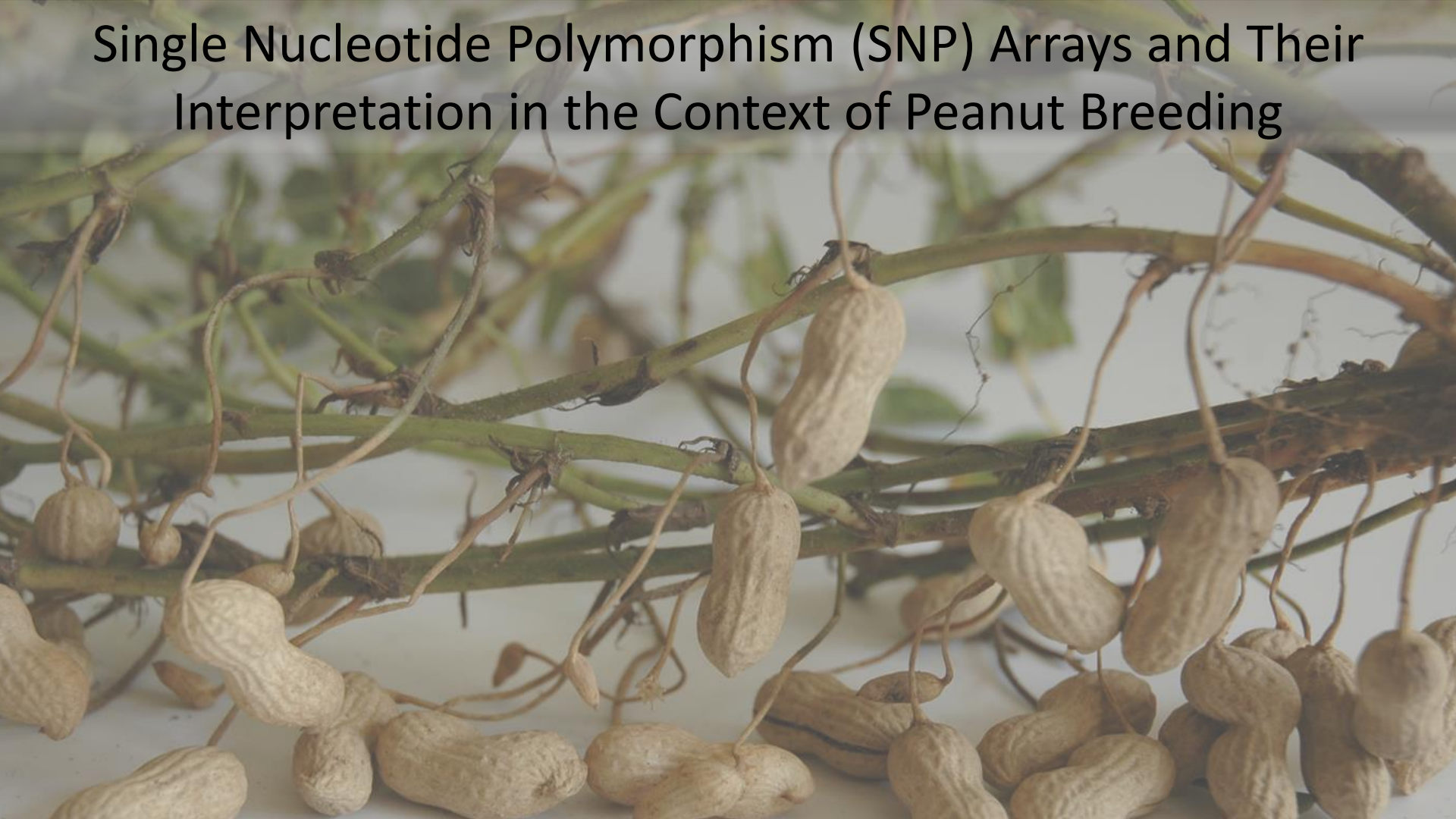


# Single Nucleotide Polymorphism (SNP) Arrays and Their Interpretation in the Context of Peanut Breeding



Ye Chu, Josh Clevenger, Walid Korani, David Bertoli, Soraya Bertoli, Davis Gimode, Stephanie Botton, Carolina Chavarro, Chandler Levinson, Larissa Arrais Guimaraes, Kathleen Marasigan, Bárbara Müller, Scott Jackson, Daniel Fonceka, Tom Isleib, Tom Stalker, Corley Holbrook, Peggy Ozias-Akins



**Institute of Plant Breeding,  
Genetics and Genomics**  
*College of Agricultural & Environmental Sciences*  
**UNIVERSITY OF GEORGIA**

AAGB 2018, 12-15 Nov, Saly, Senegal

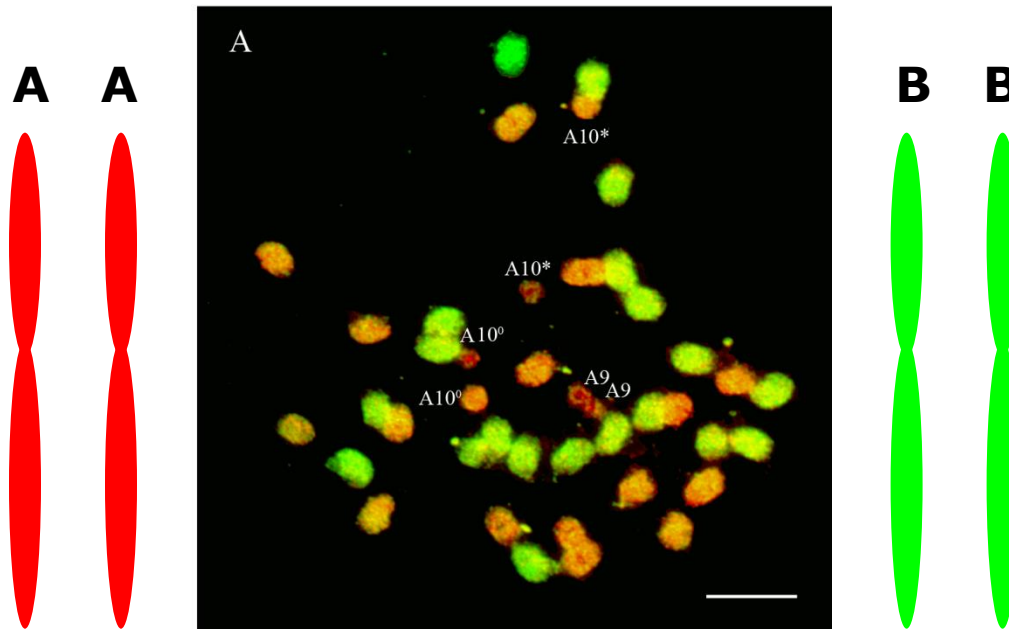
# Marker-Assisted Breeding

---

- Genetic purity
- Confirm crosses
- Whole genome mapping
- Trait association –
  - High oleic/linoleic acid ratio
  - Nematode resistance
  - Leaf spot resistance
  - Virus resistance
  - Seed size



