vs. B73
Tx303 vs. B73
Hp301 vs. B73
Mo17 vs. B73

c: maize plant carrying mutation in tiller branch 1 (tb1)
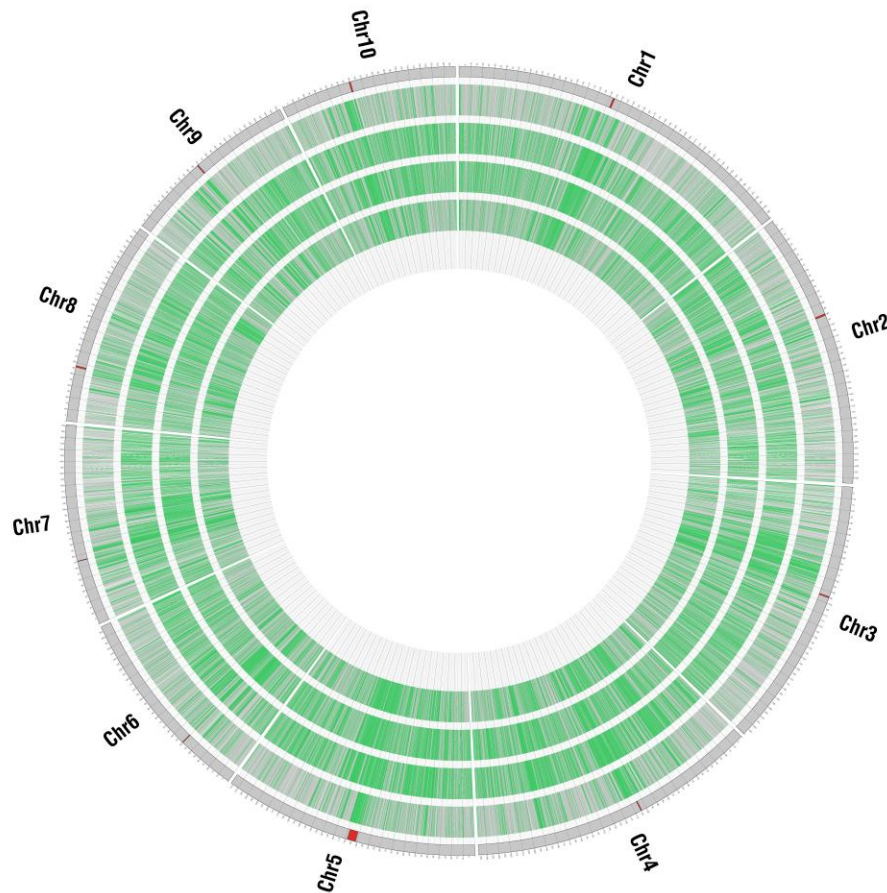
~10,000 YBP

Summary of SV

- Maize haplotypes exhibit unprecedented levels of SV (100s of genic CNV and PAVs) (supported by re-sequencing results; Lai et al., submitted)
- SV may help explain extraordinary level of phenotypic diversity in maize
- SV provides a testable hypothesis for heterosis; could enhance prediction of hybrid performance
- Deletions can actually be favorable "allele"
- Determining rates at which SV arises in maize; can this process explain some of the success of long-term selection experiments?
- Re-sequence of commercial and non-commercial lines to capture full gene inventory of maize?
- Challenge of *de novo* genome assembly and display of diversity data
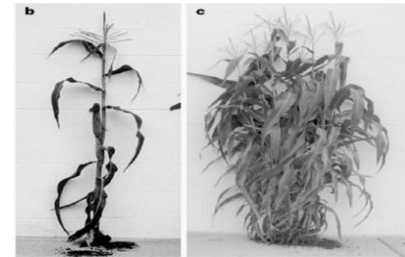
Global distribution of maize variation

Log(M/B)

Chr.1
Chr.2
Chr.3
Chr.4
Chr.5
Chr.6
Chr.7
Chr.8
Chr.9
Chr.10

tb1

y1

PM
97%, 97%
90%, 90%
75%, 70%
<75%,<70%

tb1

y1

Conserved (green) & Variable (grey) Intervals relative to B73

Outer to Inner rings:
Teosinte vs. B73
Tx303 vs. B73
Hp301 vs. B73
Mo17 vs. B73

c: maize plant carrying
mutation in tiller branch 1
(tb1)

~10,000 YBP

Origins of Low Diversity Segments

- Selective Sweeps associated with domestication

  - (~1/2 dozen major loci)

- Selective Sweeps associated with exploitation of major "QTL"

- Genetic Drift (recent IBD)-different heterotic groups?

