Population genomic and genome-wide association studies of agroclimatic traits in sorghum

Geoffrey Morris, Univ. of South Carolina

NSF-BREAD Sorghum Project Team

Population genomic and genome-wide association studies of agroclimatic traits in sorghum

Geoffrey P. Morris1,*, Purna Rama1,*, Anjali P. Deshpande1,*, C. Thomas Hoit1, Tushar Shi1, Heidi B. Upadhyay2, Owen Para Lasso2,*, Rachel J. Brown3, Christine R. Ackey2, Bryce S. Mitchell2, James Wrenn2,*, Jeffrey C. Glazebrook1,*, Edward S. Budihardjo1,*, and Stephen Rounsaville2*

1Department of Plant Science, University of California, Davis. 2. NSF-BREAD Sorghum Project. 3. USDA-ARS. *Corresponding author. *Present affiliation: Florida A&M University, Tallahassee, Florida. *Present affiliation: Florida A&M University, Tallahassee, Florida.

A broader question on GWAS strategy

• “The dramatic increase of sequencing capacity… would eventually make it feasible to conduct routine GWAS…” (Yu and Garrick)
• What will “routine GWAS” look like?
• How much population genomics needs to be integrated into each GWAS?

“Genotype once, phenotype often”

Two case studies of GWAS in sorghum

• The genome-wide SNP map:
  - 265,000 SNPs in ~1000 accessions (330 for GWAS)
  - LD decay within 10-200 kb, mean SNP density of 1 SNP/2.7kb
  - Highly heterogenous landscape of genome diversity
• Two traits:
  1. Plant architecture: inflorescence branch length (Morris et al. 2013)
  2. Grain quality: Presence of condensed tannins (in prep.)

Branch length as agroclimatic trait

• Dense panicles allow higher yields, but open panicles reduce losses under humid conditions
• Highly-associated with population structure (geographic origin and botanical races), so challenging case for GWAS
• Genetics of inflorescence branch length has been studied in model grasses
GWAS on panicle branch length

Conclusions: Branch length GWAS

- Assuming the goal of GWAS is generate candidate genes worthy of further study...
- Identify candidate genes including homologs of known floral regulators
- Able to detect associations with “rare” alleles
- Global panel with mixed linear models seem to be effective

Mapping grain tannin in sorghum

- Flavonoid pathway is among the best understood in all of plant biology, and well-conserved among plants
- Two classic loci (B₁ and B₂) are found on Chr2 and Chr4
- Tannin1 (B₂) was recently cloned (Wu et al. 2012)
- About the trait:
  - Tannins are secondary metabolites involved in plant defense and seed dormancy
  - Grain tannins impart bitterness and astringency, and act as anti-nutrient
  - Widely segregating in many populations
  - Mapped as binary trait (tannin vs. nontannin, data from Wu et al. 2012)

GWAS on grain tannin

Parallel evolution may lead to synthetic associations

- Evidence of multiple haplotypes carrying functional polymorphisms for the same gene in sorghum
  - Shattering1 (Lin et al. 2012)
  - Tannin1 (Wu et al. 2012)
  - Maturity1 (Murphy et al. 2011)

Close-up on Tannin1 locus

- MLMs do not increase the significance of tan1-a allele relative to nearby peaks

Orozco et al. 2010/Lin et al. 2012
Conclusions: Tannin GWAS

- Parallel evolution, which aids comparative genomics, is a double-edged sword for association genomics
  - ✔ Candidates from model plants are very informative
  - ✔ Allelic heterogeneity at key loci likely causes synthetic associations
- Global association panels may only yield Mb-level resolution due to synthetic associations

Routine GWAS?

- Use global panel for:
  - First pass mapping of QTL
  - Characterizing trait distributions relative to population structure
- Design custom regional mapping panels based on trait distributions
  - Genotyping underway for 1600 additional sorghum landraces

Thank you!

- More on Sorghum BREAD: