Exploring the genetics of adaptive traits in cotton with a next-generation platform

Michael Gore, Pedro Andrade-Sanchez,
David D. Fang, Andy French, Doug Hunsaker,
Jesse Poland, and Jeff White

October 11, 2012
ICG1 Research Conference

Rationale

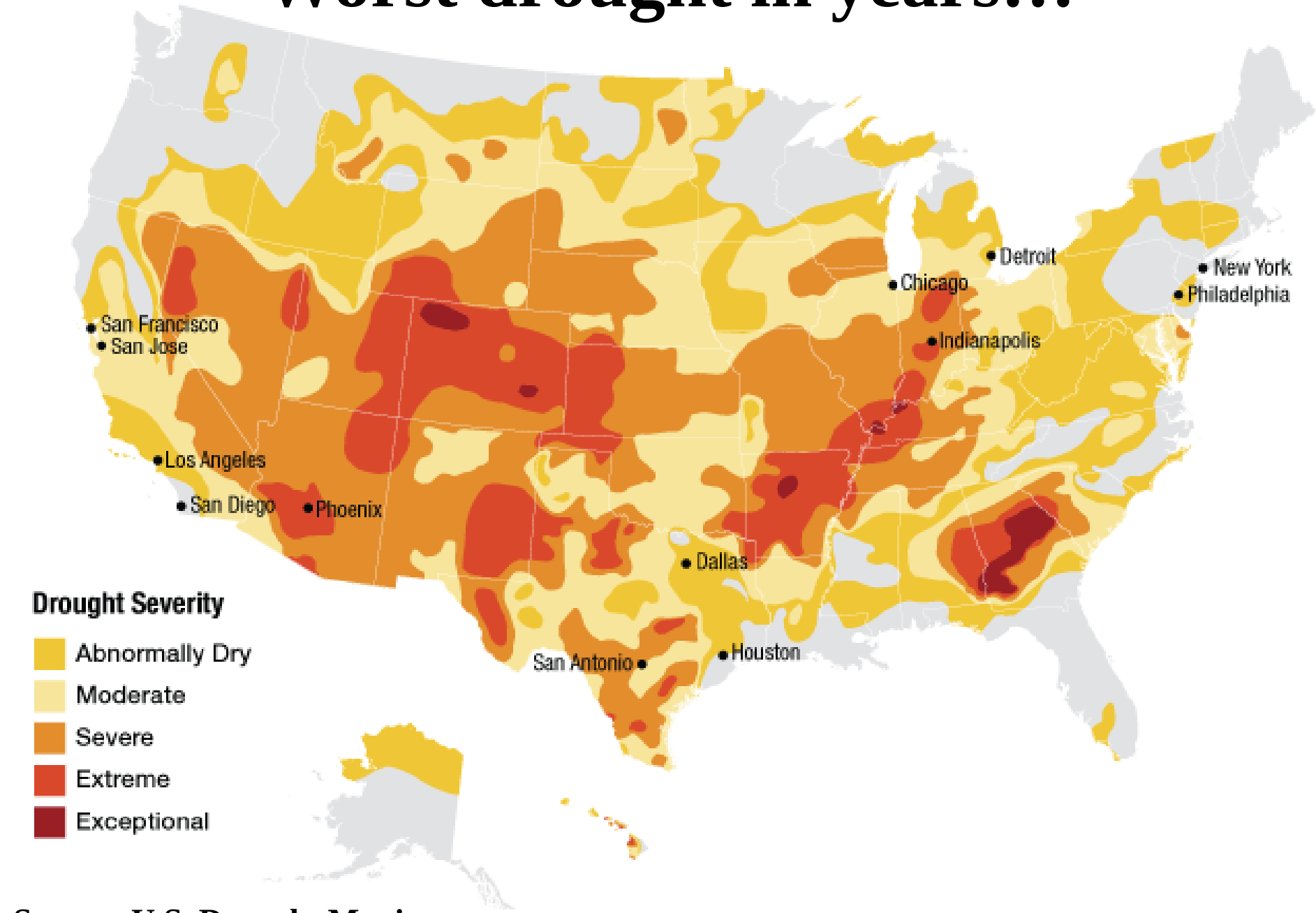
- Challenge to increase yield in the face of climate change and diminishing water resources

- Genetic improvement via modern plant breeding is the most sustainable and economic approach to meet this challenge

- Development of superior heat and drought tolerant cultivars has been slow and difficult

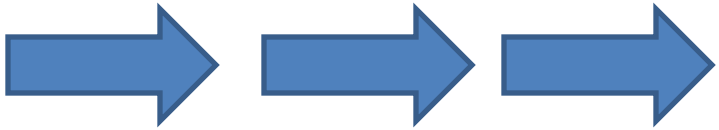
- Breeding progress could be improved by development of new ways to connect **phenotype** to **genotype**

“Worst drought in years…”



Source: U.S. Drought Monitor

How do we connect genotype to phenotype for complex adaptive traits?



Research Objectives

- To develop a genotyping-by-sequencing (**GBS**) approach for multiplex genotyping of a cotton RIL population

- To develop a field-based, proximal remote sensing approach for high-throughput phenotyping (**HTP**) of adaptive traits in a cotton RIL population

Genotyping-By-Sequencing

"…massively parallel sequencing of multiplex reduced-representation genomic libraries."

"massively parallel sequencing" = sequencing on Illumina HiSeq platform

"multiplex" = using DNA barcode (unique 5-10bp)

- DNA sequence synthesized on the adapter
- pool 48-384 samples together

"reduced-representation" = use restriction enzyme to capture only the low-copy portion of the genome flanking restriction sites

- methylation-sensitive restriction enzymes

Elshire et al. 2011 *PLoS ONE* 6(5): e19379

GBS: Sample Genotyping Costs

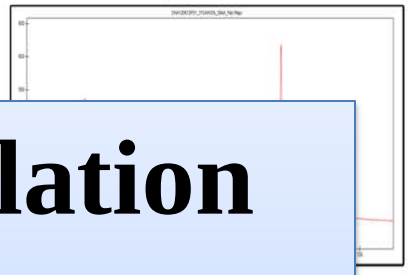
J. Poland

1. Quantify Genomic DNA

2. Normalize
3. Add Restriction Enzymes

5. Pool Samples

6. PCR
7. Clean and Size Library



TM-1×NM24016 RIL population

TM-1: genetic standard for *G. hirsutum*

NM24016: an elite *G. hirsutum* line with considerable (1/3) but stable introgression from several *G. barbadense* lines

HA

"HA

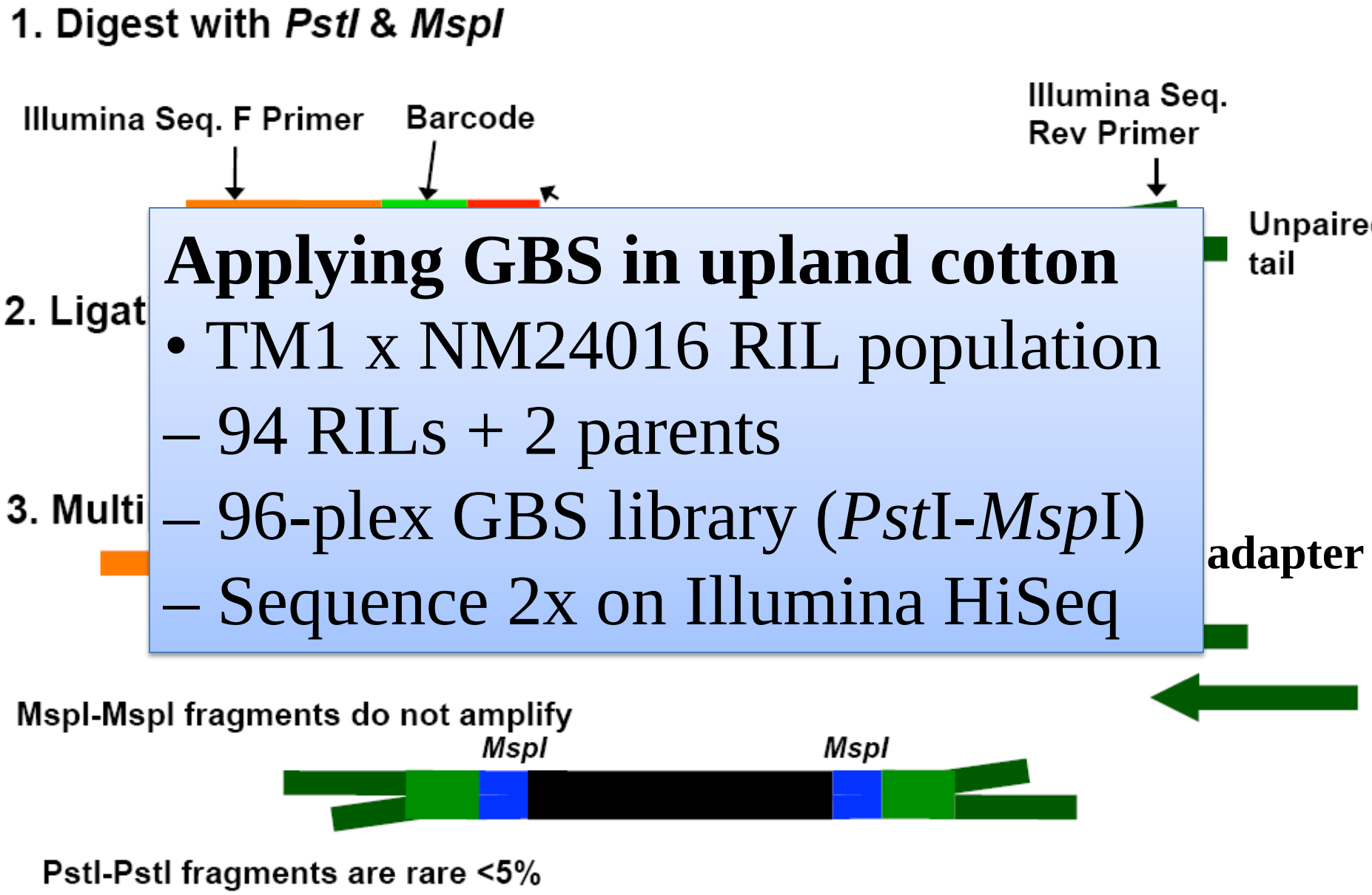
(Internal)

