The genetic architecture of seed composition in soybean is refined by genome-wide association scans across multiple populations

Justin N. Vaughn, *Zenglu Li Group, UGA*

IDENTIFICATION OF LOCI ASSOCIATED WITH PROTEIN AND OIL CONTENT IN SOYBEAN

Song Q, . . ., Cregan P

GENOME-WIDE ASSOCIATION STUDIES OF 6 AGRONOMIC TRAITS IN SOYBEAN

Wen Z, . . ., Wang D

A GENOME-WIDE ASSOCIATION ANALYSIS FOR SEED OIL AND PROTEIN ON 13, 000 SOYBEAN ACCESSIONS

Bandillo N, . . ., Lorenz A

EXPLORATION OF QTL, GENETIC RESOURCES AND INTEGRATED OMICS TOOLS TOWARDS SEED OIL AND PROTEIN IMPROVEMENT IN SOYBEAN
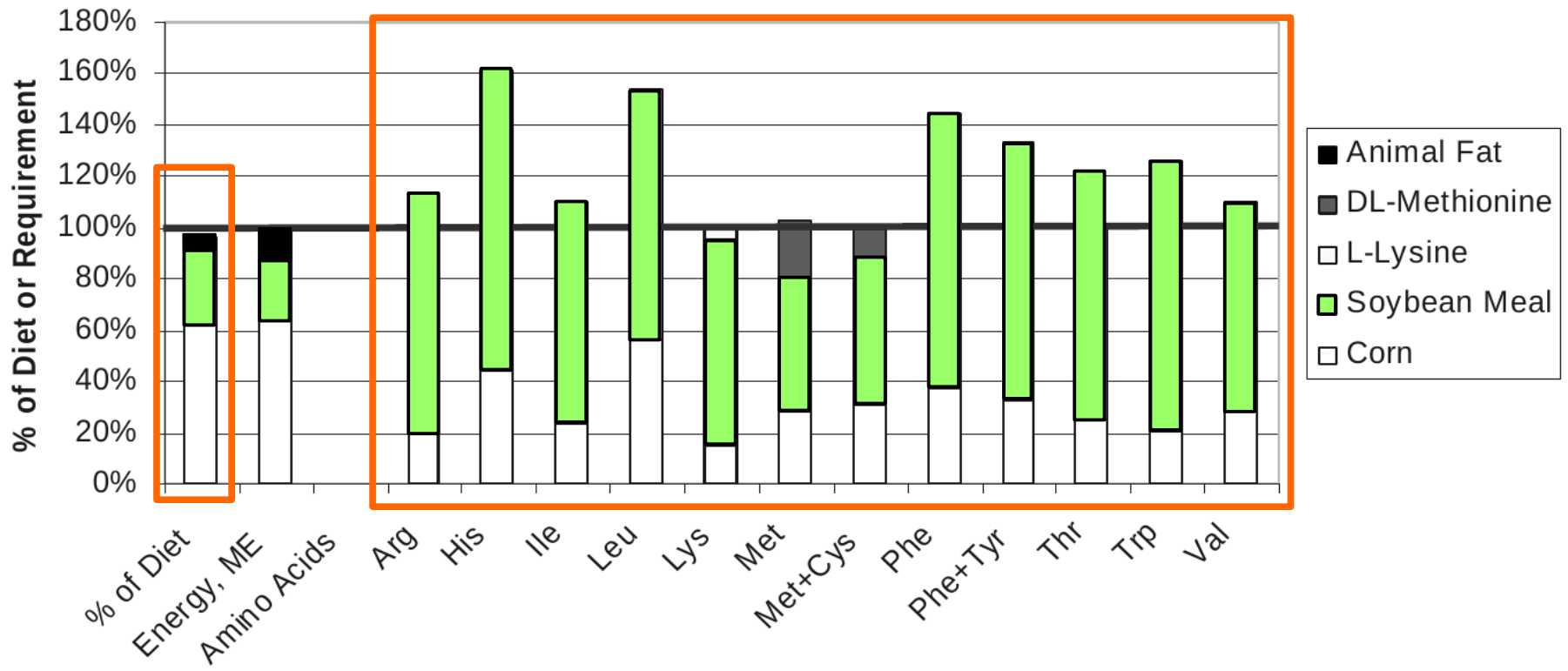
Sonah H, . . ., Nguyen H

IDENTIFICATION OF LOCI GOVERNING EIGHT AGRONOMIC TRAITS USING A GBS-GWAS APPROACH AND VALIDATION BY QTL MAPPING IN SOYBEAN
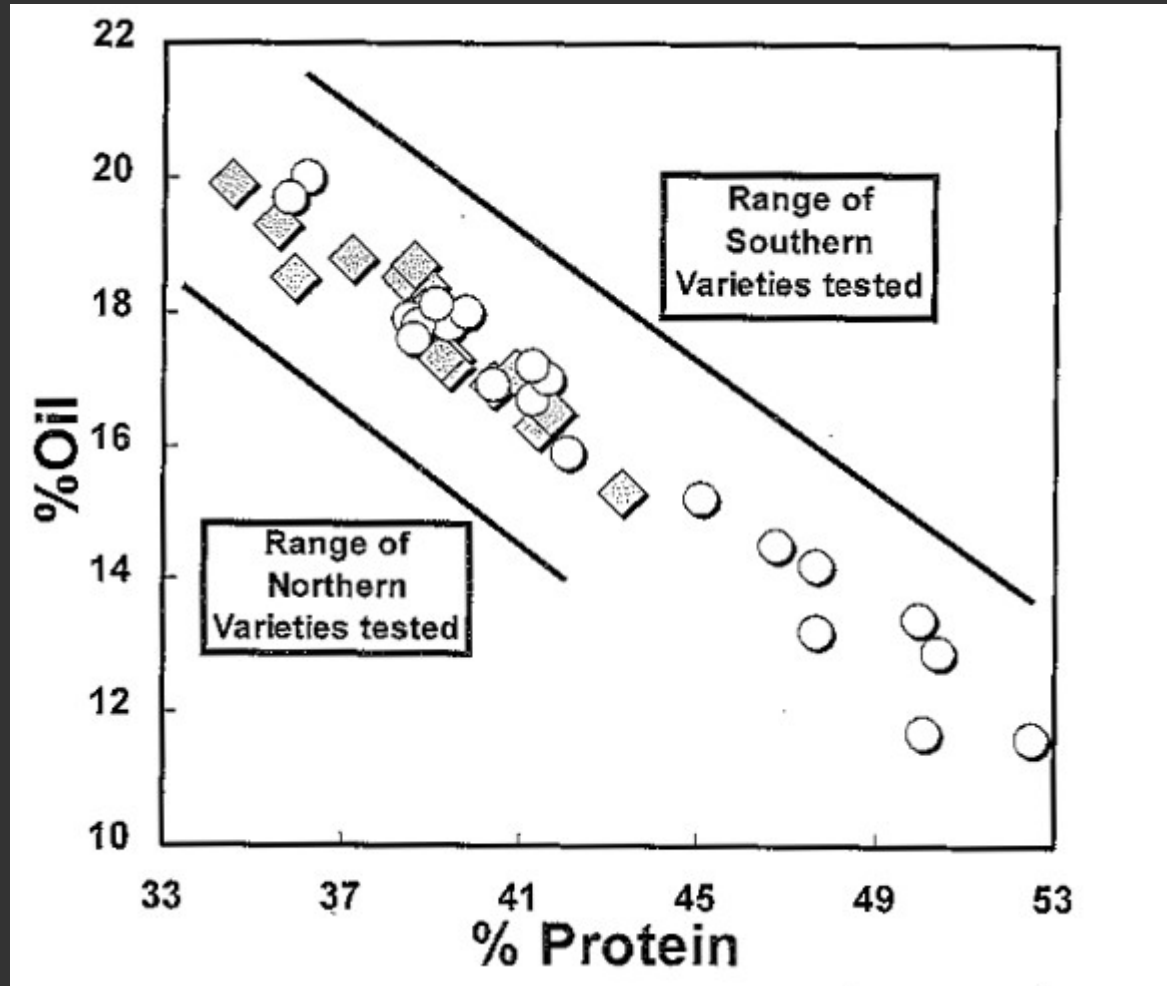
Sonah H, . . ., Belzile F

Chicken shall not live on corn alone. . .

