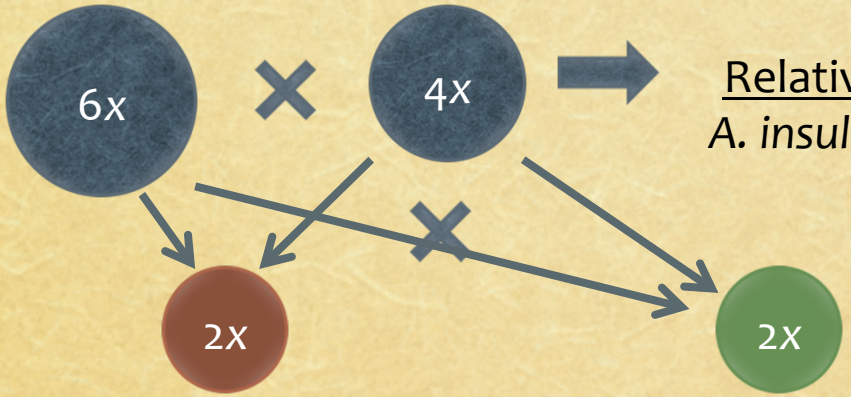
What does CslF6 gene sequence tell us about genome relationships in polyploid *Avena*?

Epistasis and cryptic genetic variation

Rick Jellen, Melissa Fogarty, Eric Jackson, Doug Brown, Veronica Cepeda-Cornejo, and Peter J. Maughan

Brigham Young University and General Mills, Inc.
Plant & Wildlife Sciences Department
Provo, Utah, USA

What did chromosome pairing analysis tell us about out genome relationships?



Relative mean chiasmata frequency:
A. insularis > *A. magna* > *A. murphyi*
(Ladizinsky, 1998)

X A-genome taxa:
Low mean frequency

X C-genome taxa:
Very low pairing frequency

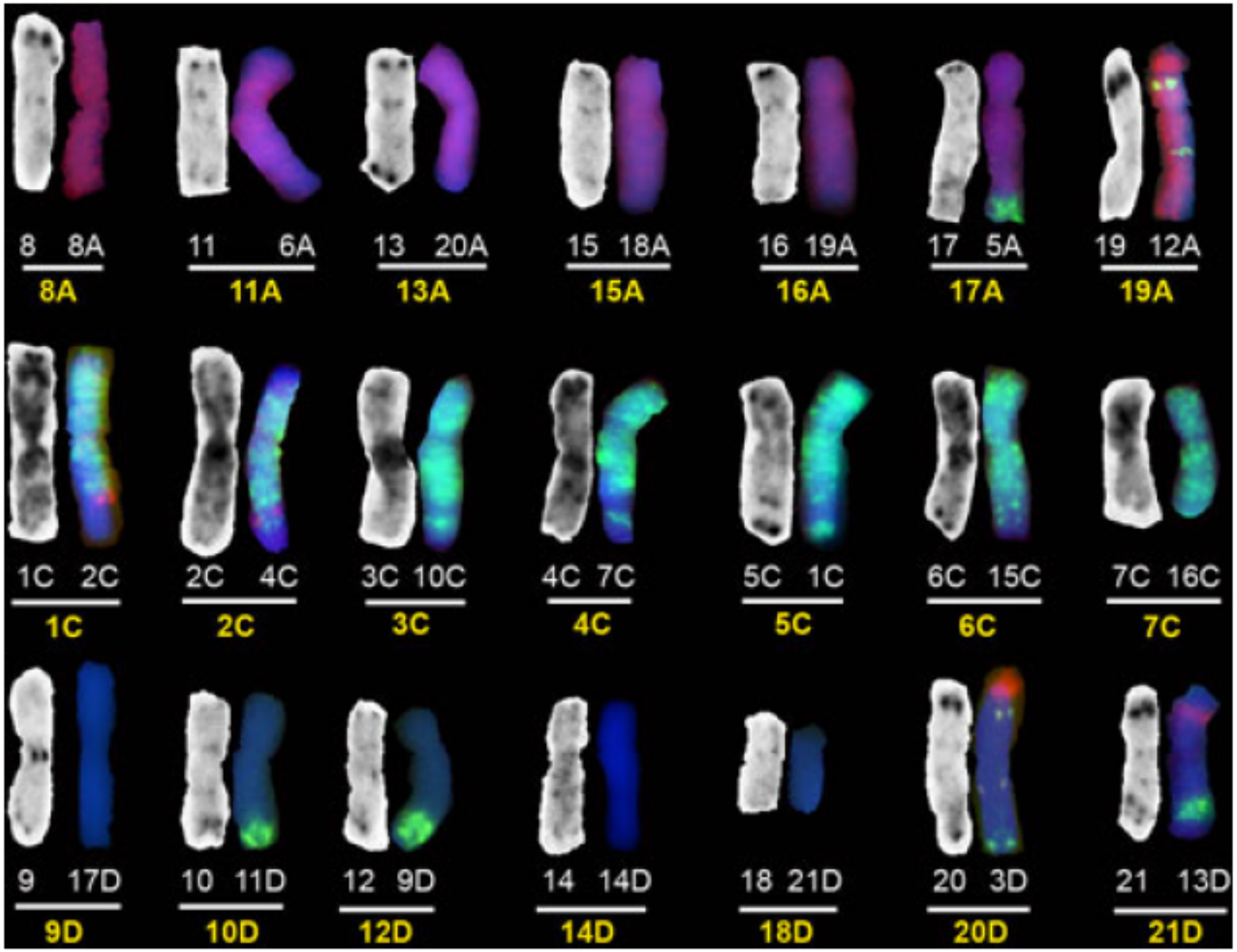
- Conclusions: *A, C, and D are poor genome designators in the hexaploids*
- Hexaploids appear to have a common origin thru 4x *A. insularis*

