

Association analysis in a maize diversity panel

Marilyn Warburton, Alain Charcosset, James Gethi, Pichet Grudloyma, Esther Khosa, Wanchen Li, Cosmos Magorokosho, Jose Luis Araus, Yunbi Xu, Jonathan Crouch, Kevin Pixley, Natalia Palacios, Debra Skinner, Luz George, Jianbing Yan, Marianne Banziger, Gary Atlin, Jose Crossa, Ed Buckler, Tim Setter, Maria Salas, Sharon Mitchell, Torbert Rocheford, Cathy Bermudez, Javier Betran, Philippe Monneveux, Jean Marcel Ribaut, Mark Sawkins,

Project Rationale

- ▶ Association Analysis seeks to find statistical associations between a change in a specific piece of DNA and a change in the phenotype of the plant
- ▶ GCP project 13 seeks to find genes improving drought tolerance
- ▶ GCP project 14 includes an association mapping test to find genes improving pro-vitamin A content in maize grains

Activities (DAM)

- ▶ **Germplasm selection:** 350 inbreds (all testcrossed with CML312, 3 precocity groups)
- ▶ **Field trials:** (5 sites, 2 treatments, 2 reps, 2 years, 13 – 15 phenotypes measured)
- ▶ **Metabolite analysis:** (3 tissues, 2 time points, 2 reps, 2 years, 5 metabolites measured on 350 inbreds; 1 year measured on 100 hybrids (tails))
- ▶ **Candidate gene selection:** (102 to date, 3 – 5 SNPs per gene, included in panel of 1536 SNPs)

Results (DAM)

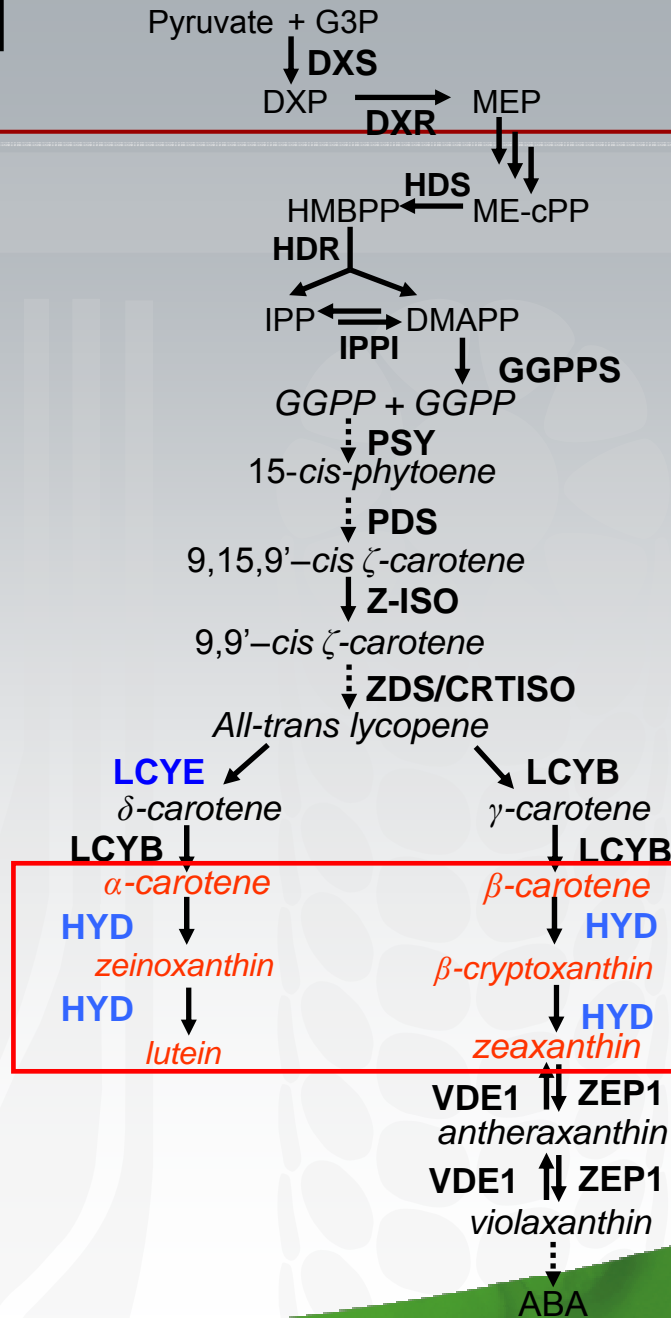
- ▶ Phenotypic data available for the 2006-2007 field season (5 sites, 2 treatments, 2 reps)
 - 1 field site looks excellent! (good correlation between reps, big difference between stressed and non-stressed treatment plants)
 - 2 field sites look good (good correlation between reps, smaller difference between stressed and non-stressed treatment plants, especially early plants)
 - 1 field site looks OK (lower correlation between reps, smaller difference between stressed and non-stressed treatment plants, especially early plants)
 - 1 field site may not be used this year (oops). A learning experience!

