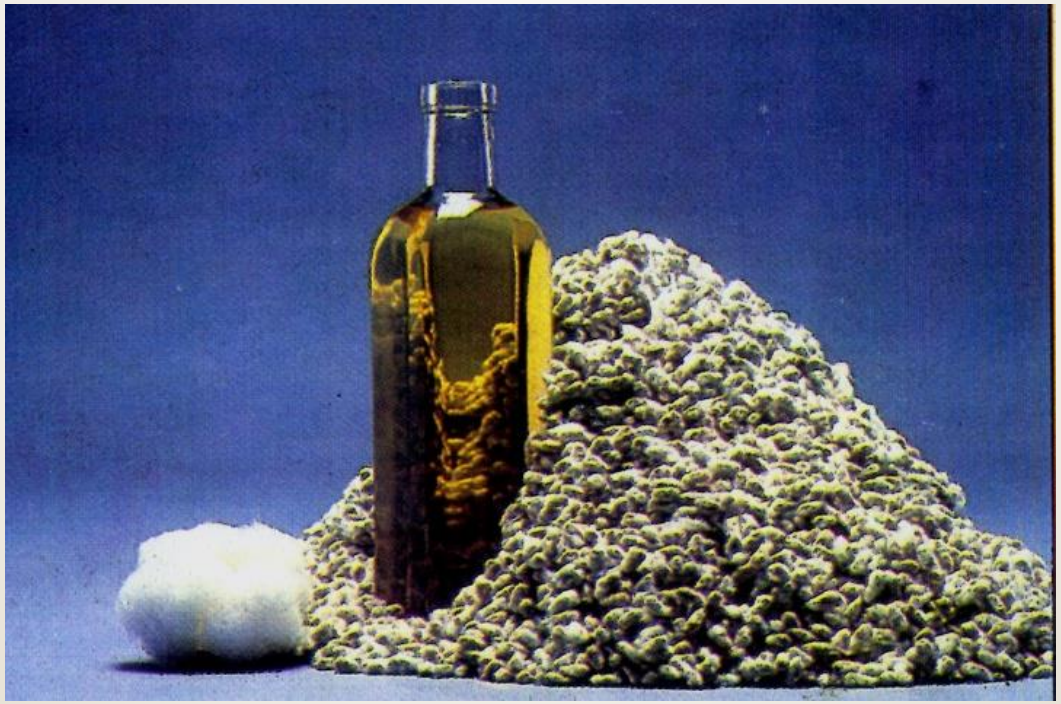
Capturing Genetic Variation in Cottonseed Constituents to Enhance Value

Kent D. Chapman
University of North Texas
BioDiscovery Institute and Department of Biological Sciences
Denton, TX



**Seed yield is ~1.5X fiber but less than half of seed is processed-

Place to add value- Oil>>protein

2019 Cotton Breeder's Tour- Cotton Incorporated
College Station, TX-- July 22, 2019

UNT UNIVERSITY OF NORTH TEXAS EST. 1890

BIODISCOVERY INSTITUTE

Capturing Genetic Variation in Cottonseed Constituents to Enhance Value

Variation in Cottonseed Oil and Protein Content- Non-destructive measurements by Time-Domain ^{1}H-NMR. Quantifying variation across *Gossypium* germplasm and genotypes

Compositional Differences-- Identification of a naturally-occurring mutant allele for increased oleic acid in cottonseeds.

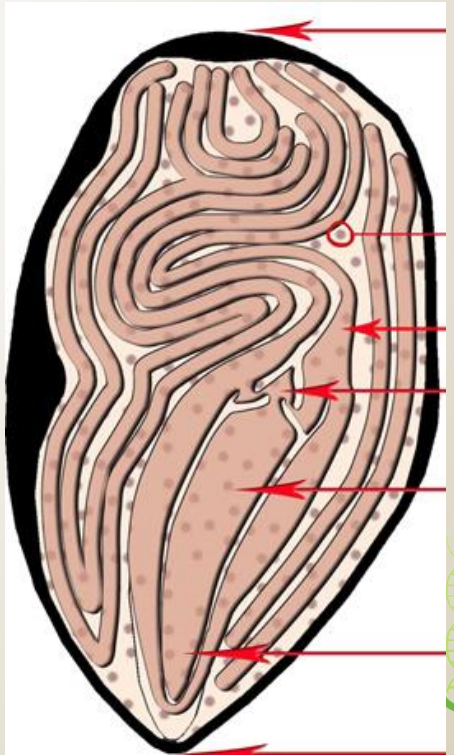
A Potential Future with a Holistic Focus on Seed Products?
Targeted and untargeted strategies-- Ultra-low gossypol stacked with altered seed oil traits (high oleic, low saturate FA, cyclic FA, Vitamin E) or GWAS studies with NAM populations to identify important loci. Identity preservation considerations.

BIODISCOVERY INSTITUTE

Synthesis of Seed Products during Development

Oil and protein are synthesized in the embryo of the seed during development and maturation, primarily in the cotyledons of cotton seed (bulk of seed mass at maturity)

Fiber elongated from the seed coat surface during seed development and fills fruit (boll) pubs.caes.uga.edu/caespubs/pubcd/B1252.htm



BIODISCOVERY INSTITUTE

Quantifying Seed Oil and Protein Content-

Non-destructive, time-domain ^{1}H-NMR

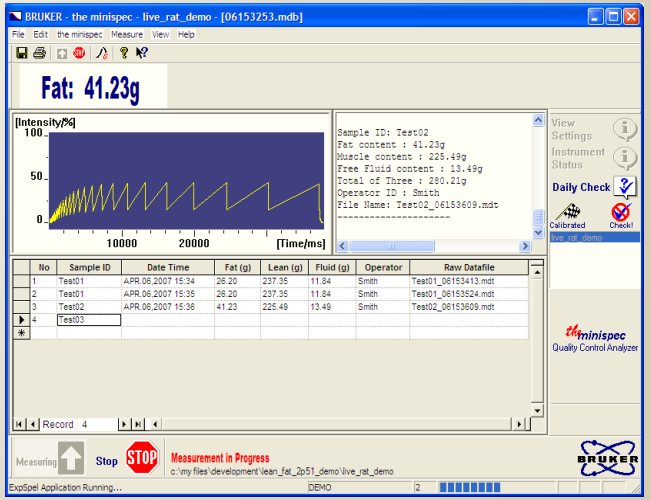
Patrick Horn

In 2009-2011 worked with scientists at Buker Optics to develop NMR-based assay for quantification of protein and oil in ~3g cottonseed (single seed for oil).
Horn et al., 2011, JAOCS (AOCS method for oil AOCS Cd 16b-93)

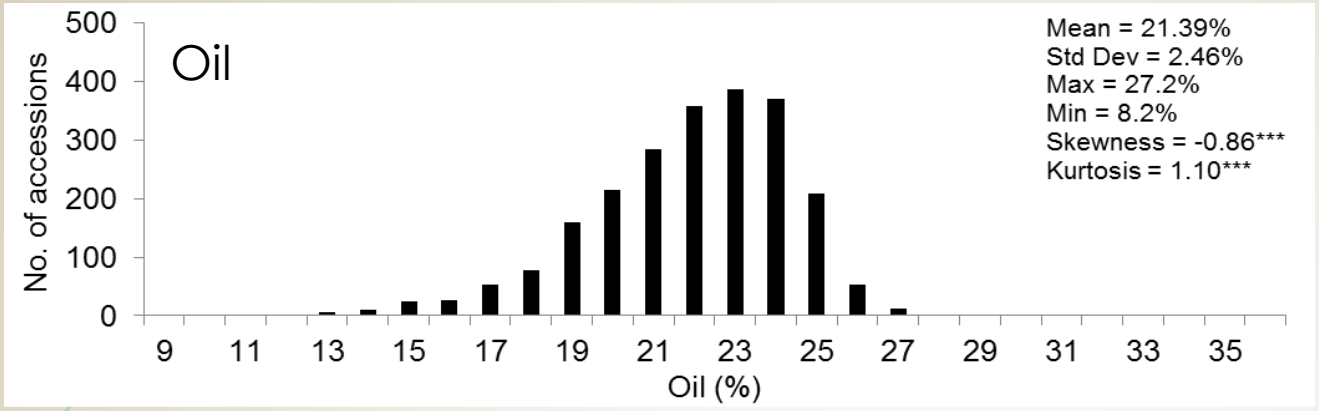
Since then, we have worked with the cotton breeding community to provide analyses where possible (**T. Campbell,** J. Snider, C. Main, and many others)

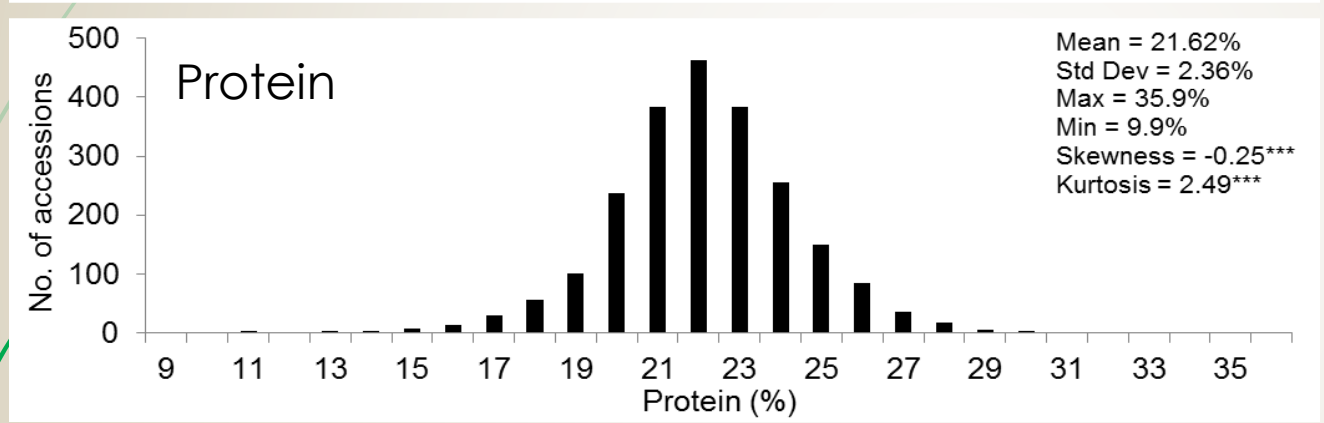
New service facility being set up by Cotton Inc to quantify seed oil (and protein) for public and private breeders, like is done already for fiber.