What do molecular data tell us about genome relationships in *Avena*?

- Repetitive sequences: A, C, and D repeats are "apparent" and "identifiable" at the whole-chromosome level

Sanz et al. (2010) *TAG* 121:1541-1552

PINK = A-genome sequence; GREEN = C-genome sequence



What do molecular data tell us about genome relationships in *Avena*?

- Repetitive sequences: A, C, and D repeats are "apparent" and "identifiable" at the whole-chromosome level
- Single-gene sequences: COMPARE HOMOEOLOGOUS SEQUENCES WITH THEIR HOMOLOGS IN DIPLOIDS

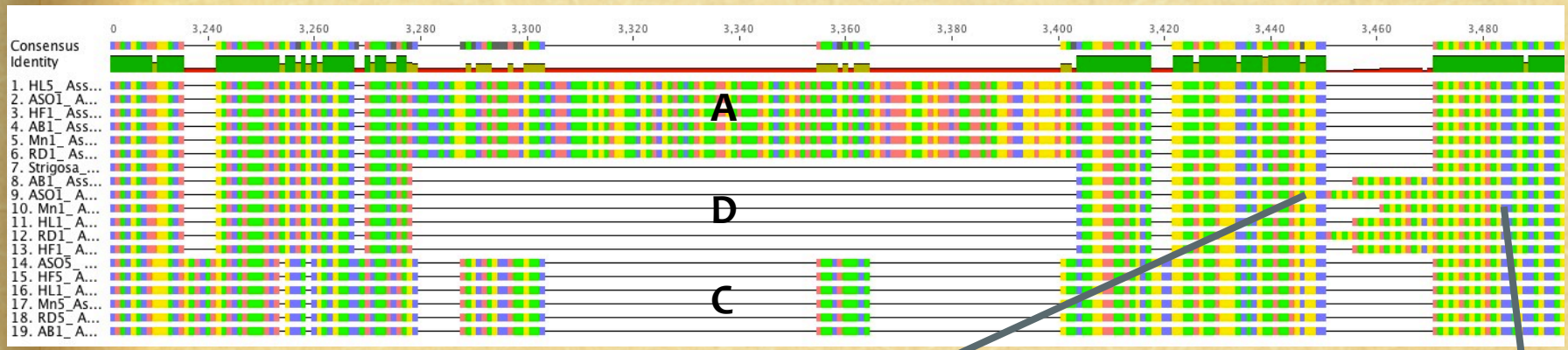
Why CslF6?

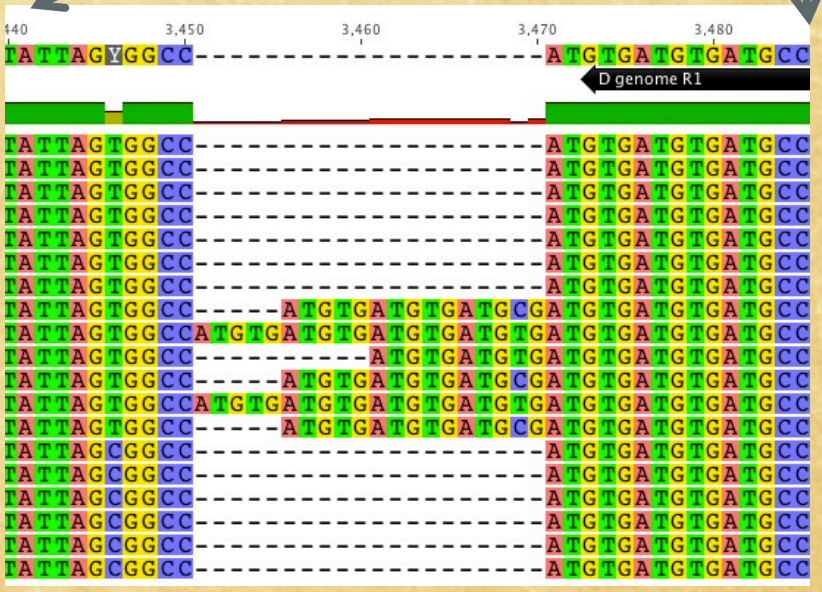
- *CslF* genes promote β-glucan production in transformed *Arabidopsis* plants (Burton et al. 2006)
- Expression of CslF6 is greater than all other CslF genes in barley (Burton et al. 2008)
  - CslF6 transcripts predominate in oat and barley EST database
- Overexpression of CslF6 in barley increased β-glucan (Burton et al. 2011) levels while RNAi in wheat dramatically decreased β-glucan (Nemeth et al. 2010)
- *Beta glucan less* (*bgl*) mutants in barley have a point mutation resulting in the change of a highly conserved amino acid in *CslF6* (Taketa et al. 2012)

