HOS 6236 Molecular Marker Assisted Plant Breeding

Fall 2017

Last Class:

QTL analysis methods

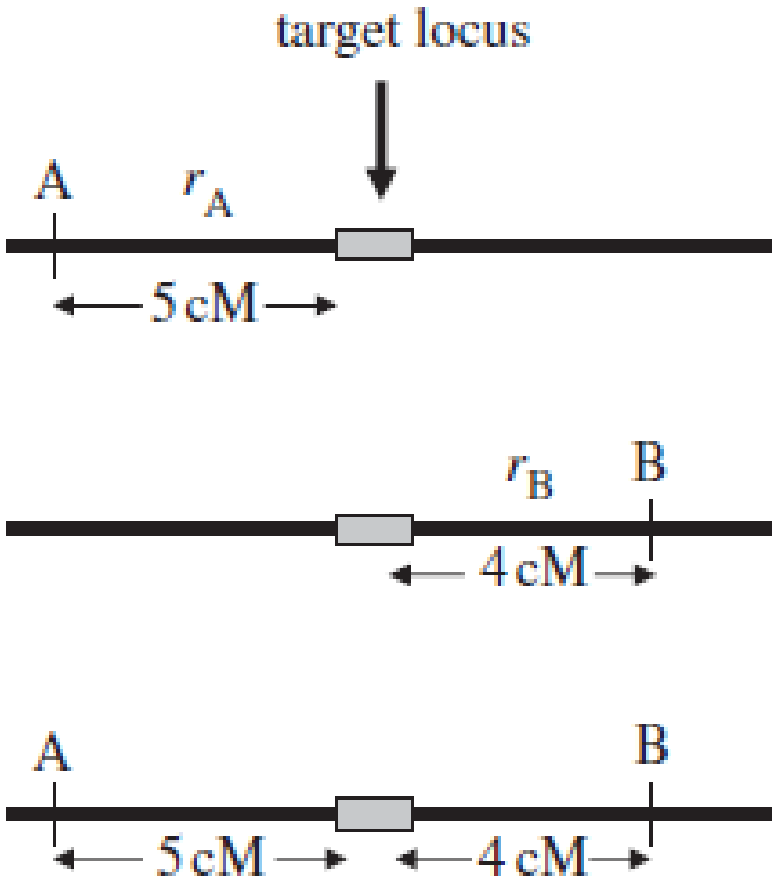
Todays Class:

More about QTL analysis, hands-on activity

Steps for a QTL analysis

1. Create or find a suitable population
2. Genotype with molecular markers
3. Use markers to build a linkage map
4. Phenotype for trait of interest (and more)
5. **Use the linkage map with the phenotypic data to determine whether markers are correlated to traits**; Number of QTLs, the amount of variation and position on genome

Single vs. Flanking Markers



reliability for selection

using marker A only:
$1 - r_A = \sim 95\%$

using marker B only:
$1 - r_B = \sim 96\%$

using both marker A and B:
$1 - 2\,r_A r_B = \sim 99.6\%$

Collard and Mackill, 2007

Permutation Testing

- LOD threshold for QTL is commonly determined by permutation testing
  - Shuffle phenotypic data
  - Keep genotypic data constant
  - Repeat 1000 or more times
  - Determine a significance level (α)

Permutation Testing

- If $\alpha = 0.05$, then with 1000 permutation tests, we would falsely declare a QTL-marker association 50 times
  - Genome-wide LOD score is calculated for each random iteration of phenotypic data
  - For 1000 permutation, the 950^{th} largest LOD score becomes the LOD threshold value