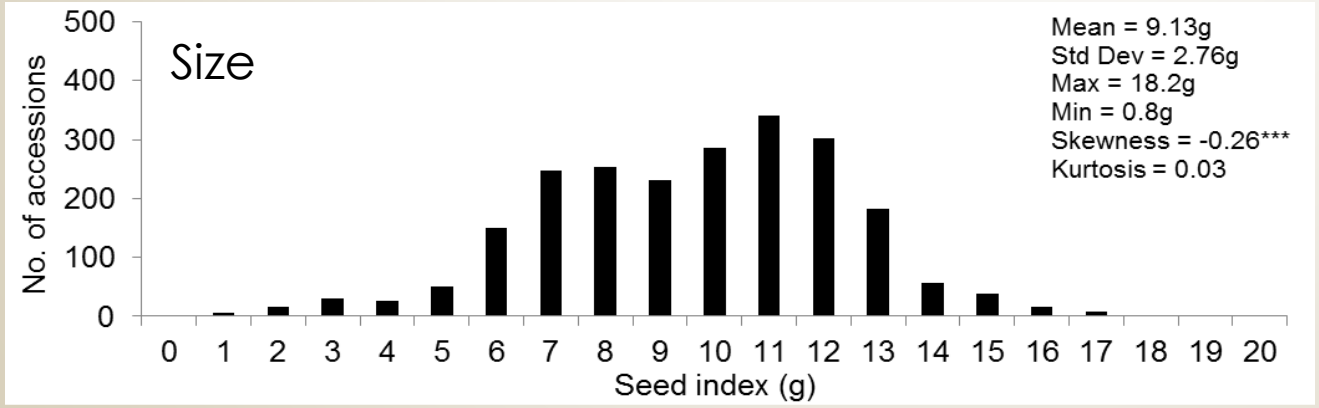
Also used this methodology to survey variation across U.S. *Gossypium* germplasm collection (with **L. Hinze**, R. Percy).



Seed Characteristics for 2250 accessions (US *Gossypium* Collection)







Distribution of oil, protein, and size for 2,256 cotton accessions in the *Gossypium* Diversity Reference Set.

Protein-- 10% to 36%

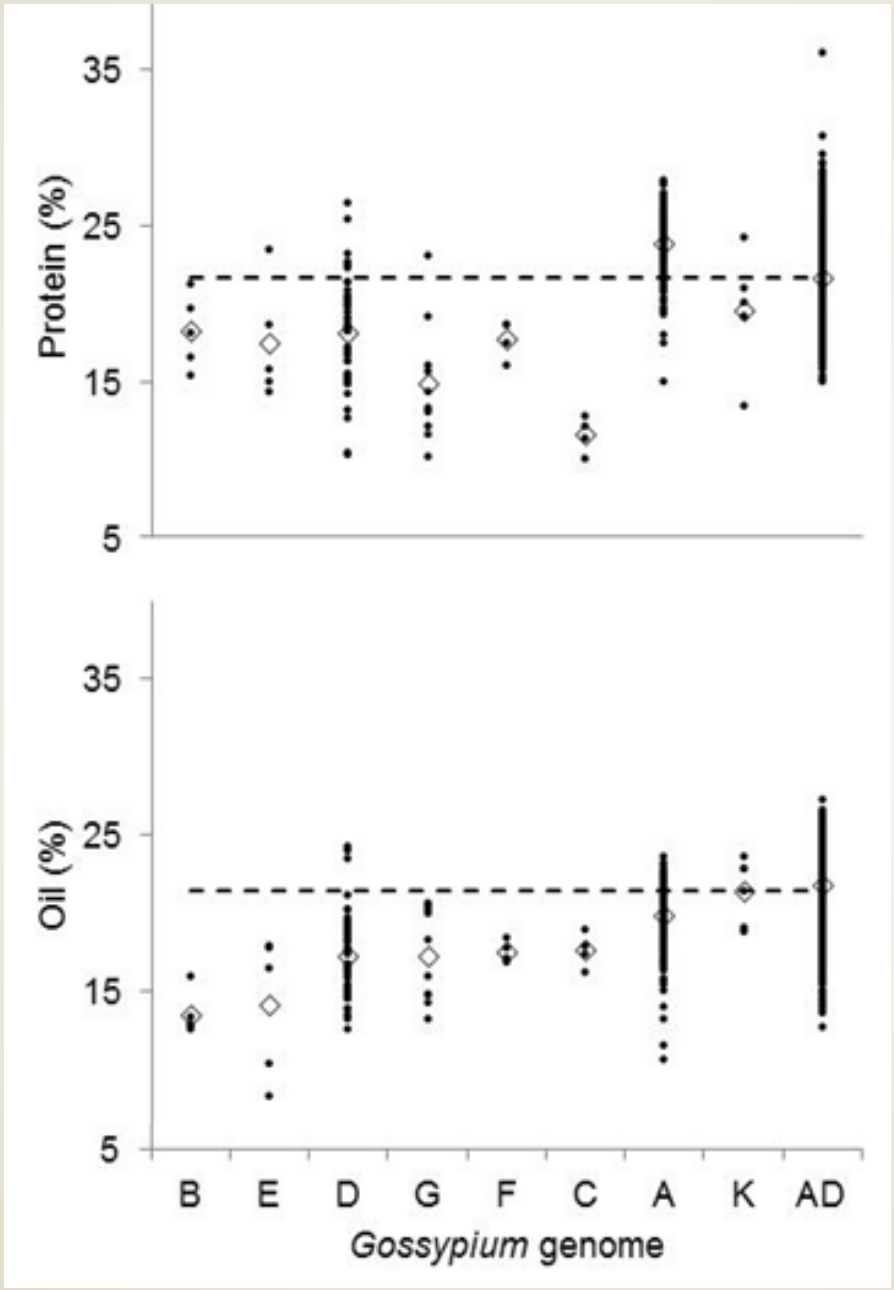
Oil – 8% to 27%

Size– 1g/100 to 18g/100

BIODISCOVERY INSTITUTE

Hinze et al., Crop Science, vol. 55 2015

Seed Characteristics for 2250 accessions (US *Gossypium* Collection)



Distribution of protein and oil by genome in the *Gossypium* Diversity Reference Set