Genotypes Sequenced

- 22 *A. sativa* cultivars
- 1 synthetic hexaploid (Ladizinsky's *A. strigosa* x *A. magna domestica*)
- 2 *A. magna*
- *A. strigosa, A. wiestii*
- *A. ventricosa*
- *A. canariensis*
- *A. insularis* (partial)

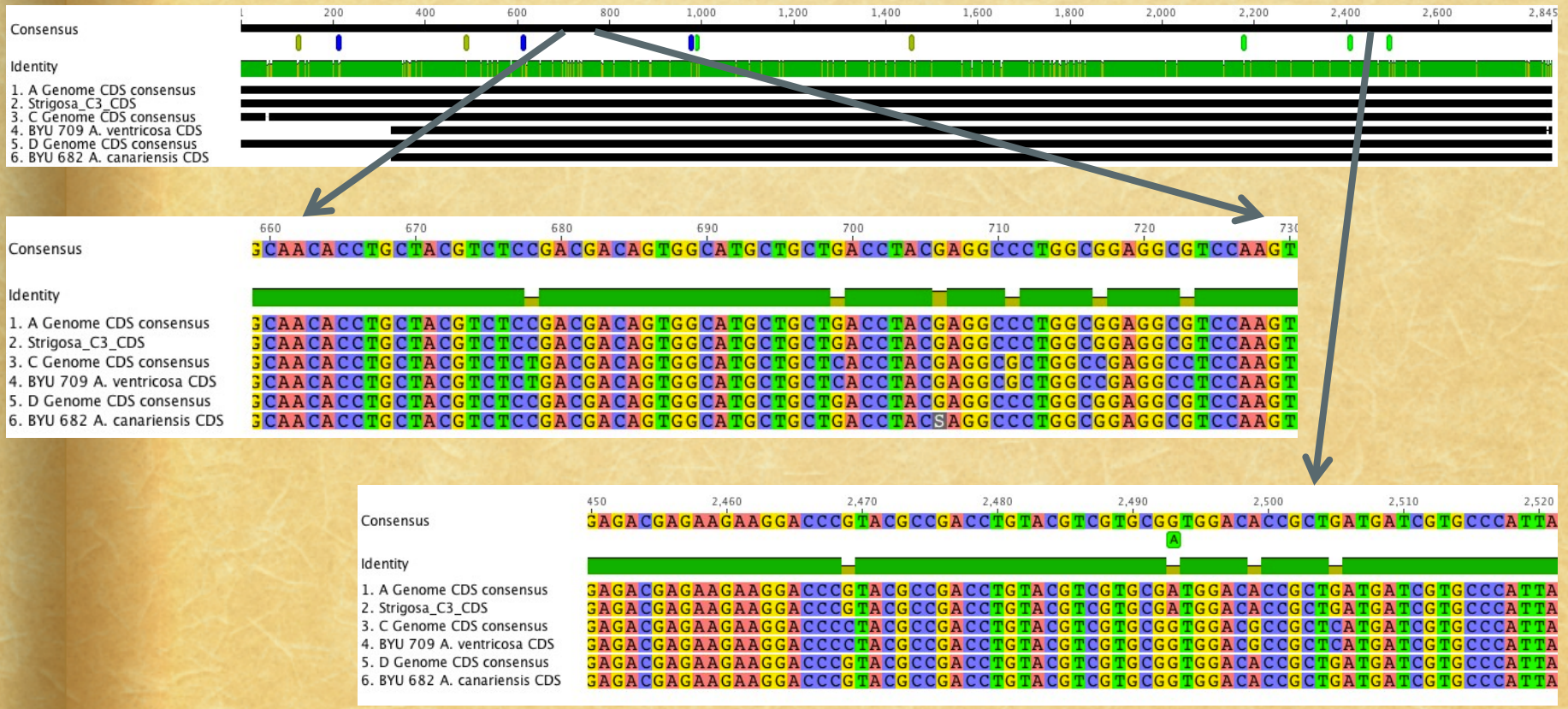
CslF6 in oat: subgenome identity is retained at the DNA sequence level





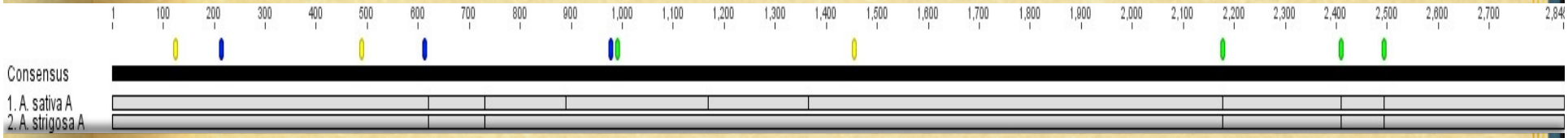
CslF6 in oat: subgenome identity is retained at the DNA sequence level

