

Transcriptomic approaches for the identification of drought-responsive genes in cotton



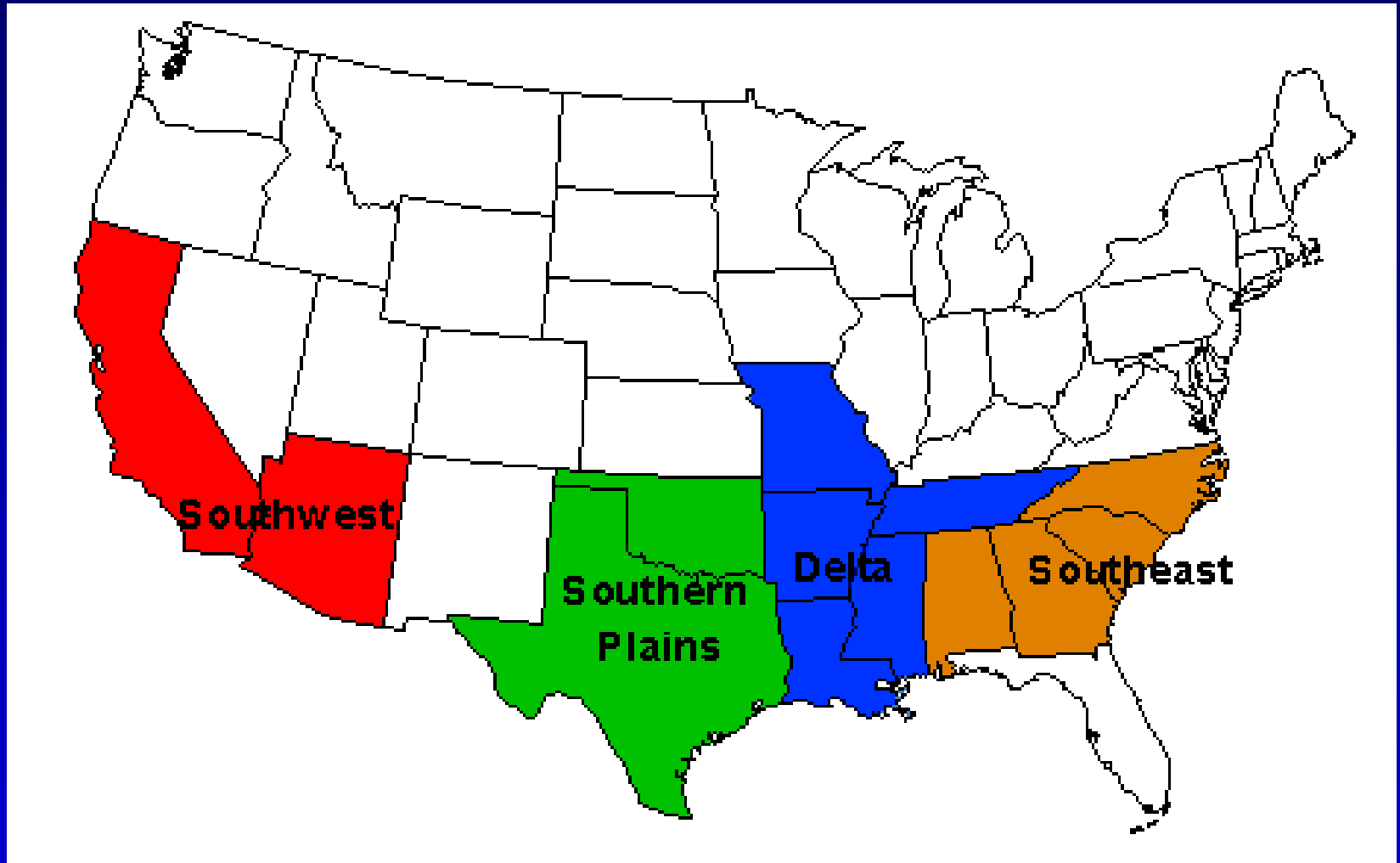
Wonkeun Park^{1,2}, Joshua A. Udall³, Justin T. Page³, Megan Bowman¹, Philip J. Bauer¹, and B. Todd Campbell¹