Wide **range is captured in the tetraploid genomes-**

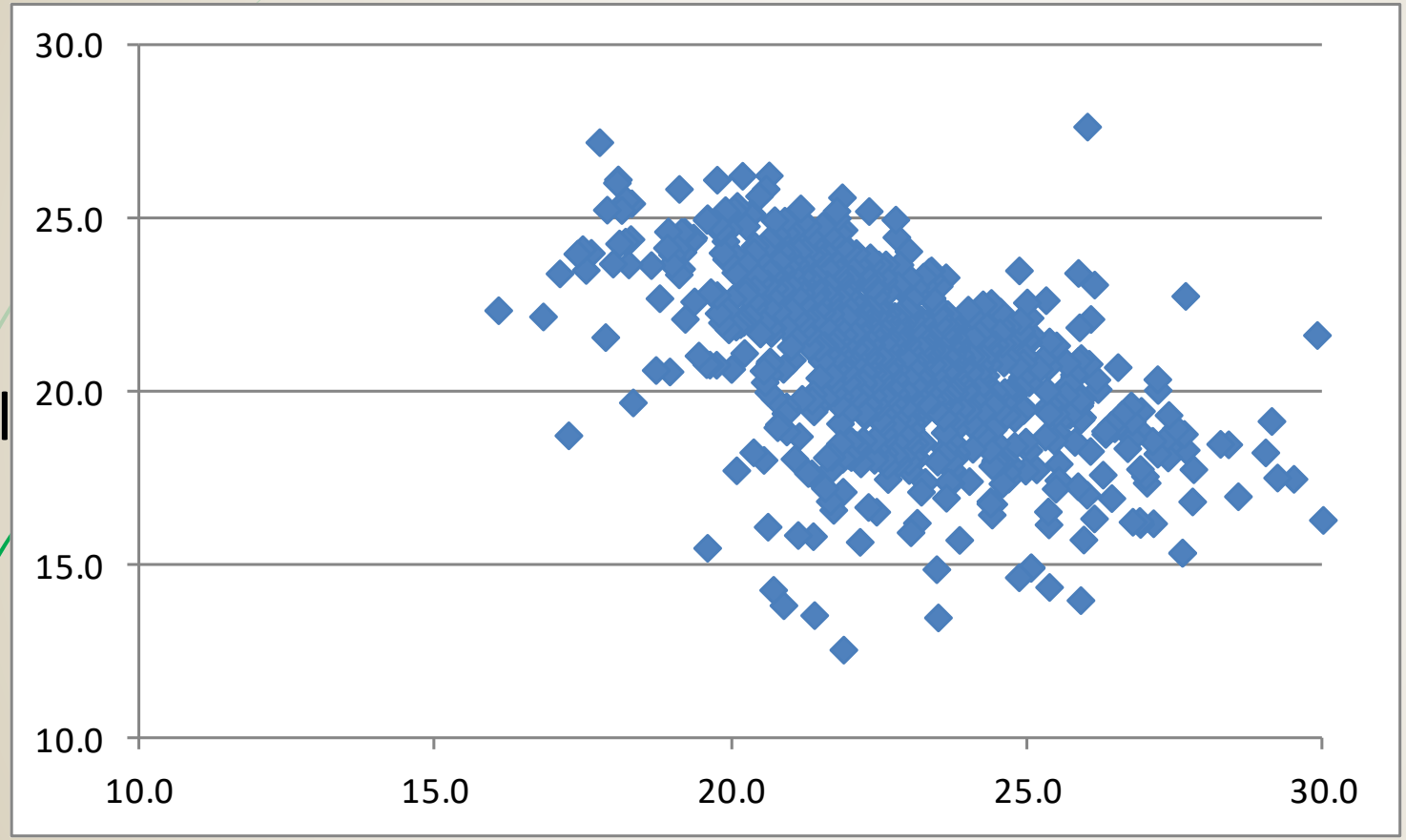
Significant variation in seed oil and protein exists within most contemporary breeding programs.

Hinze et al., Crop Science, vol. 55 2015

BIODISCOVERY INSTITUTE

Relationship of Seed Oil to Protein-

Gossypium hirsutum (AD-genome) "Texas" accessions



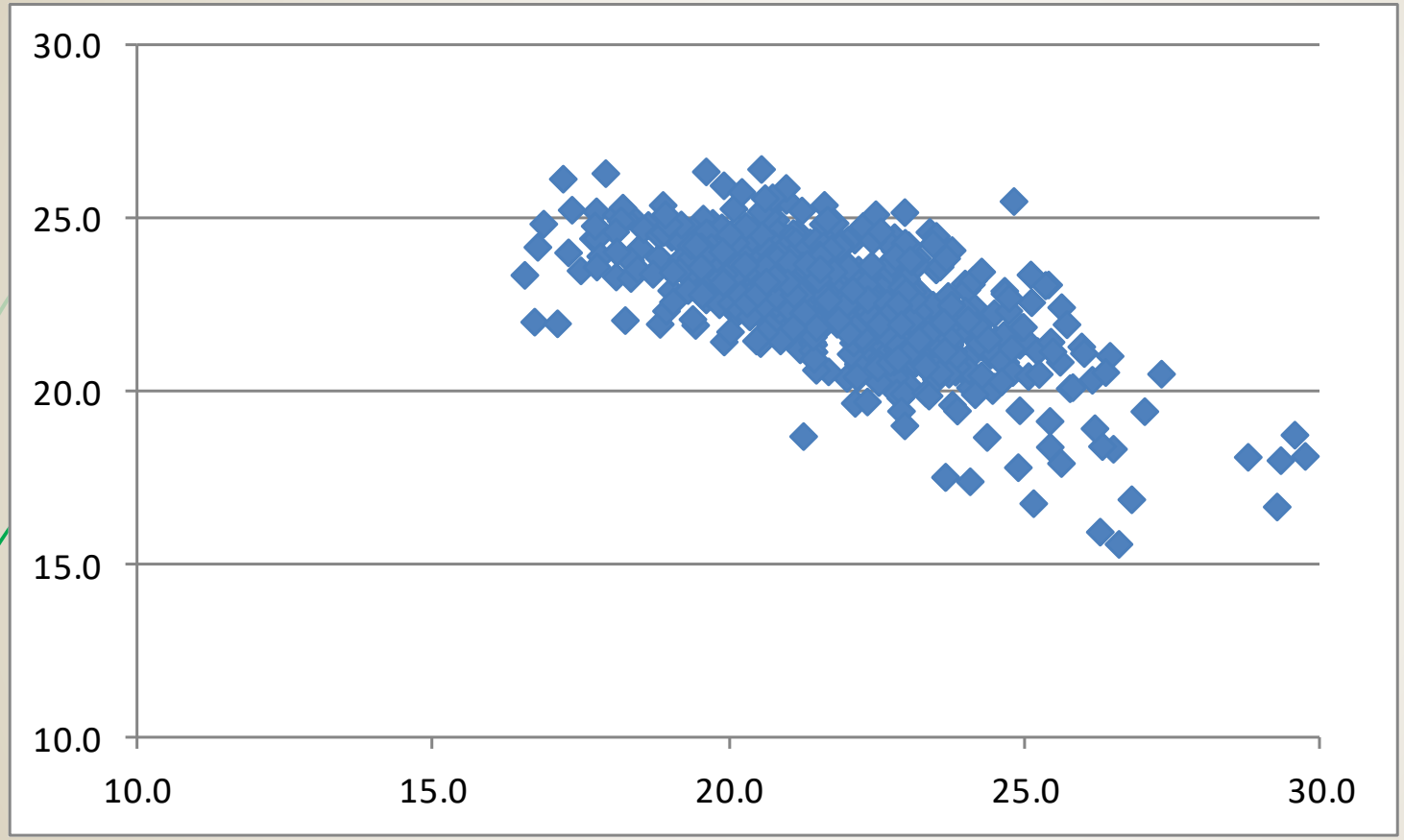
Strong **negative correlation between protein and oil**– often seen and reported by others. **Outliers evident**

Variation suggests opportunities for selection.

BIODISCOVERY INSTITUTE

Relationship of Seed Oil to Protein-

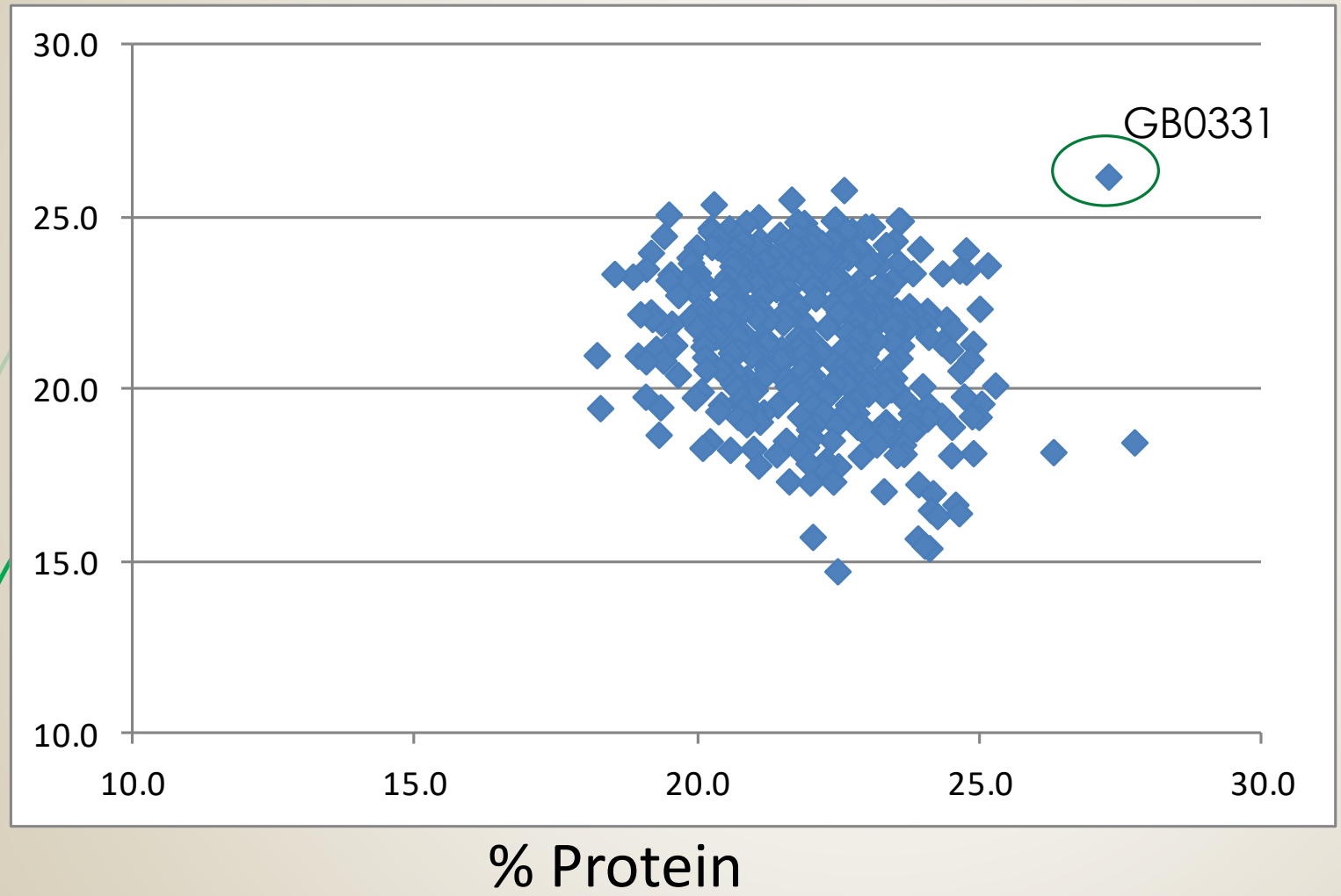
Gossypium hirsutum (AD-genome) "Stoneville" accessions



Strong **negative correlation between protein and oil**– often seen and reported by others.