Filter SNPs

erence

J. Poland

GBS: Library Construction



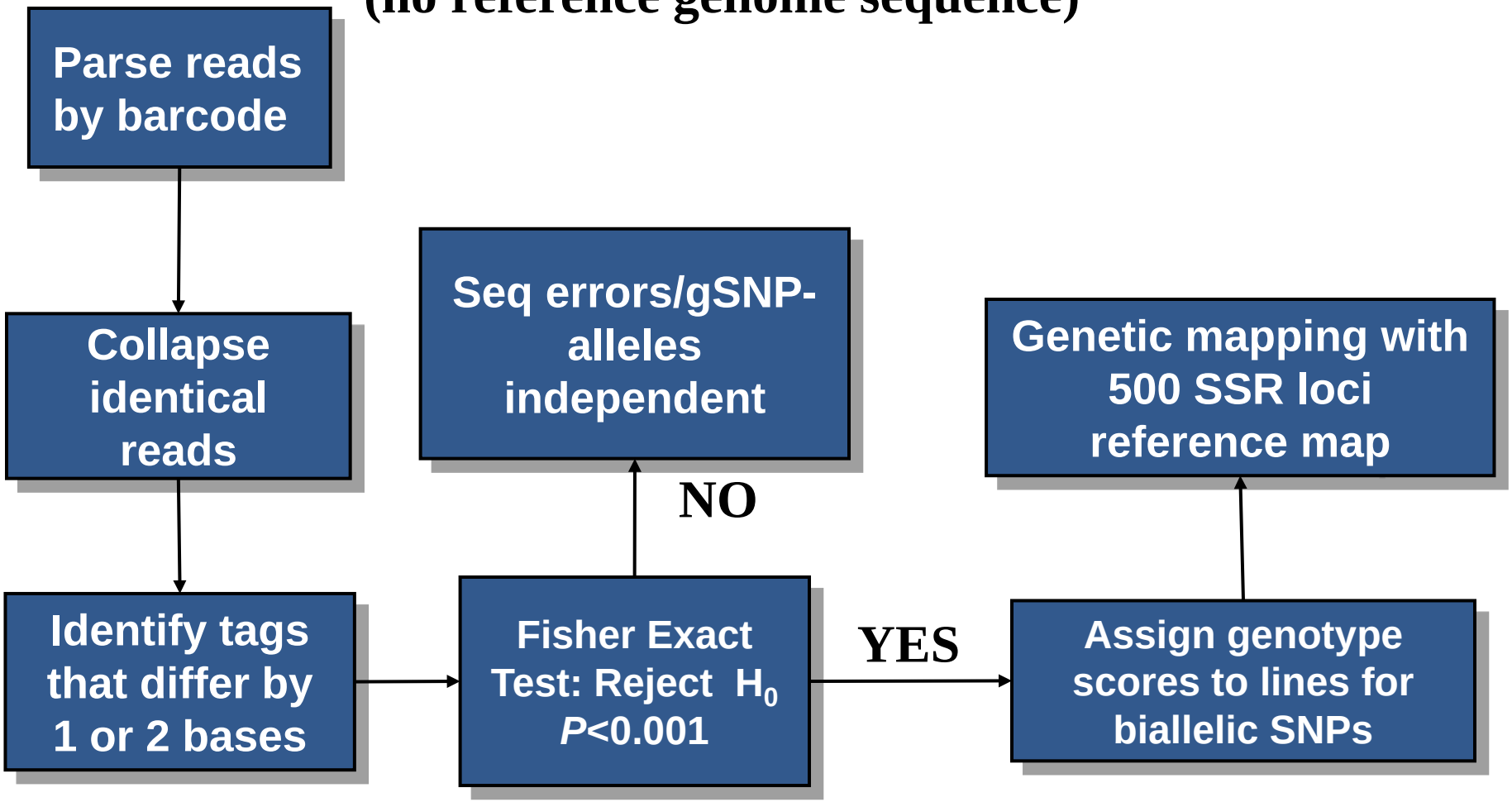
Applying GBS in upland cotton

- TM1 x NM24016 RIL population
  - 94 RILs + 2 parents
  - 96-plex GBS library (*Pst*I-*Msp*I)
  - Sequence 2x on Illumina HiSeq

Poland et al. 2012 *PLoS ONE* 7(2): e32253

GBS: Bioinformatics Pipeline

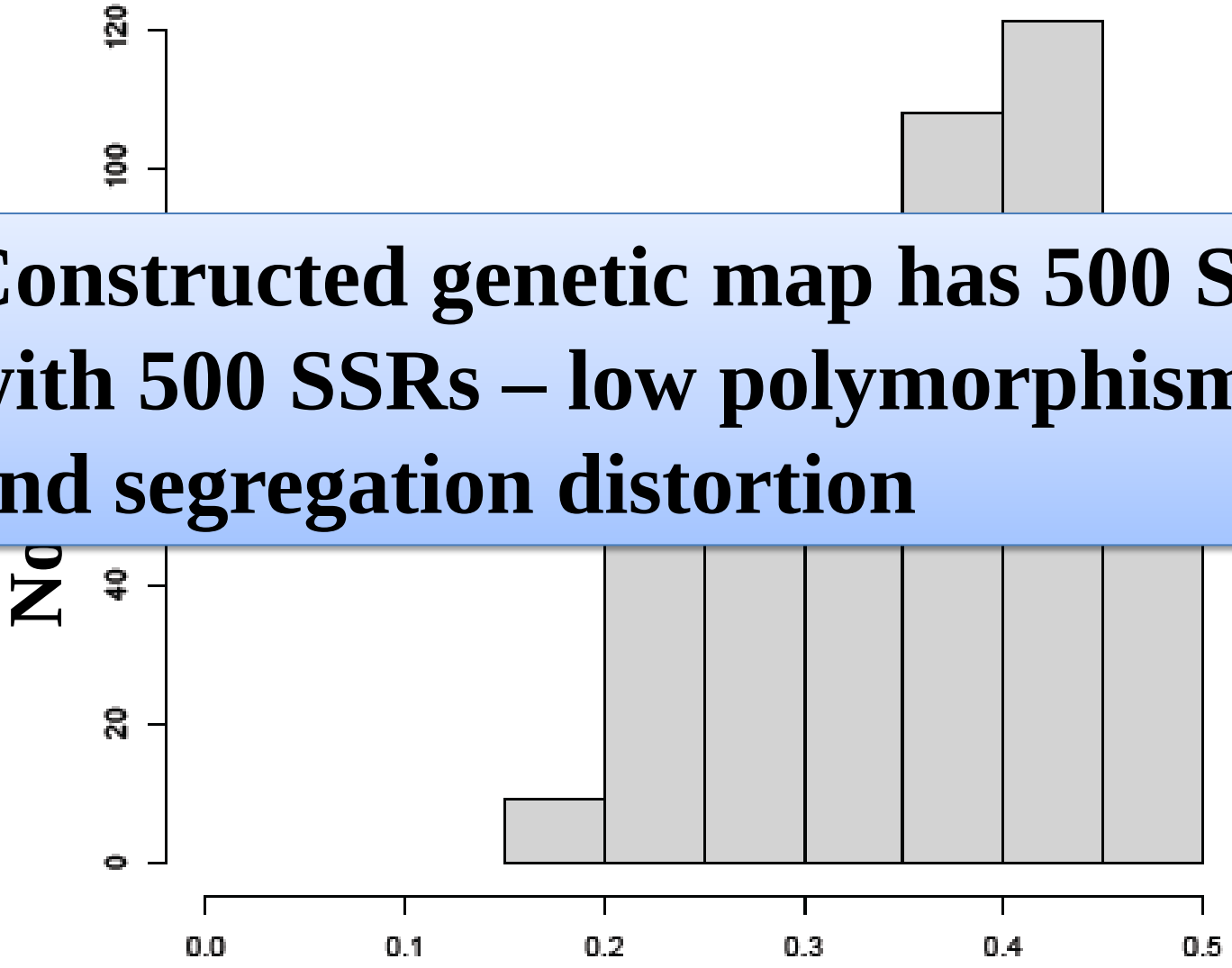
(no reference genome sequence)



Poland et al. 2012 *PLoS ONE* 7(2): e32253

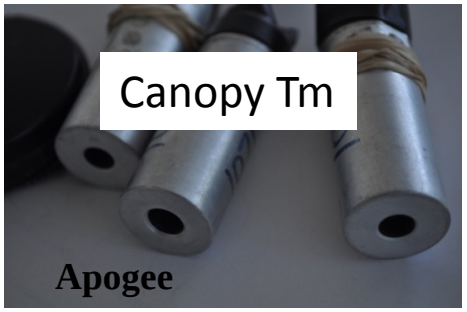
www.maizegenetics.net/tassel

GBS: SNP Calls for RILs



Constructed genetic map has 500 SNPs with 500 SSRs – low polymorphism rate and segregation distortion