A and D are more similar to each other than to C
A sequence most similar to *strigosa-wiestii*, D to *canariensis*

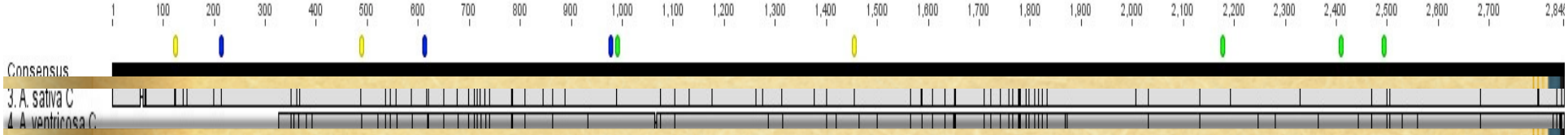


CslF6 homoeolog-specific SNPs

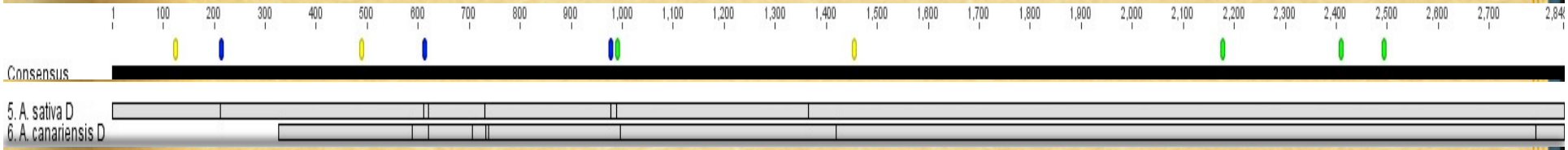
Avena sativa (AACCDD) vs. *A. strigosa* (AA)



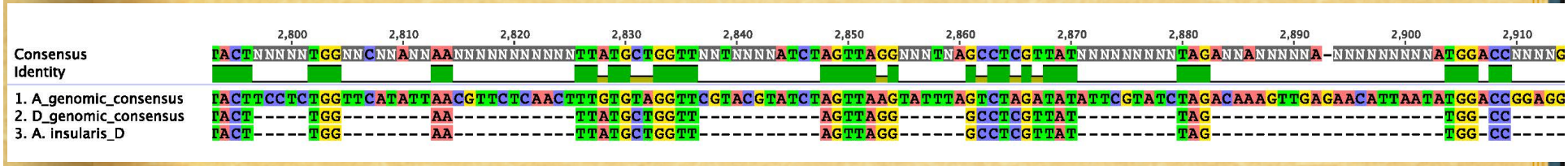
Avena sativa (AACCDD) vs. *A. ventricosa* (CC)



Avena sativa (AACCDD) vs. *A. canariensis* (DD)



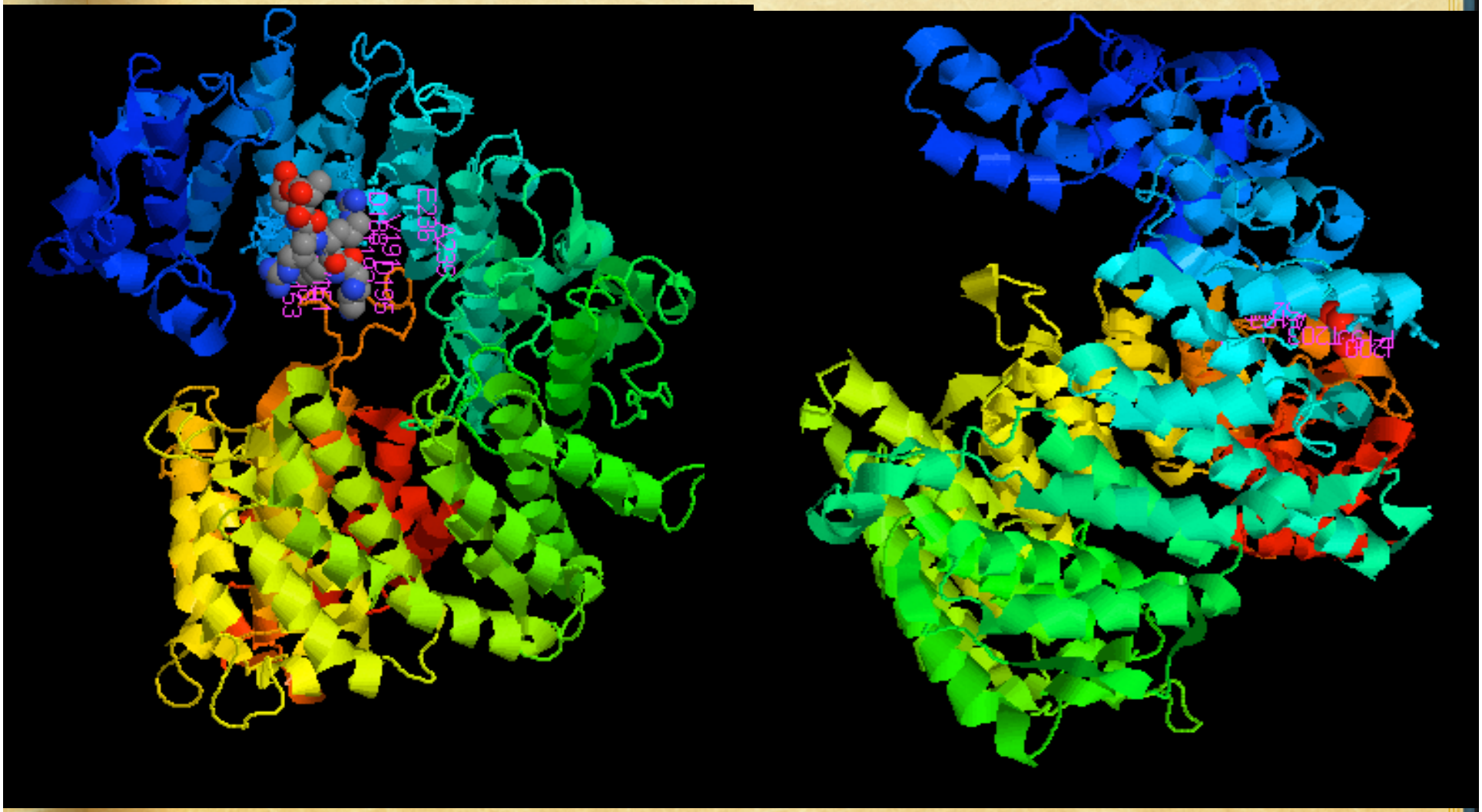
Avena sativa (AACCDD) vs. *A. insularis* (DD)