# Peanut is an allotetraploid species with two subgenomes

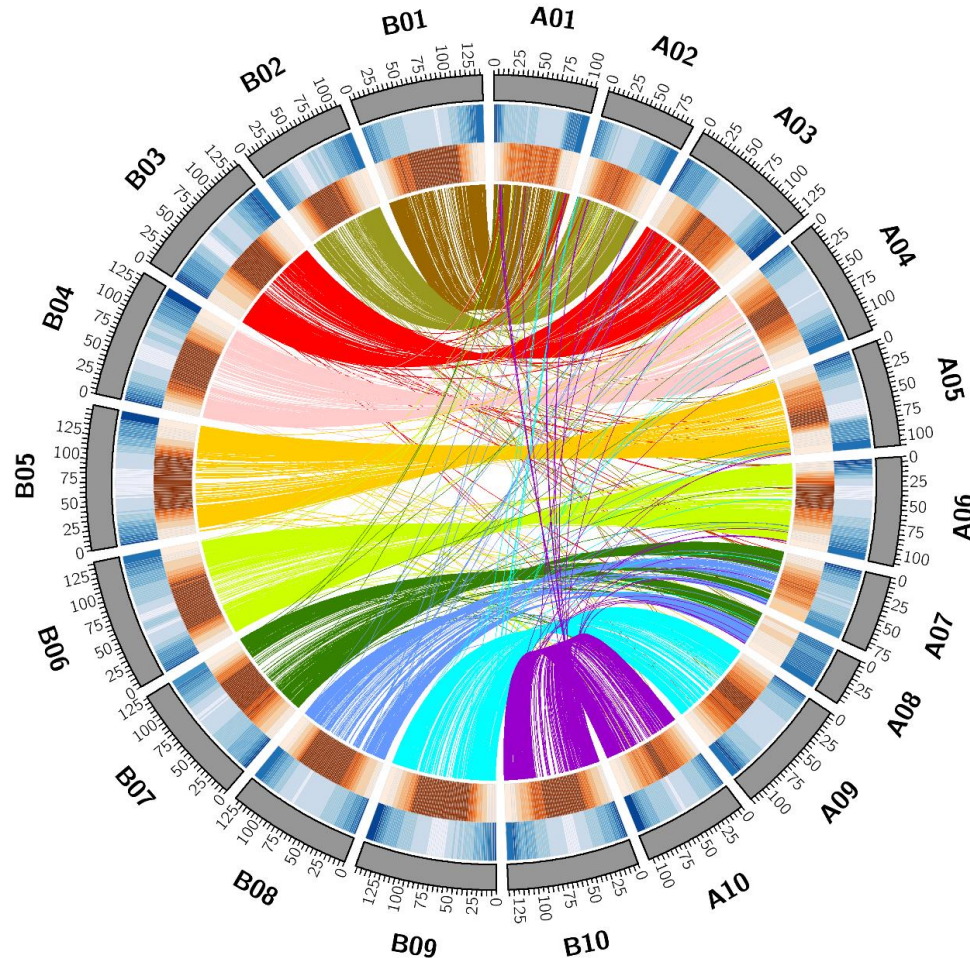


- A subgenome (*A. duranensis*)
- B subgenome (*A. ipaensis*)

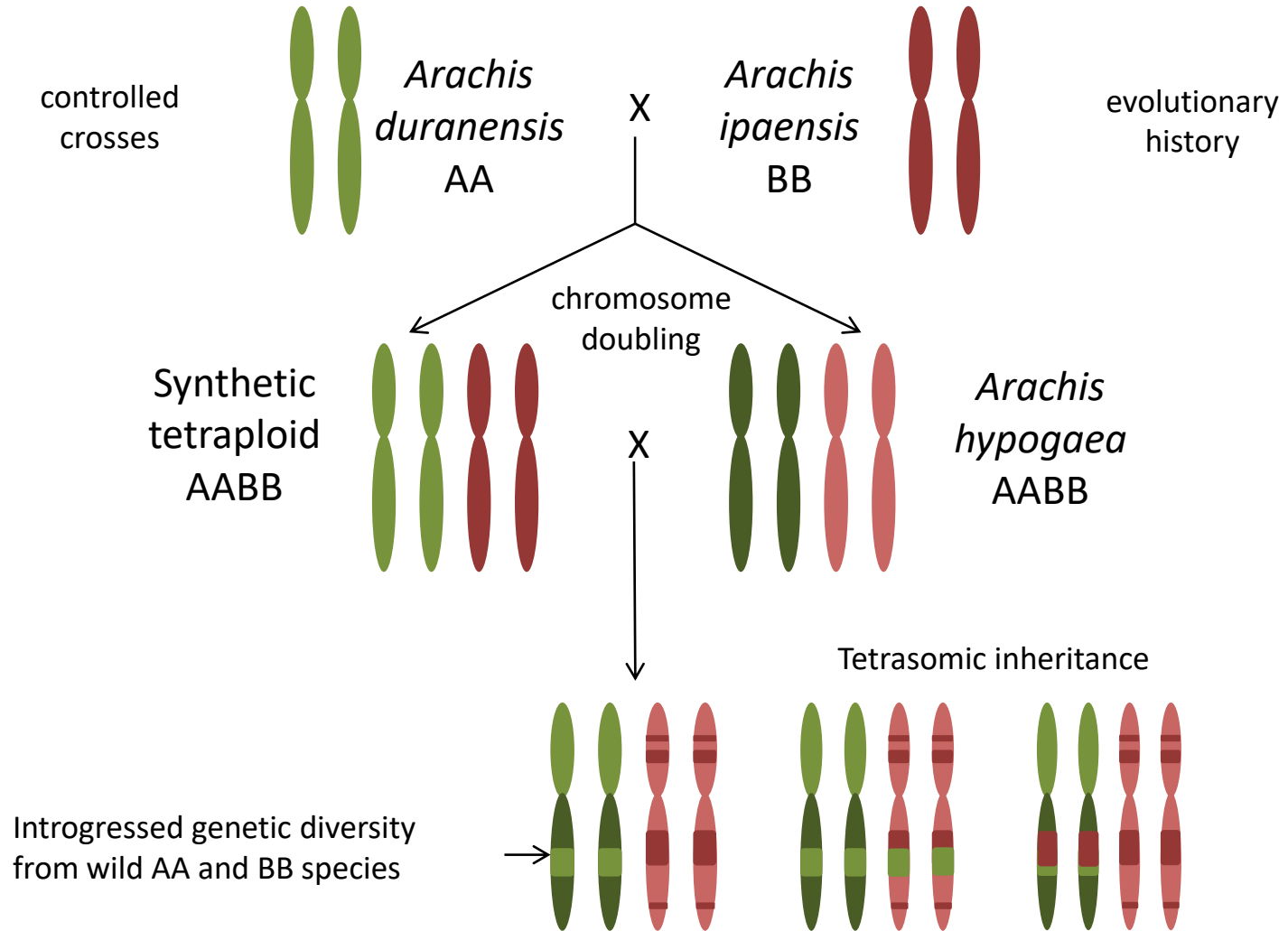
Ramos *et al.*, 2006. *Mol Gen Genomics*. 275: 578-592.



# Remarkable Similarity Between A & B Genomes and Tetraploid Subgenomes



# Expanding Genetic Diversity



# Computational Challenge to Identify True SNPs



sub genome 1

reference



sub genome 2

TCCGCGGGCCACCTCACGGCGGGACGTCCTCGACG

TCCG**A**GGCCACCT**T**ACGGCGGGACGTCCTC**A**ACG

TCCG**A**GGCCACCT**T**ACGGCGGGACGTCCTC**A**ACG

TCCGCGGGCCACCTCACGGCGGGACGTCCTCGACG

TCCG**A**GGCCACCT**T**ACGGCGGGACGTCCTC**A**ACG

TCCGCGGGCCACCTCACGGCGGGACGTCCTCGACG

TCCGCGGGCCACCTCACGGCGGGACGTCCTCGACG

TCCGCGGGCCACCTCACGGCGGGACGTCCTCGACG

TCCG**A**GGCCACCTCACGGCGGGACGTCCTC**A**ACG

TCCGCGGGCCACCTCACGGCGGGACGTCCTCGACG

TCCG**A**GGCCACCTCACGGCGGGACGTCCTC**A**ACG



# Peanut SNP Arrays

---

Developed Axiom\_Arachis1 and Arachis2 SNP arrays (former PhD students Josh Clevenger and Walid Korani applied novel computational pipelines and machine learning)

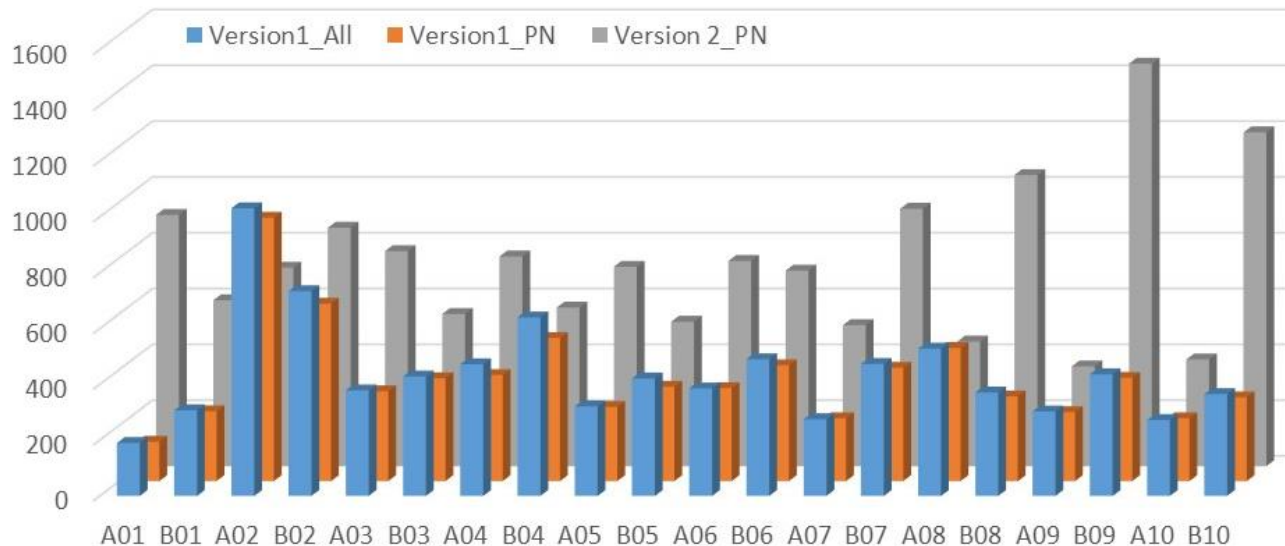
- 58,233 and 48,000 total features, respectively, with 15,287 and 30,539 polymorphic in diverse tetraploids
- 54,564 and 33,763 markers, respectively, that produced high quality, polymorphic clusters between diploid species
- 47,116 polymorphic markers between cultivated and interspecific hybrids