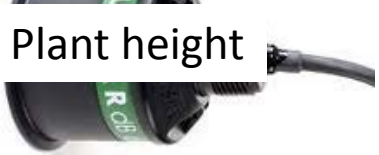
HTP: Sensors, Platform, and Vehicle



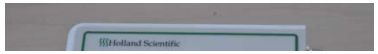
Infrared thermometer

Pulsar db 3-m

Plant height

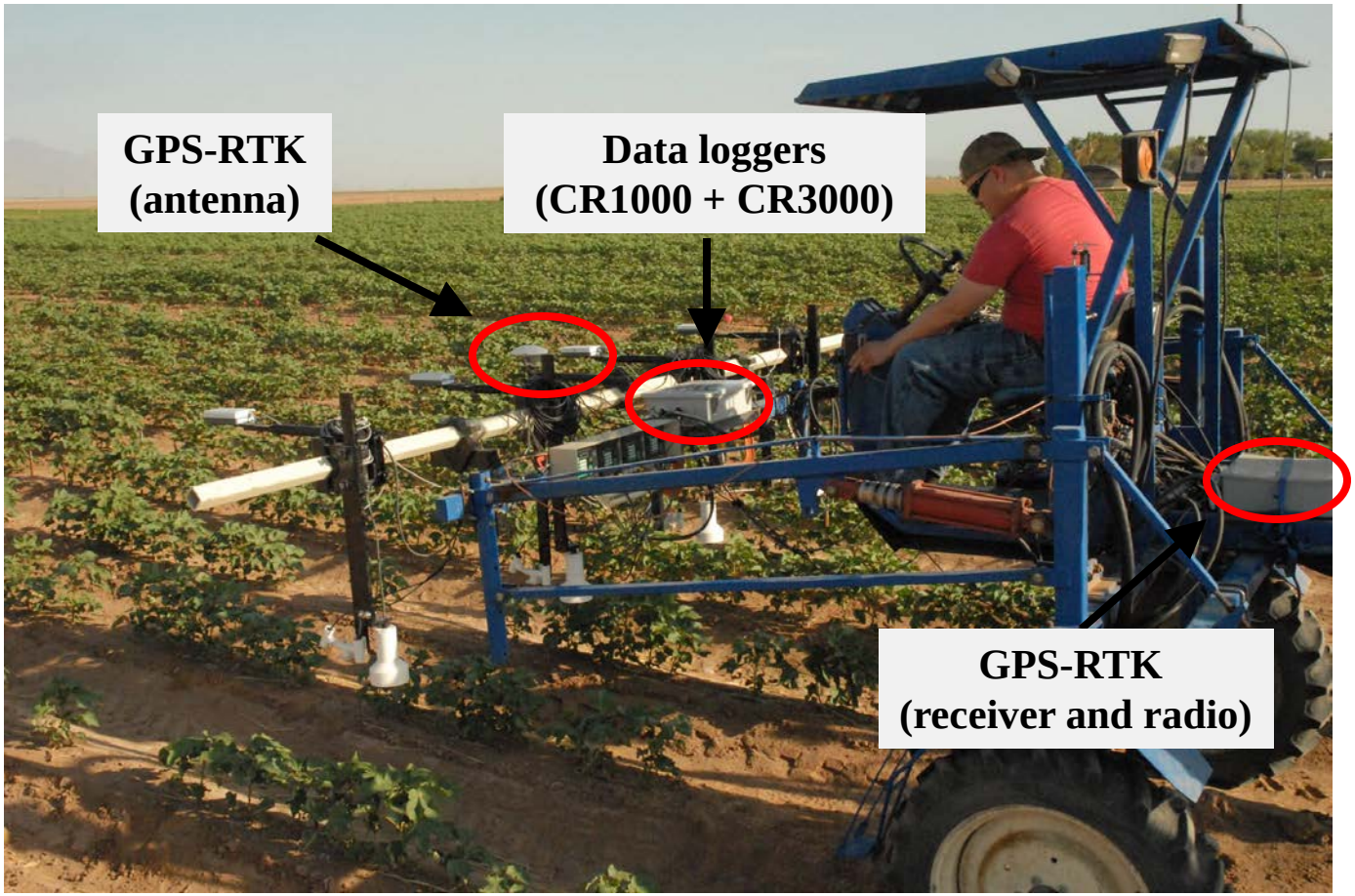


Ultrasonic Transducer



Vegetation Indices

Multi-spectral crop canopy sensor

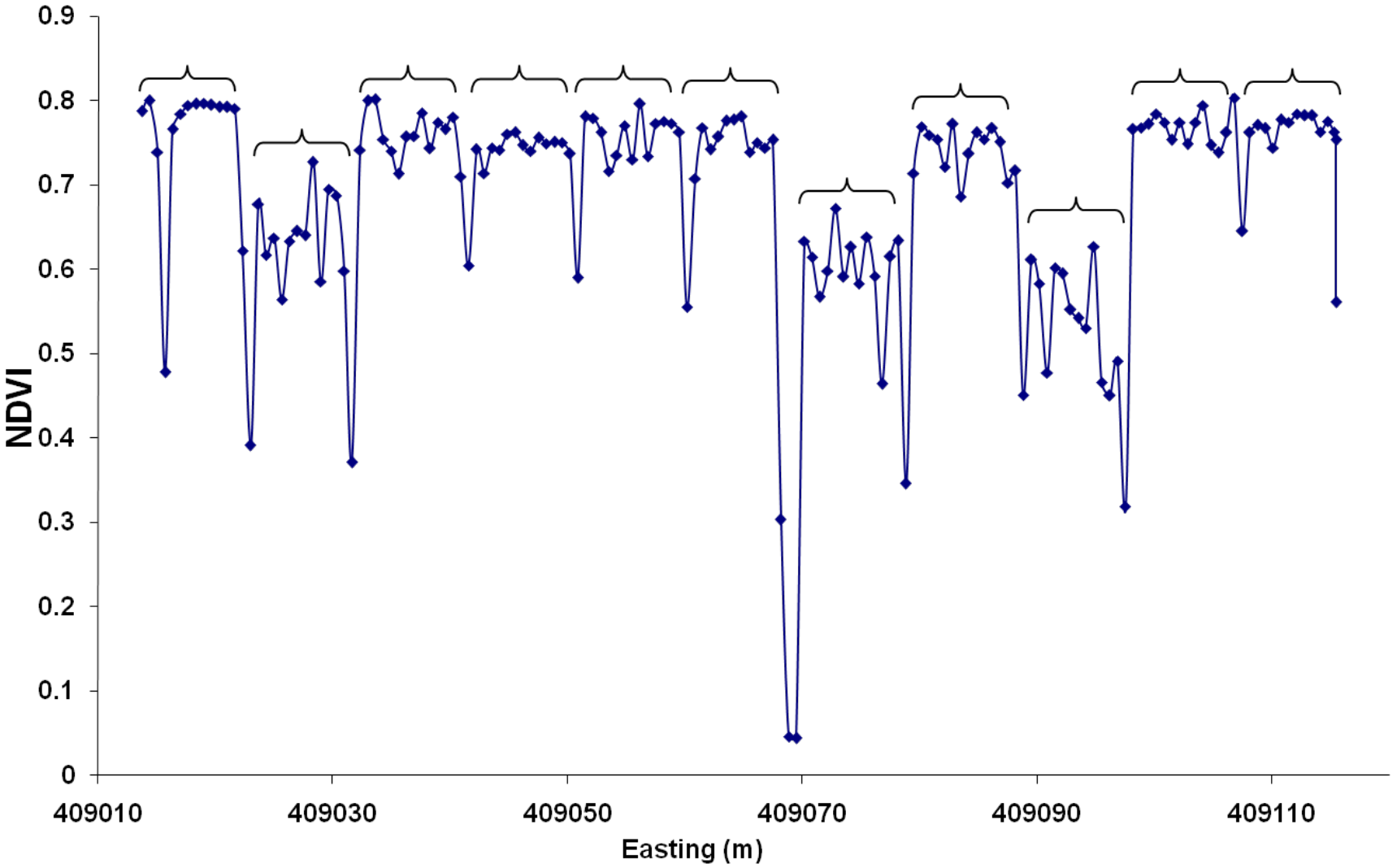


High-clearance tractor

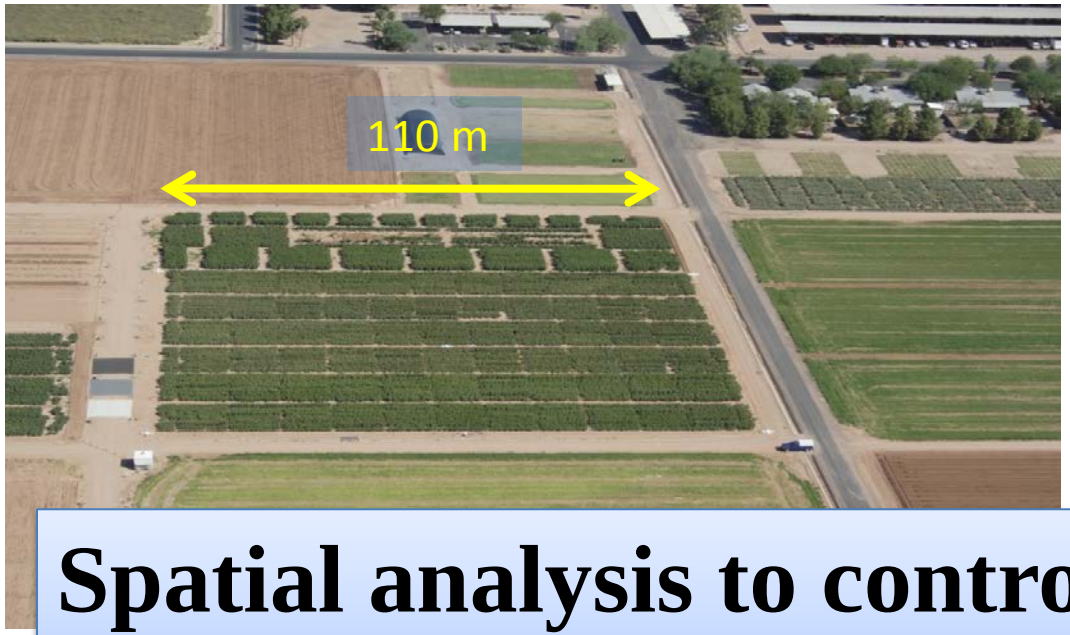
Average speed of 2.82 km/h

1 data point/meter (1 Hz)

HTP: Geoprocessing



HTP: Canopy Temperature



Central Arizona: clear skies, very limited rain, high temperatures

Population: TM-1xNM24016 of 94 RILs (*Gossypium hirsutum* L.)