Summary of Genomic Distribution of Diversity

- Large numbers of putative selective sweeps (low diversity regions) in breeding lines

- Did we fix best *allele*? Best allele for "future agriculture"? (value of allele mining)

- Fixation of QTL *regions* via breeding activities may limit our ability to develop "next generation" crops (argues for precision breeding)

- Consider targeted re-introduction of genetic diversity

Corn Grain Yields
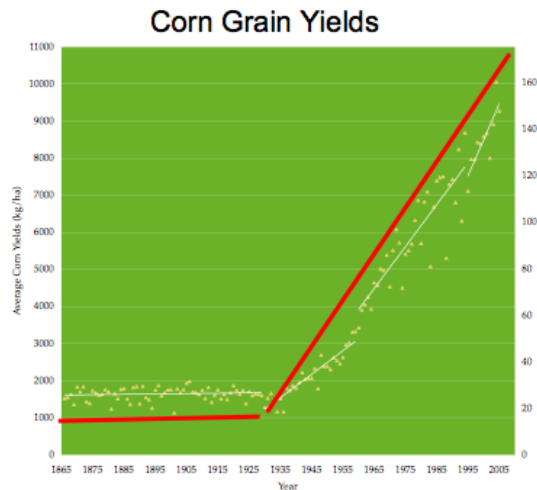
Troyer (2006) Crop Sci. 46:528-543

$$\frac{R\&D\ \$}{Unit\ Genetic\ Gain}$$

Time
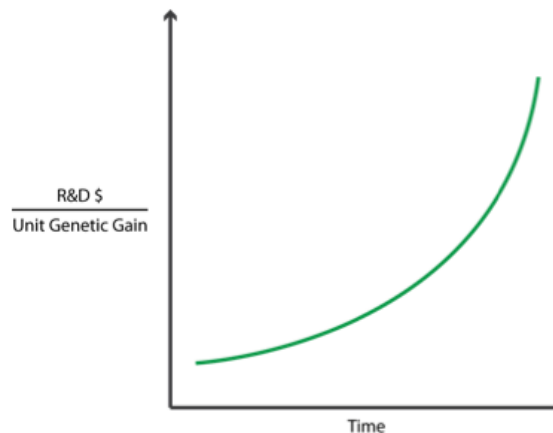
Can maize help mitigate global climate change?

McCarty Glacier, 30 Jul 1909* vs. 11 Aug 2004*

Kenai Fjords National Park, Alaska

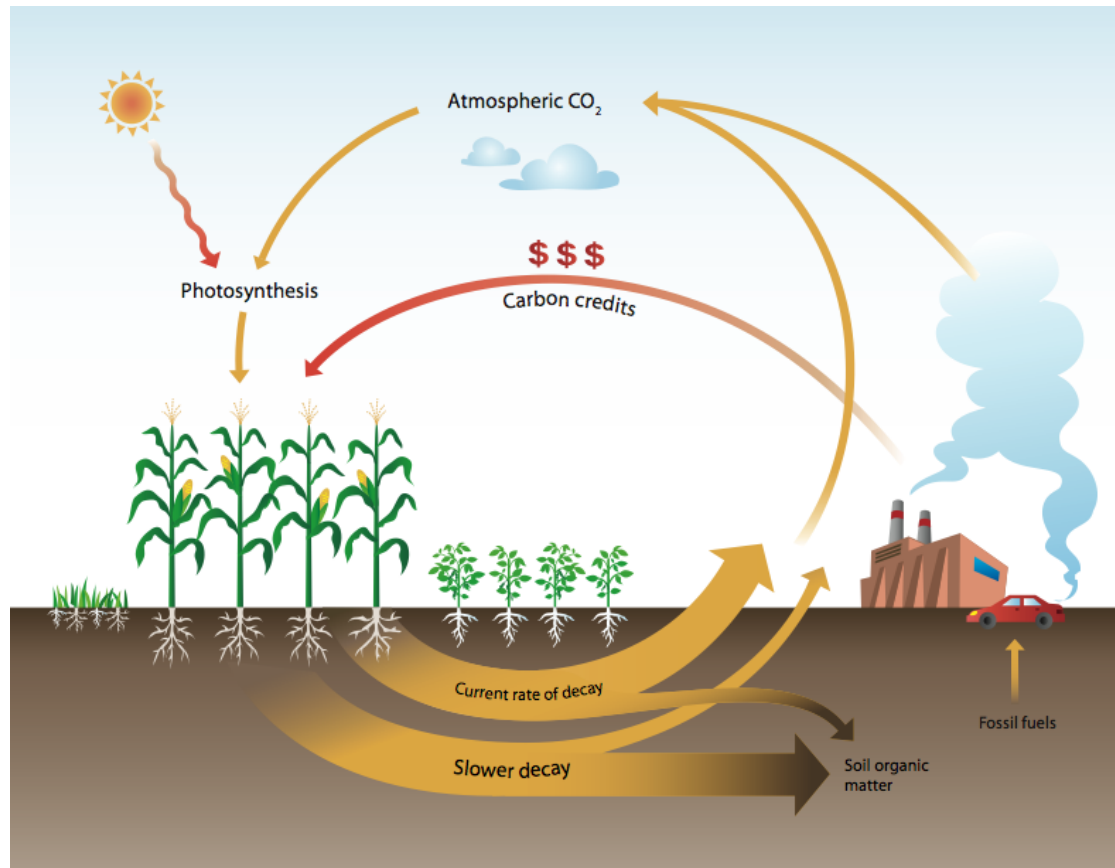# Center for Carbon-Capturing Crops

Goal: to produce crops with biomass that is more "resistant" to microbial degradation

–Alter composition to reduce rate of decay, e.g., increase concentration of compounds that have long half lives in soil.

Decomposition of Biomass-Derived Carbon

lignin-like
30%

soluble
30%

cellulose-like
40%

lignin-like

1 week 1 month 10 years 25 years

Mean residence time

Hadas et al. Soil Biology & Biochemistry 36 (2004): 255-266; Bahri et al. Soil Biology & Biochemistry 40 (2008): 1916-1922.