Energy and Amino Acid Contributions as a Percentage of Dietary Requirement for Growing Broiler Chickens
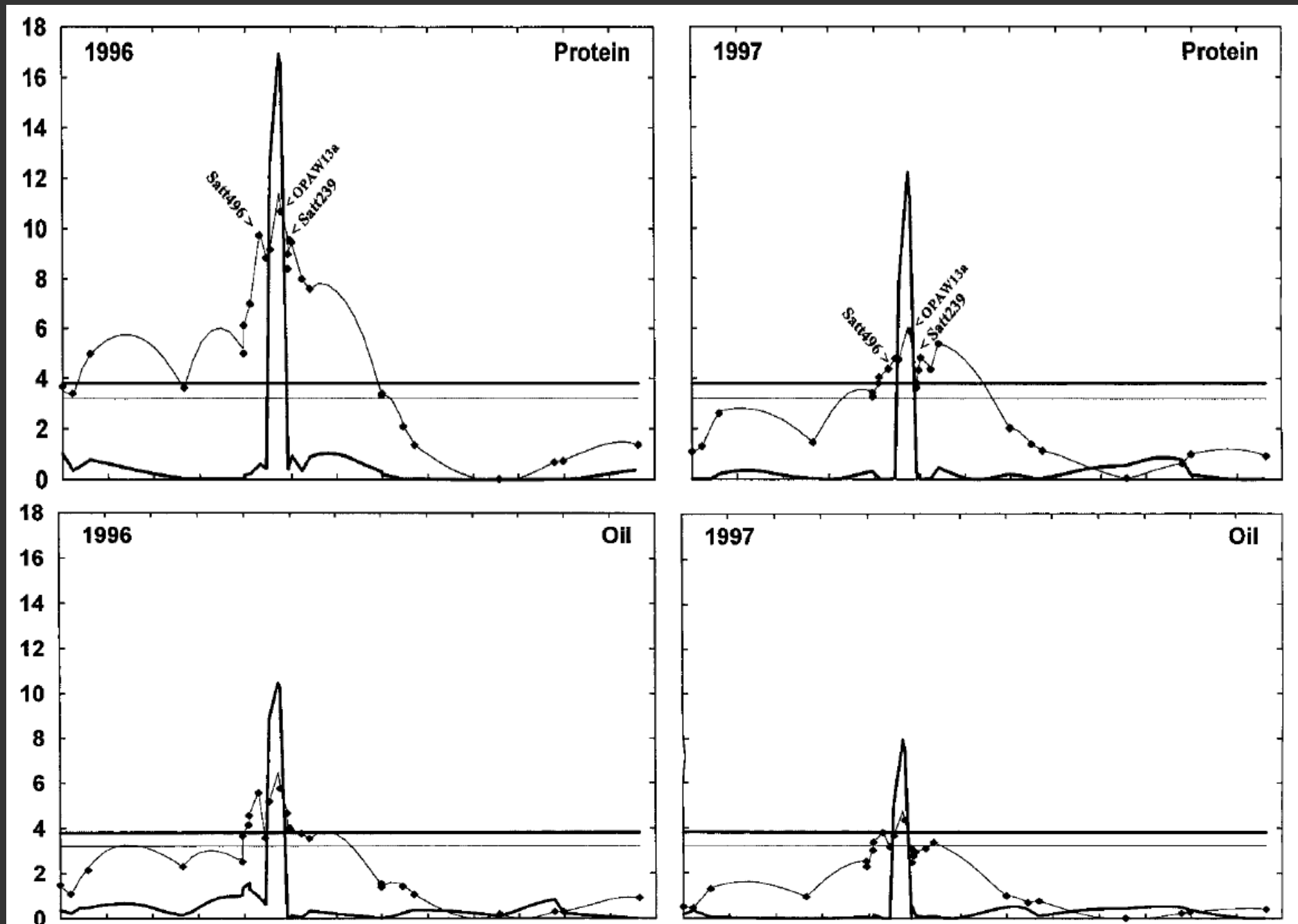
Protein and oil content are negatively correlated . . .

Wilson 1978, ed. Boerma and Specht

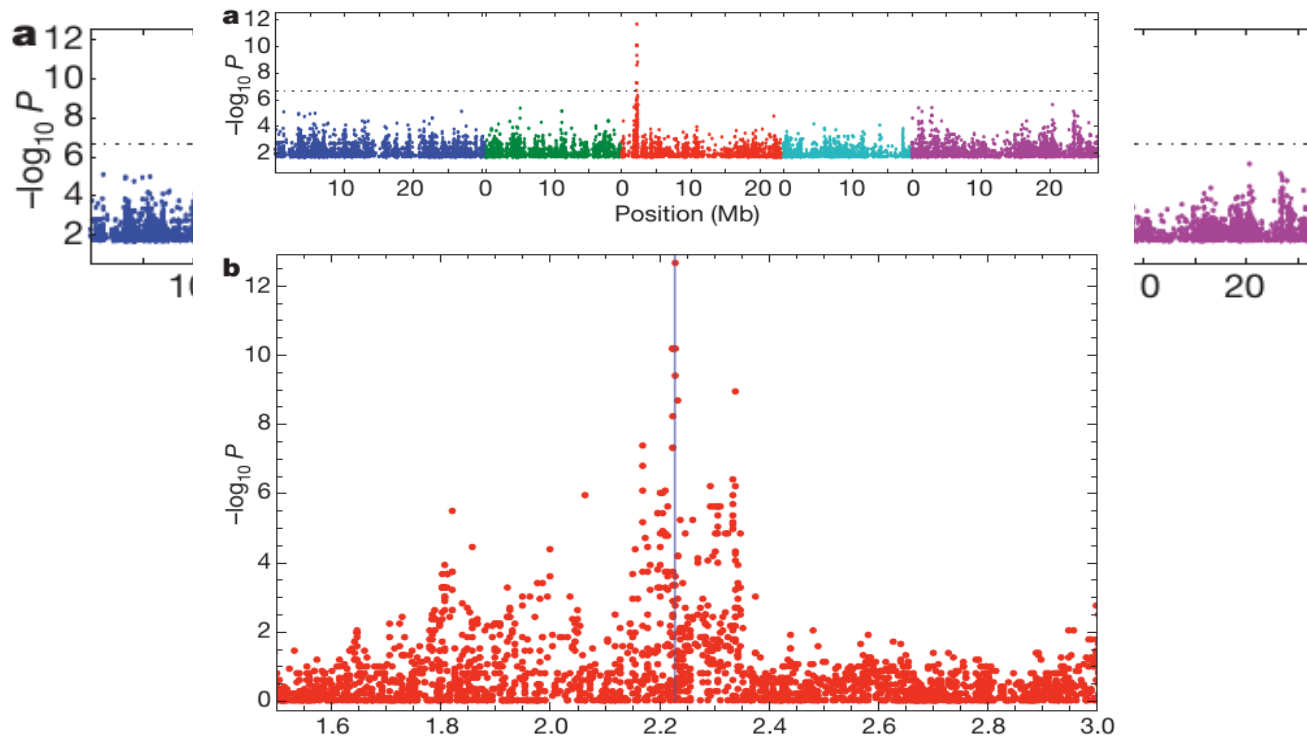# Linkage group I (chromosome 20) QTL for high protein . . .