QTL Map

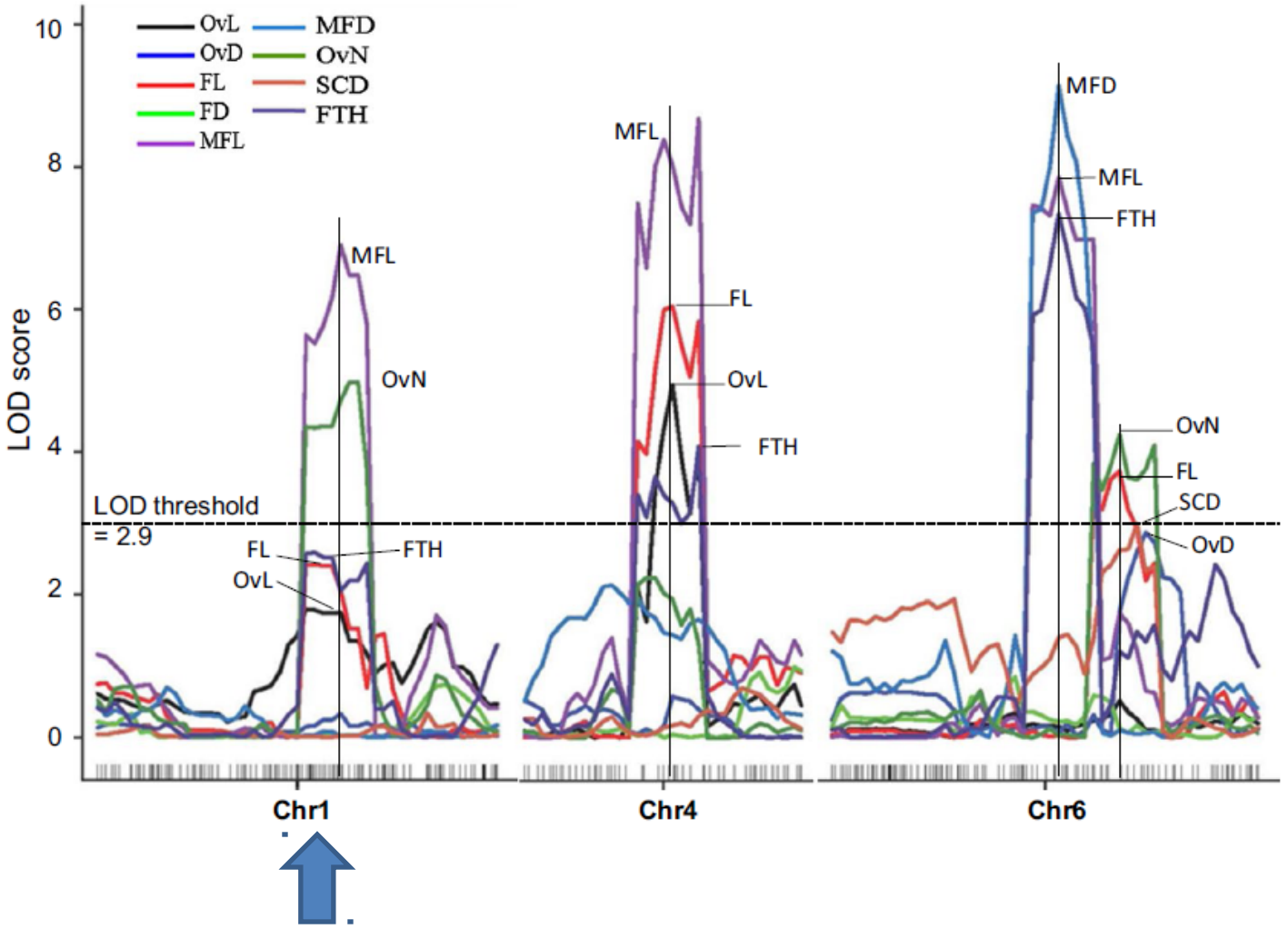