BIODISCOVERY INSTITUTE

Relationship of Seed Oil to Protein-
Gossypium barbadense accessions

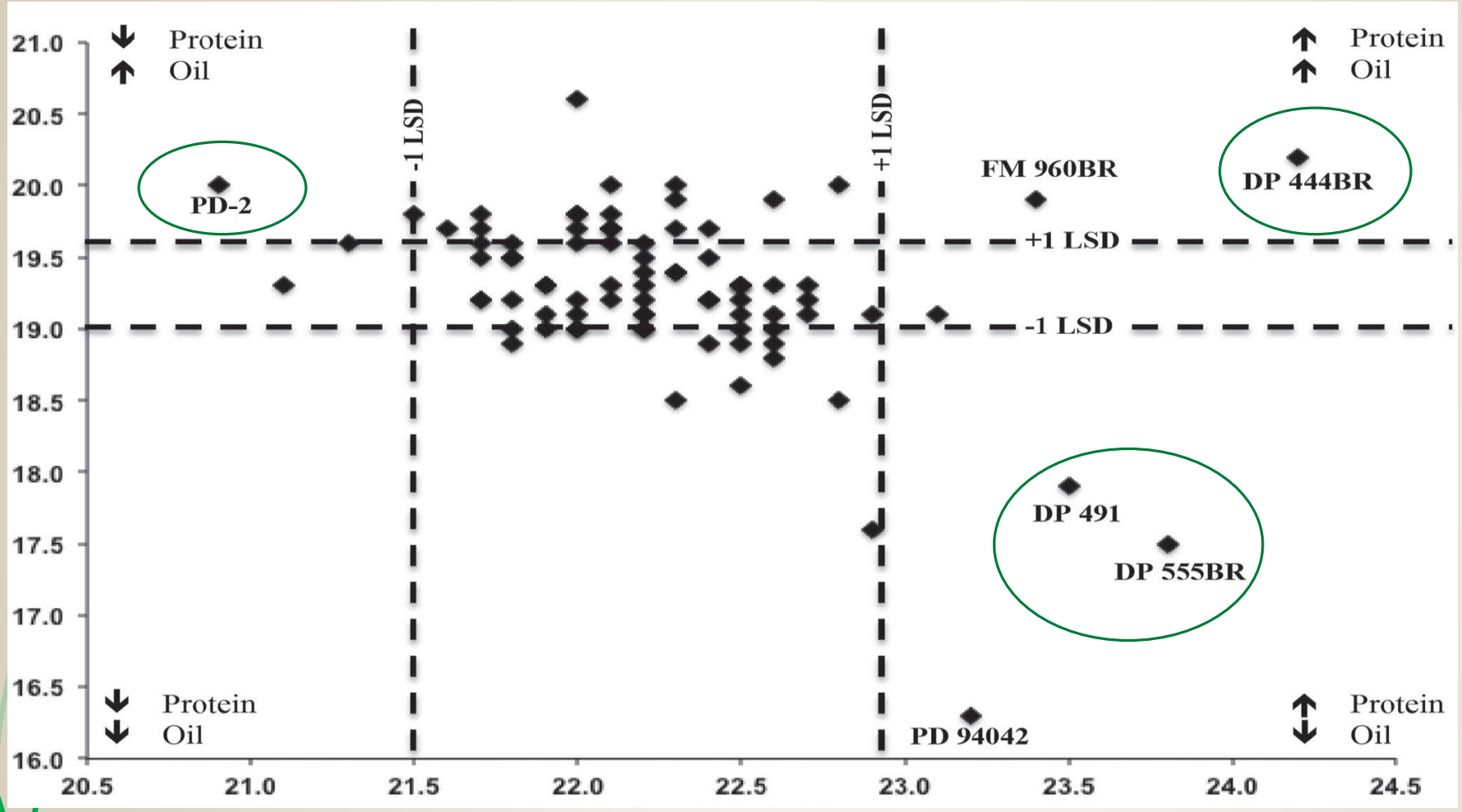


Less correlation between protein and oil–

But variation/outliers suggests opportunities for selection- e.g. GB0331 (high oil/ high protein)

BIODISCOVERY INSTITUTE

G.hirsutum Elite Varieties in the Pee Dee Collection



82 varieties, 2-5 locations over three years.

Strong **negative correlation** between protein and oil.

Also **significant environmental component**, especially for protein.

** Outliers suggest relationship can be disrupted

BIODISCOVERY INSTITUTE

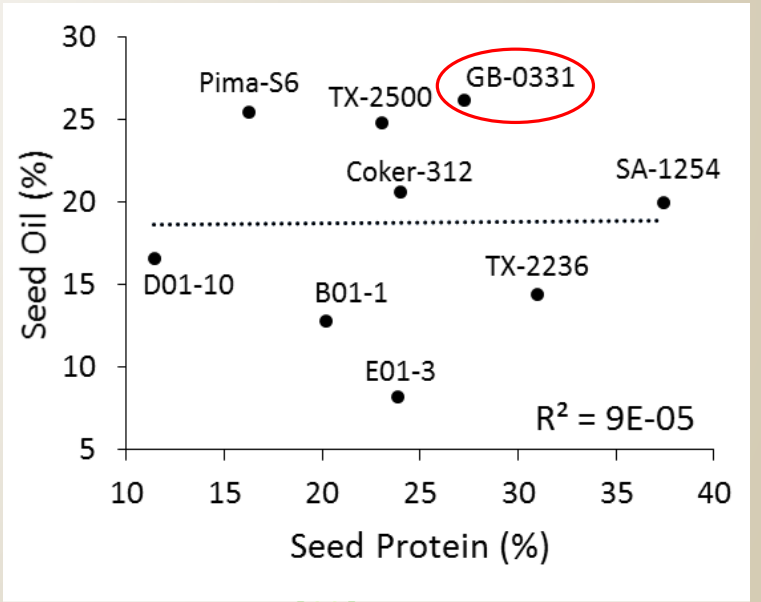
Hinze et al., Crop Science, vol. 55 2015

Examination of some outliers for oil/ protein content (US Collection)

Sample ID	Species	Genome/Race	Collected from	Seed Size (mg/seed)	Oil (%)	Protein (%)
B01-1	*G. anomalum*	B1	Africa	23.0 ± 1.5	12.8 ± 0.3	20.2 ± 0.3
D01-10	*G. thurberi*	D1	Mexico/Arizona	20.0 ± 3.0	16.6 ± 0.3	11.4 ± 0.3
E01-3	*G. stocksii*	E1	Arabia	23.8 ± 2.8	8.2 ± 0.3	23.8 ± 0.3
GB-0331	*G. barbadense*	(AD)2	South America	119.5 ± 10.1	26.2 ± 0.6	27.3 ± 0.8
PIMA-S6	*G. barbadense*	(AD)2	South America	129.9 ± 16.7	25.5 ±0.3	16.3 ± 0.4
SA-1254	*G. hirsutum*	(AD)1	Central/North America	114.8 ± 12.9	20.0 ± 0.4	37.4 ± 2.1
TX-2236	*G. hirsutum*	(AD)1	Central/North America	68.1 ± 8.1	14.4 ± 0.1	31.0 ± 0.4
TX-2500	*G. hirsutum*	(AD)1	Central/North America	83.6 ± 7.1	24.8 ± 0.1	23.0 ± 0.5
Coker-312	*G. hirsutum*	(AD)1	Central/North America	110.0 ± 6.8	20.6 ± 0.3	24.0 ± 1.5

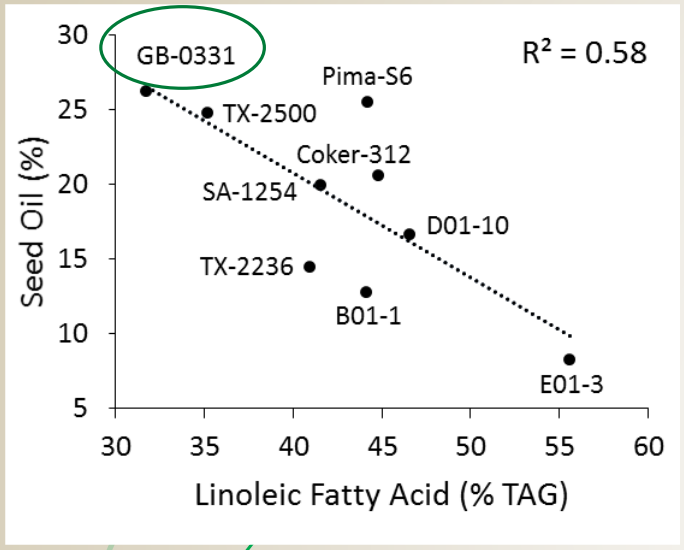
Genotypes with low oil/ high protein; high oil/ low protein; **high oil/high protein**.