Predicted protein folding change

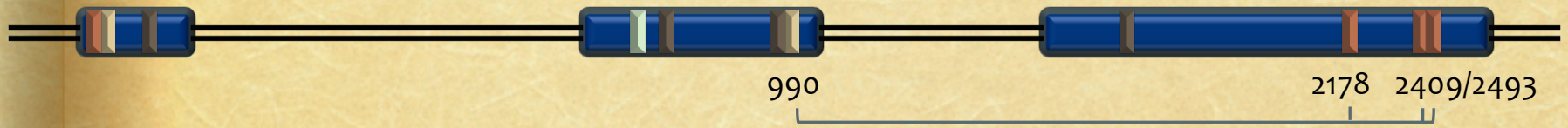
AA and DD Genome CslF6

CC Genome CslF6



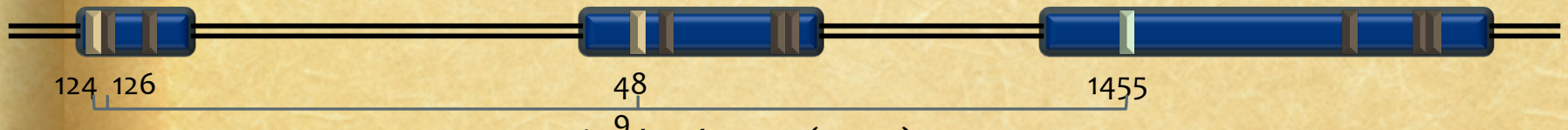
CslF6 homoeoallele-specific expression SNPs

A Genome



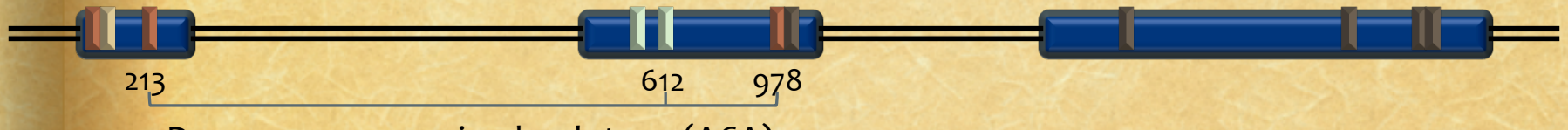
A genome expression haplotype (TAAA)

C Genome



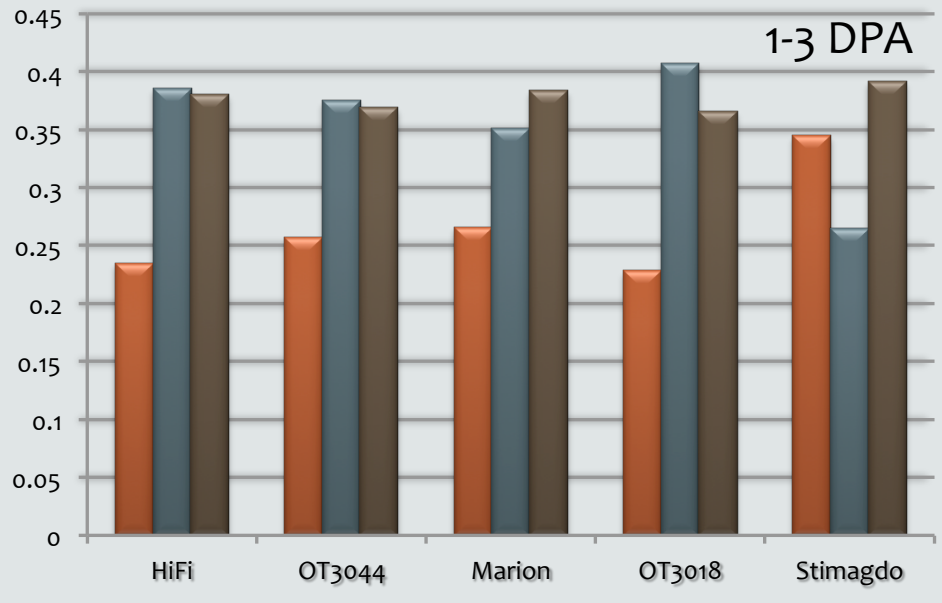
C genome expression haplotype (TGTC)