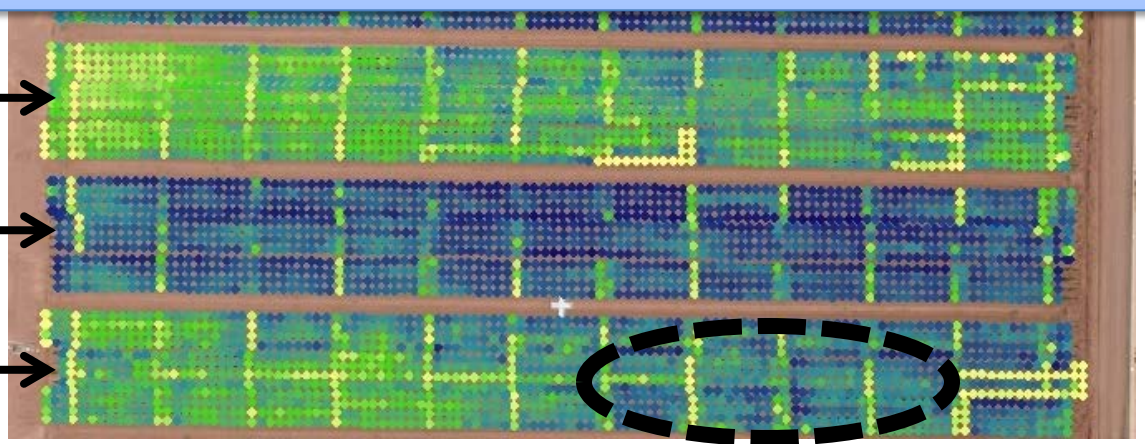
Treatment: 100 and 50% ET (2 reps) drip irrigation

Spatial analysis to control for soil variation when calculating BLUPs

Dry Rep2: 33-49ºC

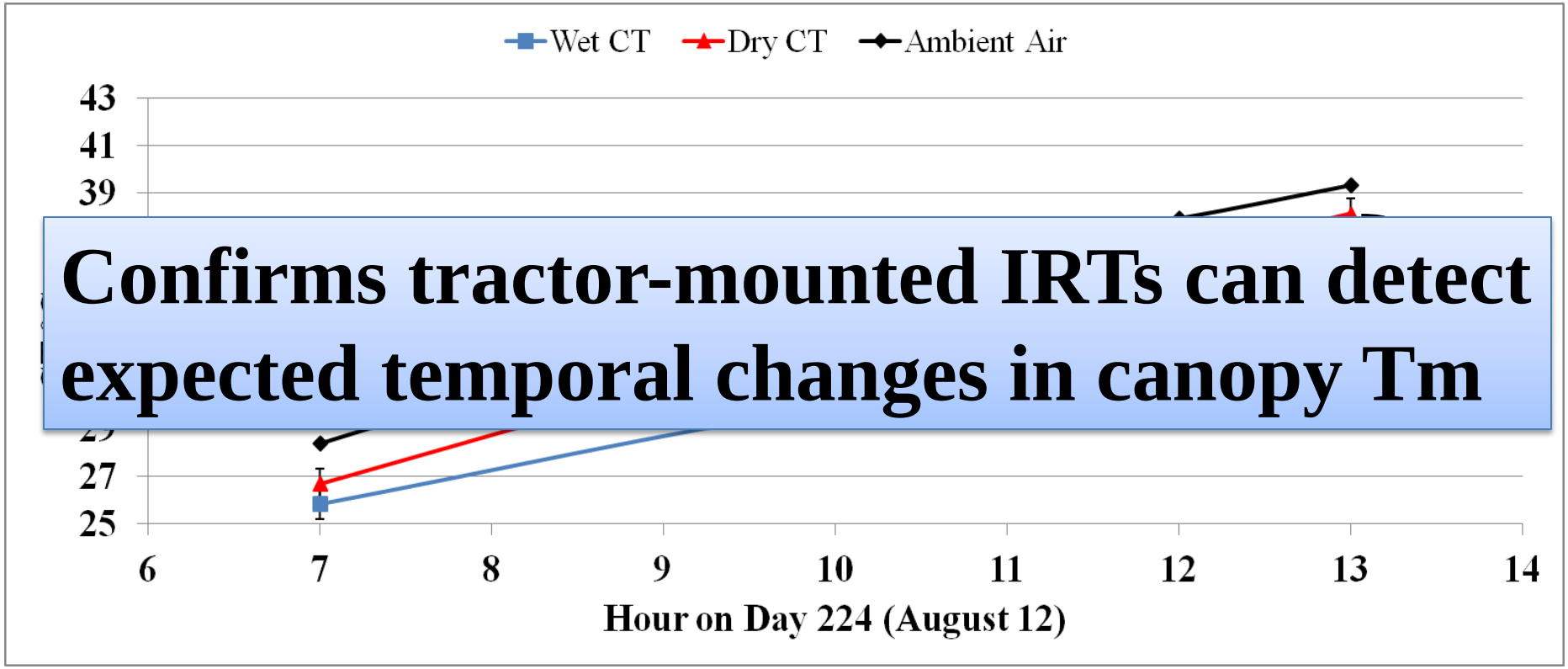
Wet Rep1: 29-36ºC

Dry Rep1: 31-46ºC



1pm on day 224 (12-Aug)

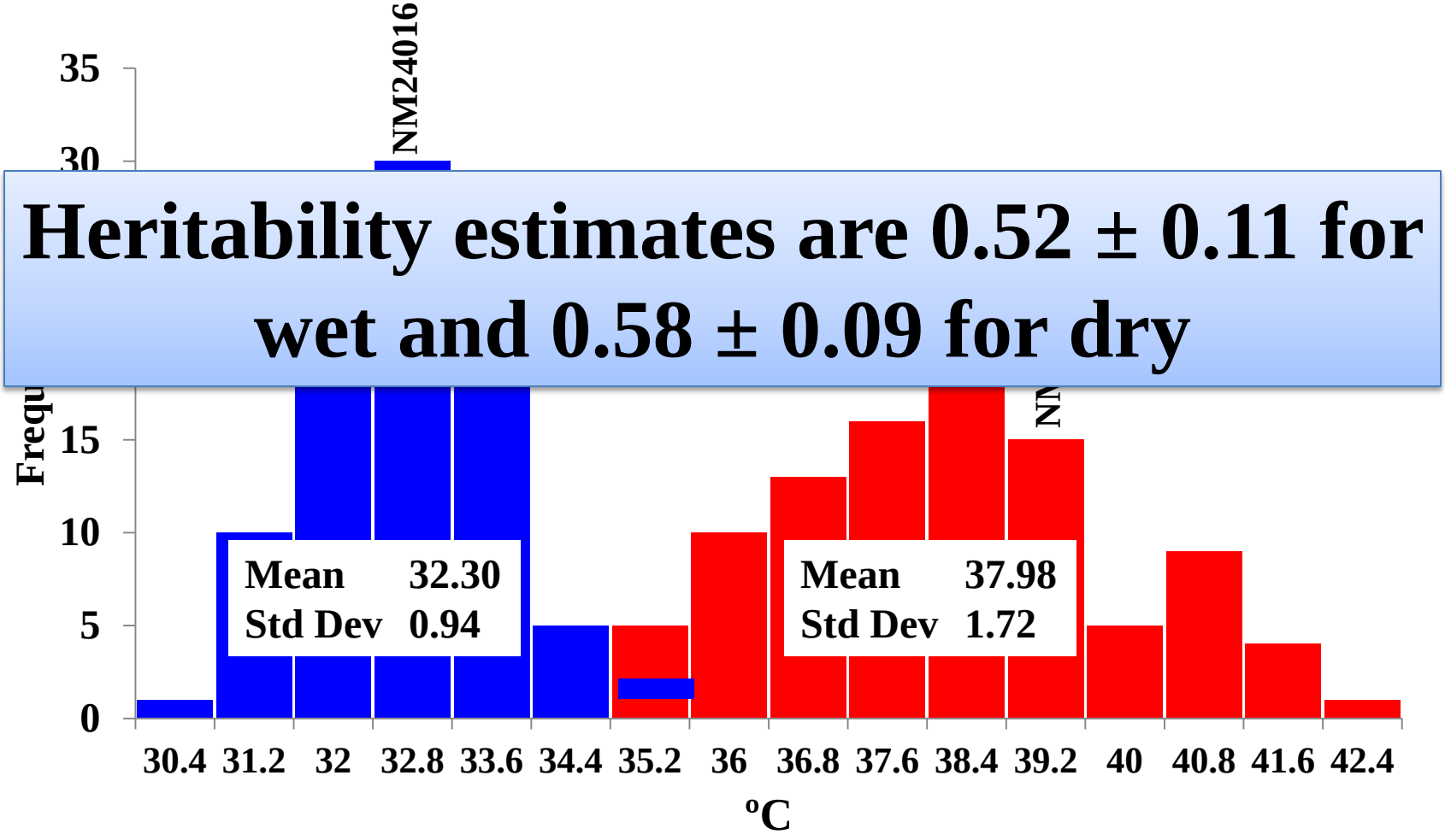
Significant Time-by-Treatment Interaction for Canopy Temperature