** Outliers suggest relationship can be disrupted



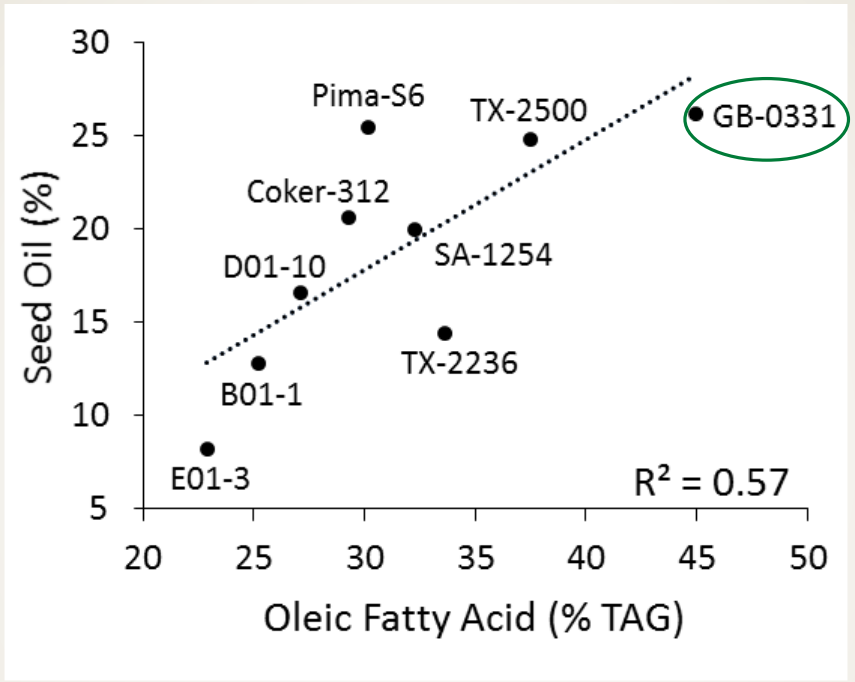
Sturtevant et al., Planta vol. 245, 2017

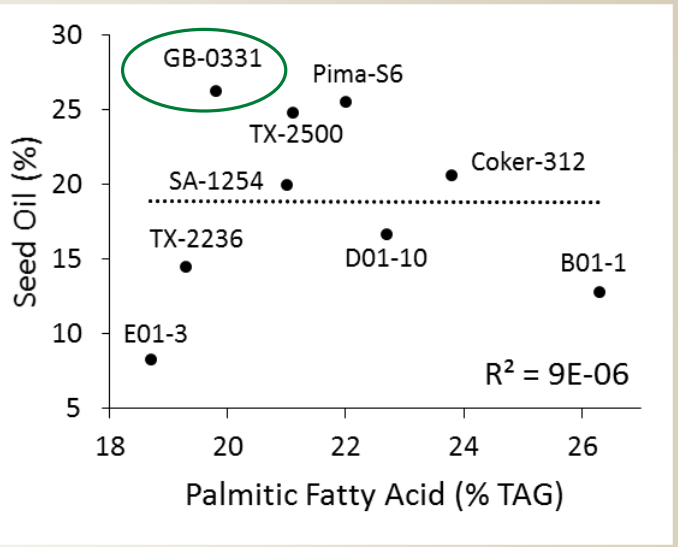
Examination of Some Outliers for Oil *Composition* (US Collection)



C18:2

"High- Oleic (C18:1)"





C16:0

** Genetic variation in Oleic Acid content (at expense of Linoleic)-
suggested variation in Fatty Acid Desaturase 2 (FAD2)

18:1 $\xrightarrow{\textbf{FAD2}}$ 18:2

BIODISCOVERY INSTITUTE

Sturtevant et al., Planta vol. 245, 2017

Molecular Identification of a Mutant Allele of FAD2-1 in GB-0331

a)

b)

GAGTGCCTTTACGTCGAGGCTGACGTTGGTGGTGGTGGTAGC**TGTTAAGAAATGGCTTAAAATTGGTAGTGGATTGTTGT**
TGTTGGTAGTGGGTTGTTGGTGGTAGTGGATTAGTGGATGGTTGTTGGTGTCCATTTTCAACCACAAATCTTTGCCAAAG
TTTTGGACAATTATGTCCTTGAACTCTCTTATAAATAGAGAGGTTCATTAGCCATATTCATCATCCCAAACCAAGAGAGA
GCAAAGCTTGTTCTTTGAAAGCTAGGATTTTAGCTTTCGGGTTTTCTATAGGGGTTGAGAGTTGTGAGGTTCTCGGGTTG
TGTCTTGAGTGTAAAACACTTGTAATCTTCATCTTGTTATAGTGAAATTTCTTTTCGCCTCTGCCCGTGGACGTAGGCAT
TAAAGCCGAACCACGTAAATCCTTGTGTTCACTTTATTTTTCGTTTCGGTCAATTTACTTGTAGTCATATCGGAGTTCTC
GAATCGATCCTTTCCGCAACAAATTGG

c)

```
FAD2-1-1D   MDAGGRMPIDGIKEENRGSVNRVPIEKPPFTLGQIKQAIPPHCFRRSLLRSFSYVVHDLCLASLFYYIATSYFHFLPQPF
fad2-1-1D   MGAGGRMPIDGIKEENRGSVNRVPIEKPPFTLGQIKQAIPPHCFRRSLLRSFSYVVHDLCLASLFYYIATSYFHFLPQPF
            *.************************************************************************

FAD2-1-1D   SYIAWPVYWVLQGCILTGVWVIAHECGHHAFSDYQWVDDTVGLILHSALLVPYFSWKISHRRHHSNTGSMERDEVFVPKP
fad2-1-1D   SYIAWPVYWVLQGCILTGVWVIAHECGHHAFSDYQWVDDTVGLILHSALLVPYFSWKISHRRHHSNTGSMERDEVFVPKP
            ********************************************************************************

FAD2-1-1D   KSKLSCFAKYFNNPPGRVLSLVVTLTLGWHMYLAFNVSGRYYDRLASHYNPYGPIYSERERLQVYISDAGIVAVIYVLYK
fad2-1-1D   KSKLSCFAKYFNNPPGRVLSLVVTLTLGWHMYLAFNVSGRYYDRLASHYNPYGPIYSERERLQVYISDAGIVAVIYVLYK
            ********************************************************************************

FAD2-1-1D   IAATKGLAWLLCTYGVPLLIVNAFLVLITYLQHTHSALPHYDSSEWDWFRGALSTIDRDYGVLNKVFHNITDTHVAHHLF
fad2-1-1D   IAATKGLAWLLCTYGVPLLIVNAFLVLITYLQHTHSALPHYDSSEWDWFRGALSTIDRDYGVLNKVFHNITDTHVAHHLF
            ********************************************************************************

FAD2-1-1D   STMPHYHAMEATKAIKPILGKYYPFDGTPIYKAMWREAKECLYVEADVGGGGSKGVFWYRNKF
fad2-1-1D   STMPHYHAMEATKAIKPILGKYYPFDGTPIYKAMWREAKECLYVEADVGGGGSC---------
            ************************************************
```

The mutant *fad2-1D-1* sequence was only identified in the GB-0331 accession (not Pima S6 or GB0332)