# Axiom\_Arachis2 Array SNP Distribution Across Genome

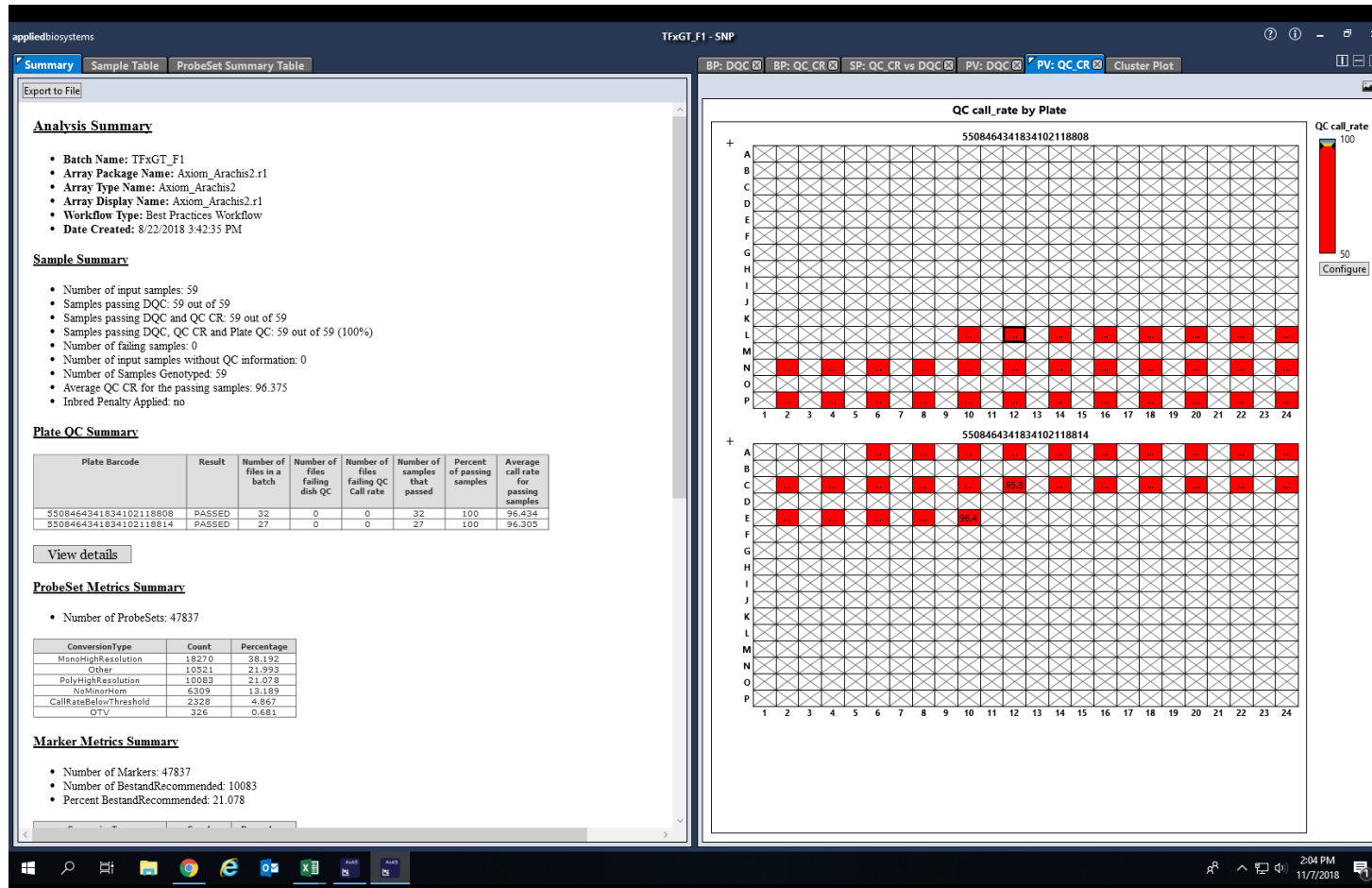
- 48,000 total markers
- 4,489 simple markers from v1
- 7,861 interspecific markers from v1
- 35,650 new markers