Confirms tractor-mounted IRTs can detect expected temporal changes in canopy Tm

Treatment	***P*<0.05**
Time	***P*<0.0001**
Treatment*Time	***P*<0.0001**

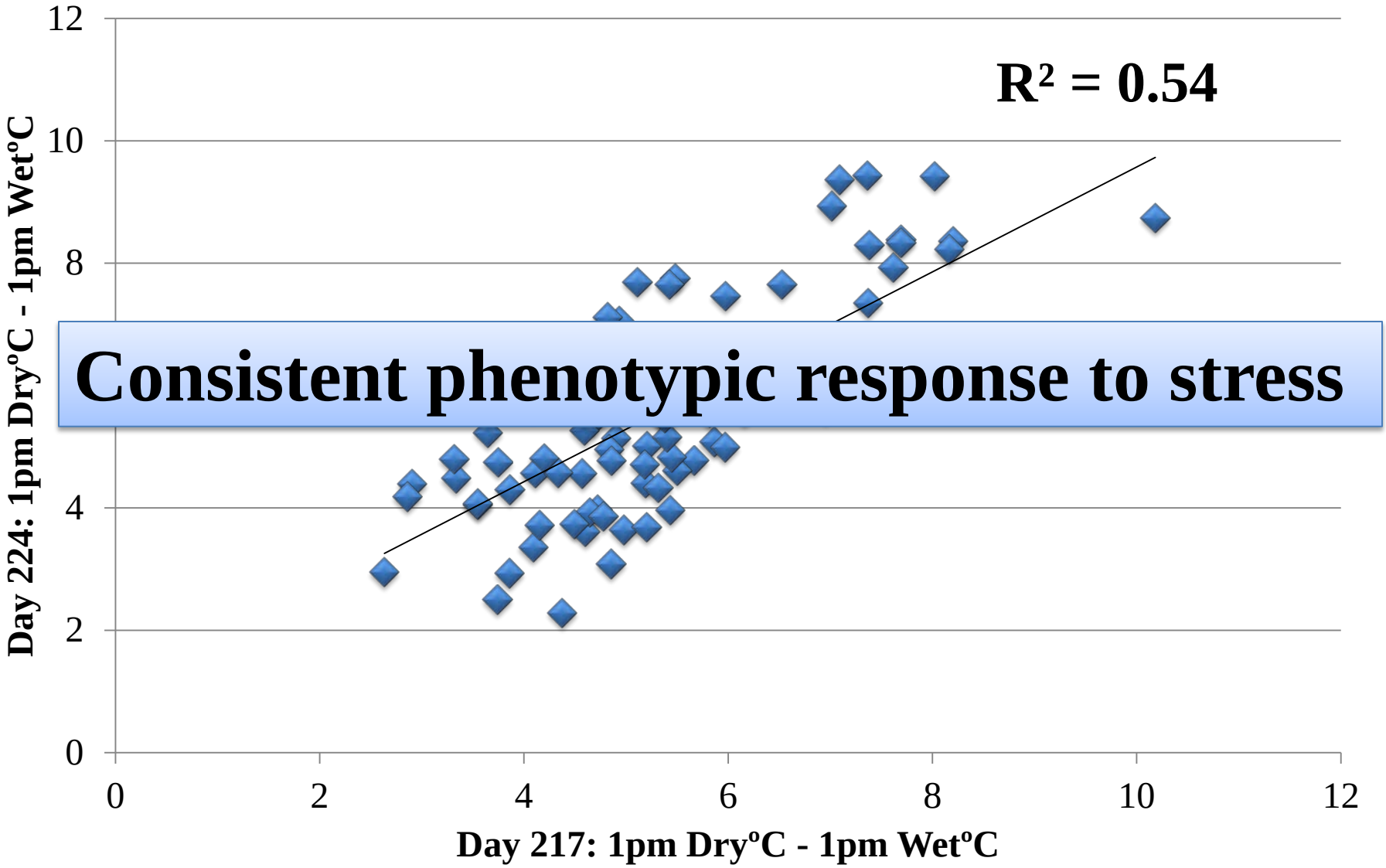
Phenotypic Variability of Canopy Temperature



Heritability estimates are 0.52 ± 0.11 for wet and 0.58 ± 0.09 for dry

Wet and Dry Plots at 1 pm on Day 224

Repeatability of Δ in Canopy Tm

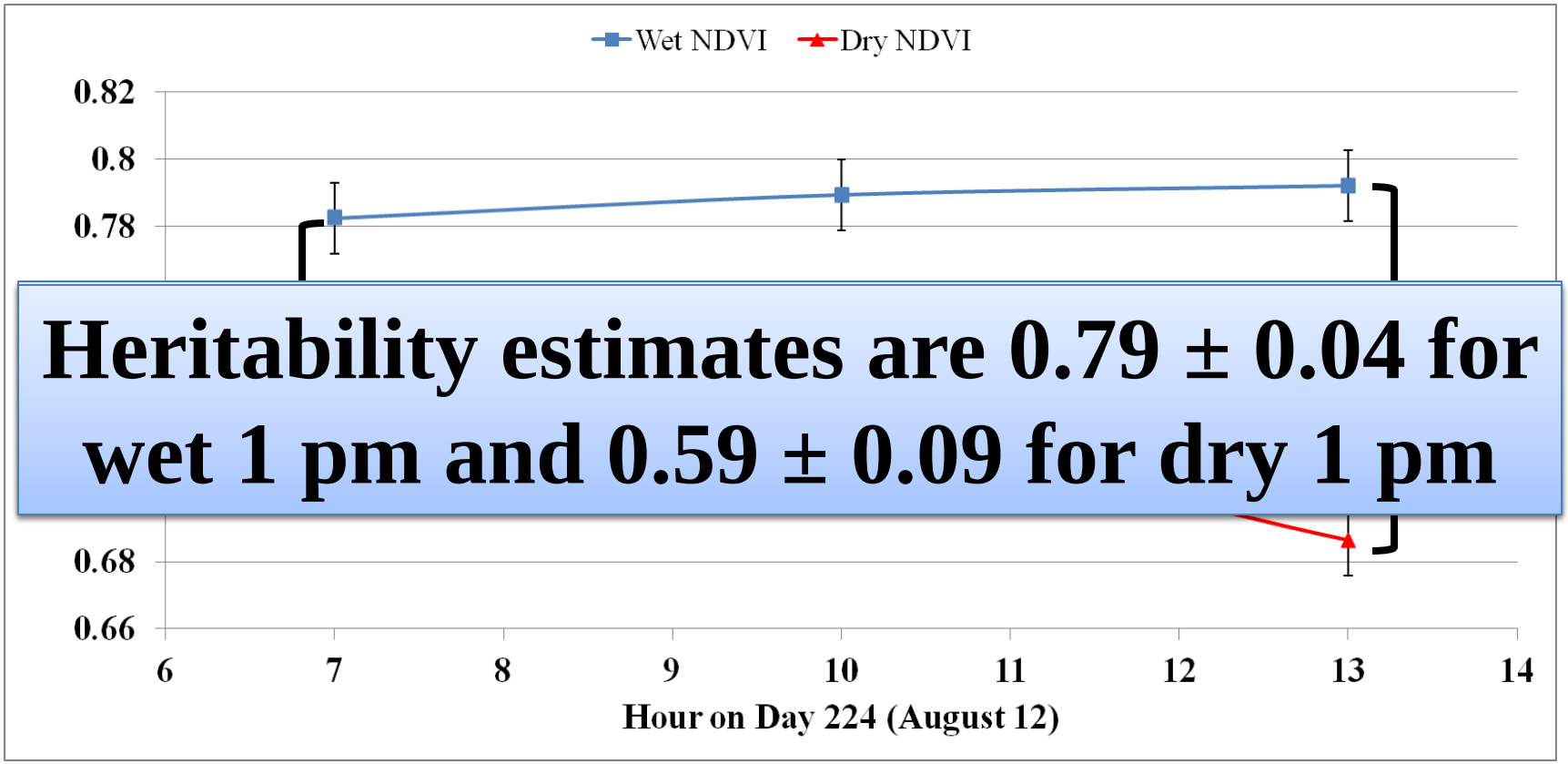


HTP: Normalized Difference Vegetation Index (NDVI)

NDVI = (NIR-red)/(NIR+red)
NIR 820 nm
Red 670 nm

Wilting Index (WI)
WI=[($NDVI_{pm}$-$NDVI_{pm}$)/$NDVI_{pm}$]