Results (DAM)

- ▶ SNP discovery in 102 drought candidate genes completed, and the Drought SNP Array created
 - 1000-3000 bases sequenced per gene (as much coding sequence as possible but also anchored within an UTR to have a greater probability of finding SNPs)
 - 3 – 5 SNPs per drought gene
 - (Illumina format) containing 1536 SNPs (866 “drought” SNPs and 670 “high information”)
- ▶ Metabolic data available from 350 inbred lines, planned for 100 hybrid tails
- ▶ Now... just need to combine data sets!!!

CAM

Pro-Vitamin A Candidate Gene Tested in a diverse panel of inbred lines

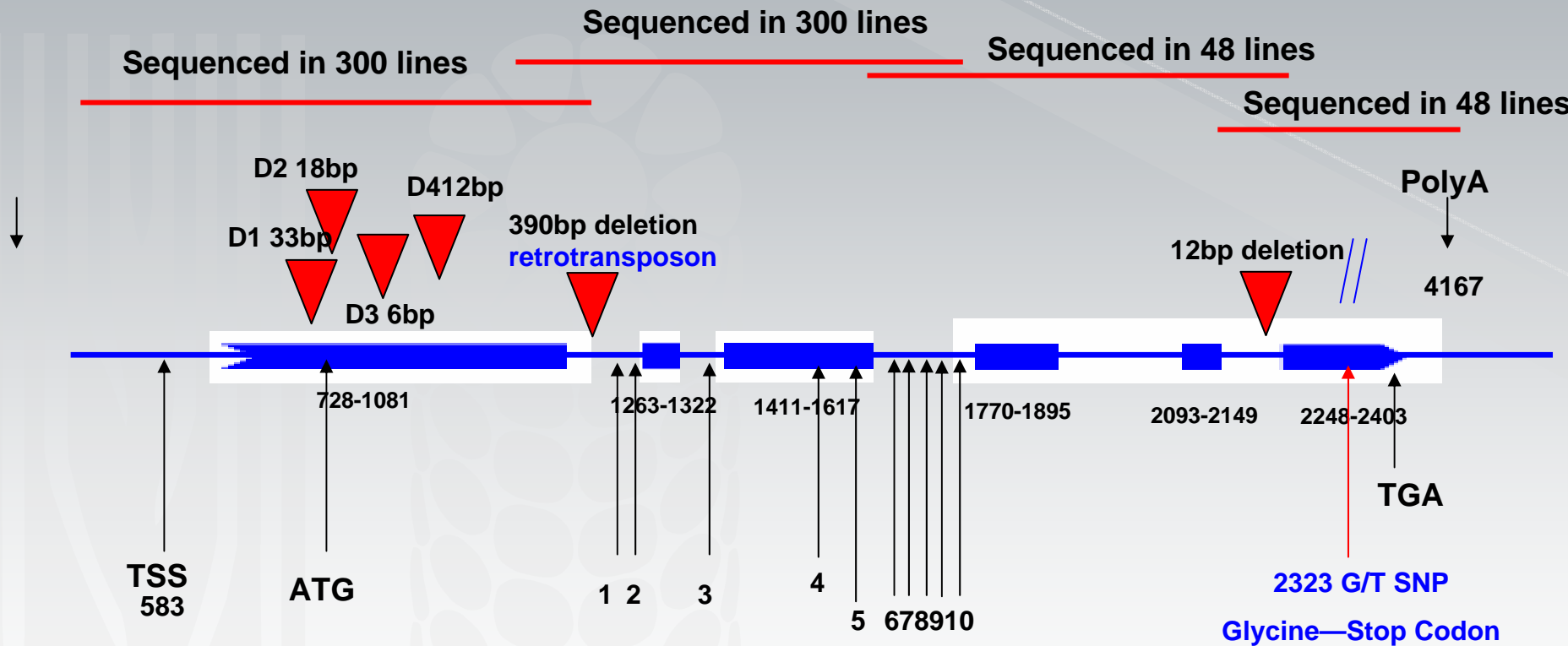


- 1) A key gene related to Pro-VA synthesis
- 2) Possible drought related gene involved in ABA pathway

Results (CAM)