¹ USDA-ARS, Florence, SC

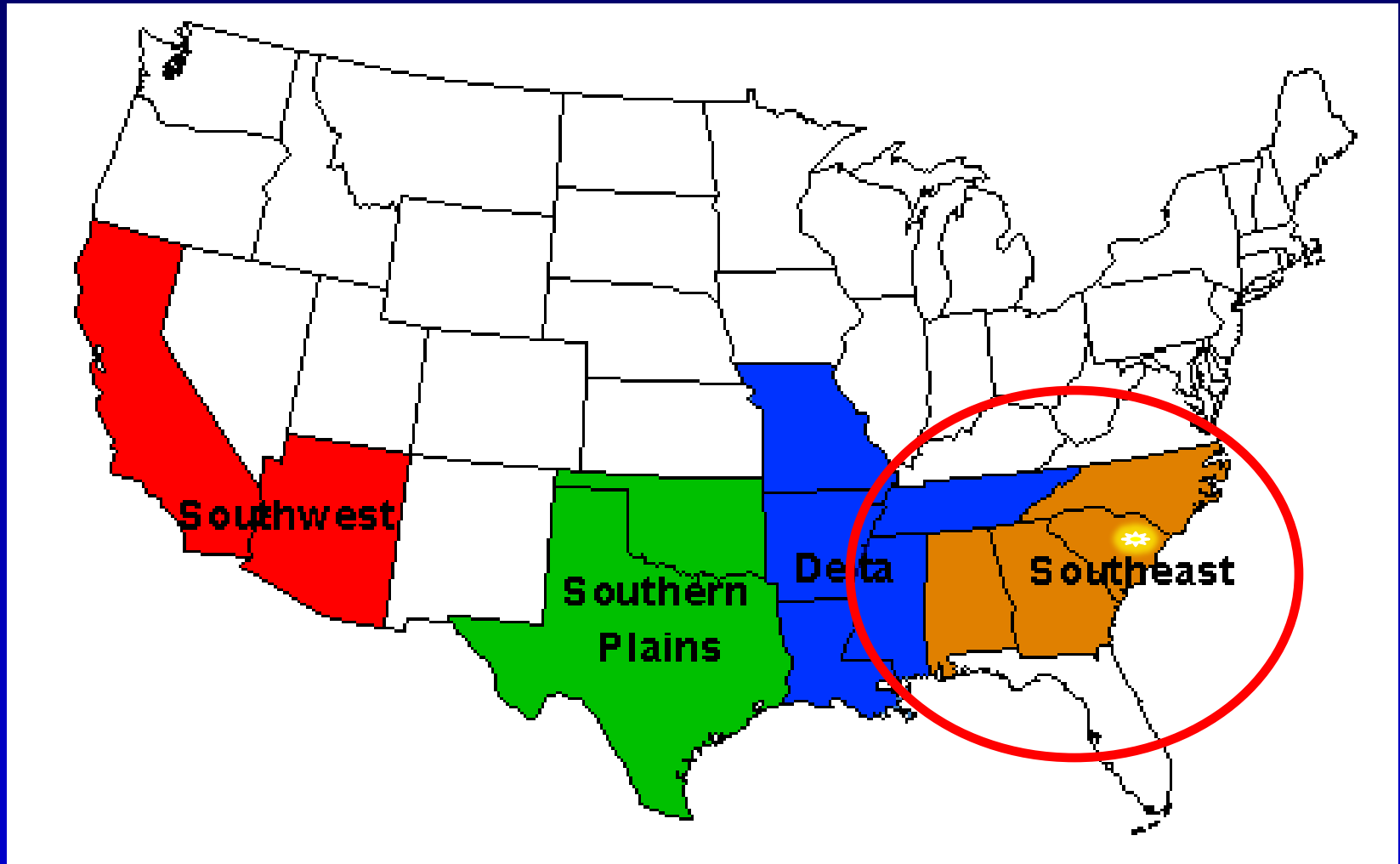
² Clemson University

³ Brigham Young University

US Cotton Production



US Cotton Production



Drought in the Southeast US

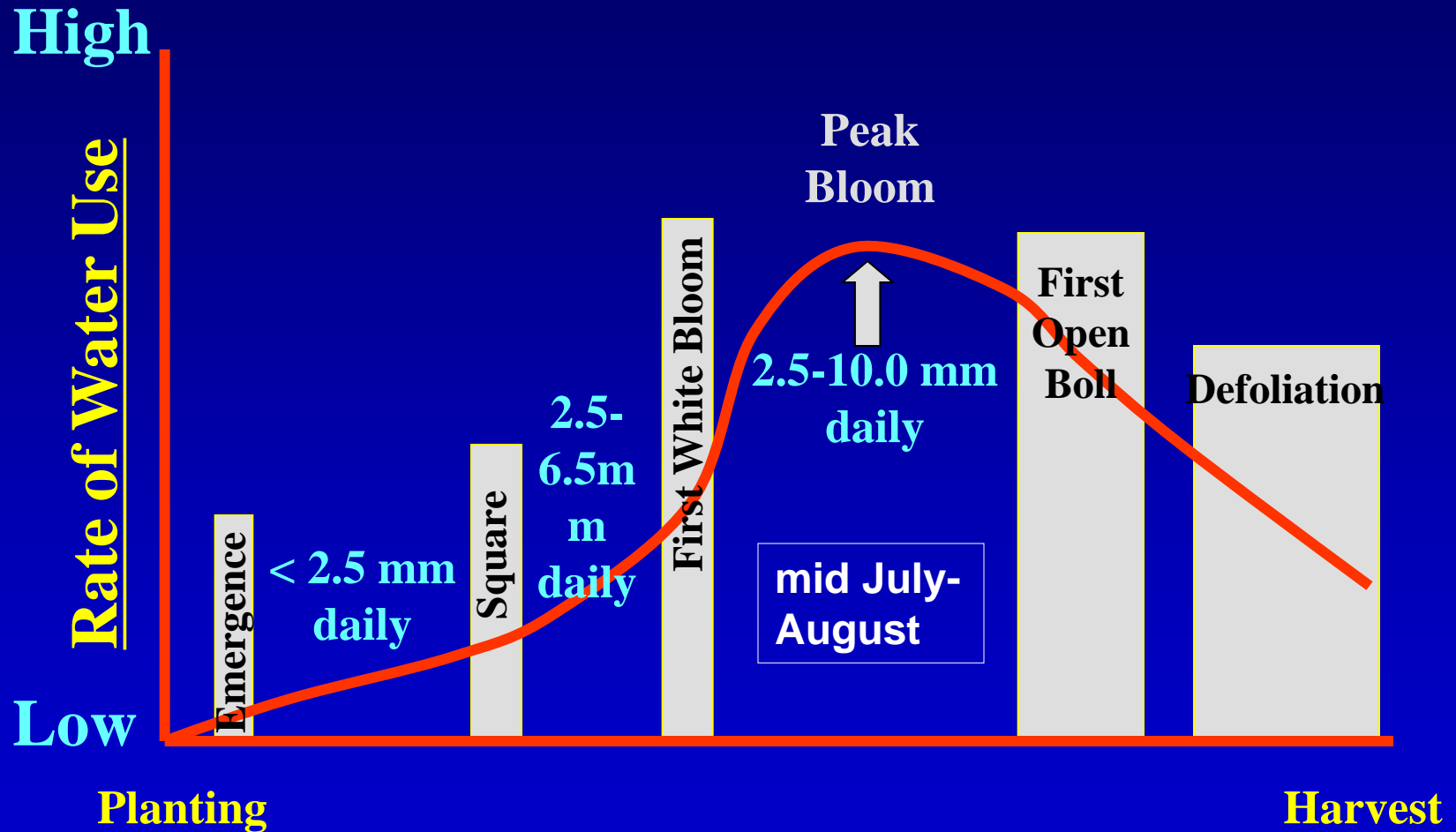
Sandy soils

Hold 2.5 cm of water per
30.5 cm of soil



Variable precipitation
50% chance of 20
continuous days without
rain each season

Water Use in Relation to Plant Development



Goal: Improve the productivity of cotton under intermittent drought conditions

Research Approaches:

- **Gene discovery**
- **Germplasm screening**

Candidate Gene Discovery - In the field

Candidate Gene Discovery - In the field



Courtesy of Greg Constable, CSIRO

Candidate Gene Discovery - In the field



- Ability to maintain high levels of photosynthesis under water deficit

(Nepomuceno, Oosterhuis, and Stewart, 1998)

Courtesy of Greg Constable, CSIRO

Candidate Gene Discovery - In the field



Courtesy of Greg Constable, CSIRO

- Ability to maintain high levels of photosynthesis under water deficit

(Nepomuceno, Oosterhuis, and Stewart, 1998)

1. Osmotic adjustment
2. ROS protection
3. Ability to capture soil water

Candidate Gene Discovery - In the field



12 Total Samples

- 2 Water treatments**
- 3 Biological replicates**
- Root and leaf**



Candidate Gene Discovery - In the field



12 Total Samples

- 2 Water treatments
- 3 Biological replicates
- Root and leaf



Physiological Traits

- Leaf water potential
- Root hydraulic conductance

Candidate Gene Discovery - In the field



12 Total Samples

- 2 Water treatments
- 3 Biological replicates