Genome-wide association (GWA) can reveal genetic architecture with high resolution . . .

MG
PROTEIN
OIL

min ▮▮▮ max

■ South Korea (745) ● Uganda (17)   ● V (675)
■ Japan (53)   ● North Korea (13) ● VI (31)
● China (50)   ● IV (6)

B

United States Department of Agriculture
Agricultural Research Service

Germplasm Resources Information Network

Phenotype and genotype data

Genotypic data was derived from the large-scale effort to genotype the USDA Soybean Germplasm
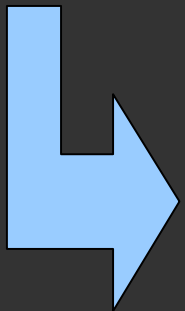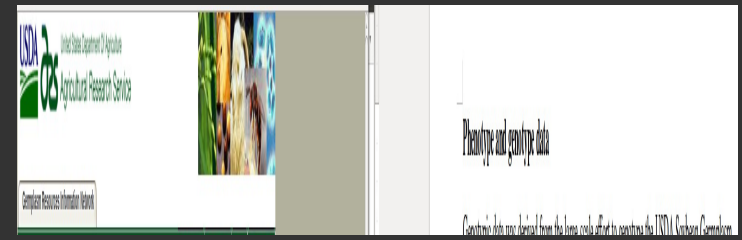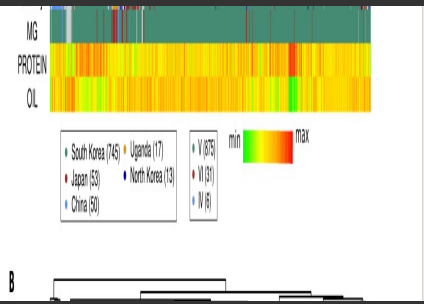
Genotype Data			Phenotype Data
Genotyped	NOT Genotyped	Genotyped	Berry Number
Low LD SNP	Functional SNP	High LD SNP	
G	T	C	15
A	T	C	14
G	T	C	13
A	T	T	12
A	T	C	11
G	A	T	10
G	A	C	9
A	A	T	8
G	A	T	7
A	A	T	6
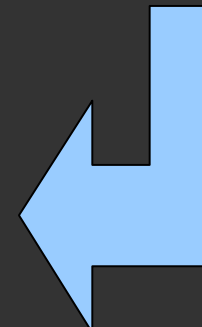
Myles et al. 2009

Protein & oil content and population structure . . .
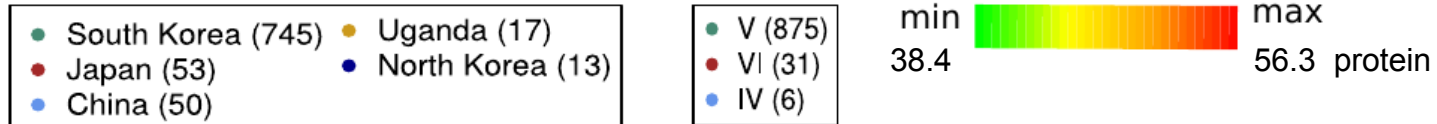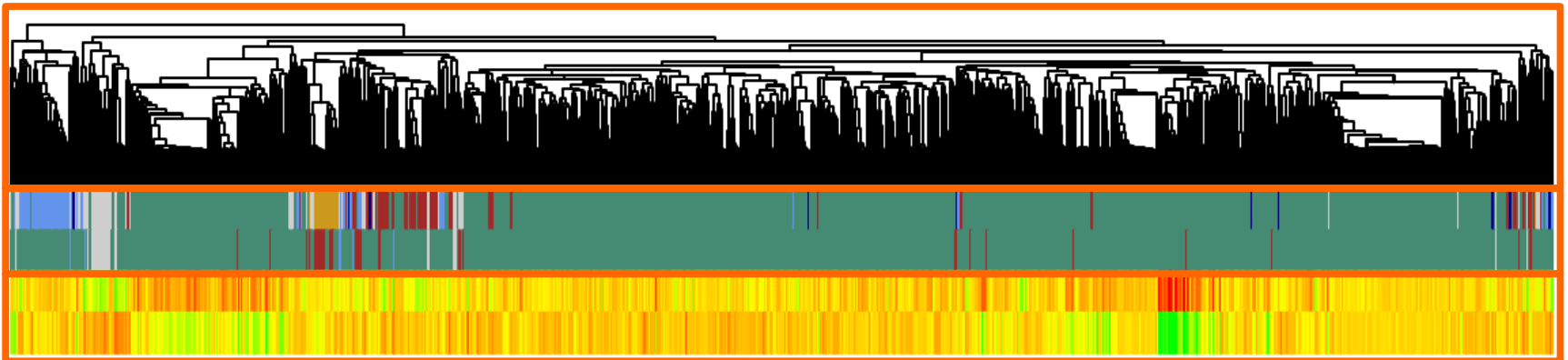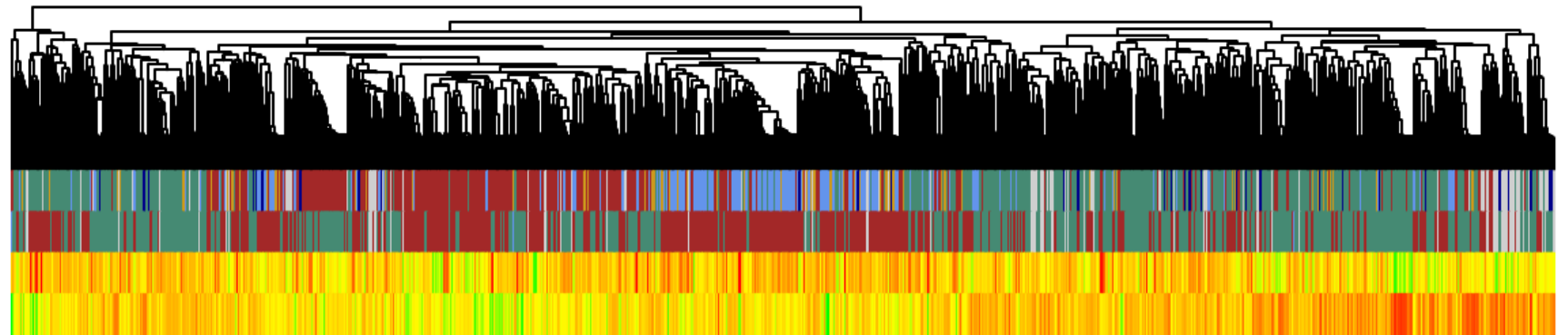
Phenotypic simulations for GWA in MS-2000 population . . .