30,539 *A. hypogaea* markers (double v1)  
33,763 diploid wild species markers

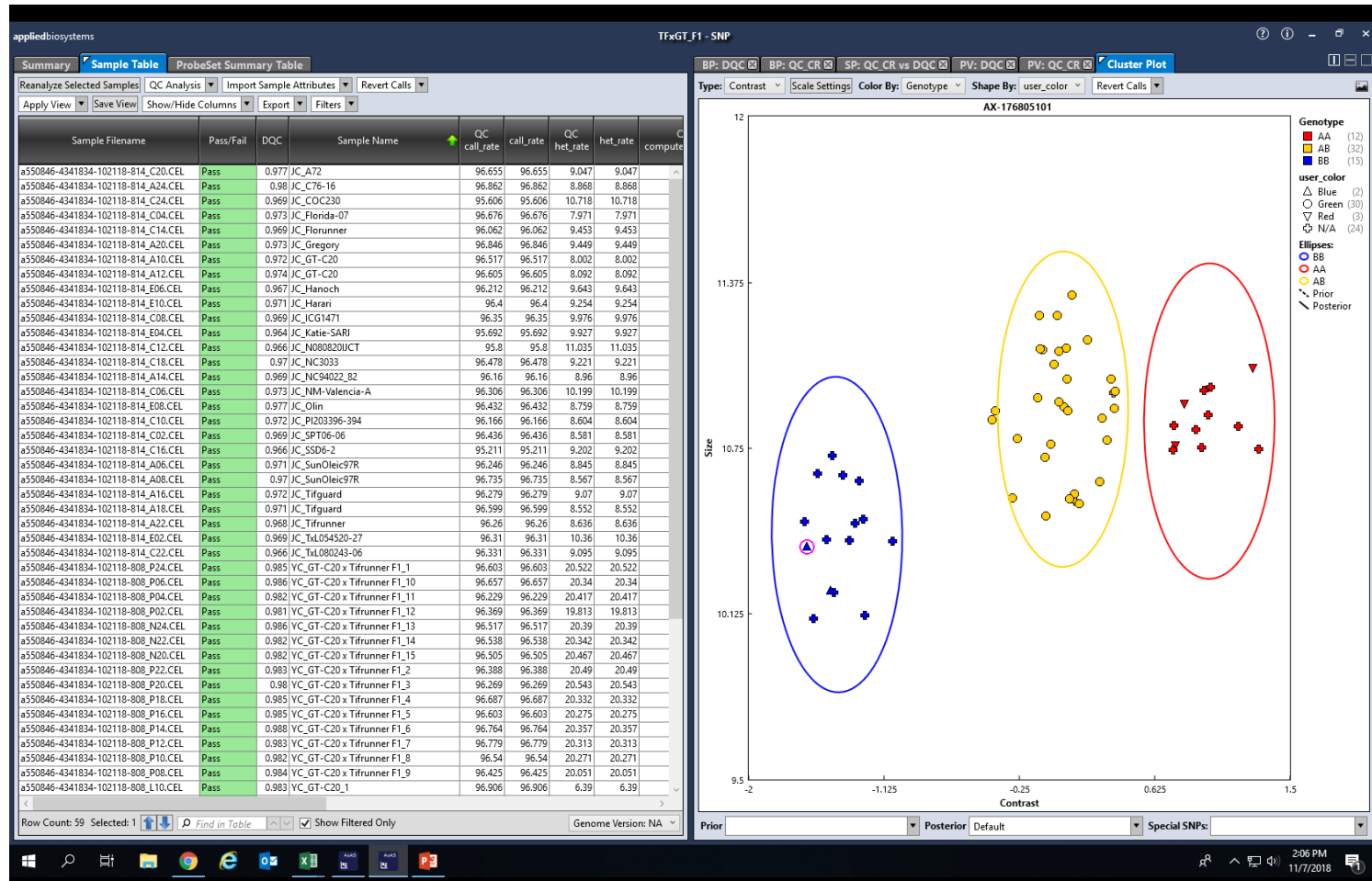




# Axiom Analysis Suite

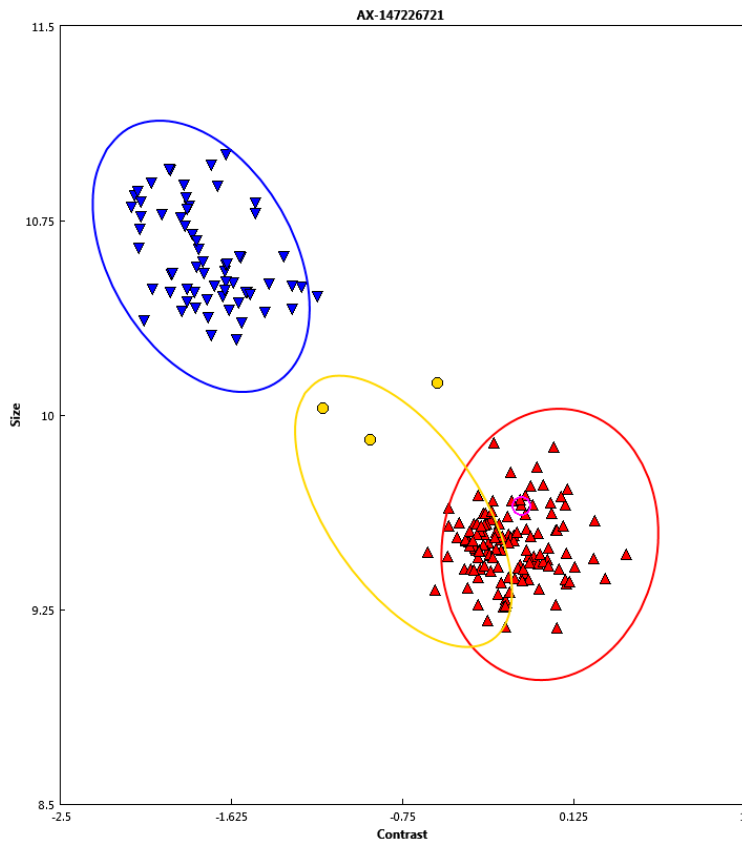


# Axiom Analysis Suite

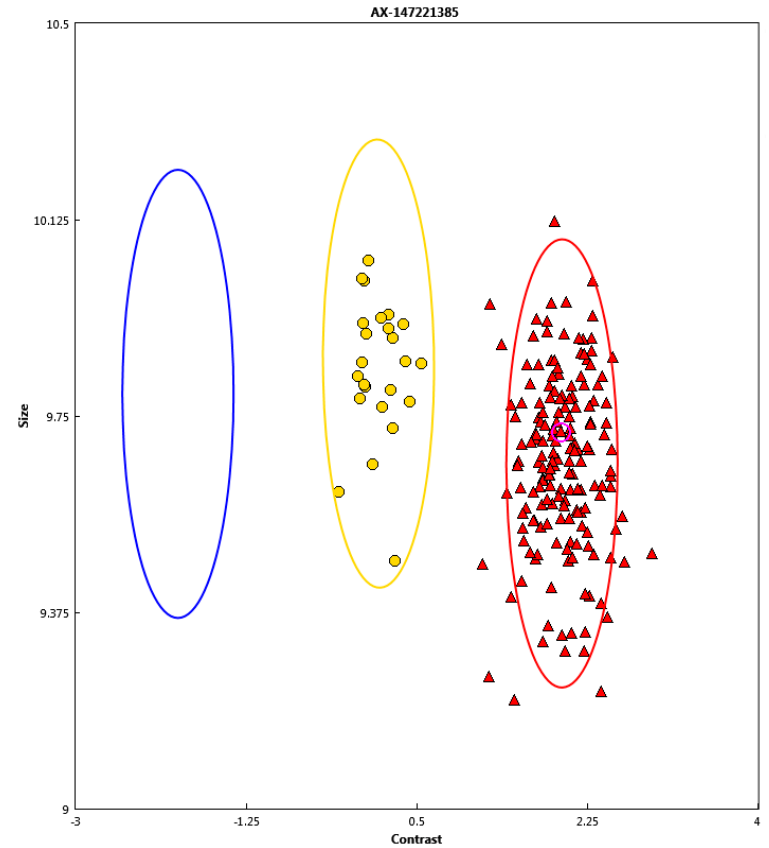