- Root and leaf

| Treatment | Plant | Leaf Water Potential (MPa) |
|--------------|-------|----------------------------|
| Well-watered | 1 | -1.60 |
| | 2 | -1.35 |
| | 3 | -1.45 |
| Rainfed | 1 | -2.20 |
| | 2 | -2.70 |
| | 3 | -2.85 |

Candidate Gene Discovery - In the field



12 Total Samples

- 2 Water treatments
- 3 Biological replicates
- Root and leaf

| Treatment | Plant | Leaf Water Potential (MPa) |
|--------------|-------|----------------------------|
| Well-watered | 1 | -1.60 |
| | 2 | -1.35 |
| | 3 | -1.45 |
| Rainfed | 1 | -2.20 |
| | 2 | -2.70 |
| | 3 | -2.85 |

- 2.00
critical

Candidate Gene Discovery - In the field



12 Total Samples

- 2 Water treatments
- 3 Biological replicates
- Root and leaf

| Treatment | Plant | Leaf Water Potential (MPa) |
|--------------|-------|----------------------------|
| Well-watered | 1 | -1.60 |
| | 2 | -1.35 |
| | 3 | -1.45 |
| Rainfed | 1 | -2.20 |
| | 2 | -2.70 |
| | 3 | -2.85 |

- 2.00
critical

- No differences in root hydraulic conductance

Candidate Gene Discovery – RNA seq



Sequencing work flow

- Barcoded Illumina RNA TruSeq cDNA libraries (12)
- Quality and concentrations assessed with Agilent Bioanalyzer and qPCR
- Single lane of 50bp HiSeq



Candidate Gene Discovery – RNA seq



Sequence analysis work flow

- Reads trimmed with sickle (Najoshi)
- Reads mapped to *Gossypium raimondii* 2.1 reference (Paterson) using GSNAP (Wu and Nacu, 2010)
- Reads categorized to A_T or D_T genome using polyCat (Page, Gingle, and Udall)
- Bioconductor 2.10 DESeq (Anders and Huber, 2010) with 5% FDR



