Selective Genotyping for Marker Assisted Selection Strategies for Soybean Yield Improvement

Ben Fallen

THE UNIVERSITY of TENNESSEE

Plant Sciences

INTRODUCTION

- Plant biotechnology in plant breeding offers new possibilities for:
  - increased productivity
  - crop diversification and production
  - developing a more sustainable agriculture
- One promising technique is molecular markers
  - The location of major loci is now known for disease resistance, tolerance to abiotic stresses and quality traits
  - Types of markers: RFLPs, SCARs, STS, SSRs and more recently SNPs

RATIONALE

- The genetic gain is ~1% a year in soybean
- The world population is expected to double by 2050 (U.S. Census Bureau, 2010)
- MAS for yield could greatly improve our understanding the genetic mechanisms of seed yield and increase breeding efficiency

PREVIOUS RESEARCH

- Many QTLs have been identified for quantitative traits
- Few have been confirmed in subsequent studies
- Even fewer have been utilized for MAS
- Most yield QTLs are population specific

OBJECTIVES

- SNPs associated with high yield are favorable for selecting high yielding lines across environments
- MAS can distinguish low yielding lines from high yielding lines
- Phenotypic selections differ from genotypic selections

$F_{5:9}$-DERIVED POPULATION OF ESSEX X WILLIAMS 82

- **Essex**
  - genetic background of many southern lines
  - gray pubescence
  - purple flowers
  - group V maturity
  - average protein and oil
  - average height and yield
  - susceptible to SDS
- **Williams 82**
  - genetic background of many northern lines
  - tawny pubescence
  - white flowers
  - group III maturity
  - average protein and oil
  - mild resistance to SDS

EXPERIMENTAL DESIGN

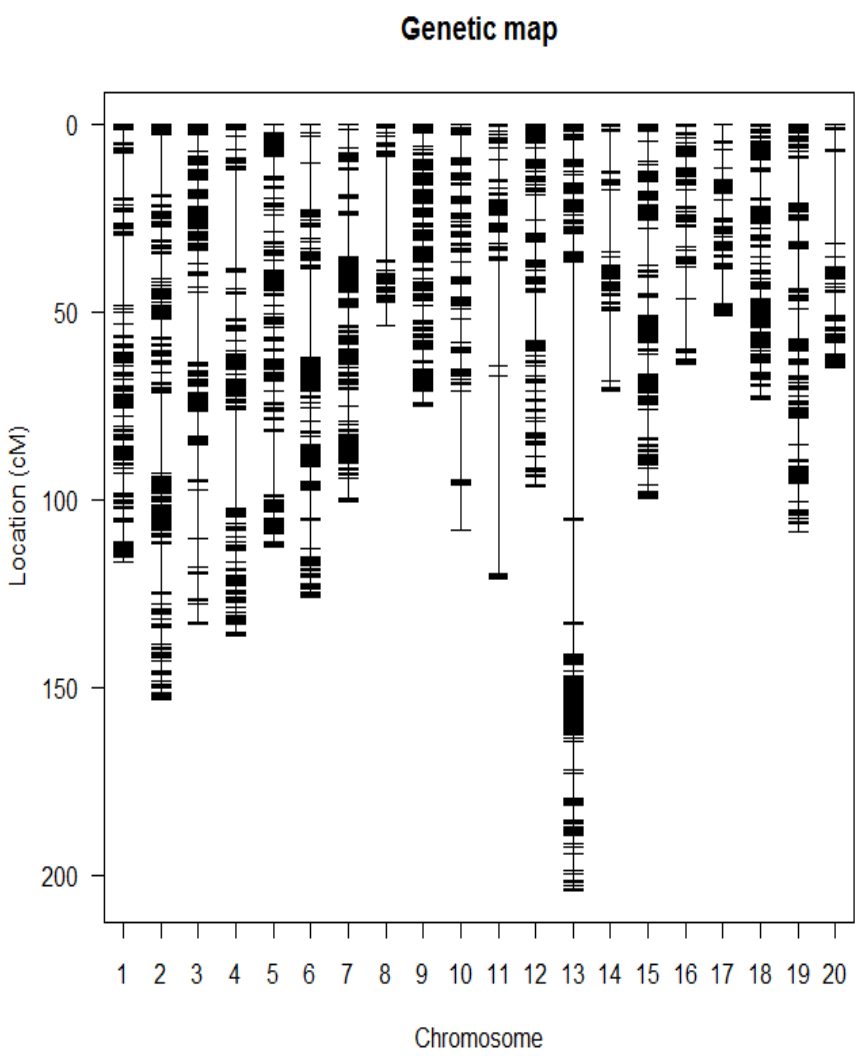
- ***Group A***
  - ***218 RILs***, 3 checks (IA3024, IA3023, LD00-3309) and the two parents grown in Knoxville, TN in 2010
- ***Group B***
  - ***221 RILs***, 3 checks (IA4005, LD00-3309,LD00-2817P) and the two parents grown in Knoxville, TN in 2010
- ***Group C***
  - ***216 RILs***, 3 checks (LD00-2817P, TN09-008 and 5002T) and the two parents grown in Knoxville, TN in 2010
- ***Group D***
  - ***220 RILs***, 3 checks (5002T, 5601T, Osage) and the two parents grown in Knoxville, TN in 2010