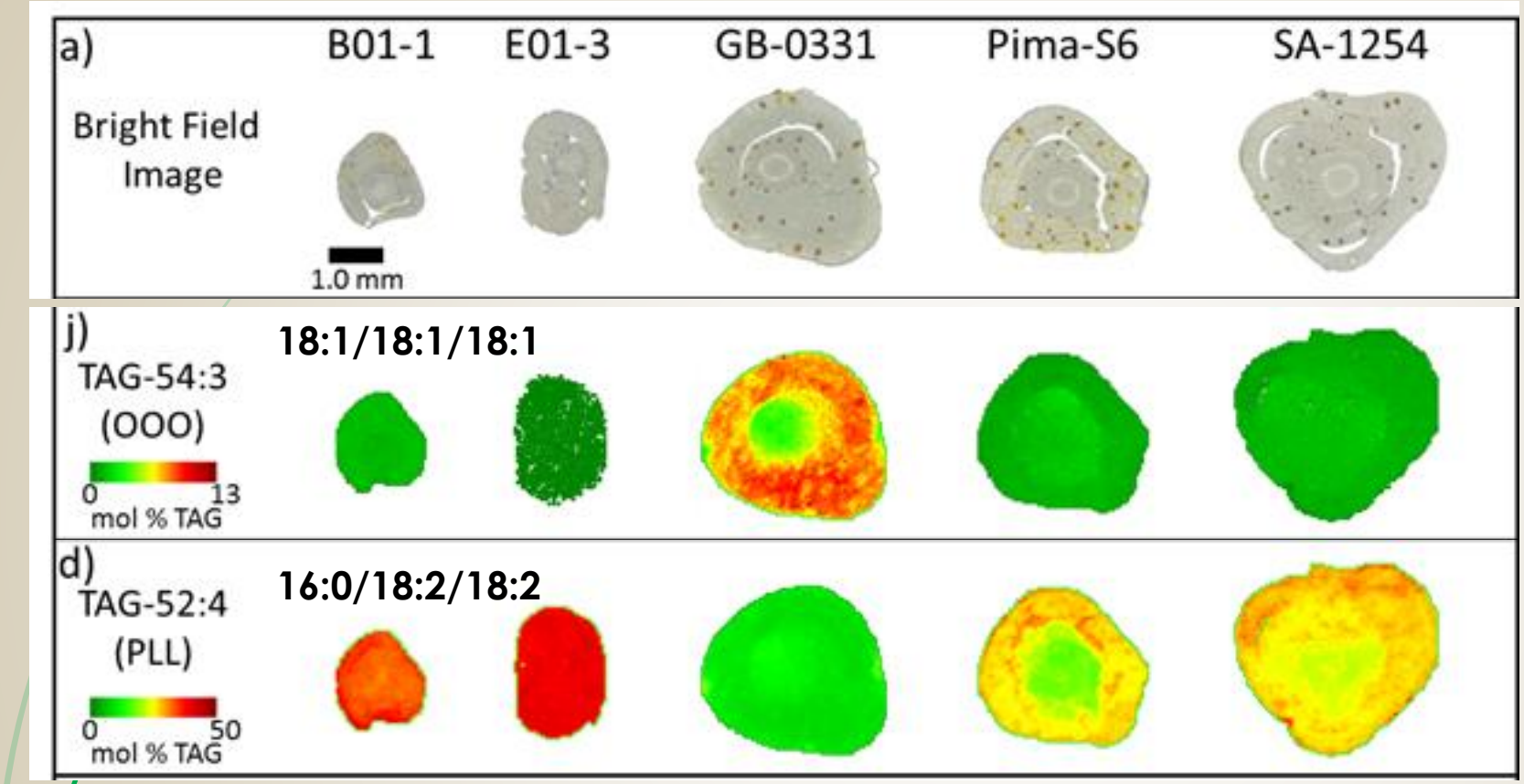
Mutation is a 90 bp insertion (red font). Patent issued 2019. Also found by Michael Dowd and Jay Shockey- USDA-ARS (back to back publications; *Planta*, 2017).

Mutation results in a truncation at the C-terminus– gene is expressed but normal subcellular targeting is disrupted.

Lori Hinze- USDA-ARS, crossing into non-photoperiodic germplasm.

BIODISCOVERY INSTITUTE

Examination of Some Outliers for Oil Distribution (US Collection)



Mass Spectrometry Imaging-

Tissue-specific distribution of individual oil molecules

***Variation Exists in Composition and Distribution within the Embryo** – Complexity of Seed Oil Metabolism

BIODISCOVERY INSTITUTE

Sturtevant et al., Planta vol. 245, 2017

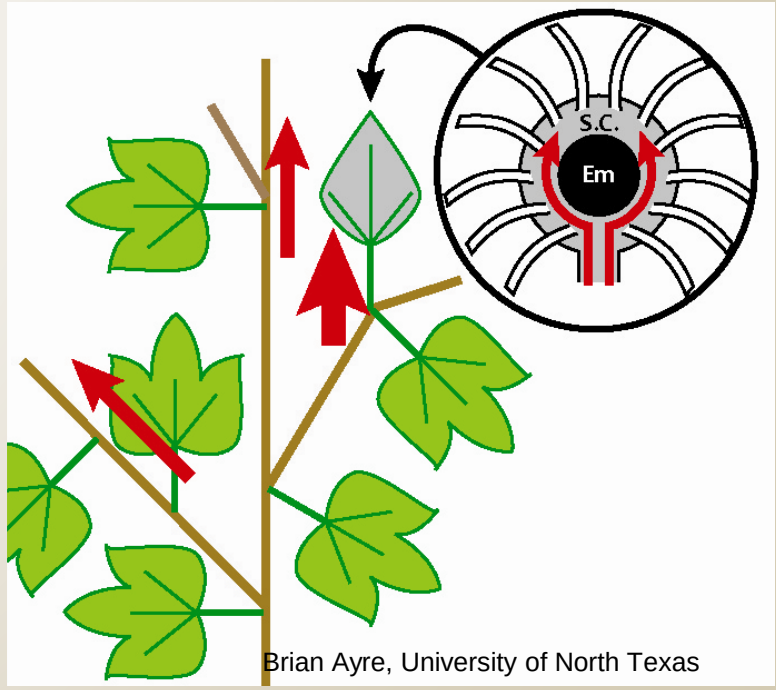
A Potential Future with a Holistic Focus on Seed Products?

Targeted strategies--- Ultra-low gossypol stacked with altered seed oil traits (high oleic oil, low saturate oil, cyclic FA, Vitamin E)-- conventional transgenics (many examples available)– gene editing approaches.

Untargeted Strategies- Marker and Genome-wide association studies appearing in the literature– also NAM (nested association mapping) populations offer opportunities to identify important loci.

More research needed- to understand 1) partitioning of resources between embryo (seed products) and seed coat (fiber) 2) complexity of synthesis and distribution pathways within the seed.

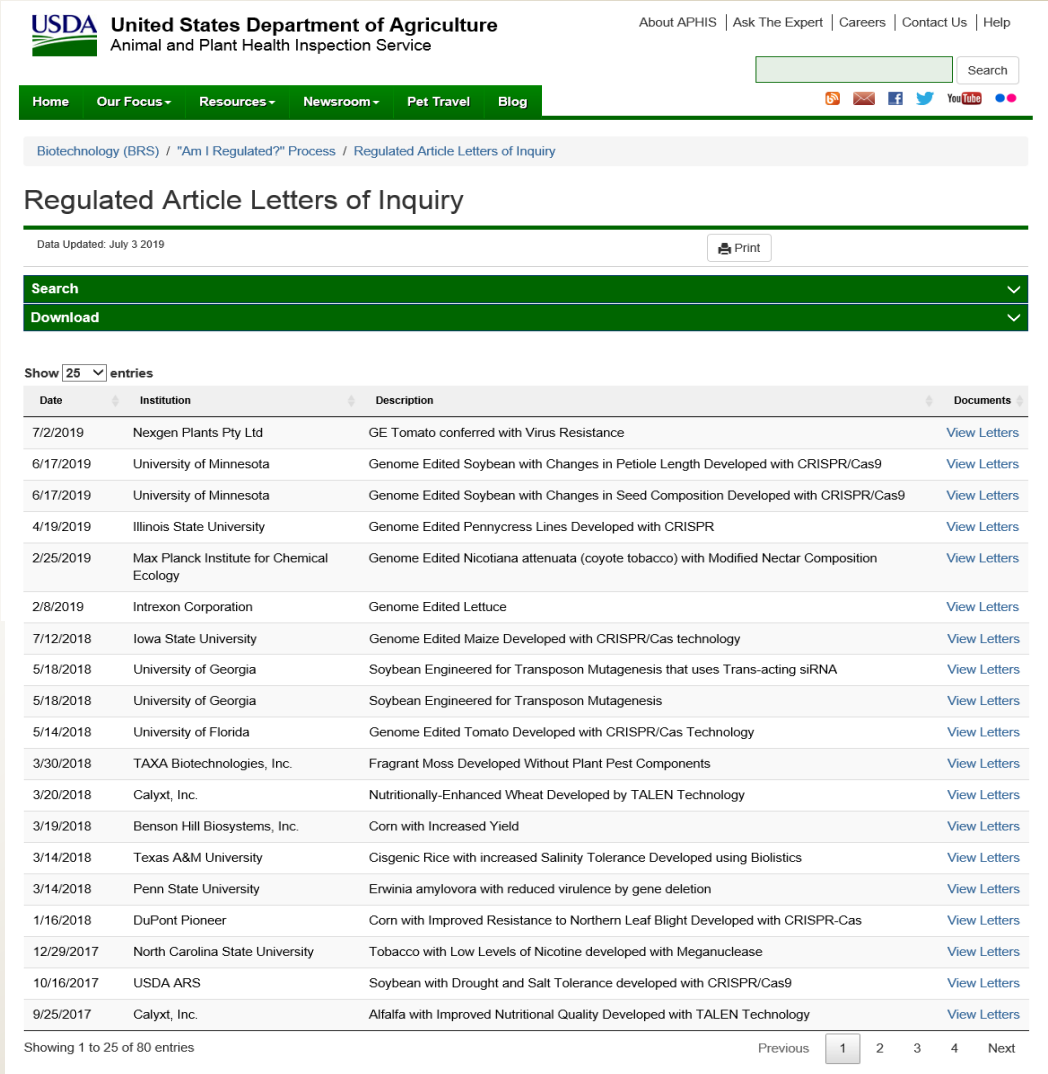
Identity Preservation of Seed Traits- To capture added value, needs to be a means to segregate traits during production.



Brian Ayre, University of North Texas

Identity preservation will be required to capture enhanced cottonseed value

- In the U.S., gene editing requires minimal regulatory cost since USDA-APHIS process is established.
- Most of the world still considers gene editing to be GMO thus unlikely that a company will incur the cost for global GMO regulatory approval of a cottonseed trait
- If growers, ginners and processors want gene-enhanced cottonseed value, the protein and whole seed must be kept out of specific export markets (EU) and risks need to be addressed.