Candidate Gene Discovery – RNA seq

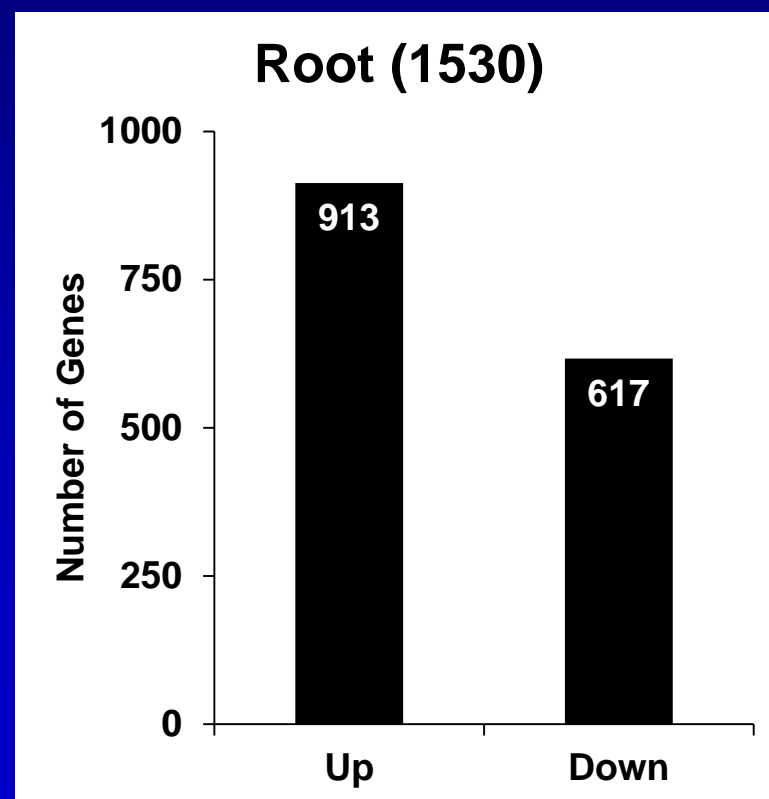
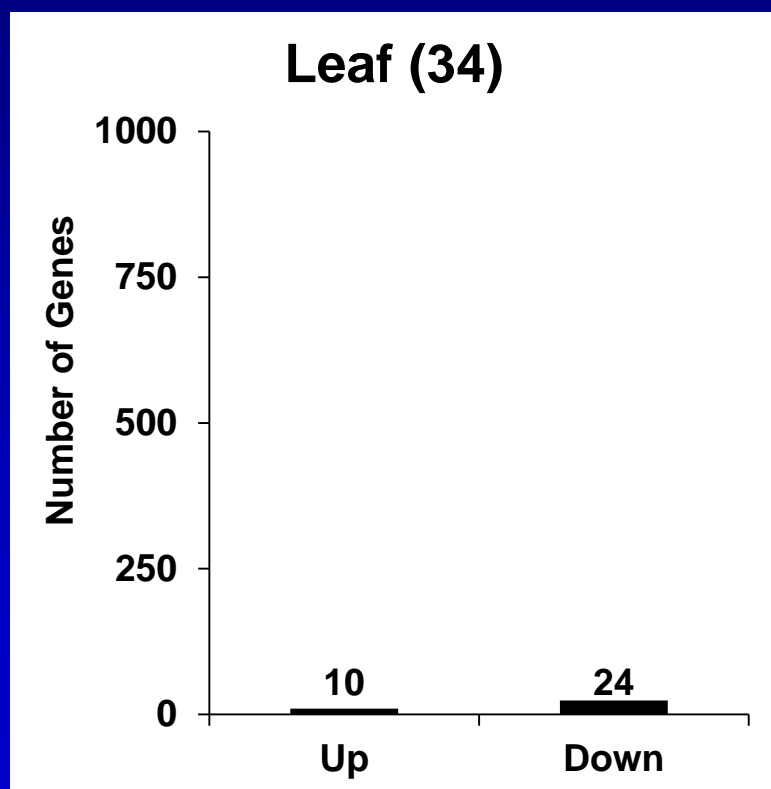


Approximate sequencing statistics

- 308 million reads
- 290 million trimmed reads
- 150 million mapped reads
- 148 million reads mapped to annotated genes
- 34,000 genes surveyed