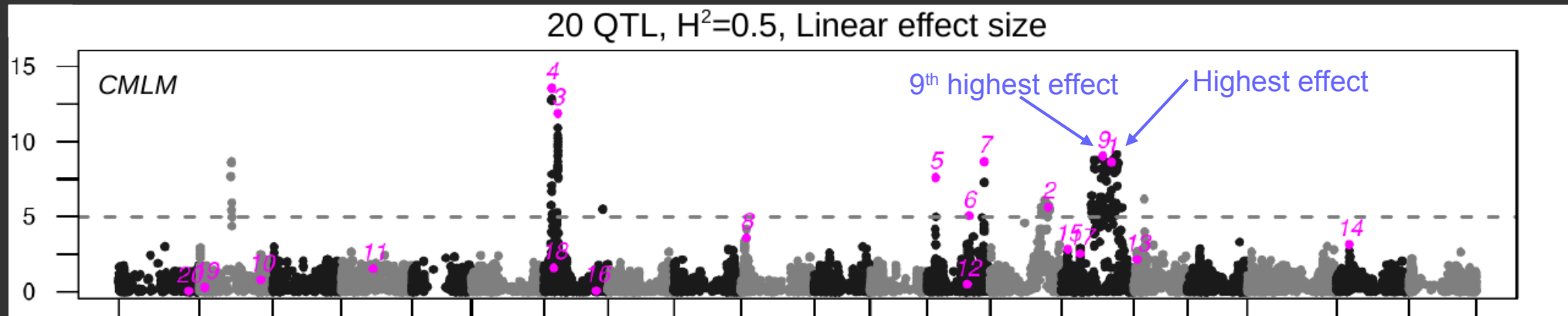| #QTL	4				20				200	
Effect Distr.	linear		uniform		linear		uniform		linear	
H^2	0.95	0.5	0.95	0.5	0.95	0.5	0.95	0.5	0.95	0.5
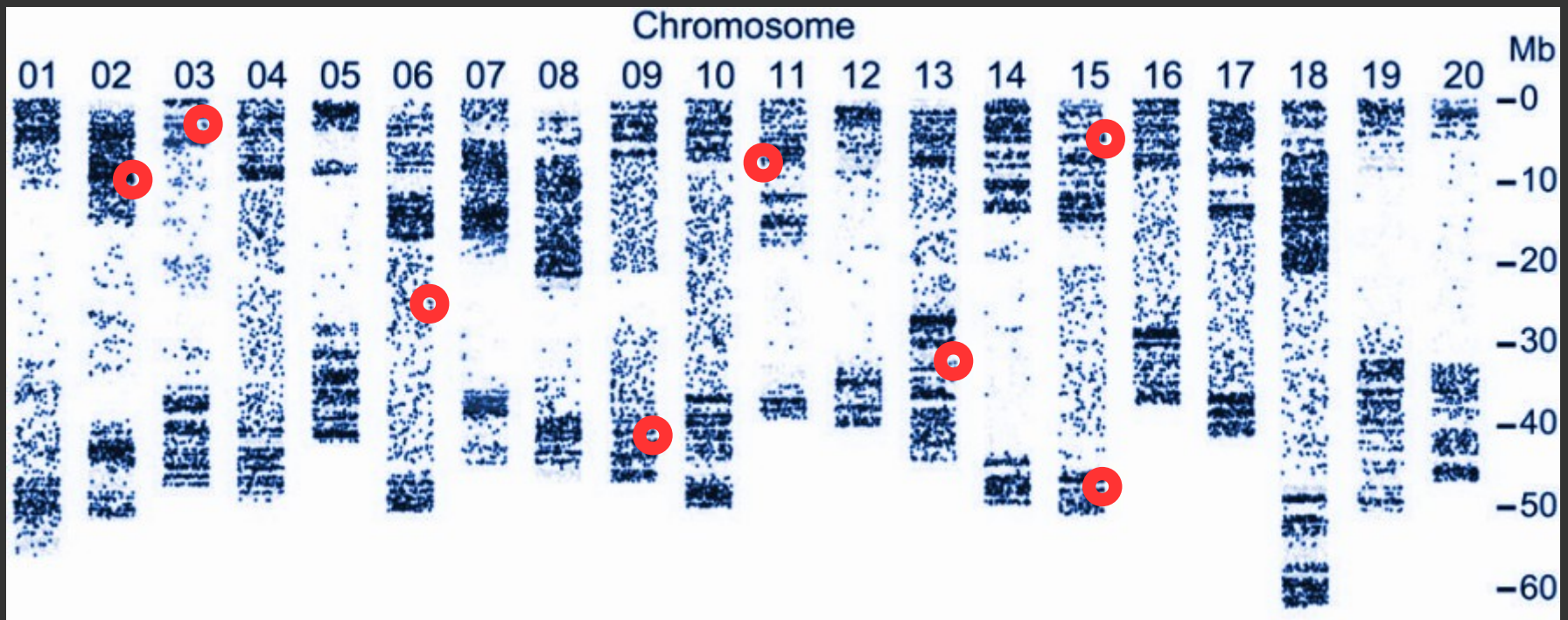
Modified from Haun et al. 2011. Plant physiol..

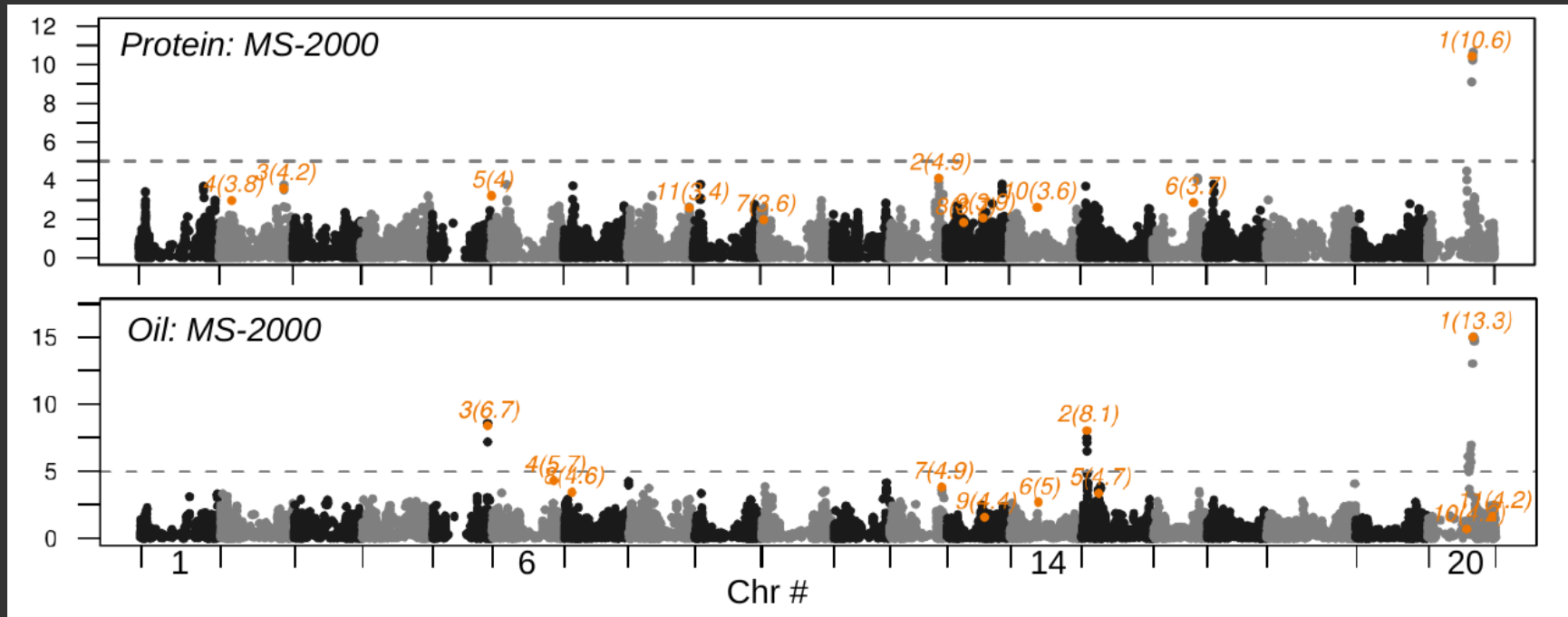# Simple architecture "easy" to predict . . .