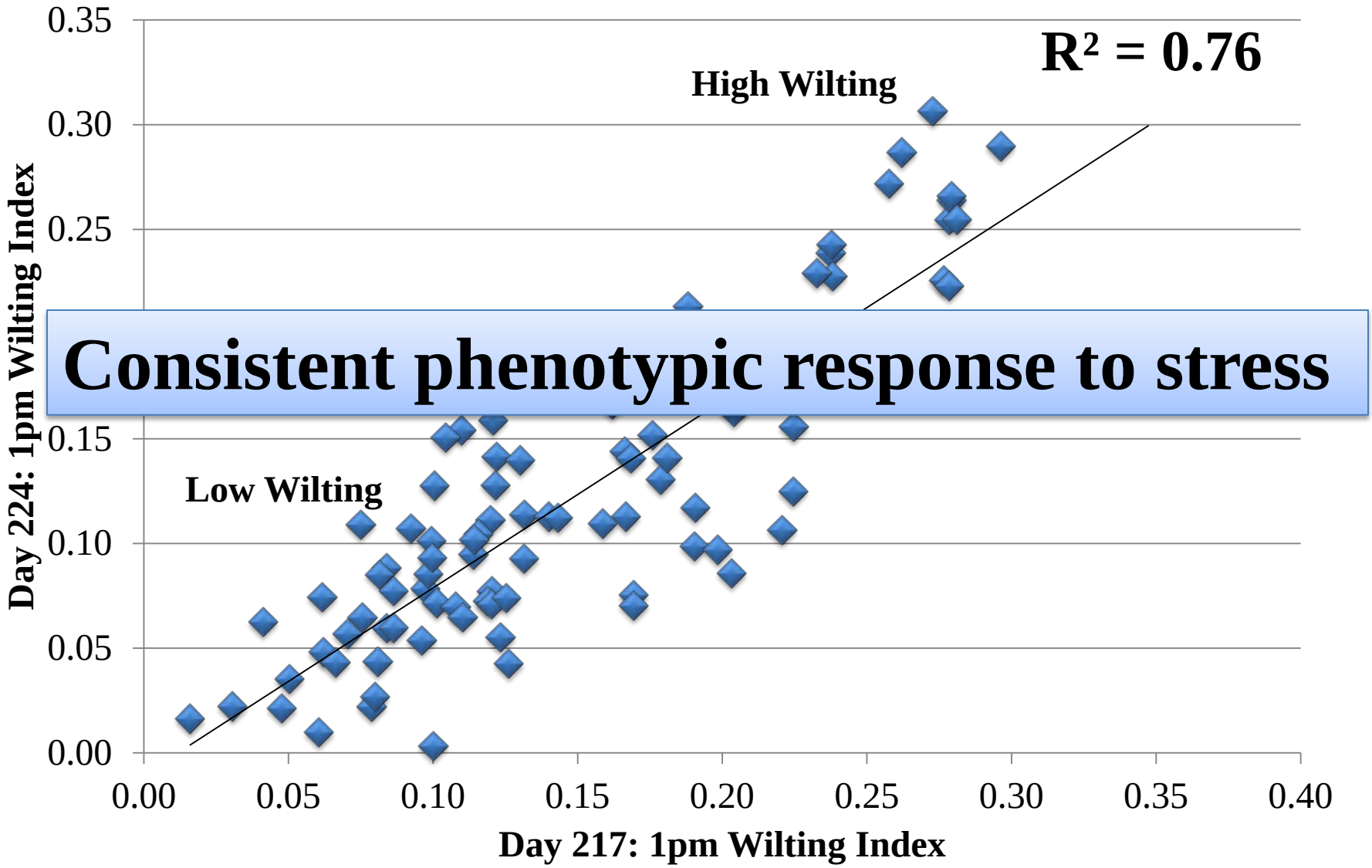
Significant Time-by-Treatment Interaction for NDVI



Heritability estimates are 0.79 ± 0.04 for wet 1 pm and 0.59 ± 0.09 for dry 1 pm

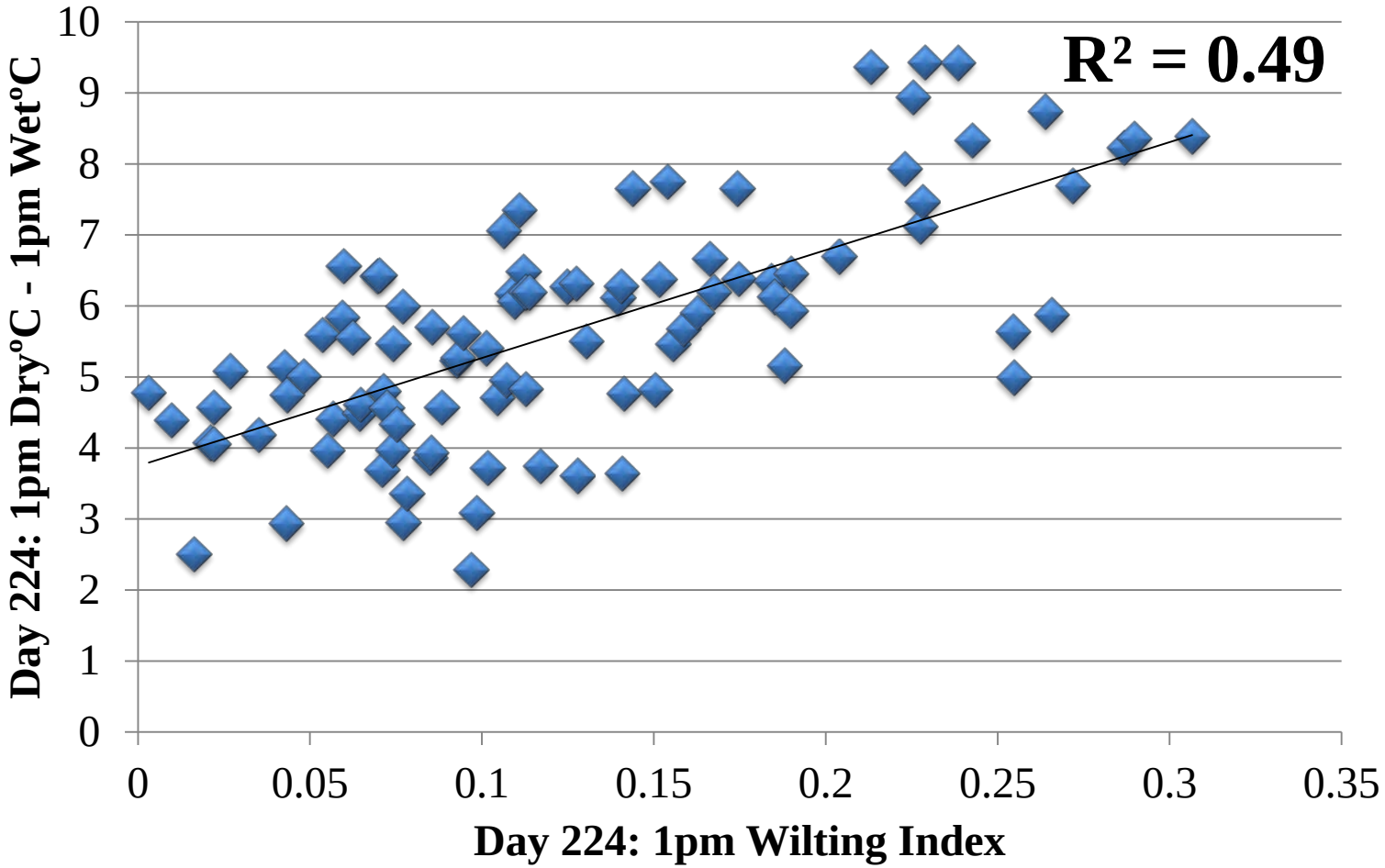
Treatment	***P*<0.05**
Time	***P*<0.0001**
Treatment*Time	***P*<0.0001**