# Axiom SNP Calls

Poly High Res

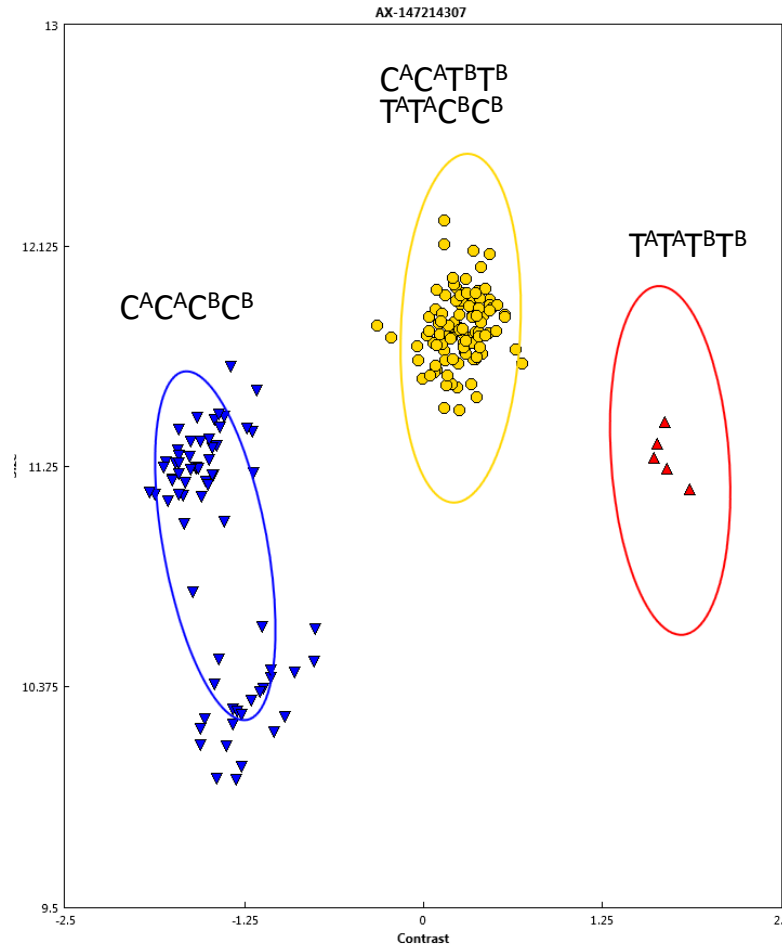


No Minor Homo



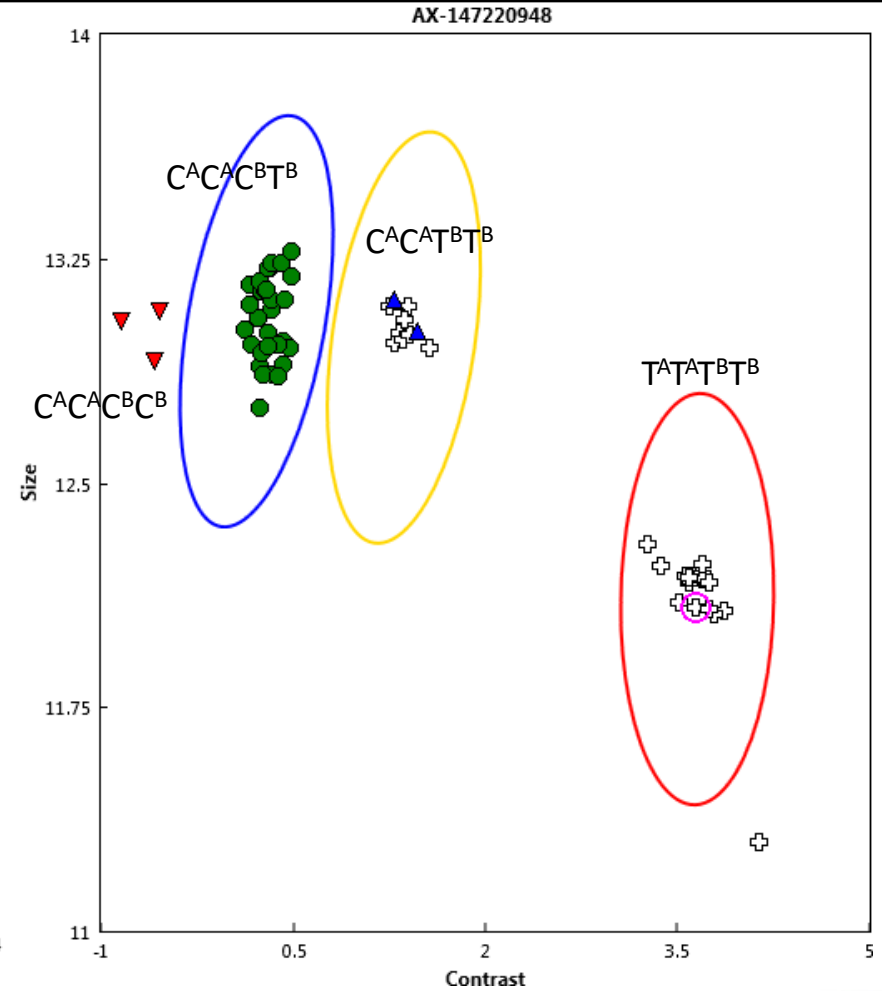
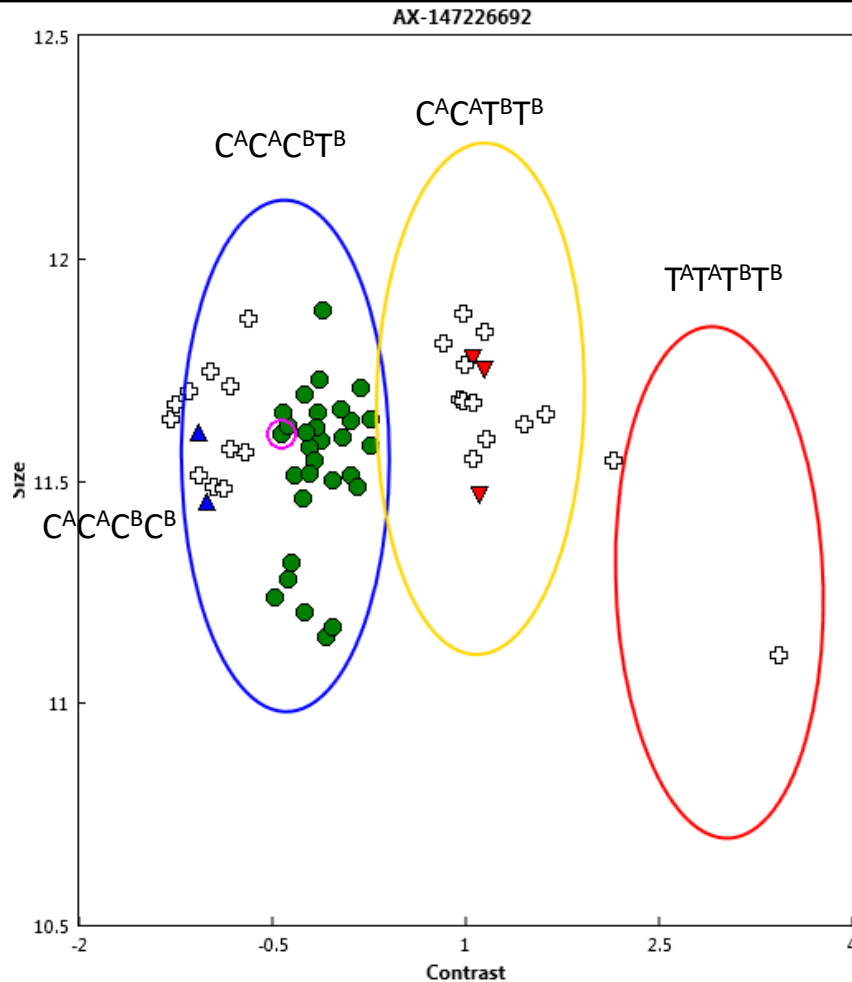
# Genotype Clusters for Inbred Peanut Lines