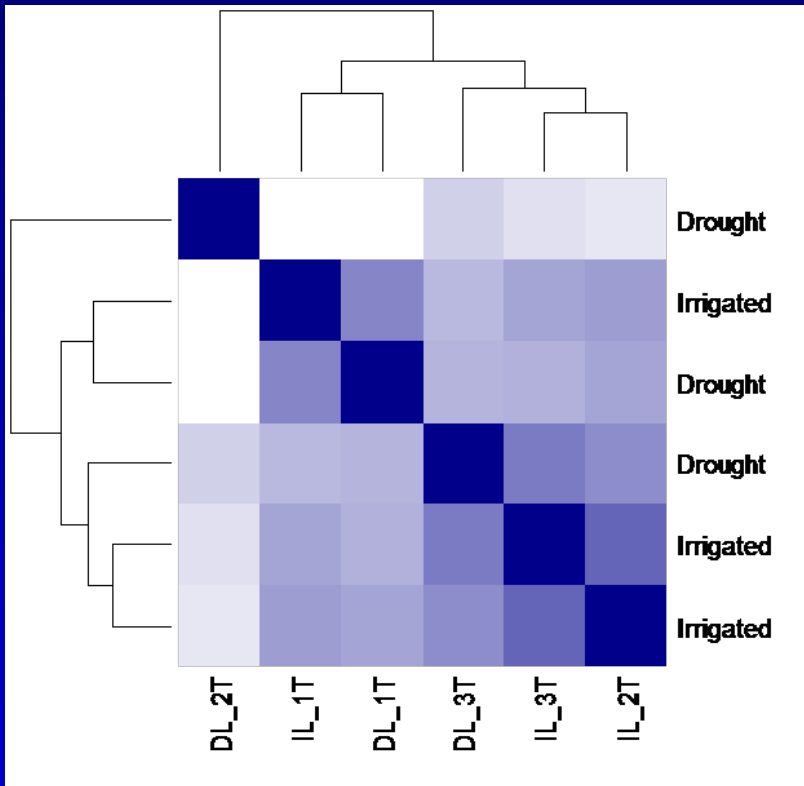


Number of Genes Differentially Expressed In Response to Drought

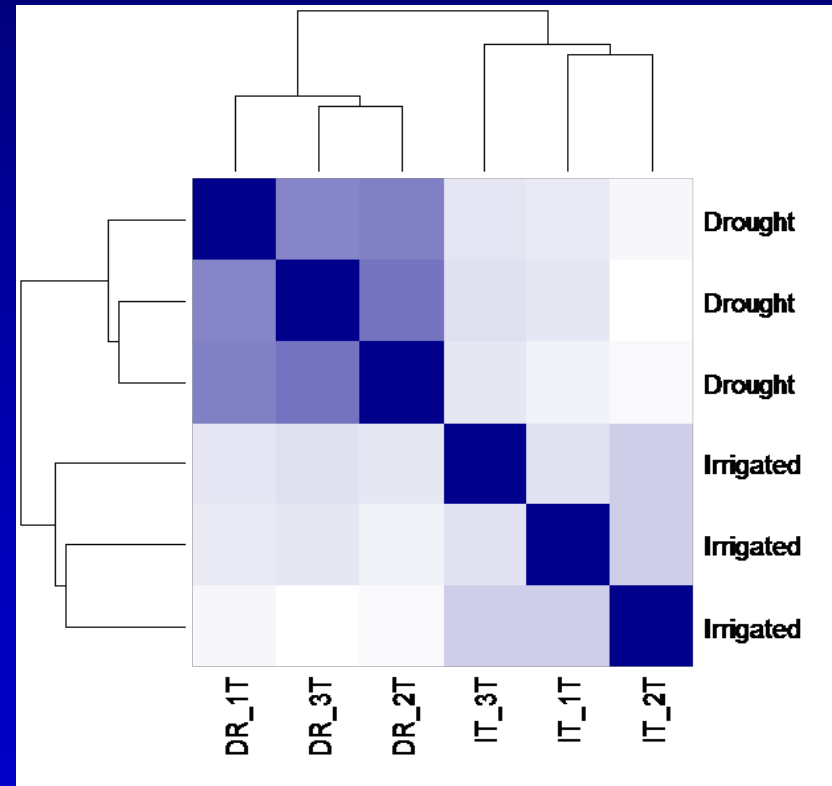


Biological Replicate Variation

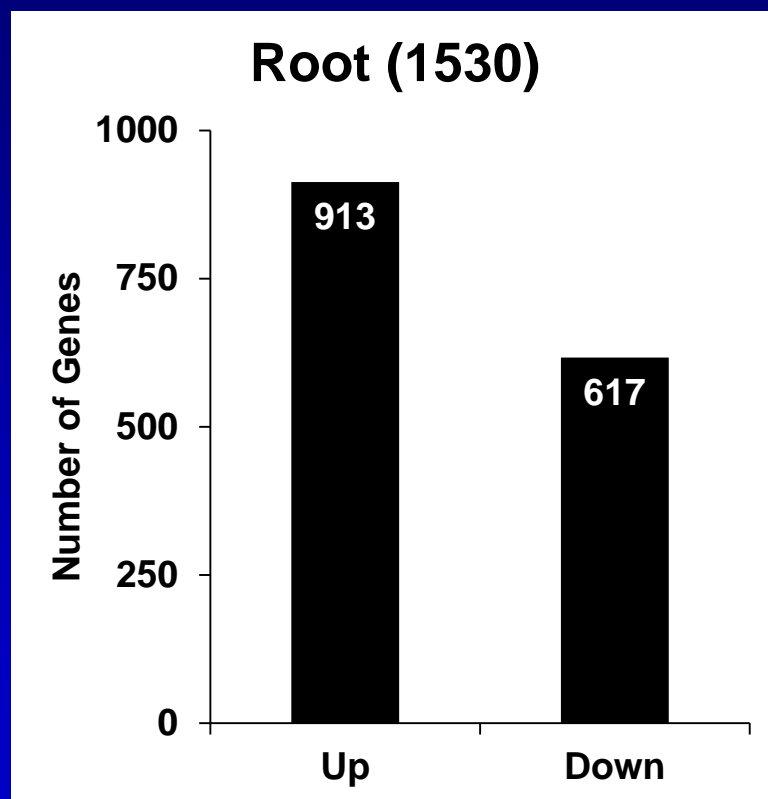
Leaf Heat Map



Root Heat Map



Number of Genes Differentially Expressed In Response to Drought



Sub-genome read categorization for root

Up-regulated genes

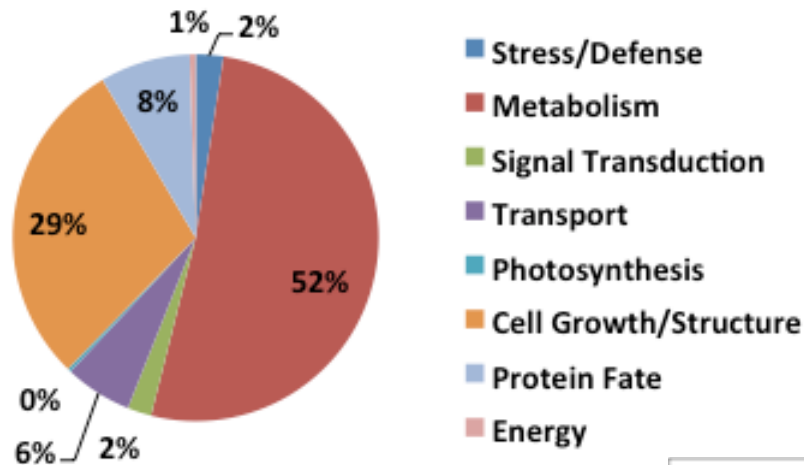
| Total Genes | A _T both | D _T both | A _T drought, D _T wet | D _T drought, A _T wet | A _T only | D _T only | Non e |
|-------------|---------------------|---------------------|--|--|---------------------|---------------------|-------|
| 913 | 407 | 315 | 72 | 74 | 2 | 3 | 40 |
| | 44.6% | 34.5% | 7.9% | 8.1% | 0.2% | 0.3% | 4.4% |