Repeatability of Wilting Index

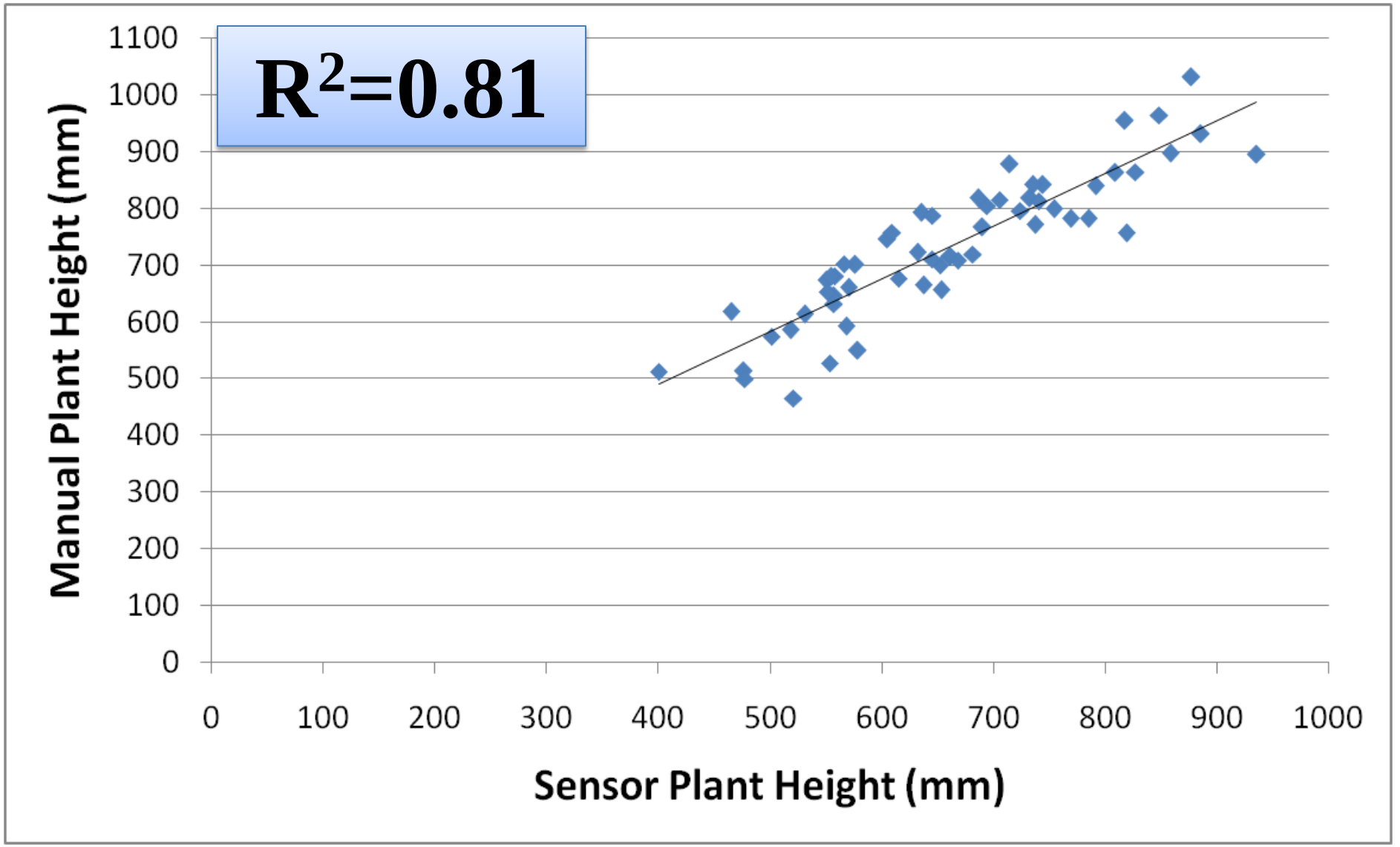


Consistent phenotypic response to stress

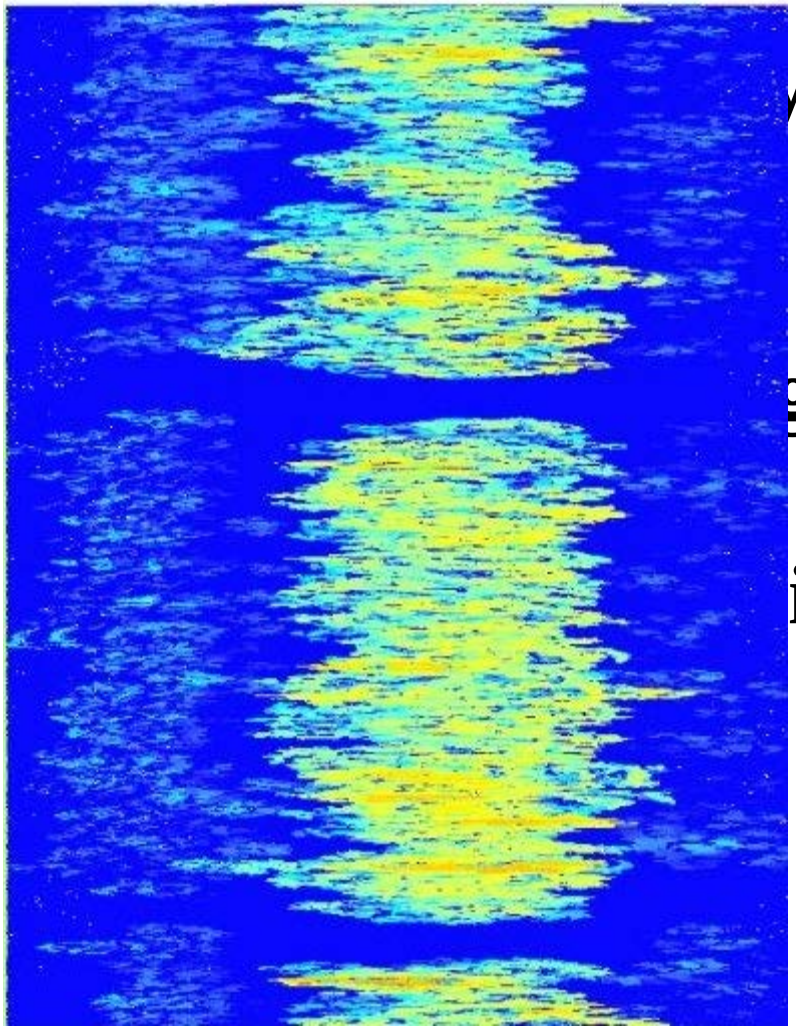
Relationship between Δ Canopy Tm and Wilting Index

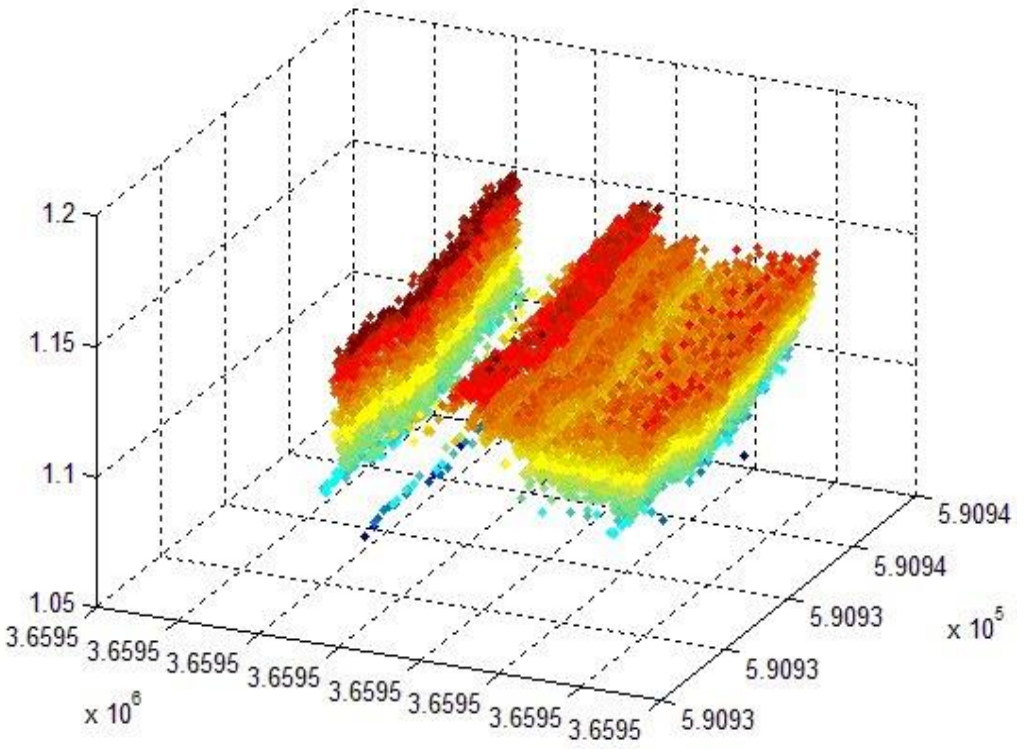


HTP: Plant Height



Optical remote sensing with Light Detection And Ranging (LIDAR) for





Bob Strand and Andy French

Next Steps for HTP

- Time-related QTL analysis of weekly NDVI, canopy temperature, and plant height in RIL population

- HTP for screening 1,000 cotton cultivars and day-neutral landrace lines for stress tolerance

- HTP to help develop cotton germplasm with superior stress tolerance by genomic selection

- Investigate GxE in multiple crops with HTP for many different phenotypes (disease, yield…)

Conclusions

• Constructed a GBS genetic map for tetraploid cotton RIL population without a need for a reference genome or downstream SNP assays

• Developed a field-based HTP approach to rapidly phenotype 100s to 1000s of plots for several canopy traits

• Canopy temperature and wilting index are moderately to highly heritable as well as repeatable on a temporal scale

Acknowledgements

Gore Lab

Alex Lipka

Joel Gilley

Kristen Harbour

Virginia Moreno

Sara Wyckoff

U of A

Pedro Andrade-Sanchez

John Heun

NMSU

Jinfa Zhang

U.S. ALARC

Andy French

Doug Hunsaker

Mike Salvucci

Kelly Thorp

Jeff White

USDA-ARS

Jesse Poland

Richard Percy

David Fang

*Pst*I *Msp*I

1) Ligation

```
   Forward Adapter    Barcode      Genomic DNA      Reverse Y-Adapter
5' CACGACGCTCTTCCGATCTXXXXXXTGCAGNNNN...NNNNCCGAGATCGGAAGAGCGGGGACTTTAAGC              3'
3' GTGCTGCGAGAAGGCTAGAYYYYYYTGCACNNNN...NNNNNNGGCTCTAGCCTTCTCGCCAAGTCGTCCTTACGGCTCTGGCTAG 5'
```

Exact match to reverse primer
(but not complement)