(Multi-locus)
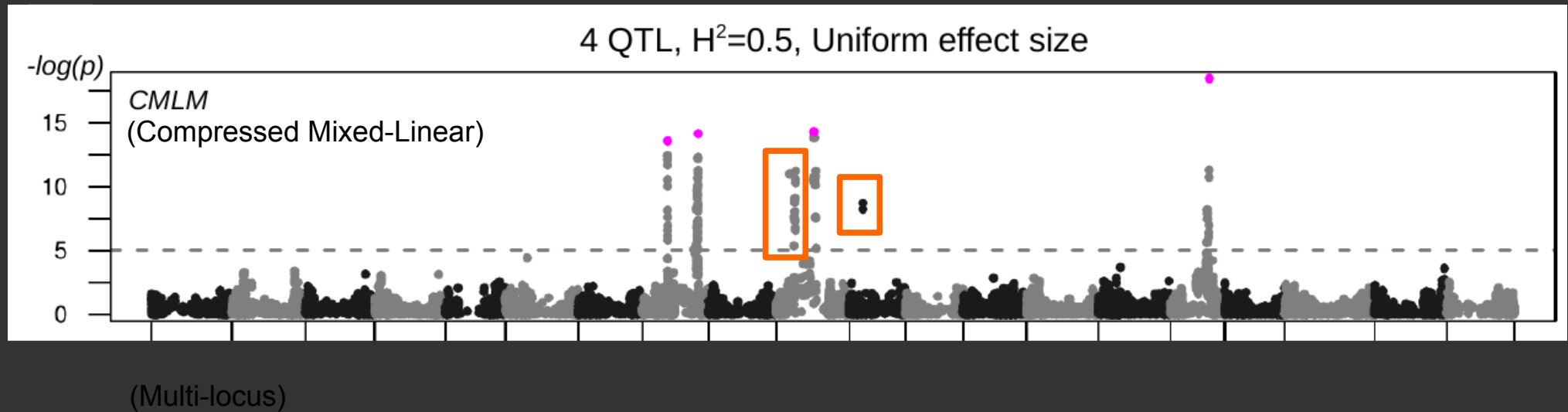
Complex architecture requires more critical evaluation . . .
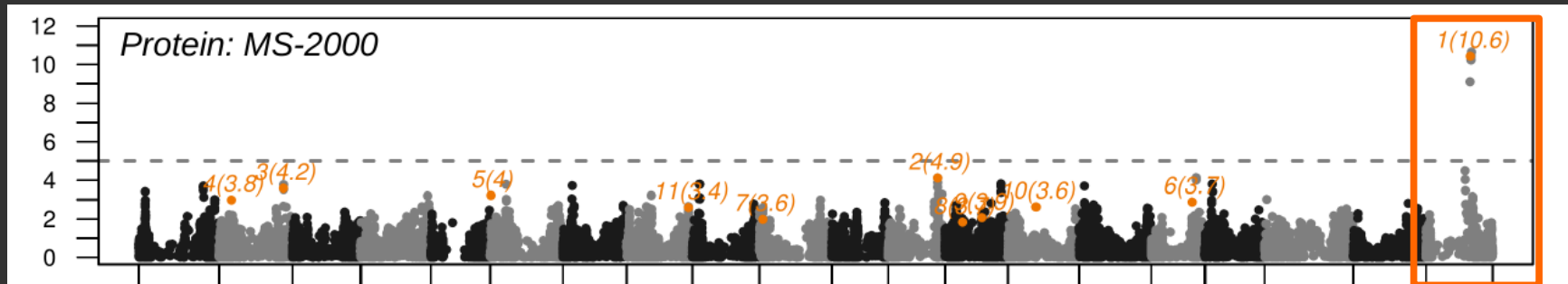
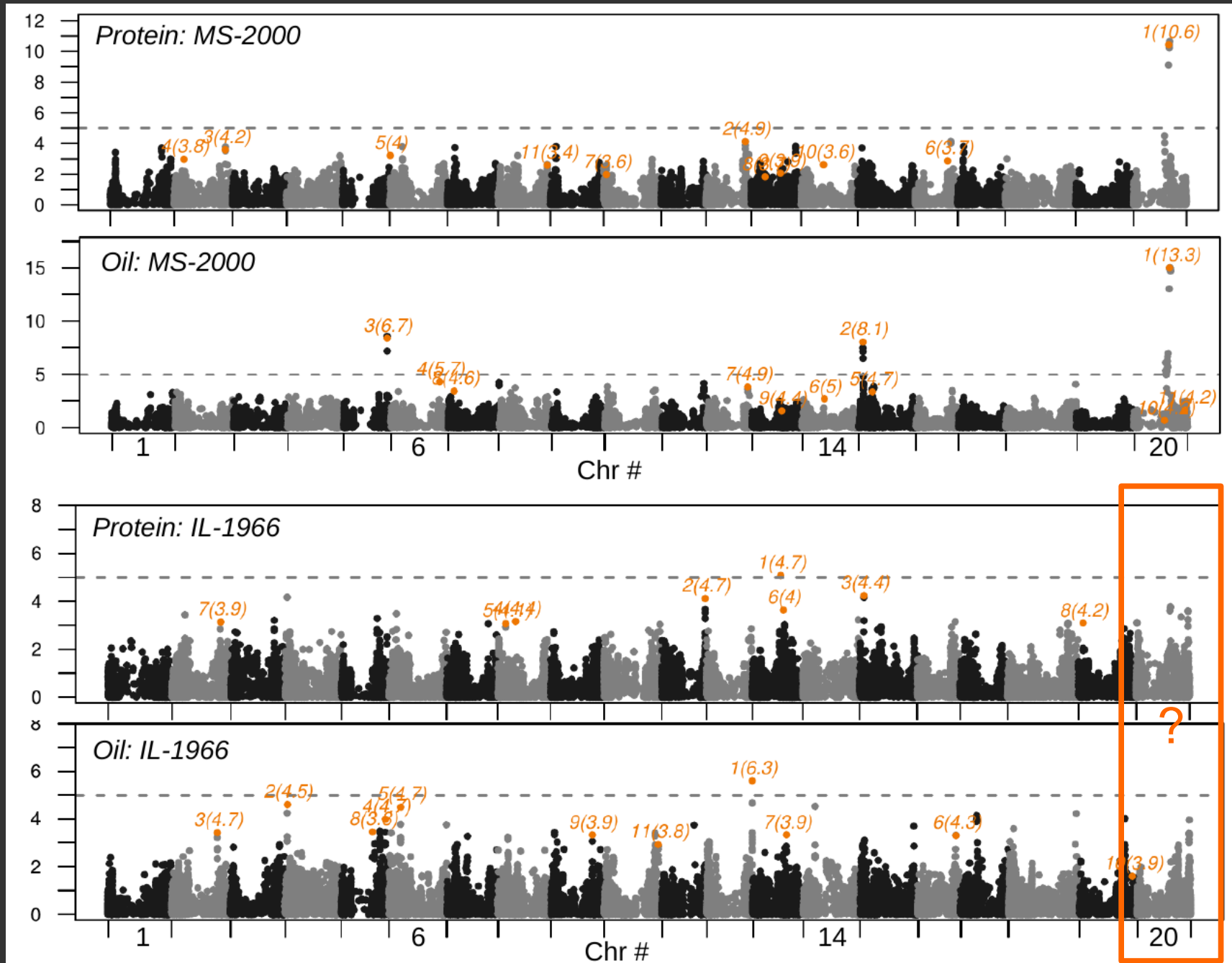20 QTL, H²=0.5, Linear effect size