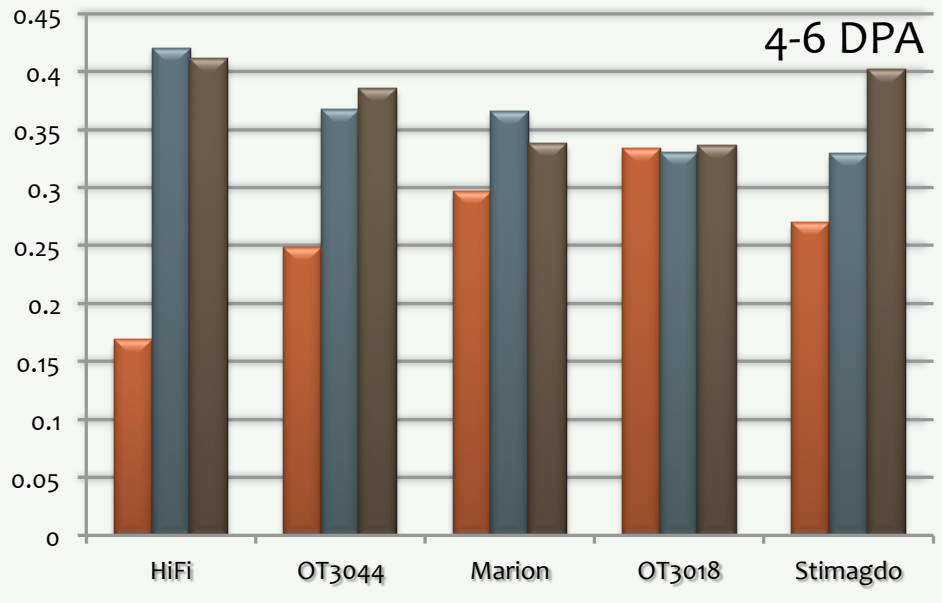
D Genome

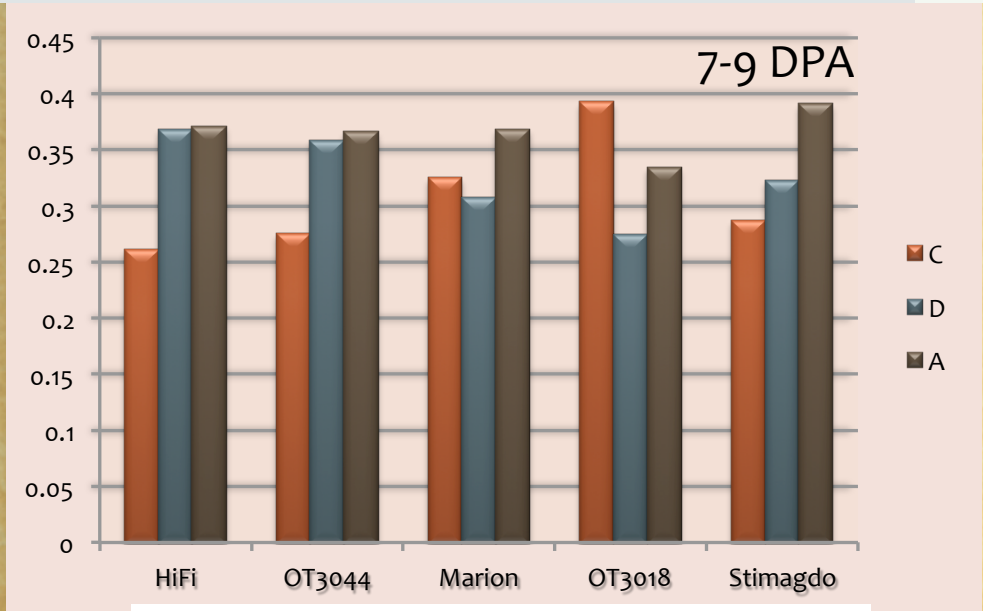


D genome expression haplotype (ACA)

Homoeologous loci exon expression







← Higher Lower →

NOTE higher expression of C-genome locus in lower beta-glucan lines

Dominant Negative Epistasis?