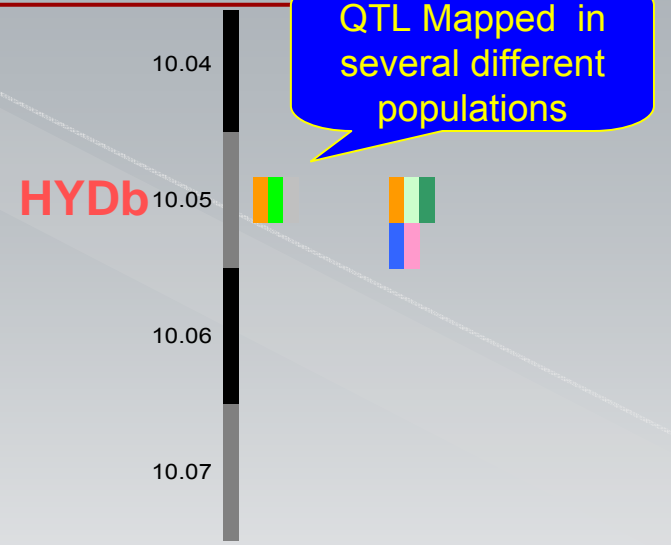
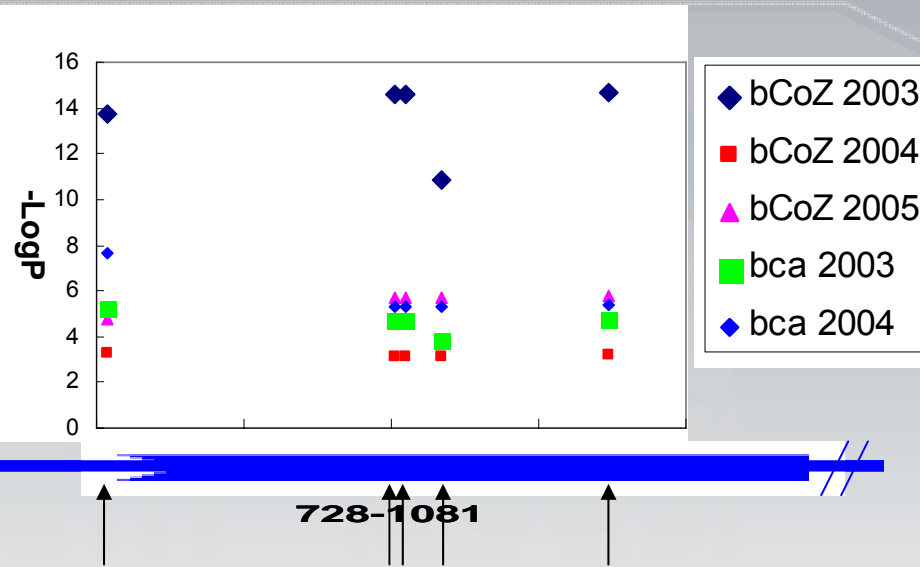
HYDb gene structure

661-3052 has been sequenced



5 major polymorphisms found in Promoter and 1st exon region;
one SNP causing premature stop codon at 2323

CAM Association Results



bCoZ=beta carotene over zeaxanthin
bca=beta carotene

- Five major SNPs/Indels associated with phenotype strongly had been found in promoter and 1st exon region
- Mutated HYDb gene loses the function that may stop/reduce the synthesis of zeaxanthin and increase the content of beta carotene
- PCR markers now available for MAS

Data Format (what analyzed by):	Release Date:
ID, pedigree, origin, etc of drought and proA association mapping sets in excel	September 2007
Genetic relationship data analyzed in ICIS and Structure	December 2007
SNPs in the 1536 drought SNP chip and in the proA associated genes	December 2007
SNP data analyzed by Illumina proprietary software (csv)	December 2007
Phenotypic Data analyzed in SAS and Excel	December 2007 and June 2008
Association data currently analyzed in Tassel and SAS	June 2008

Links with other projects

- ▶ Drought association mapping project linked to association mapping efforts underway at Cornell University, GCP ADOC (*A dataset on allele diversity at orthologous candidate genes in GCP crops*), and GCP15, (*Determination of a common genetic basis for tissue growth under water-limited conditions across plant organs and genomes*).
- ▶ Carotenoid association mapping project linked to association mapping efforts underway at Cornell University, Harvest Plus projects ongoing at CIMMYT and the University of Illinois, a USAID linkage grant on allele mining and ongoing QTL mapping of proA at China Agriculture University.

Product delivery/impacts on users

- ▶ List of candidate genes for drought tolerance
- ▶ Drought SNP array (Illumina format): 3 – 5 SNPs per gene, 866 “drought” SNPs and 670 “high information” SNPs on array
- ▶ Markers that explain over 80% of the phenotypic variation in the germplasm sets studied to date for provitamin A including PCR based assays
- ▶ Haplotypes of inbred lines from CIMMYT and HarvestPlus for provitamin A for planning crosses

Thank you