USDA United States Department of Agriculture
Animal and Plant Health Inspection Service

About APHIS | Ask The Expert | Careers | Contact Us | Help

Search

Home | Our Focus | Resources | Newsroom | Pet Travel | Blog

Biotechnology (BRS) / "Am I Regulated?" Process / Regulated Article Letters of Inquiry

Regulated Article Letters of Inquiry

Data Updated: July 3 2019

Print

Search

Download

Show 25 entries

Date	Institution	Description	Documents
7/2/2019	Nexgen Plants Pty Ltd	GE Tomato conferred with Virus Resistance	View Letters
6/17/2019	University of Minnesota	Genome Edited Soybean with Changes in Petiole Length Developed with CRISPR/Cas9	View Letters
6/17/2019	University of Minnesota	Genome Edited Soybean with Changes in Seed Composition Developed with CRISPR/Cas9	View Letters
4/19/2019	Illinois State University	Genome Edited Pennycress Lines Developed with CRISPR	View Letters
2/25/2019	Max Planck Institute for Chemical Ecology	Genome Edited Nicotiana attenuata (coyote tobacco) with Modified Nectar Composition	View Letters
2/8/2019	Intrexon Corporation	Genome Edited Lettuce	View Letters
7/12/2018	Iowa State University	Genome Edited Maize Developed with CRISPR/Cas technology	View Letters
5/18/2018	University of Georgia	Soybean Engineered for Transposon Mutagenesis that uses Trans-acting siRNA	View Letters
5/18/2018	University of Georgia	Soybean Engineered for Transposon Mutagenesis	View Letters
5/14/2018	University of Florida	Genome Edited Tomato Developed with CRISPR/Cas Technology	View Letters
3/30/2018	TAXA Biotechnologies, Inc.	Fragrant Moss Developed Without Plant Pest Components	View Letters
3/20/2018	Calyxt, Inc.	Nutritionally-Enhanced Wheat Developed by TALEN Technology	View Letters
3/19/2018	Benson Hill Biosystems, Inc.	Corn with Increased Yield	View Letters
3/14/2018	Texas A&M University	Cisgenic Rice with increased Salinity Tolerance Developed using Biolistics	View Letters
3/14/2018	Penn State University	Erwinia amylovora with reduced virulence by gene deletion	View Letters
1/16/2018	DuPont Pioneer	Corn with Improved Resistance to Northern Leaf Blight Developed with CRISPR-Cas	View Letters
12/29/2017	North Carolina State University	Tobacco with Low Levels of Nicotine developed with Meganuclease	View Letters
10/16/2017	USDA ARS	Soybean with Drought and Salt Tolerance developed with CRISPR/Cas9	View Letters
9/25/2017	Calyxt, Inc.	Alfalfa with Improved Nutritional Quality Developed with TALEN Technology	View Letters

Showing 1 to 25 of 80 entries

Previous 1 2 3 4 Next

Kater Hake & Tom Wedegaertner, Cotton Inc

https://www.aphis.usda.gov/aphis/ourfocus/biotechnology/am-i-regulated

Identity Preservation for Enhanced Seed Value- System already in place can be adapted to capture value and maintain stewardship

COVERY
UTE

Kater Hake & Tom Wedegaertner, Cotton Inc

Concluding Remarks

> Cottonseed production on a mass basis exceeds fiber by about 50%, yet less than half of the seed is processed for added value.

> Considerable genetic variation exists in seed oil (and protein) content and composition. Likely other valuable components as well. Largely untapped opportunities.

> Seed trait development will require more focused research, more attention in selection (measurements), identity preservation for commercial success.

> Time is right- genomics resources increasingly available for cotton, promise of gene editing for targeted trait development, recent deregulation of ultra-low gossypol seed (new potential markets for meal), and some oil traits already available (GMO and non-GMO).

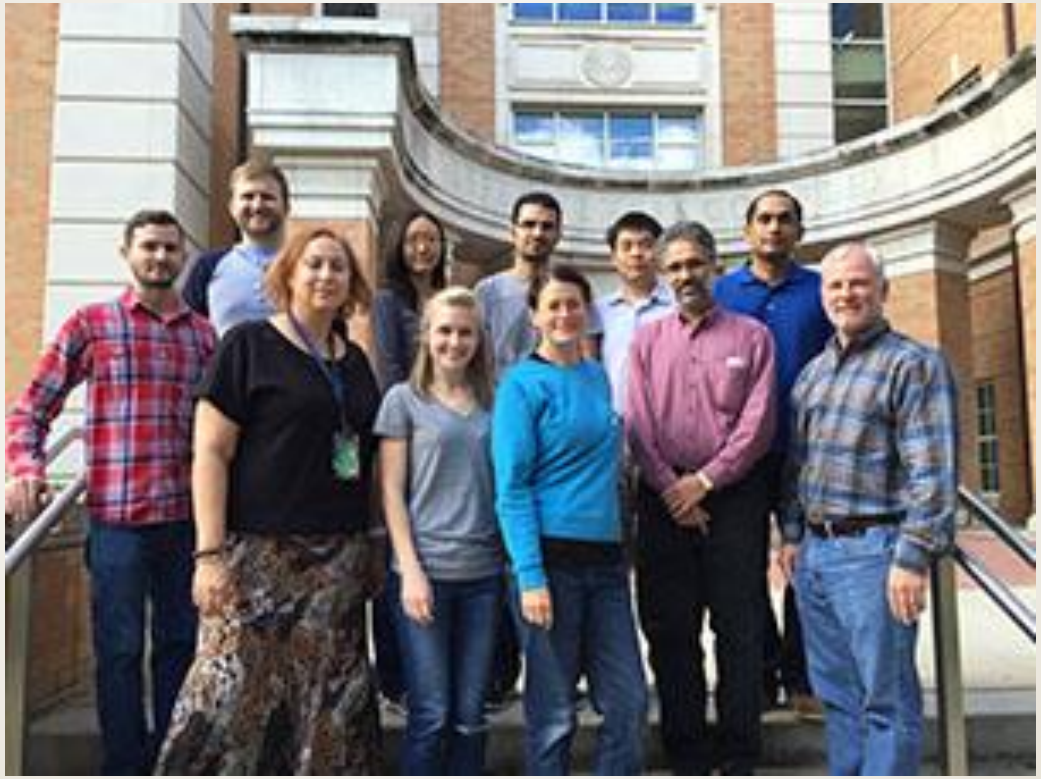
BIODISCOVERY INSTITUTE

Acknowledgements

University of North Texas
Patrick Horn, Drew Sturtevant,
Chris Kennedy, many undergrads