- "Traffic jam" model: semi-functional "cryptic" C-genome-encoded *CslF6* subunits slow down holoenzyme function
- Reduce the conversion of (1,4) → (1,3) linkages
- Greater C-genome expression →→→ less beta-glucan
- *Could this be altered by selective breeding?*
  - *Mutagenesis – knock out* CslF6-c
  - *Natural* CslF6-c *null mutants?*

Importance of "cryptic" *GBSSI* alleles: example from allotetraploid goosefoot

- *GBSSI* controls amylose (helical starch) production
- Single-gene dominant trait
- Starch reduction may be related to increased soluble fiber
- Homozygous mutants are *waxy*
  - Examples: sticky rice, *waxy* maize and barley
- Goosefoot, *Chenopodium berlandieri*, *2n=4x=36*

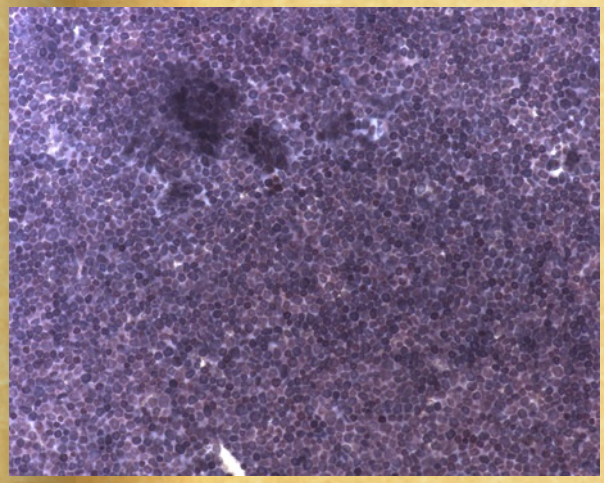
Seed goosefoot = *quinoa*

Vegetable goosefoot = *huauzontle*

Huauzontle (right)
Quinoa (below)

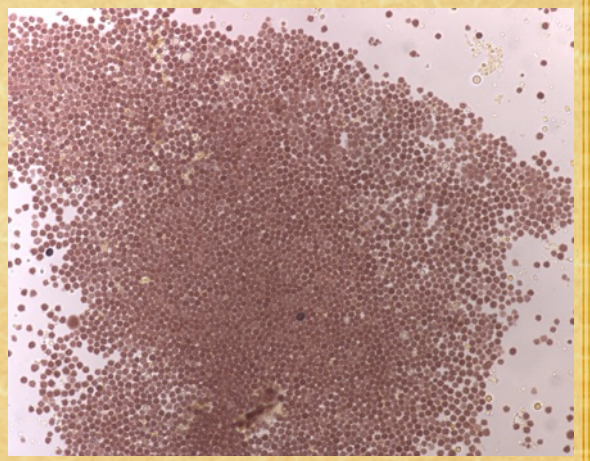
Phenotypic screening for *waxy*

- Phenotypic screen: I_2KI stains the amylose helix blue
- No *waxy* phenotypes identified in quinoa
- Found in three landrace populations of huauzontle from Mexico State



Nonwaxy
(amylose +)

waxy (amylose -)



Allelic Variation in *Chenopodium GBSSI*

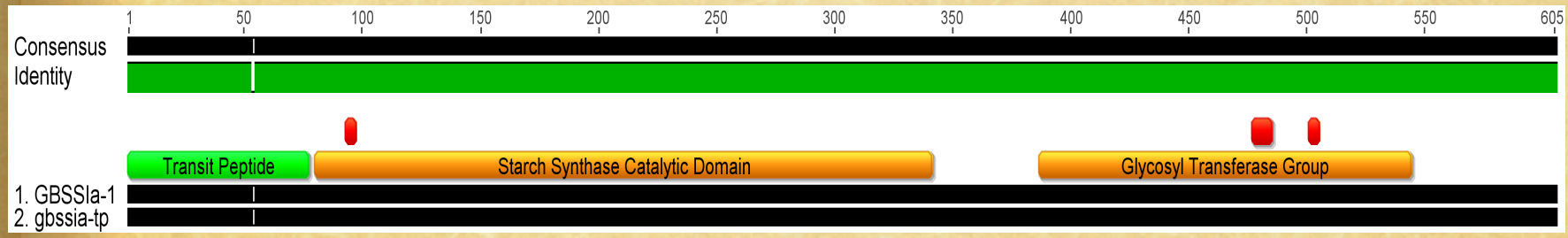
- 19 alleles in 18 accessions
- One putative null A-genome allele
- Two putative null B-genome alleles
- One of the B-genome nulls is in quinoa
- A-genome mutant is structurally sound but the plastid targeting peptide has an I→T substitution