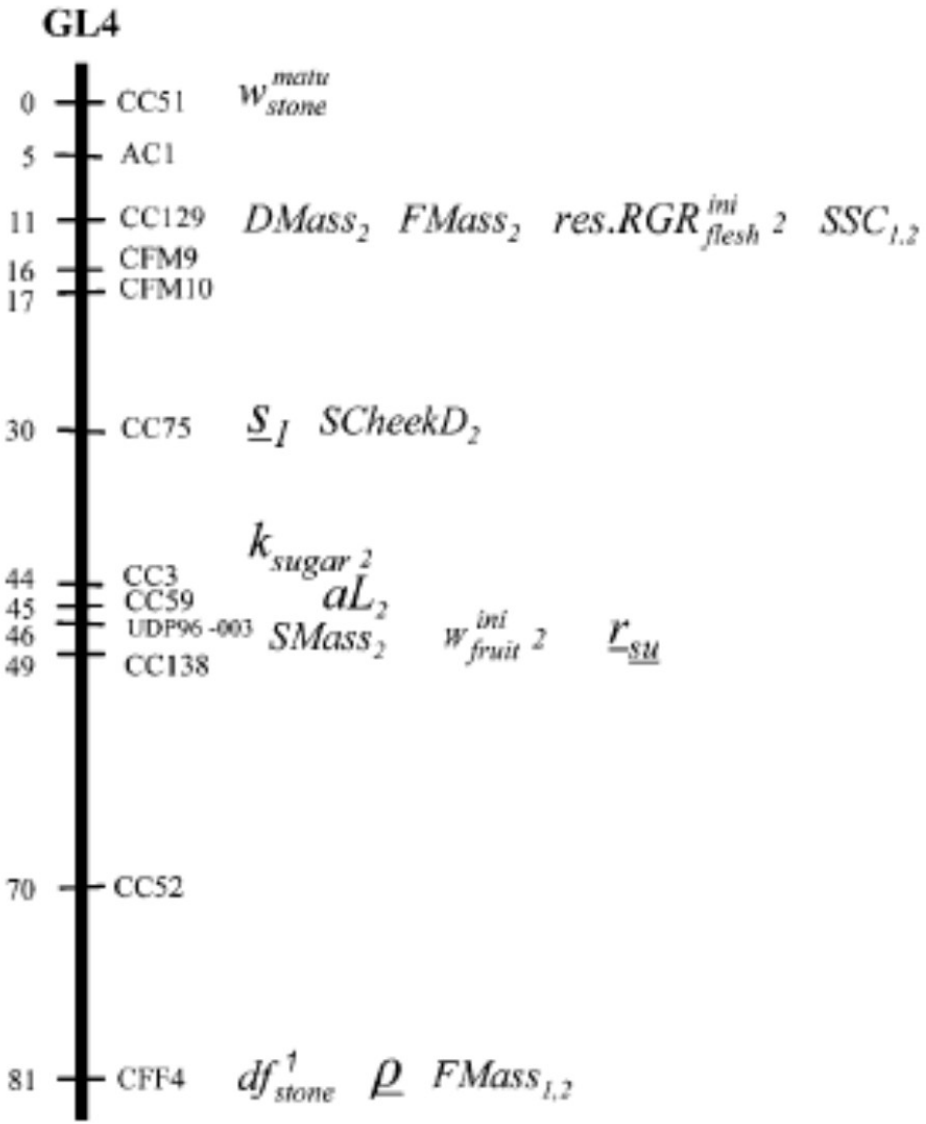