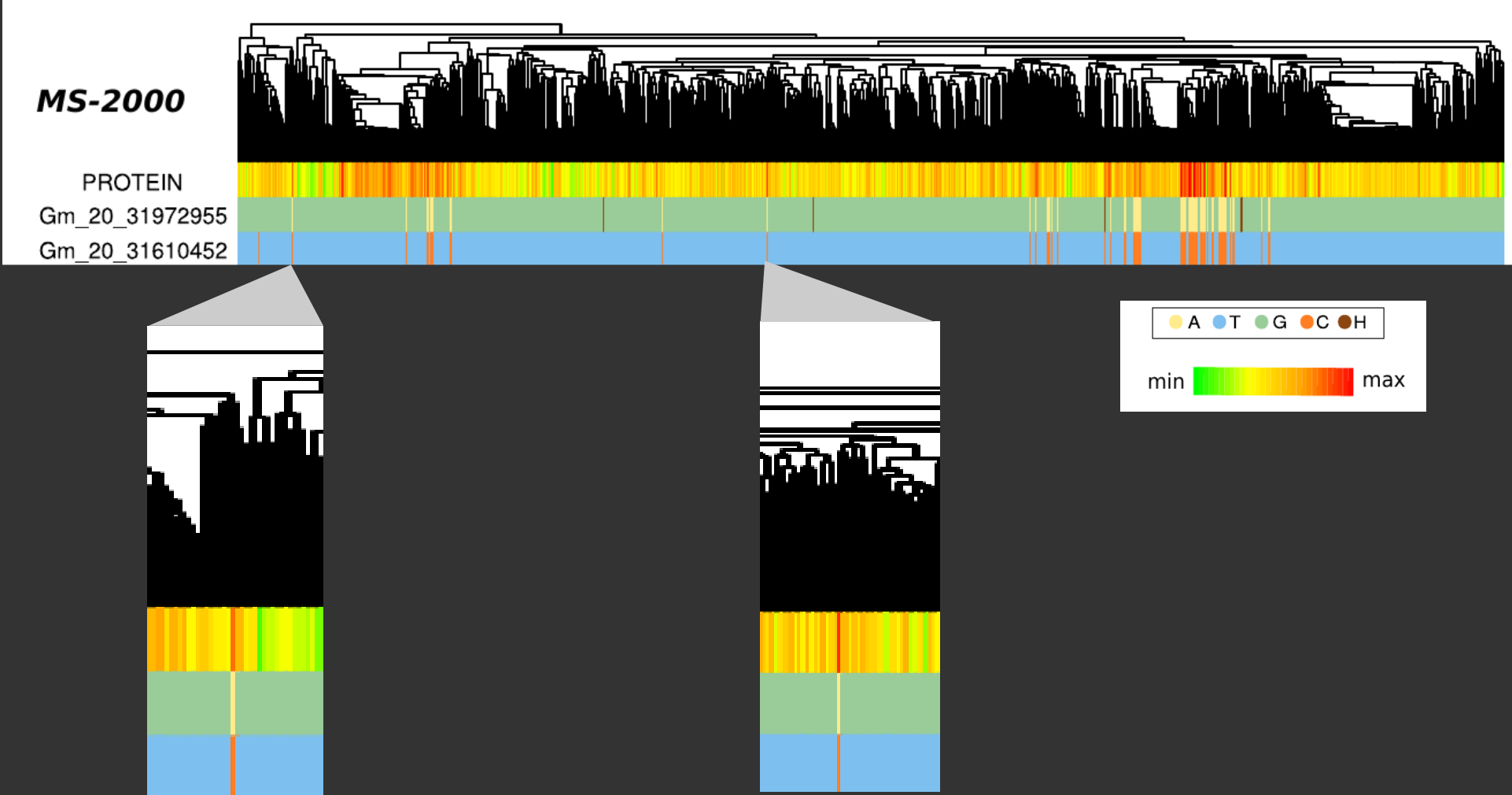# Linkage group I (chr 20) in MS-2000; what about IL-1996? . . .

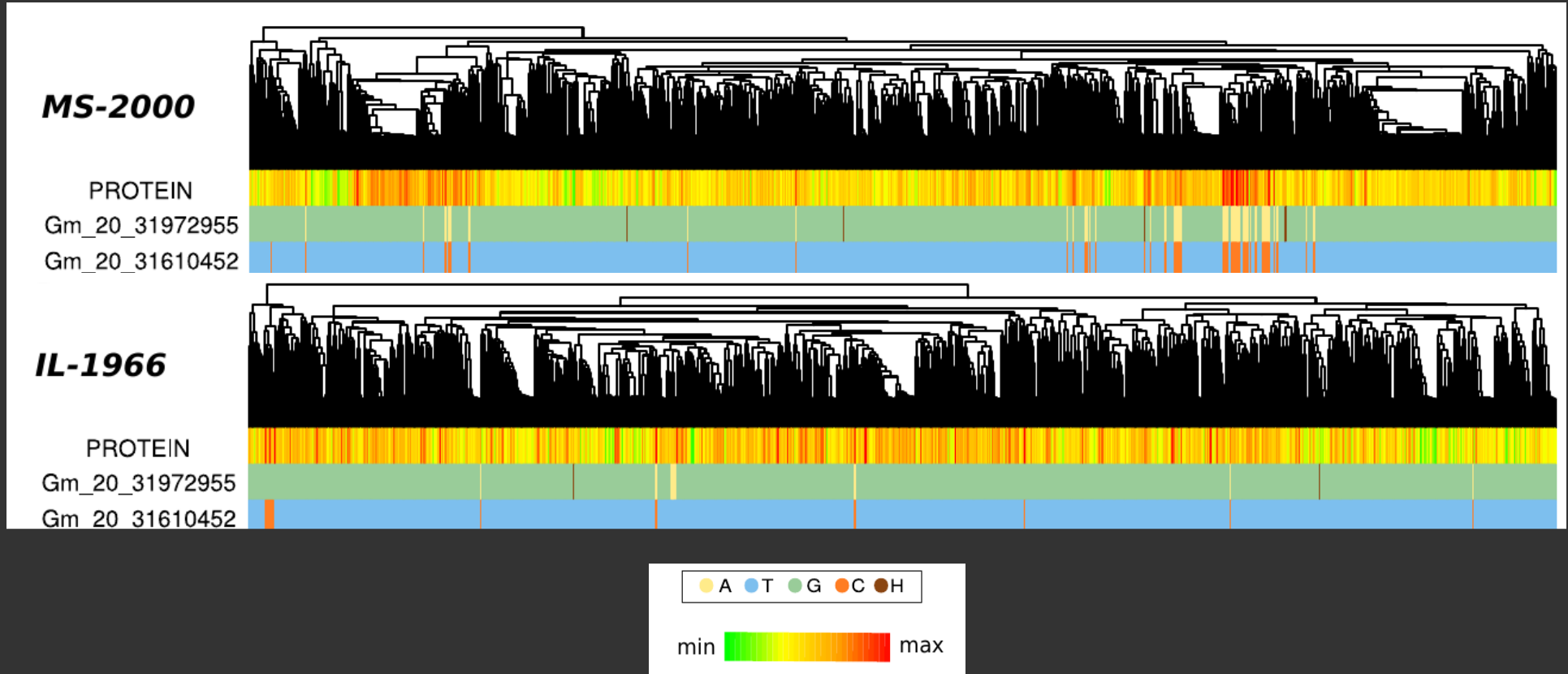# Linkage group I (chr 20) in MS-2000; what about IL-1996? . . .

Linkage group I (chr 20) in MS-2000; what about IL-1996? . . .

GWA scan by eye . . .