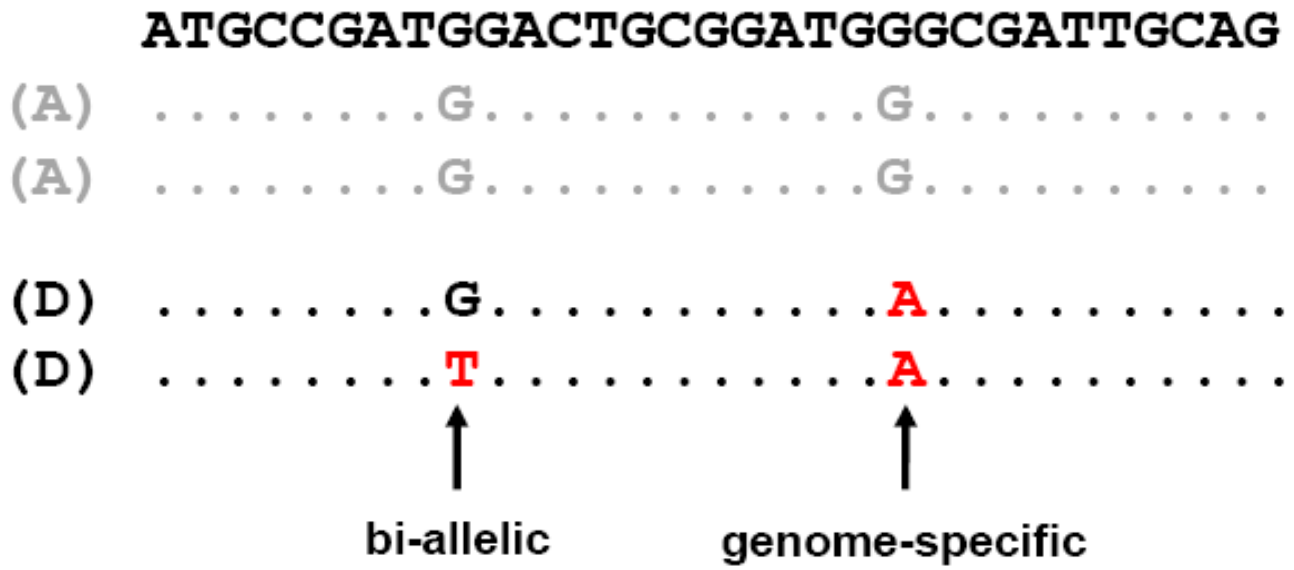
2) First PCR Cycle

```
    Forward Primer                                              PCR ⇒
5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTXXXXXXACG
                                    3' GTGCTGCGAGAAGGCTAGAYYYYYYTGCACNNNN.....

                                PCR ⇒
                .....NNNNNCCGAGATCGGAA
                .....NNNNNGGCTCTAGCCTTCTCGCCAAGTCGTCCTTACGGCTCTGGCTAG 5'
```

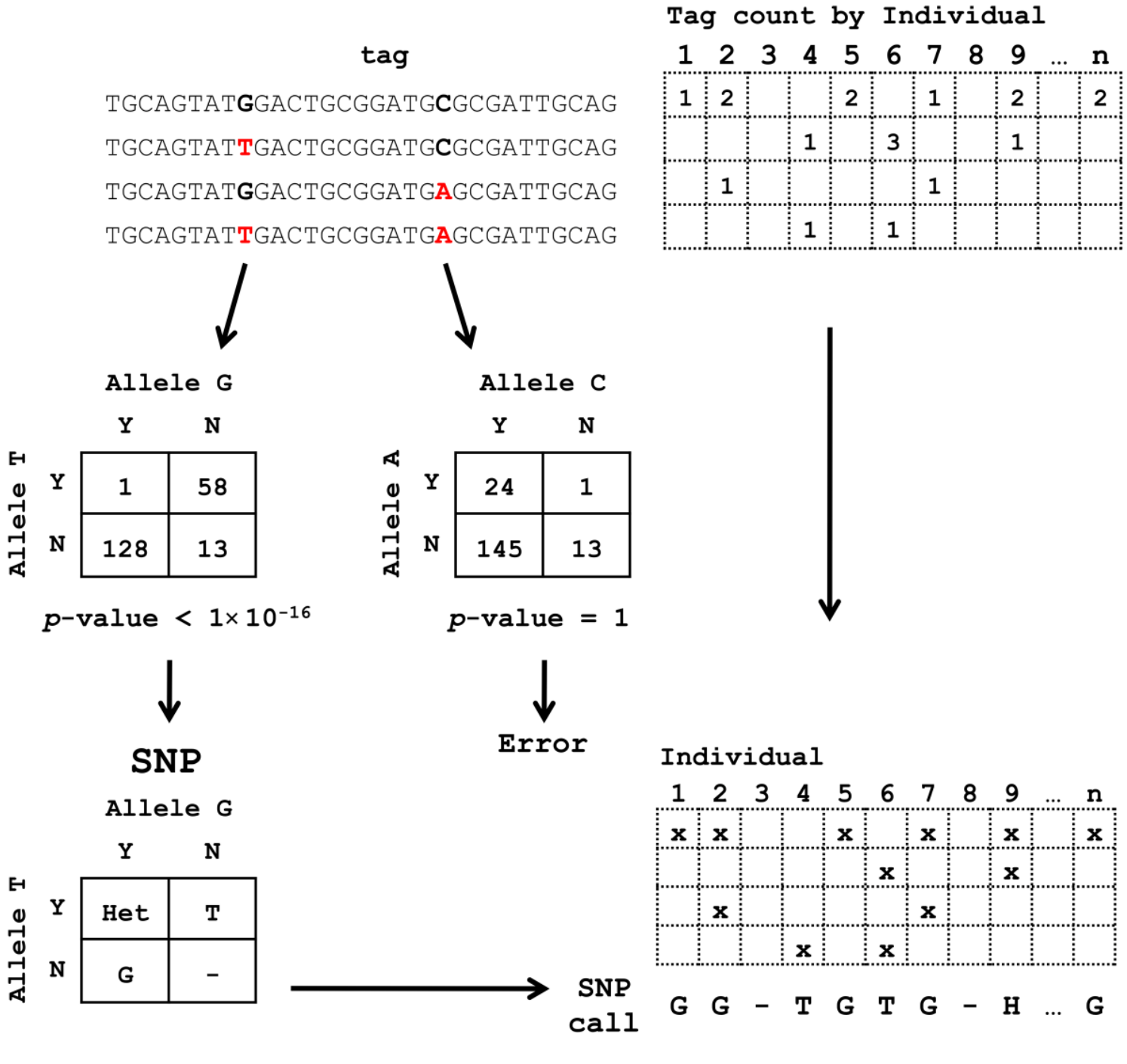
3) Second PCR Cycle

```
            5' ....NNNNNCCGAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATC  3'
                                    TCGAAGTCGTCCTTACGGCTCTGGCTAGAGCATACGGCAGAAGGACGAAC 5'
                                ⇐ PCR                                  Reverse Primer
```



Biallelic SNPs will be mutually exclusive in inbred line

Genome-specific SNPs will be heterozygous in inbred line


tag
TGCAGTATGGACTGCGGATGCGCGATTGCAG
TGCAGTATTGACTGCGGATGCGCGATTGCAG
TGCAGTATGGACTGCGGATGAGCGATTGCAG
TGCAGTATTGACTGCGGATGAGCGATTGCAG
Tag count by Individual
1 2 3 4 5 6 7 8 9 … n
1 2 2 1 2
1 3 1
1 1
1 1
Allele G
Y N
Allele T
Y 1 58
N 128 13
p-value < 1×10-16
Allele C
Y N
Allele A
Y 24 1
N 145 13
p-value = 1
SNP
Allele G
Y N
Allele T
Y Het T
N G -
Error
Individual
1 2 3 4 5 6 7 8 9 … n
SNP call
G G - T G T G - H … G

GBS: Bioinformatics Pipeline

Parse sequences to samples using barcode

www.maizegenetics.net/tassel

Collapse identical tags (matrix of counts – presence/absence)

<u>SNPs</u>
Identify tags:
1) **differing by 1 or 2 bp**
2) **Segregating in population**
3) **Not in same line**

de novo genetic map

<u>Map reads to genetic map using reference markers</u>

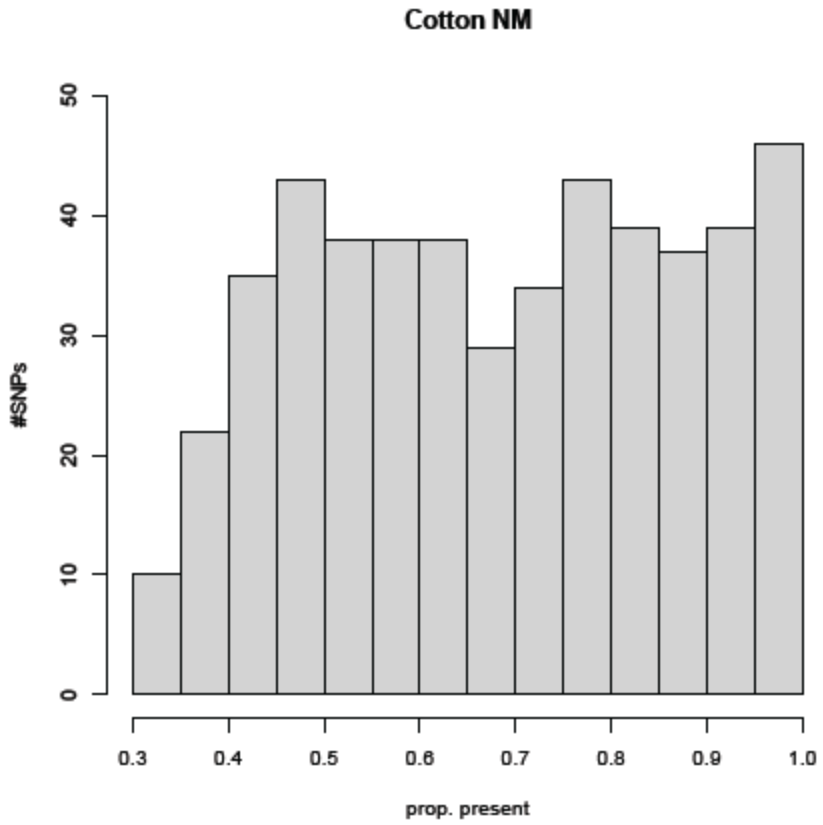
SNPs: **Fisher exact test for significance,**
BIN mapping
ML to position in interval between markers

Tags: **(Dominant markers)**
binomial probably for significance testing,
ML to position within interval

(matrix of tags/SNPs with genotype calls by line)

Poland et al. 2012 *PLoS ONE* 7(2): e32253

Marker data present



expect 3-4x coverage

Sonar sensor

Temperature sensors

30deg and nadir

Spectral sensor