Fig. 4 LOD profiles of fruit size-related QTLs detected with MQM model in the RIL population and high-density SNP maps in cucumber chromosomes 1 (*left*), 4 (*middle*), and 6 (*right*). The *dashed horizontal line* is LOD threshold for all QTLs (LOD = 2.9). *OvL* ovary length, *OvD* ovary diameter, *FL* immature fruit length, *FD* immature fruit diameter, *MFL* mature fruit length, *MFD* mature fruit diameter. *FTH* flesh thickness, *SCD* seed cavity size and *OvN* ovule number in mature fruit

Weng et al. (2015) Theor. Appl Genet. 128:1747

28.5 cM interval: Width of interval – marker density; 2 LOD confidence interval (Lynch and Walsh 1998)

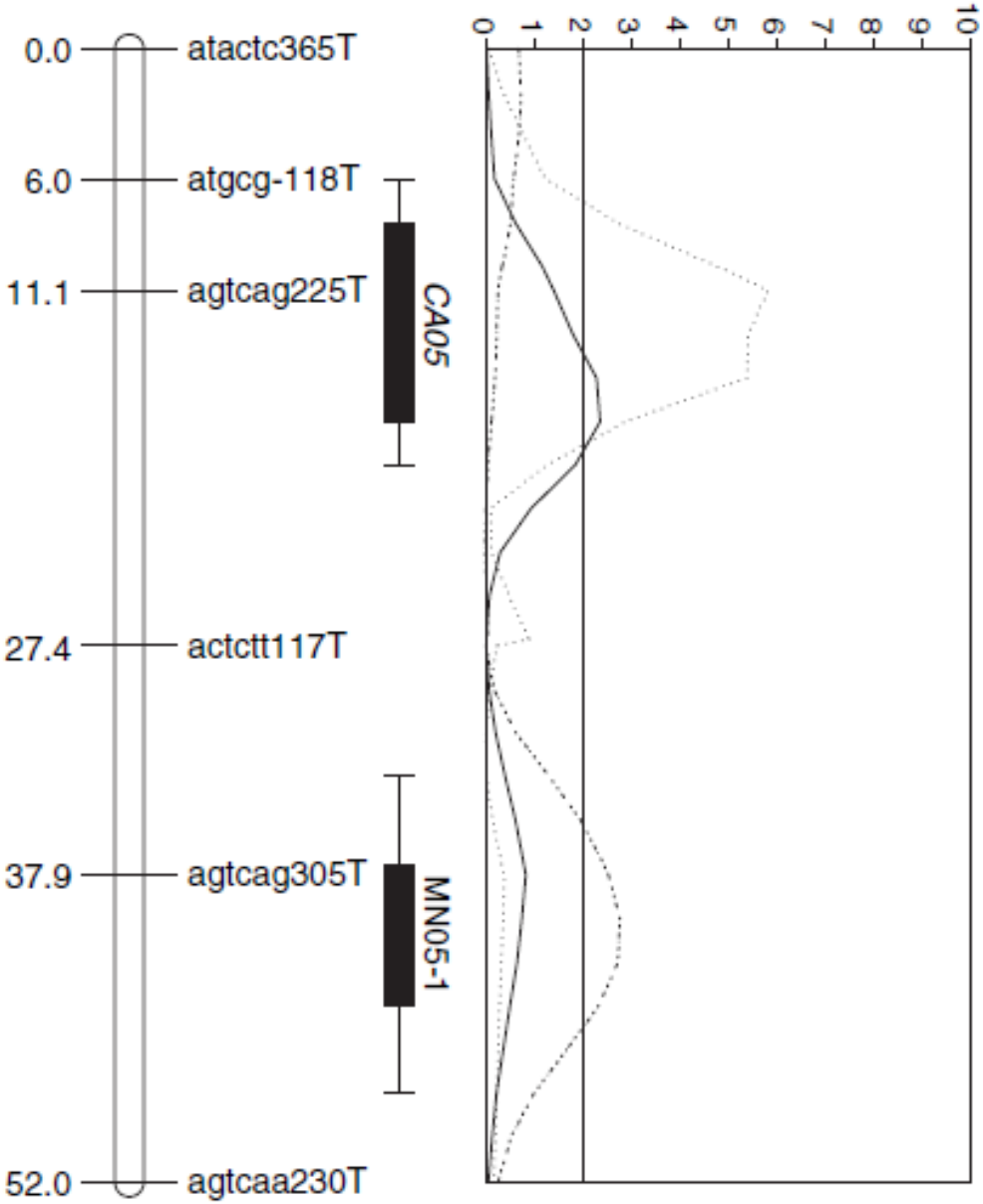
Reporting a QTL



Quilot et al., 2005

UF UNIVERSITY of FLORIDA
The Foundation for The Gator Nation

Reporting a QTL



Weebadde et al., 2008

UF UNIVERSITY of FLORIDA
The Foundation for The Gator Nation

Reporting a QTL

Table 4 Characteristics of the detected QTLs for each trait measured in different years