Allele	Individuals	Mutation
GBSSIa-1	Ollague, G205, NL6, KU2, 835 (*C. desiccatum*, 2*x*)	Reference Allele
GBSSIa-2	0654, Chucapaca	D393E
gbssia-tp	H04, H02 (Huauzontle)	Transit Peptide: I54T, I325V, V456L
GBSSIa-3	937 (*C. berlandieri* var. *boscianum*)	I325V, V456L
GBSSIa-4	803 (*C. berlandieri* var. *macrocalycium*)	I325V, V456L, M555I
GBSSIa-5	652 (*C. berlandieri* var. *zschackei*)	325V, V456L, N575D
GBSSIa-6	1101 (*C. hircinum*)	I325V, V456L, L463S
GBSSIa-7	1302 (*C. berlandieri* var. *zschackei*)	V28I, R246P, V289F, I325V
GBSSIa-8	1005 (*C. hians*, 2*x*)	K294Q
GBSSIa-9	921 (*C. standleyanum*, 2*x*)	K458R
GBSSIa-10	843 (*C. neomexicanum*, 2*x*)	S274P
GBSSIb-1	Ollague, 0654, NL6, KU2, 803, 1101	Reference Allele
GBSSIb-2	Chucapaca	R142M
gbssib-t	G205	W129X (early termination)
GBSSIb-3	H04	T74P
gbssib-Δ	H02	Deletion 64-142, A417E
GBSSIb-4	937	K55N
GBSSIb-5	652	K55N, T74P
GBSSIb-6	943 (*C. ficifolium*, 2*x*)	T531I

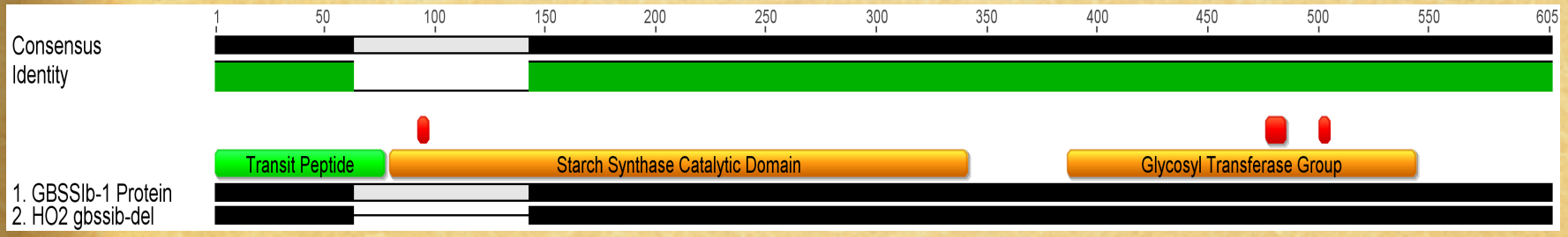
GBSSI Functionality Hypotheses

waxy phenotype only results when both genomes are homozygous for recessive alleles! (Three genomes must be homozygous null in a hexaploid like oat!)

gbssia-tp (shared by non-*waxy* accession, 'H04')



gbssib-Δ



Uncovering "unconventional" recessive alleles

- Gene encoding enzyme Z has 3 functional copies:
  - $\{Z_A Z_A Z_C Z_C Z_D Z_D\}$ **X** {any genotype} = **conventional progeny**
- Gene encoding enzyme P has 2 functional, 1 null copies:
  - $\{P_A P_A P_C P_C p_d p_d\}$ **X** $\{P_A P_A P_C P_C p_d p_d\}$ = **conventional progeny**
- Gene encoding enzyme T has 1 functional, 2 null copies:
  - $\{t_a t_a T_C T_C t_d t_d\}$ **X** $\{T_A T_A T_C T_C t_d t_d\}$ = **conventional progeny**
  - $\{t_a t_a T_C T_C t_d t_d\}$ **X** $\{T_A T_A t_c t_c T_D T_D\}$ = **seg. rare unconventional (1/64 in F_2)**

Uncovering fully-functional recessives in the presence of a *dominant negative* allele

- Gene encoding enzyme F has 3 functional copies:
  - $\{F_A F_A F_C F_C F_D F_D\}$ **X** $\{F_A F_A F_C F_C F_D F_D\}$ = **conventional progeny (all negative phenotype)**
  - $\{F_A F_A F_C F_C F_D F_D\}$ **X** $\{F_A F_A f_C f_C F_D F_D\}$ = seg. 3:1 (semi-functional:fully-functional) in F_2

- **Predict any gene encoding a multi-meric enzyme subunit would be susceptible to cryptic dominant-negative mutations in an allopolyploid**

Conclusions

- Gene discovery and mining of novel alleles is absolutely critical to making significant breakthroughs in future oat breeding
- Is one of the three genomes in oat more "degenerate" than the other two?
  - Expect the last to enter the complex genome would be the most functional: appears to be the "A"
  - "A" and "D" are most similar
  - "C" is the most heterochromatic, and full of *dispersed* heterochromatin: evidence for transposon proliferation and therefore more mutations (?)

Acknowledgments

Jeff Maughan
Gongshe Hu- USDA-ARS Aberdeen
Emir Islamovic- USDA-ARS Aberdeen

Eric Jackson – General Mills
Joe Lutz – General Mills
Ryan Brown - General Mills

Robert Campbell- USDA-ARS Aberdeen
Irene Shackelford- USDA-ARS Aberdeen
Kathy Esvelt Klos – USDA-ARS Aberdeen
Gerard Lazo- USDA-ARS Albany

Robert Reid –UNC Charlotte
Jess Schleuter – UNC Charlotte
Mike Wang- DHMRI

Nick Tinker- AAFC ECORC Ottawa
Yung-Fen Huang – AAFC ECORC Ottawa