Down-regulated genes

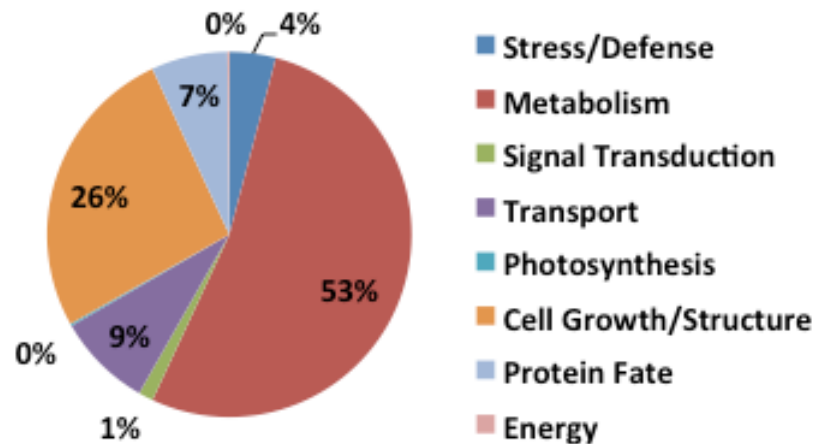
| Total Genes | A _T both | D _T both | A _T drought, D _T wet | D _T drought, A _T wet | A _T only | D _T only | Non e |
|-------------|---------------------|---------------------|--|--|---------------------|---------------------|-------|
| 617 | 225 | 217 | 50 | 54 | 5 | 5 | 61 |
| | 36.5% | 35.2% | 8.1% | 8.8% | 0.8% | 0.8% | 9.9% |