Heterozygote class is not a true het, but detection of A/B subgenomes

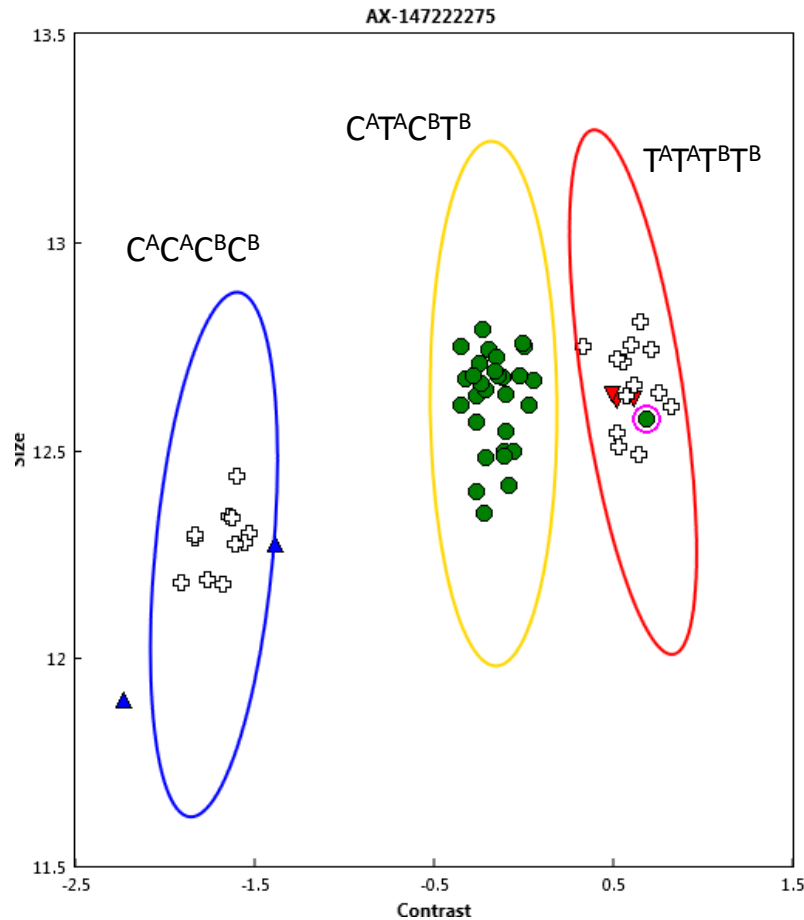


# Clustering Dependent on Analysis Group

## Examples from Biparental Cross and Multiple F1s



# Clustering Dependent on Analysis Group

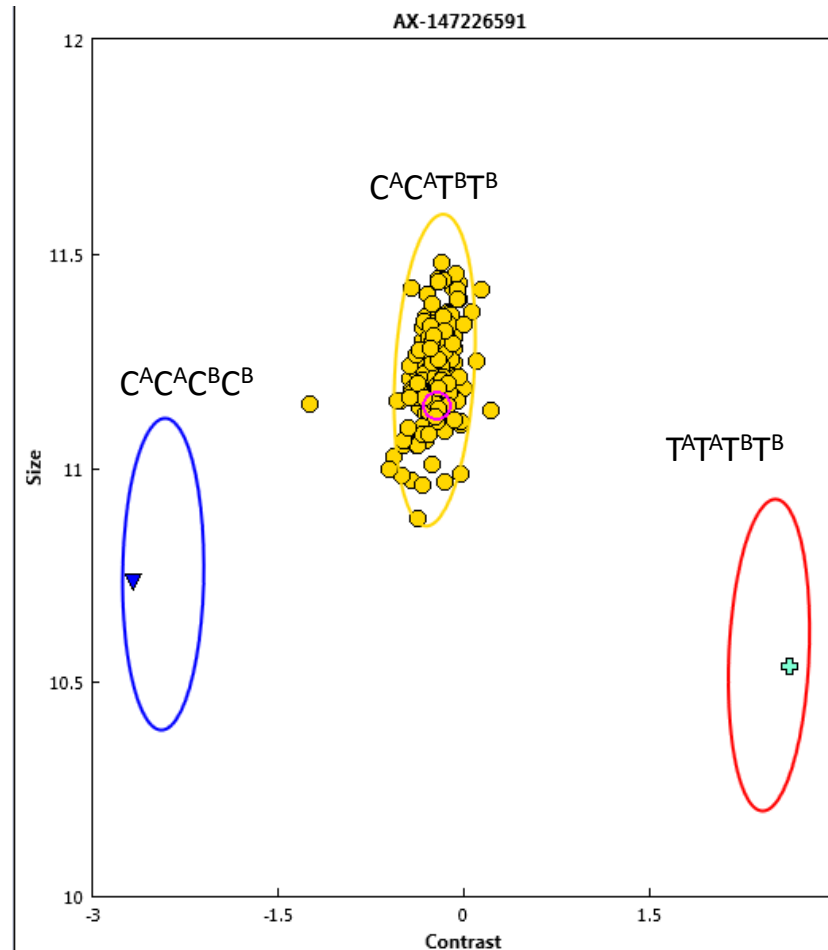


Gene conversion  
Tetrasomy  
Parental offtype

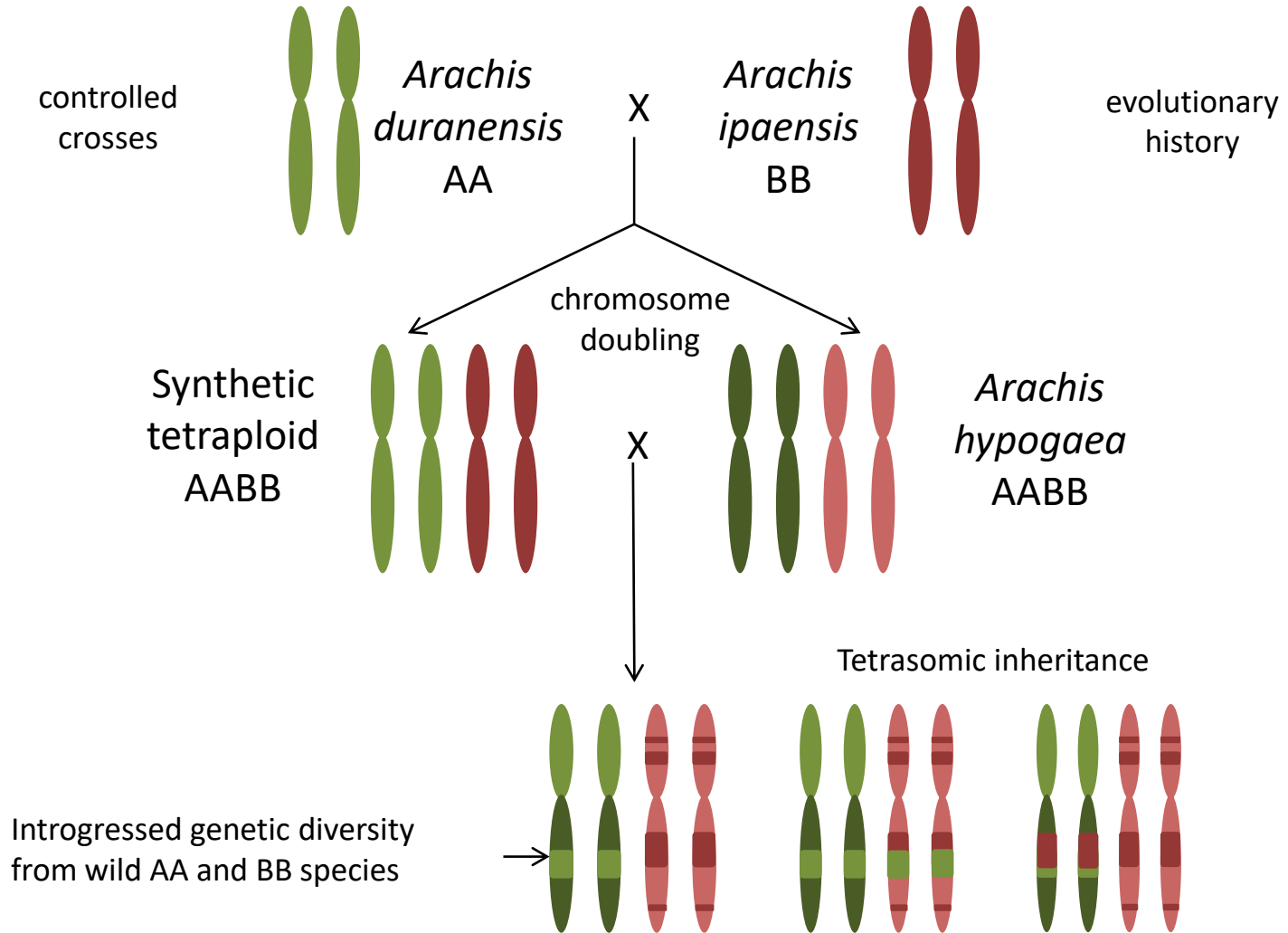