USDA-ARS
Todd Campbell
Michael Dowd, Jay Shockey
Lori Hinze, Richard Percy

Cotton Incorporated
Kater Hake
Tom Wedegaertner

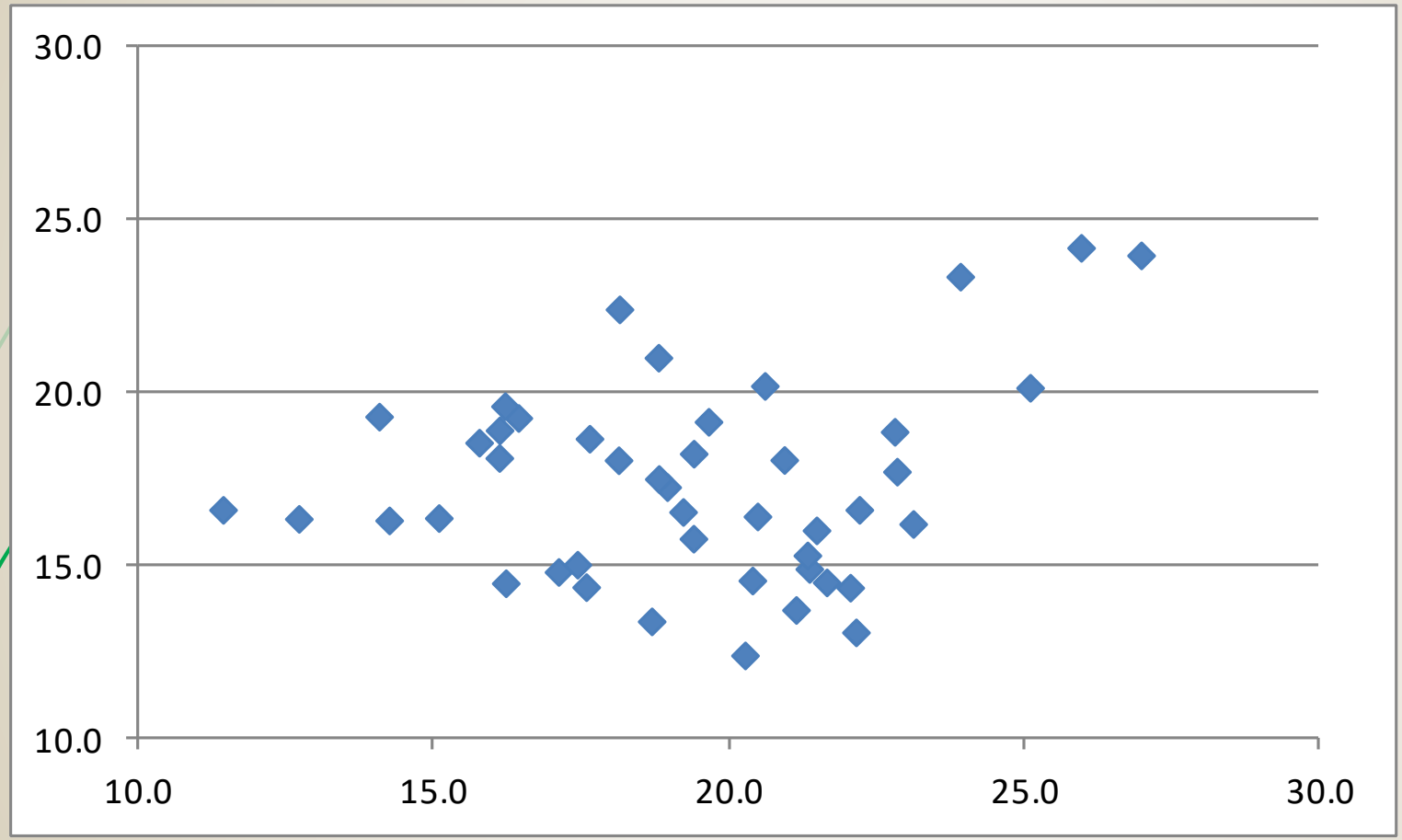


Funding
Cotton Incorporated

UNT
UNIVERSITY OF NORTH TEXAS
EST. 1890

BIODISCOVERY INSTITUTE

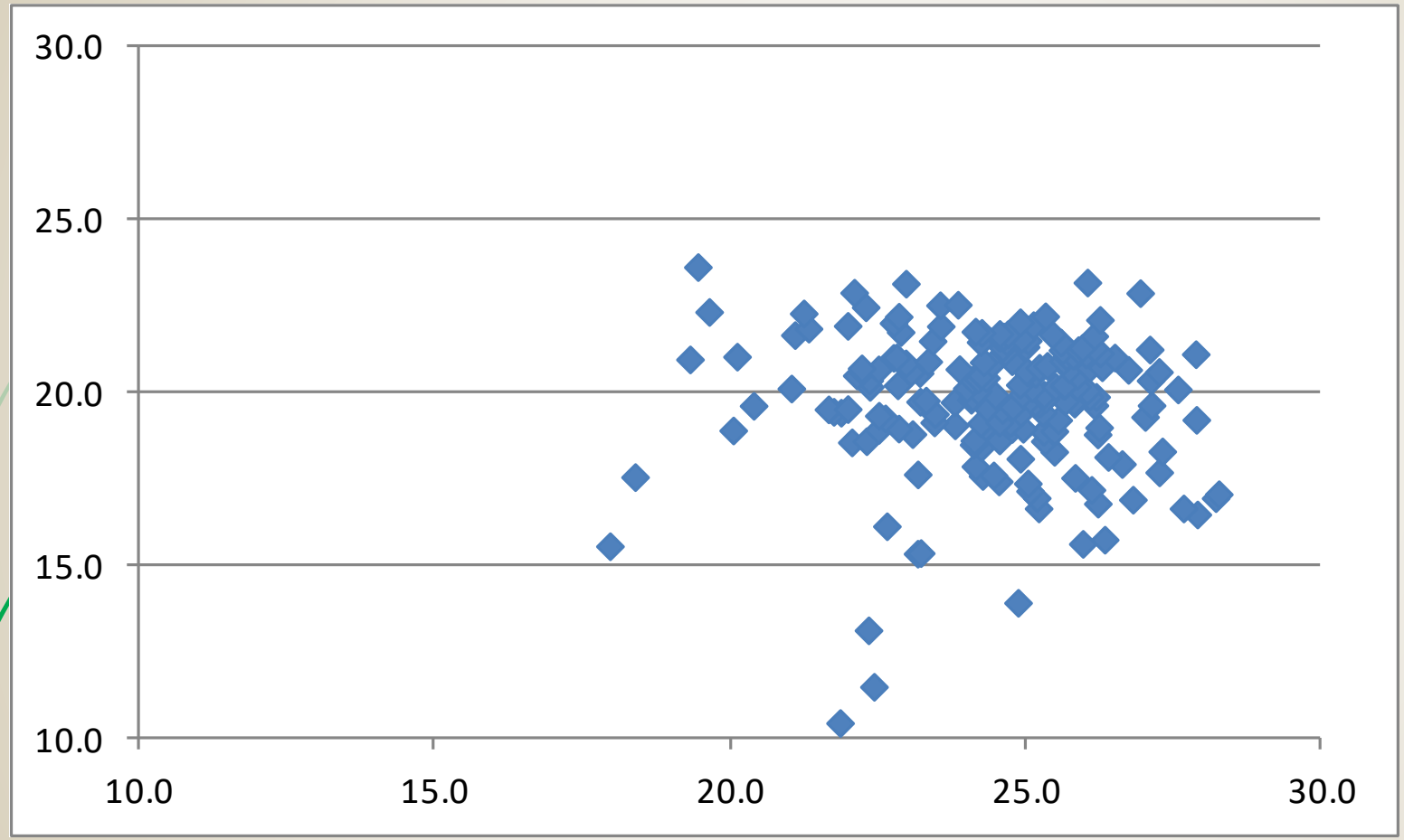
Relationship of Seed Oil to Protein-
Gossypium D-genome accessions



Less of an obvious relationship in diploid genomes– **more variation**.

BIODISCOVERY INSTITUTE

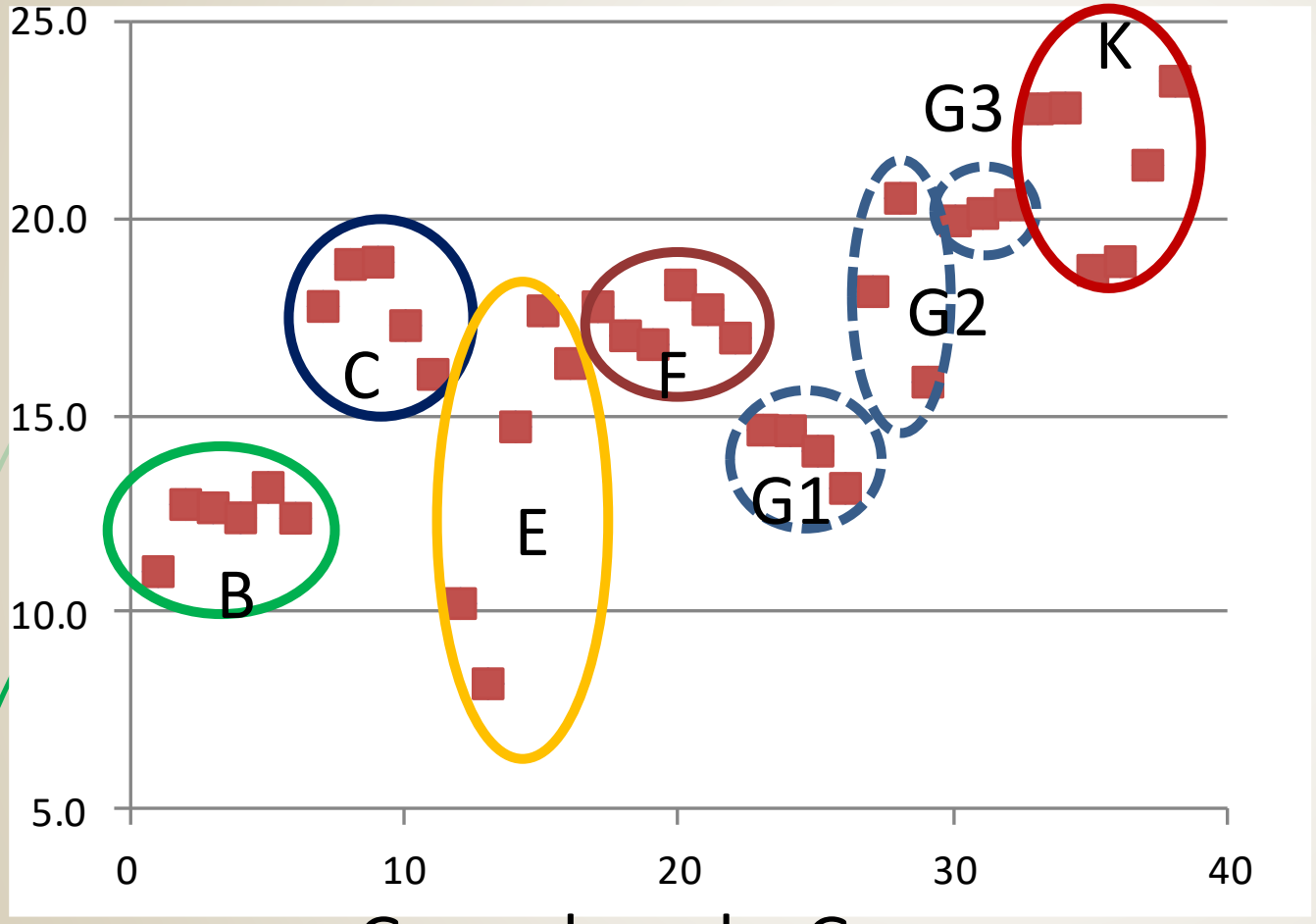
Relationship of Seed Oil to Protein-
Gossypium A-genome accessions



Less of an obvious relationship in diploid genomes– **more variation**.

BIODISCOVERY INSTITUTE

Seed Oil Content of Selected Exotic *Gossypium* Germplasm/ Genomes



Oil Content ranges from low of 8% in "E" to high of 24% in "K"

Additional opportunities?

BIODISCOVERY INSTITUTE