GWA scan by eye . . .
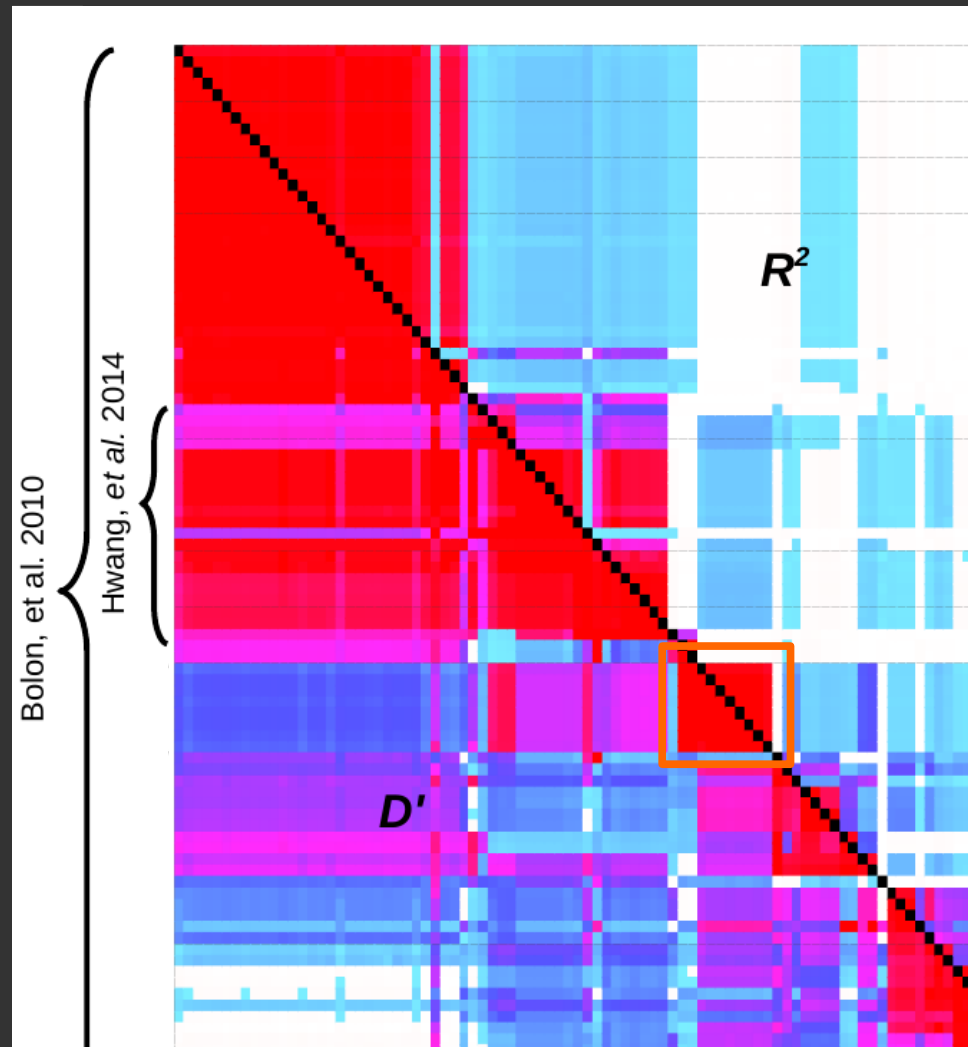

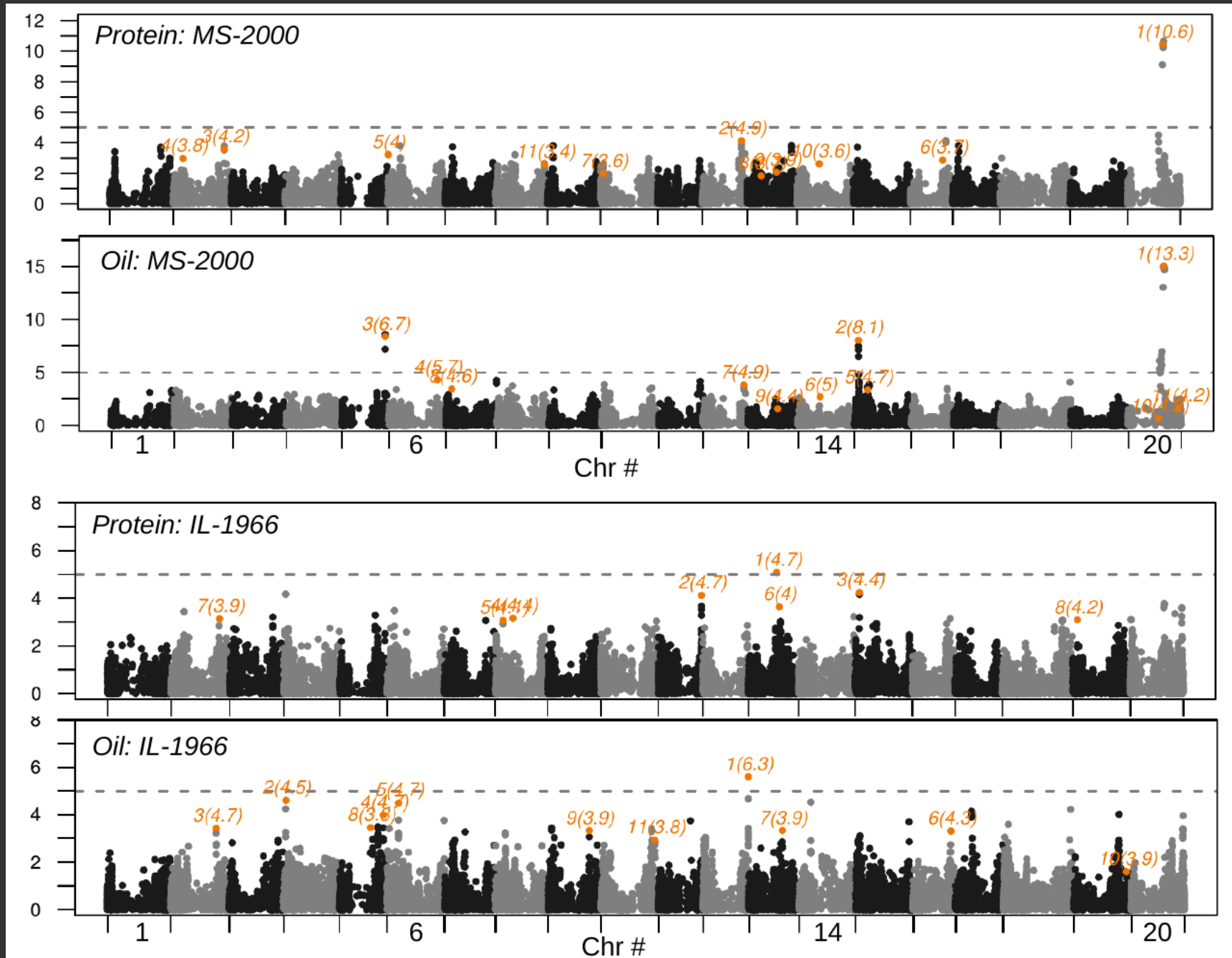**Allele has a very low frequency in many populations.**

The LG-I protein QTL is easily missed because selection has driven it to low frequencies . . .

30.5-32.2MB "Domestication" Locus
Chung et al. 2014

A GENOME-WIDE ASSOCIATION ANALYSIS FOR SEED OIL AND PROTEIN ON 13, 000 SOYBEAN ACCESSIONS

Bandillo N, . . ., Lorenz A

Heritability is ~85%, but mostly "high-hanging" fruit . . .