# RIL Population – Parents Monomorphic

## Example of Tetrasomy



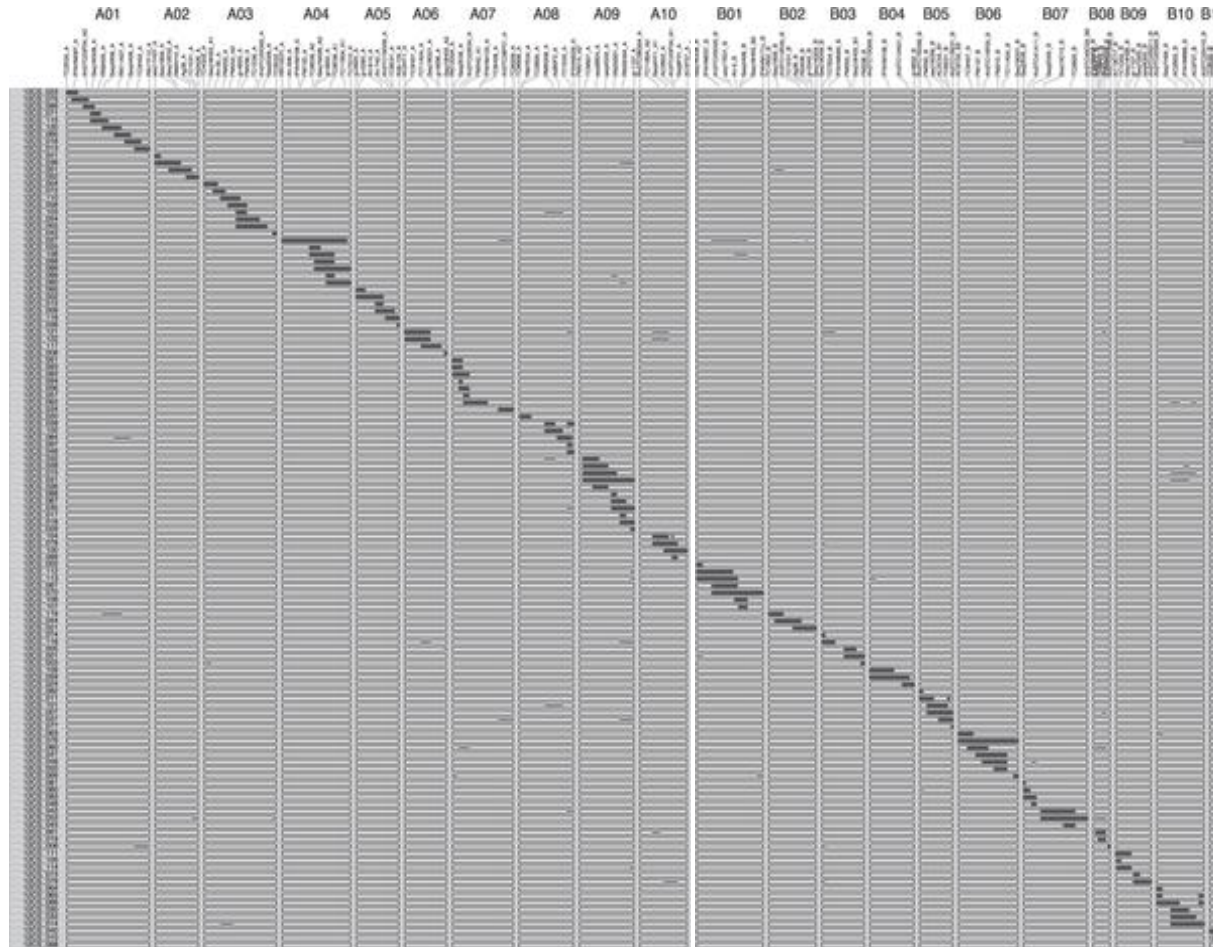
# Expanding Genetic Diversity





# *A. ipaensis* x *A. duranensis*

## Chromosome Segment Substitution Line



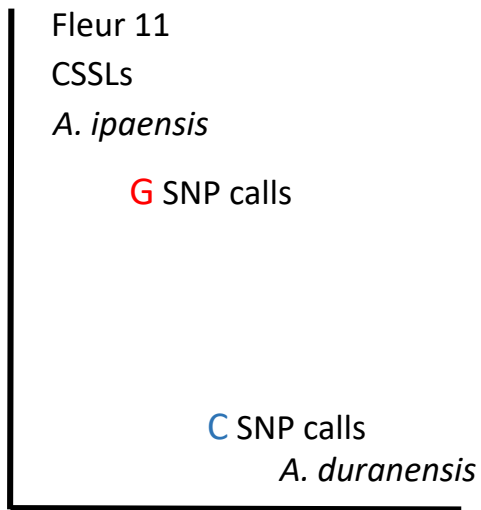
- ✓ Fleur 11 BG
- ✓ 115 SSRs
- ✓ 122 CSSLs
- ✓ Fonceka et al 2012



# *A. ipaensis* x *A. duranensis* CSSL

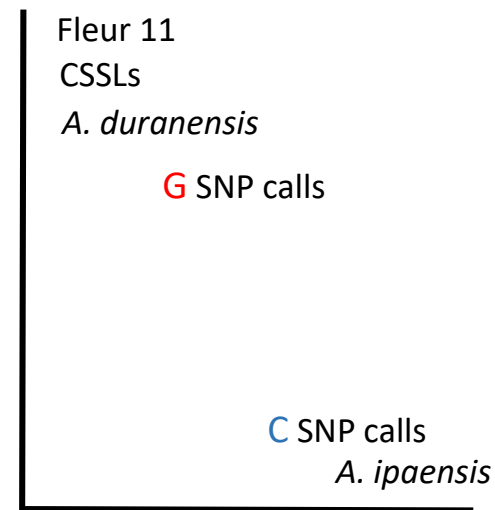
## Filter SNPs by subgenome (Davis Gimode)