Differentially expressed genes in root

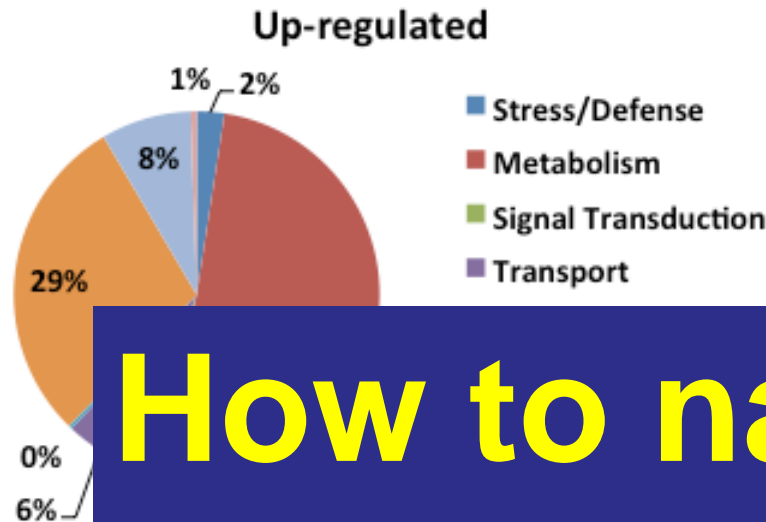
Up-regulated



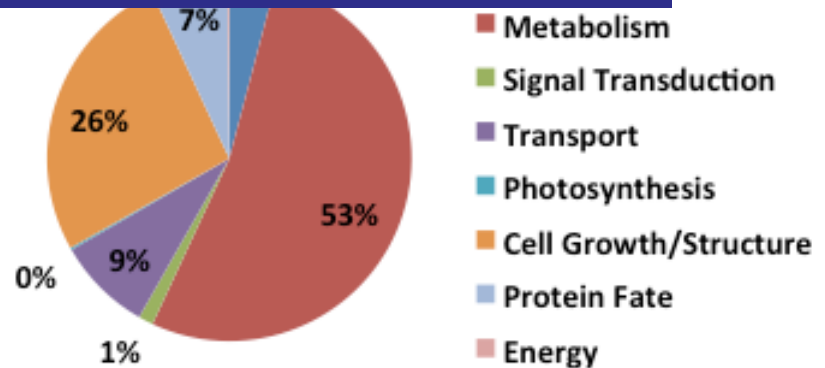
Down-regulated



Differentially expressed genes in root



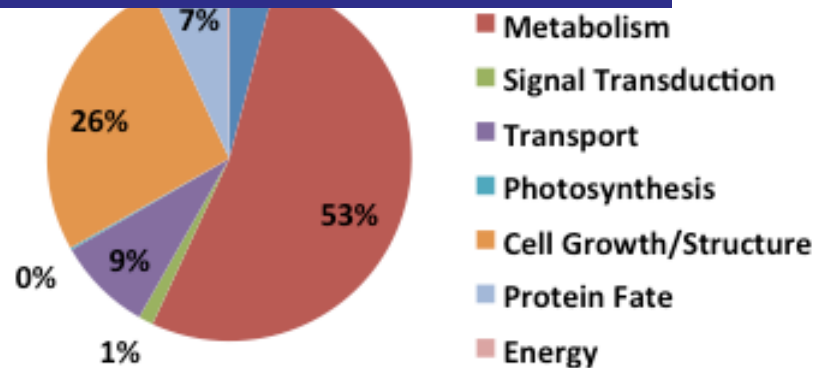
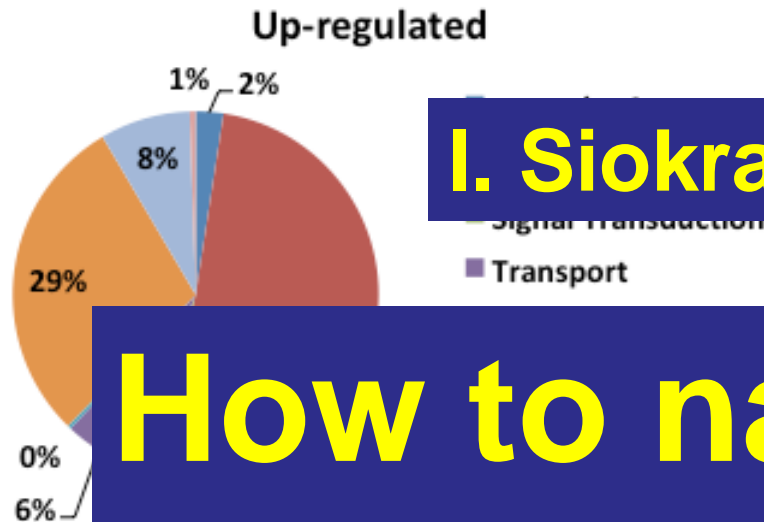
How to narrow candidate genes?



Differentially expressed genes in root

I. Siokra L-23 mechanism(s)?

How to narrow candidate genes?

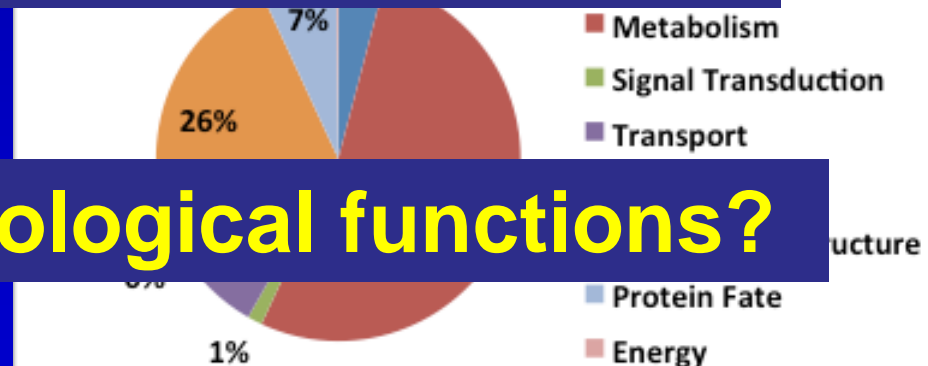
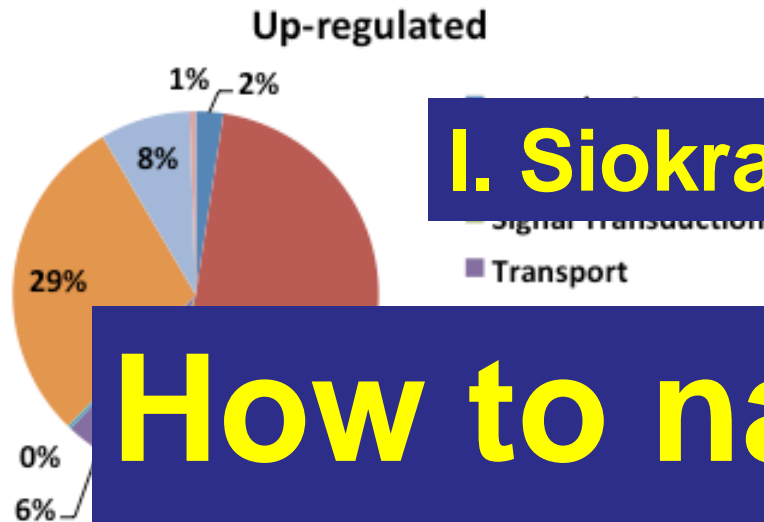


Differentially expressed genes in root

I. Siokra L-23 mechanism(s)?

How to narrow candidate genes?