Trait	Year	Linkage group[a]	Parent[b]	QTL[c]	Nearest marker	LOD score	LOD threshold[d]	R^2 (%)[e]
Number of clusters/vine	2003	19	I	*Cn3.1*	*m*CTC *e*ATC9	3.28	2.7	7
	2004	8	BP	*Cn4.1*	VVS4	4.23	2.5	10
		8	BP	*Cn4.2*	VMC7H2	4.17	2.5	10
Cluster weight	2002	12	I	*Cw2.1*	*m*CAC *e*ACA7	3.19	2.6	2.3
	2003	5	BP	*Cw3.1*	*m*CAT *e*AAG13	3.01	2.8	1.3
		16	I	*Cw3.2*	VMC1E11	2.98	2.5	4
	2004	5	BP	*Cw4.1*	*m*CAT *e*AAG13	3.32	2.7	6.7
		17	BP	*Cw4.2*	*m*CTC *e*ATG12	3.02	2.7	4
Number of berries/cluster	2002	8	BP	*Bn2.1*	*m*CAT *e*AAG4	3.72	2.6	5
		12	BP	*Bn2.2*	*m*CTG *e*ATT3	3.12	2.6	7
		17	I	*Bn2.3*	*m*CTG *e*ATC8	3.02	2.8	6
		2	I	*Bn2.4*	VVI055	3.19	2.8	1.2
	2003	5	BP	*Bn3.1*	*m*CAT *e*AAG137	4.09	2.7	4
		7	I	*Bn3.2*	*m*CAT *e*ATT1	3.01	2.7	4.5
	2004	7	BP	*Bn4.1*	VMC7A4	4.25	2.6	9
		7	BP	*Bn4.2*	*m*CAT *e*ATG15	4.25	2.6	1.5
		5	I	*Bn4.3*	*m*CAG *e*ATG15	3.32	2.2	2.8
		7	I	*Bn4.4*	VVMD7	3.02	2.2	5.7
		7	I	*Bn4.5*	VMC16F3	2.83	2.2	4
Berry weight	2002	5	I	*Bw2.1*	*m*CAT *e*ATT2	3.31	3	10
		16	I	*Bw2.2*	*m*CTA *e*AAG5	3.35	3	3
		5	BP	*Bw2.3*	VMC3B9	3.2	2.7	19
	2003	4	BP	*Bw3.1*	VMC7H3	3.13	2.5	5
		13	BP	*Bw3.2*	*m*CAG *e*AAG13	2.95	2.5	2
	2004	20	BP	*Bw4.1*	*m*CAT *e*AAG14	3.19	2.5	5.8
		20	I	*Bw4.2*	*m*CTG *e*AAG3	3.24	2.8	4.8

Fanizza et al., 2005

Estimating the QTL effect

Sample size plays a significant role in the QTL effect.