---



a

'A' genome SNPs



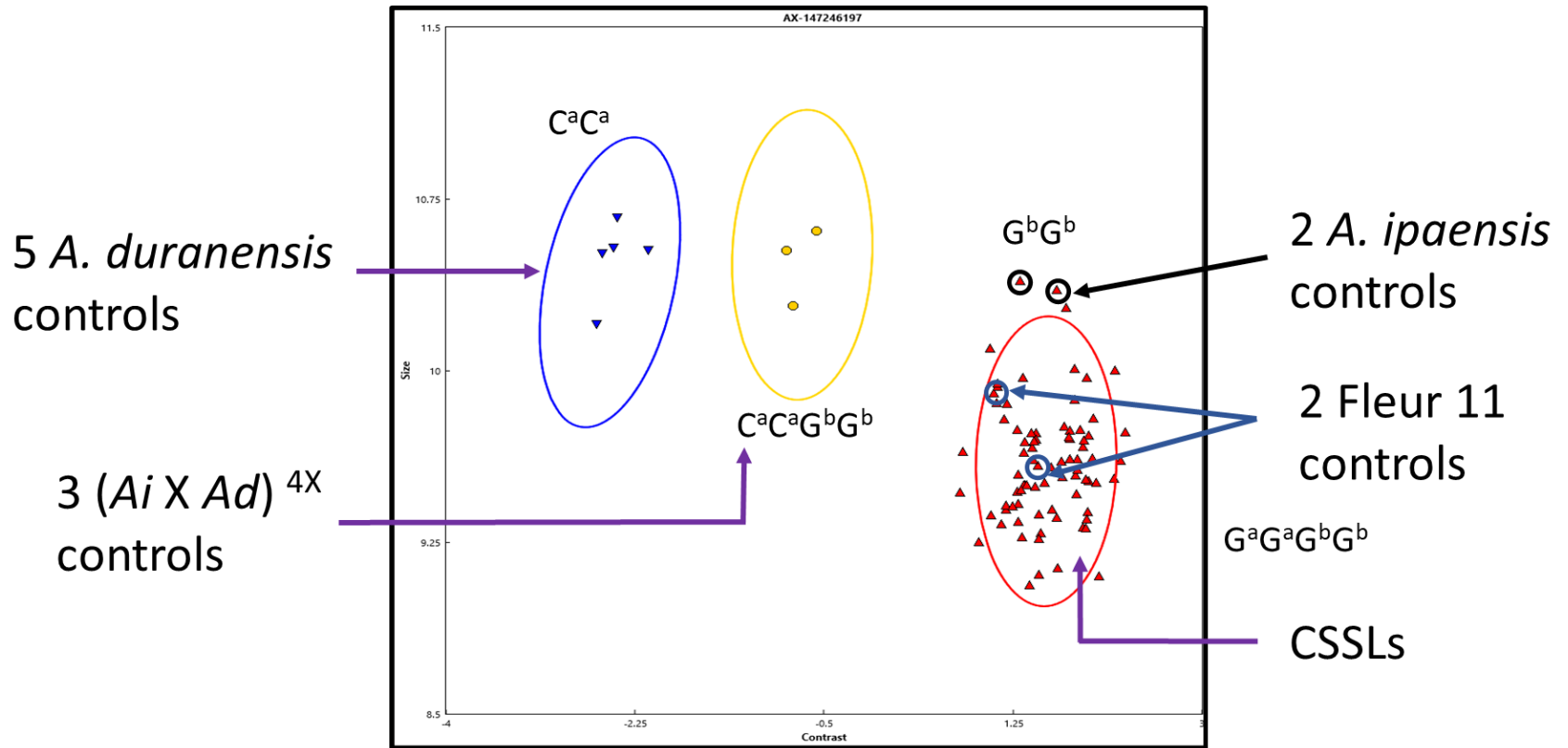
b

'B' genome SNPs



# *A. ipaensis* x *A. duranensis* CSSL

Fleur 11 background

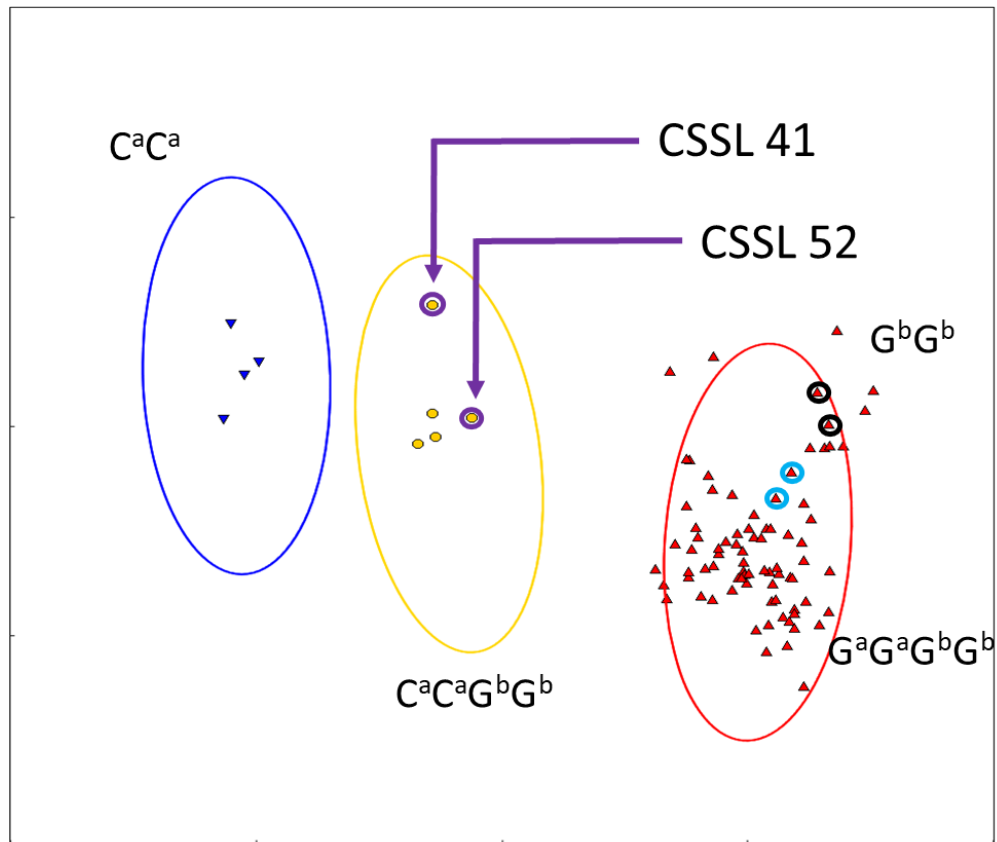


Example of expected clustering for 'A' genome SNPs



# *A. ipaensis* x *A. duranensis* CSSL

## Detecting A genome introgression

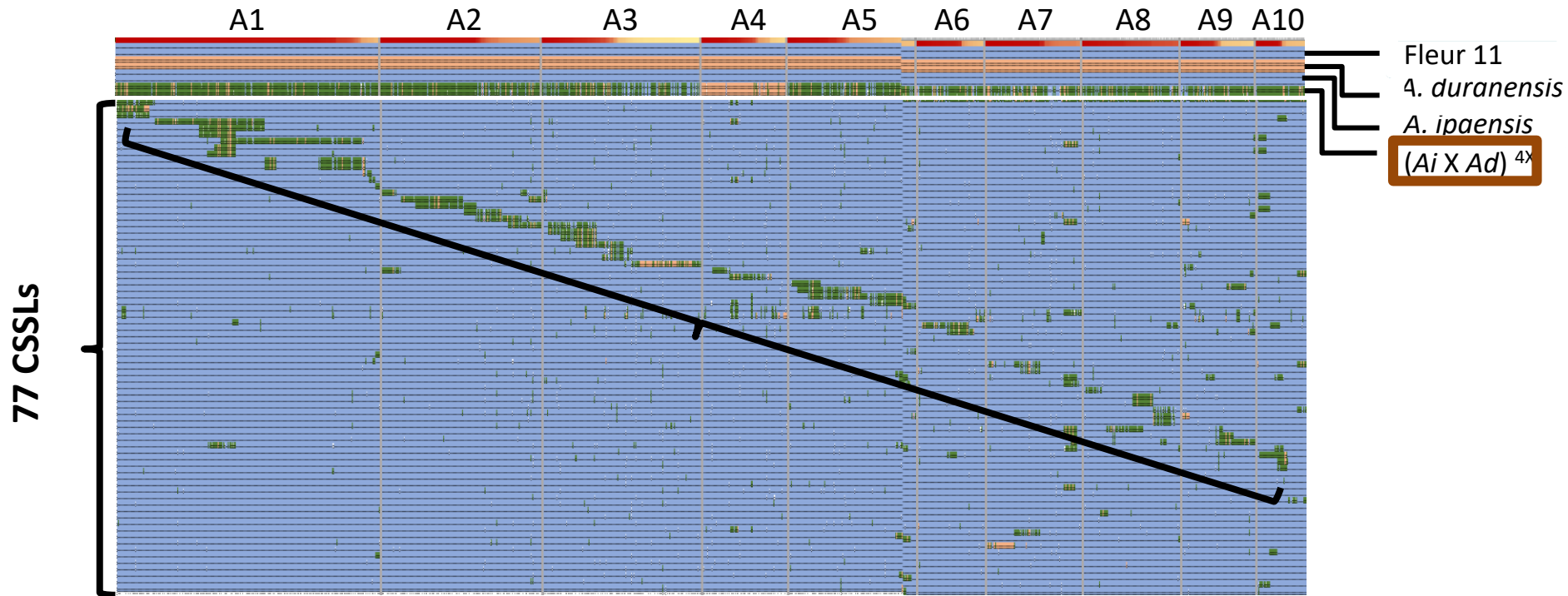


Introgressions cluster with  $(Ai \times Ad)^{4X}$



# *A. ipaensis* x *A. duranensis* CSSL

## Introgression across A genome



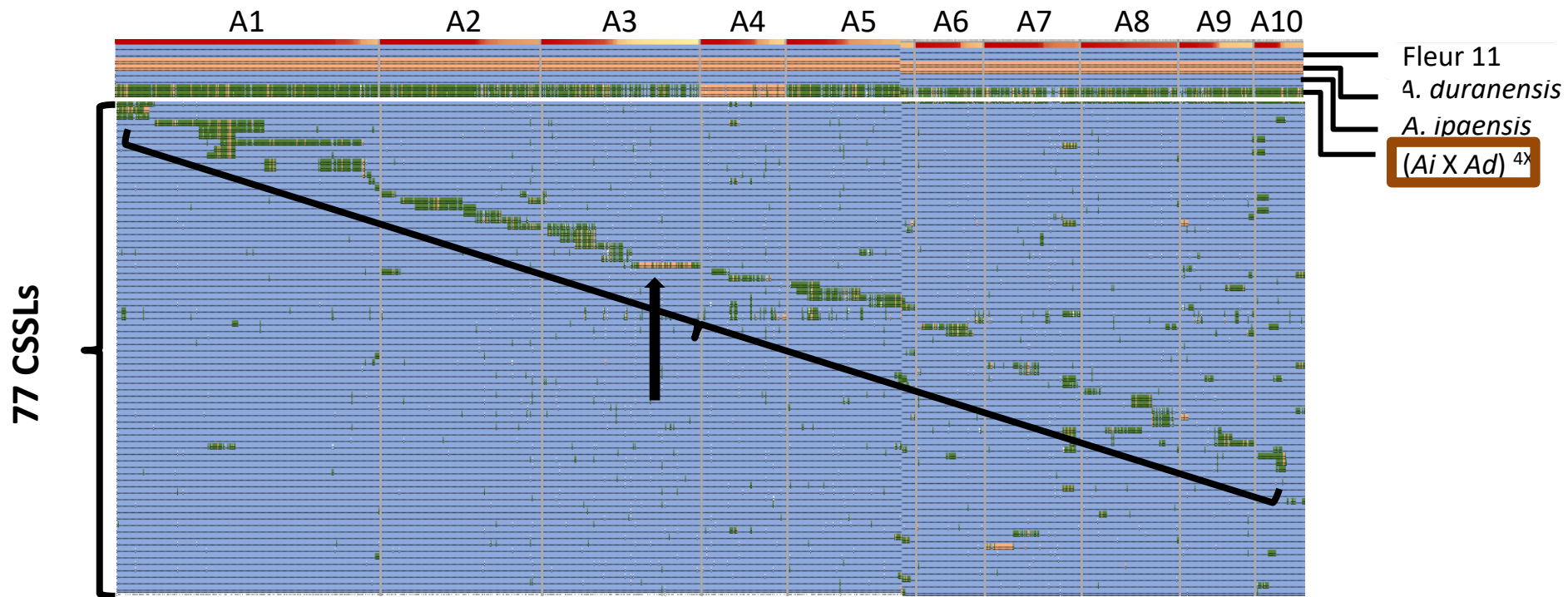
Graphical representation of *A. duranensis* genome

1,610 SNPs

Introgressions span  
the A genome

# *A. ipaensis* x *A. duranensis* CSSL

Introgression across A genome



Graphical representation of *A. duranensis* genome

1,610 SNPs

Tetrasomic region  
at arrow

# A. ipaensis x A. correntina Tetrasomy



# Summary

---

- A SNP array is available for the peanut community
  - Must process 384 samples at a time - \$28 per sample
- The array is useful for genotyping among cultivated lines and interspecific hybrids
- Manual inspection is recommended to validate Axiom calls
- Awareness of segmental allopolyploidy will inform interpretations





# Acknowledgements

## UGA Tifton

- Ye Chu
- Josh Clevenger
- Walid Korani
- Davis Gimode
- Chandler Levinson
- Samuele Lamon
- Larissa Arrais
- Albert Culbreath
- Tim Brenneman
- Renjie Cui
- Stephanie Botton
- Kathleen Marasigan
- Yingping Guo
- Barbara Muller

## NCSU

- Tom Stalker
- Wesley Hancock
- Tom Isleib

## CERAAS

- Daniel Fonceca

## UGA Athens

- Scott Jackson
- Soraya Bertioli
- David Bertioli
- Dongying Gao
- Carolina Chavarro
- Walid Korani

## USDA-ARS

- Corley Holbrook
- Corley's team
- Patty Timper

## MARS Wrigley

- Josh Clevenger