II. Other specific biological functions?



High expression level

| Top 10 based on read count | | | | | |
|----------------------------|--------------|---------|-------------|---------|--|
| Gene ID | Well-watered | Rainfed | Fold Change | P-Value | Annotation |
| Gorai.004G128700 | 17186 | 39930 | 2.32 | 0.004 | Starch Synthase |
| Gorai.006G169400 | 11640 | 2704 | 4.30 | 0.035 | H+/oligopeptide symporter |
| Gorai.013G238000 | 11321 | 23950 | 2.12 | 0.028 | Ring Finger and Zinc Finger Domain |
| Gorai.006G011500 | 10795 | 46055 | 4.27 | 0.000 | Phosphorylase |
| Gorai.005G113800 | 9732 | 34082 | 3.50 | 0.000 | Amylase |
| Gorai.006G223900 | 9429 | 20020 | 2.12 | 0.021 | Adenosylhomocysteinase |
| Gorai.011G083000 | 6983 | 17025 | 2.44 | 0.003 | Phospholipase |
| Gorai.009G187700 | 5855 | 14651 | 2.50 | 0.003 | Farnesoic acid O-methyl transferase |
| Gorai.009G218000 | 5701 | 2601 | 2.19 | 0.010 | Subtilisin/Kexin-Related Serine Protease |
| Gorai.005G244900 | 5647 | 25095 | 4.44 | 0.000 | Sugar-1-Phosphate Guanyl Transferase |

Large fold change

| Top 10 based on fold change and p-value | | | | | |
|---|--------------|---------|-------------|---------|--|
| Gene ID | Well-watered | Rainfed | Fold Change | P-Value | Annotation |
| Gorai.001G024200 | 122 | 1 | 117.88 | 4E-09 | Polyketide cyclase / dehydrase and lipid transport |
| Gorai.005G238000 | 2 | 164 | 102.03 | 3E-03 | Ethylene-responsive transcription factor ERF109-like |
| Gorai.011G250800 | 1595 | 27 | 58.54 | 3E-05 | 26S proteasome regulatory complex |
| Gorai.006G073600 | 1 | 52 | 39.74 | 9E-10 | Glycosyl hydrolase family |
| Gorai.009G207100 | 14 | 0 | 39.70 | 2E-02 | Cytochrome P450 |
| Gorai.007G145000 | 55 | 2 | 34.57 | 3E-02 | Triose-phosphate Transporter |
| Gorai.011G251800 | 74 | 2 | 32.41 | 2E-12 | 26S proteasome regulatory complex |
| Gorai.002G078800 | 1 | 22 | 30.93 | 2E-04 | Receptor-like protein kinase |
| Gorai.008G231800 | 1 | 19 | 30.70 | 5E-04 | Phosphate-transporter related |
| Gorai.006G021000 | 18 | 1 | 29.32 | 7E-04 | Serine-Threonine protein kinase |

Low expression level

Bottom 10 based on read count

| Gene ID | Well-watered | Rainfed | Fold Change | P-Value | Annotation |
|------------------|--------------|---------|-------------|---------|--|
| Gorai.002G023500 | 76 | 180 | 2.3 | 0.048 | Long-chain acyl-CoA transporter, ABC superfamily |
| Gorai.009G053200 | 1303 | 514 | 2.5 | 0.048 | Serine-Threonine protein kinase |
| Gorai.003G023600 | 29 | 4 | 6.5 | 0.048 | Cytochrome P450 |
| Gorai.009G177900 | 639 | 1278 | 2.0 | 0.049 | Mannosyl-3-phosphoglycerate phosphatase |
| Gorai.002G234400 | 1508 | 754 | 2.0 | 0.049 | CCAAT-binding transcription factor |
| Gorai.005G216600 | 137 | 292 | 2.1 | 0.049 | Sugar Transporter |
| Gorai.001G252800 | 64 | 16 | 3.9 | 0.049 | AP2 transcription factor activity |
| Gorai.009G340400 | 3 | 18 | 7.0 | 0.049 | Serine-Threonine protein kinase |
| Gorai.004G090000 | 345 | 719 | 2.1 | 0.049 | UDP-glucose 4-epimerase |
| Gorai.005G081200 | 82 | 25 | 3.3 | 0.050 | Serine-Threonine protein kinase |

Small fold change

Bottom 10 based fold change and p-value

| Gene ID | Well-watered | Rainfed | Fold Change | P-Value | Annotation |
|------------------|--------------|---------|-------------|---------|--|
| Gorai.004G183400 | 2187 | 4426 | 2.0 | 0.043 | Trehalose-6-Phosphate Synthase |
| Gorai.002G098400 | 996 | 492 | 2.0 | 0.040 | Serine O-acetyltransferase |
| Gorai.010G155900 | 1573 | 3162 | 2.0 | 0.046 | AP2 transcription factor activity |
| Gorai.002G083100 | 3886 | 7794 | 2.0 | 0.044 | Ubiquitin |
| Gorai.002G234400 | 1508 | 754 | 2.0 | 0.049 | CCAAT-binding transcription factor |
| Gorai.009G177900 | 639 | 1278 | 2.0 | 0.049 | Mannosyl-3-phosphoglycerate phosphatase |
| Gorai.013G216000 | 2630 | 5257 | 2.0 | 0.042 | Protease M17 Leucine Aminopeptidase |
| Gorai.002G191700 | 1000 | 502 | 2.0 | 0.041 | Kelch-related proteins |
| Gorai.007G188400 | 2413 | 4784 | 2.0 | 0.045 | UTP--glucose-1-phosphate uridylyltransferase |
| Gorai.005G027000 | 2804 | 5519 | 2.0 | 0.034 | 56kDa Selenium binding protein (SBP56) |

Some Observations

- More transcriptome changes in roots
- Higher error variation in leaf



Some Observations

- More transcriptome changes in roots
- Higher error variation in leaf
- Osmotic adjustment
- ROS protection
- Ability to capture soil water?



Where Are We Going??

- RT-qPCR validation



Where Are We Going??

- RT-qPCR validation
- Narrow candidate genes for future study



Where Are We Going??

- RT-qPCR validation
- Narrow candidate genes for future study
- Mine naturally occurring genetic variation



Acknowledgements