Quantity *and* quality . . .
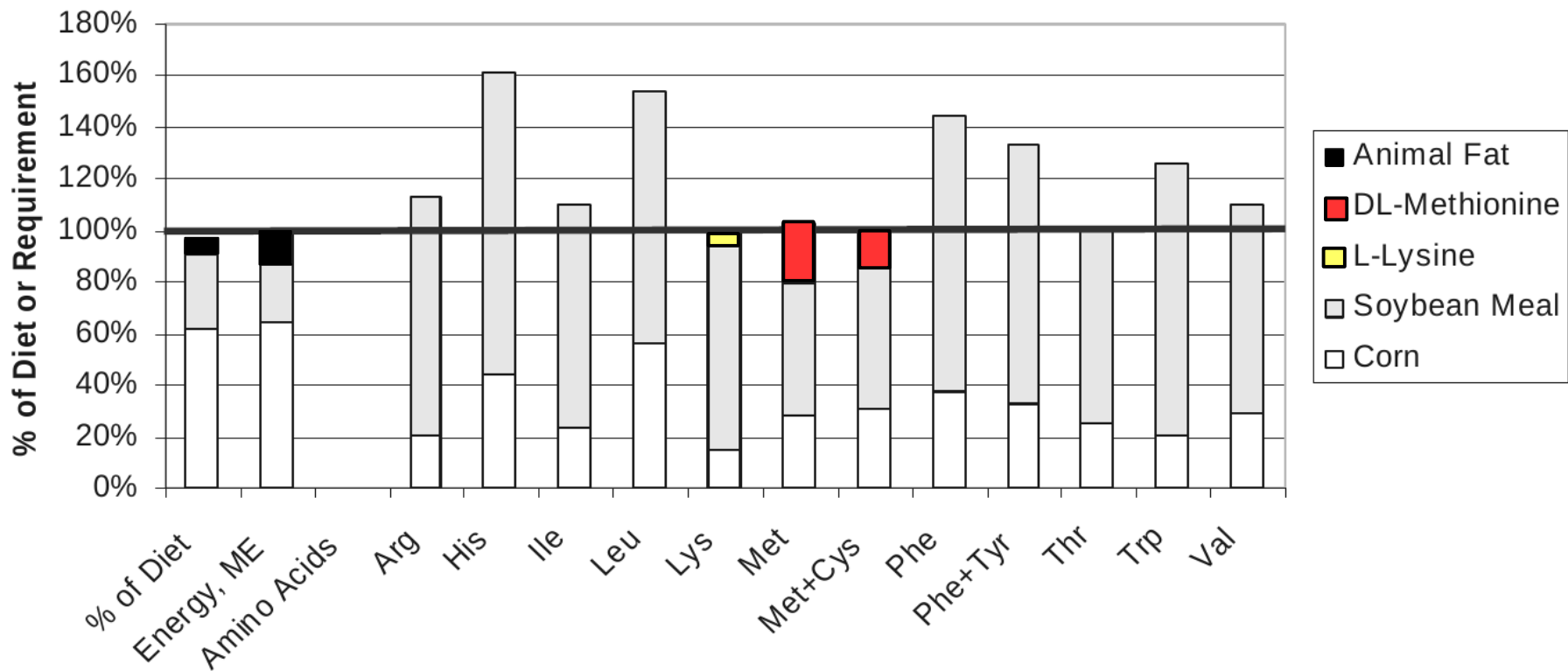
Energy and Amino Acid Contributions as a Percentage of Dietary Requirement for Growing Broiler Chickens

~$5 per ton per 10% increase of an amino acids

No overlap between GWA scan and bi-parental mapping results for cysteine and methionine . . .

Methionine: IL-1996

1(11.5)
2(8.7)
3(7.5)
7(4.4)
6(4.9)
5(4.9)
8(4.3)
9(3.9)
10(3.7)
11(3.8)
4(7.4)

Cysteine: MS-1997

1(12.3)
Near conglycinin QTL, Panthee et al. 2004
2(9.9)
3(4.3)
7(4.2)
10(3.8)
9(4.5)
4(4.5)
5(4.1)
8(4.1)
6(4.1)
11(4)
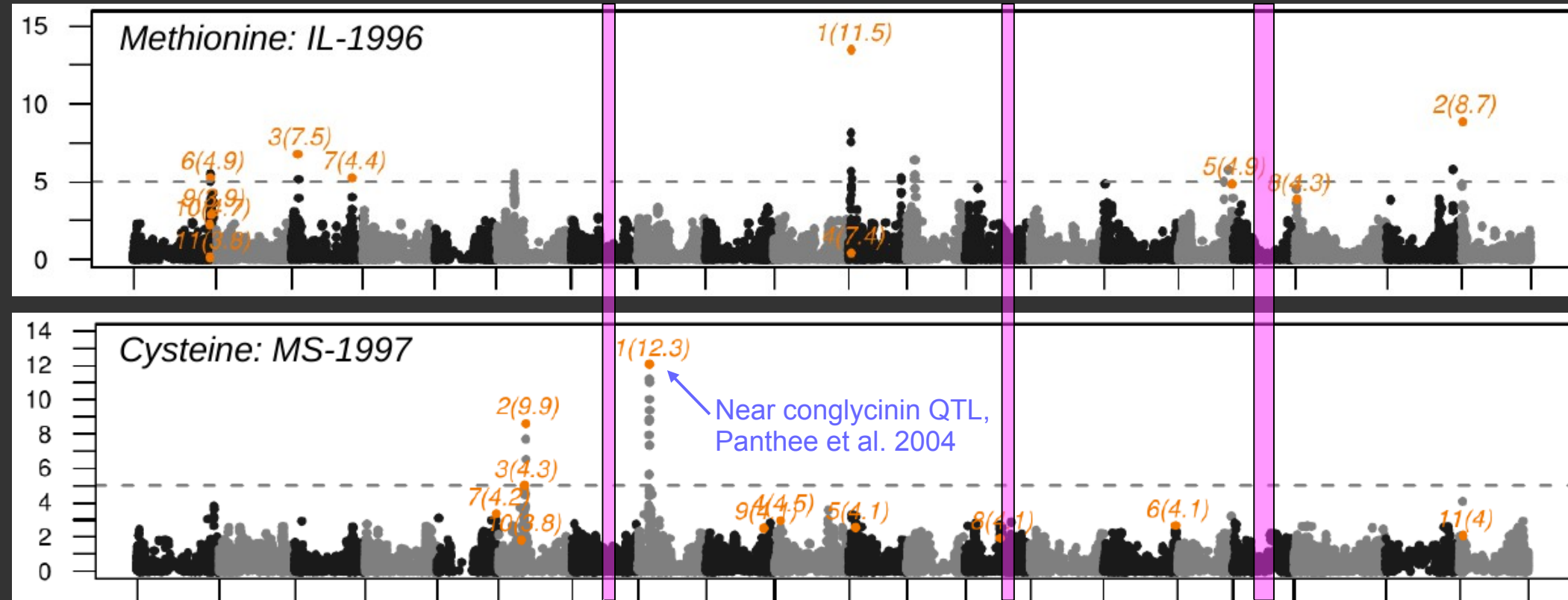
Met+Cys QTL identified in Panthee et al. 2009, Molecular Breeding

Why experimental design and genotyping platforms matters . . .

	Bi-parental cross	Nested Association Mapping (NAM)	GWA
Resolution	X	?	√
Missing alleles for a trait	Many	Some	Few

Thanks to

Zenglu Li
Li Lab
Roger Boerma
Randy Nelson
Qijian Song
Perry Cregan
Ester Peregrine
Jim Orf

Wayne Parrot
Tom Jacobs
Maria Ortega

Contact me (*jnvaughn@uga.edu*) or Zenglu (*zli@uga.edu*)