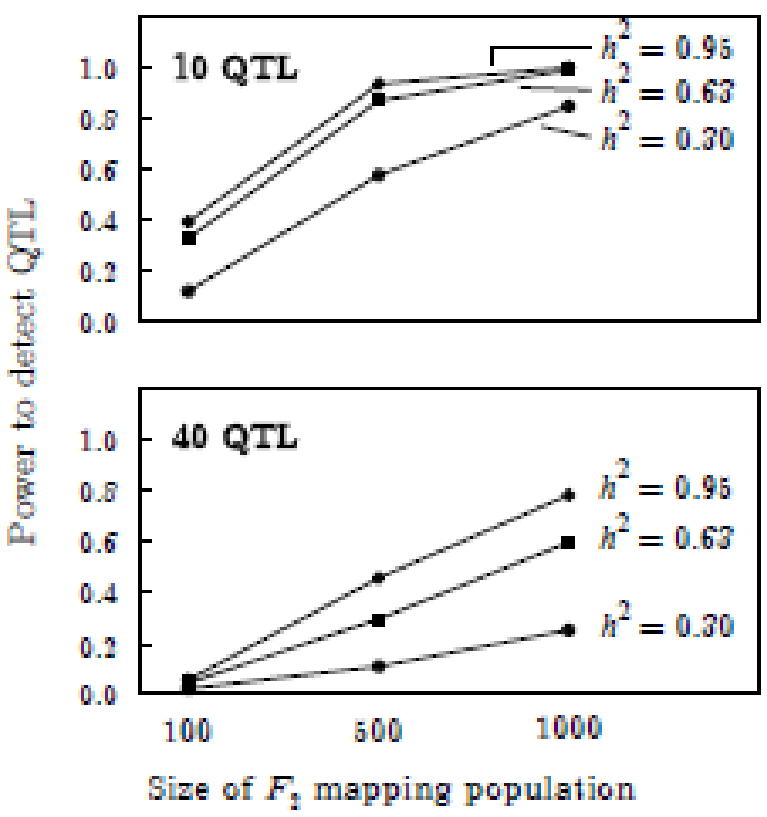
For example a small sample size may miss QTLs with small effect thus will overestimate the QTL effect of the ones that is able to identify, the “Beavis effect” (Beavis 1994, 1997)

Also, a small detected QTL effect could be either:

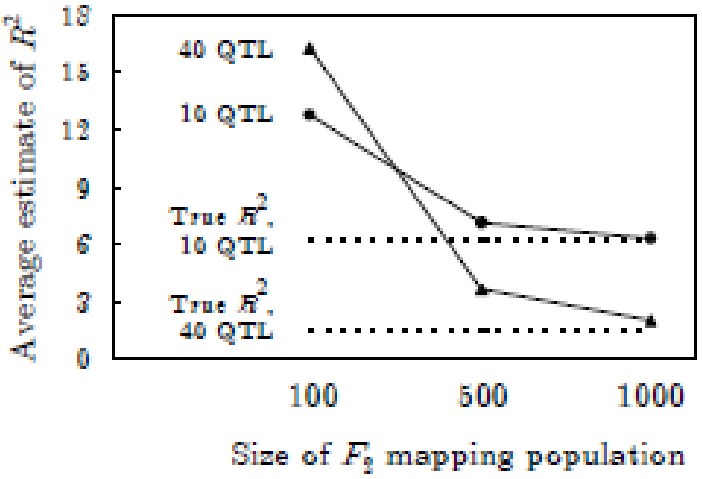
A tightly-linked QTL with a small effect

A loose linkage QTL with large effect

The Beavis Effect



- Number of QTL are underestimated with small population sizes



- Phenotypic variance associated with QTL are overestimated when small population sizes are used

Multiple Interval Mapping (MIM)

All previous methods tested by a single QTL at the time

Alternatively the test could be for multiple QTLs simultaneously. MIM consider multiple QTLs simultaneously.

Has higher precision

Can test for QTL-QTL interactions (i.e. epistasis)

Better description of the genetic architecture of a trait:

number, position, effect size, and interactions.

QTL analysis in outcrossing species

For some species inbred lines are just not possible

QTLs detected by:

- inbred-line crossing – fixed difference between lines
- outbred crosses – differences within population variation

Within population variation has less power to detect QTLs. In inbred all F1 are identical, so all are informative, and LD is maximized.

Use of population of relatives is common (i.e. half-sibs, full-sibs, etc.)

Inbred vs outbred

Not all outbred parents are **informative** (double heterozygous)

Number of alleles may be larger than 2 in outbred

Linkage phase may be different in different parents. So analysis by parent are needed.

Problems using QTLs in Breeding

1. QTLs are not the causal variant
2. QTLs are in LD with the causal variant
3. Different families can have different Marker allele – causal variant allele phase
4. Different families may have recombination among the marker and QTL
5. Different families may have epistatic effects that are not taken in account by QTLs
6. Breeding population may not have the allele discovered in the mapping population
7. There may be QTL-by-environment interaction

ALL THESE WILL BE REVIEWED IN THE MAS SECTION!