EXPERIMENTAL DESIGN

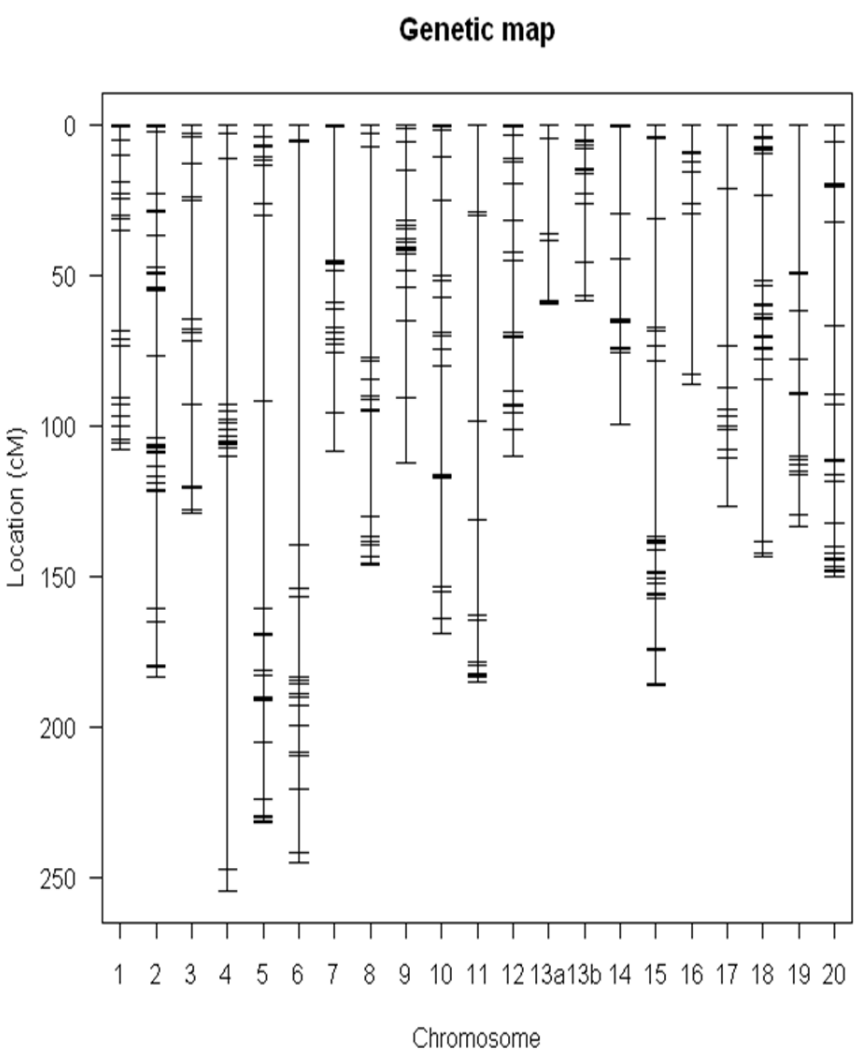
- ***Group A***
  - ***218 RILs***, 3 checks (IA3024, IA3023, LD00-3309) and the two parents grown in Knoxville, TN in 2010 and 2011 and Wooster, OH in 2011
- ***Group B***
  - ***221 RILs***, 3 checks (IA4005, LD00-3309,LD00-2817P) and the two parents grown in Knoxville, TN in 2010 and 2011 and Belleville, IL in 2011
- ***Group C***
  - ***216 RILs***, 3 checks (LD00-2817P, TN09-008 and 5002T) and the two parents grown in Knoxville, TN in 2010 and 2011 and Portageville, MO in 2011
- ***Group D***
  - ***220 RILs***, 3 checks (5002T, 5601T, Osage) and the two parents grown in Knoxville, TN in 2010 and 2011 and Plymouth, NC in 2011

EXPERIMENTAL ANALYSIS

- >50,000 SNPs (17,232 polymorphisms)
- QTL Analysis (additive effects)
  - R/qtl
  - Single factor ANOVA – SAS
- Epistatic Interactions
  - Epistacy (Holland, 1998) (additive x additive effects)
- YPM = x + A +AA



R/qtl map using 17,236 SNPs



R/qtl map using 480 SNPs

RESULTS
GROUP A: AGRONOMIC TRAITS

- Wooster, OH had an average yield (3339 kg ha^{-1}) that was significantly higher than the average yield in Knoxville, TN in 2010 (1756 kg ha^{-1}) and 2011 (1484 kg ha^{-1}).
- The yield in Knoxville, TN in 2010 averaged 52% and in 2011 averaged 44% of the yield in Wooster, OH in 2011

GROUP A: ADDITIVE EFFECTS

- R/qtl

ENVIRONMENT	MARKERS	CHR	MLG	LOC (cM)	LOD	R^2 (%)	ADDITIVE EFFECT	FAVORABLE ALLELE
Knoxville, TN 2010	Gm19_44937486_T_C	19	L	70.65	3.25	8.25	5.04	W
Knoxville, TN 2010	Gm02_707483_A_G	2	D1b	5.25	3.07	6.7	2.48	E
Knoxville, TN 2010	Gm04_48782140_G_T	4	C1	152.98	2.48	6.4	2.13	E
Wooster, OH 2011	Gm19_45198812_C_A	19	L	72.00	3.28	9.5	2.40	W
Wooster, OH 2011	Gm03_2151432_A_G	3	N	14.00	3.21	8.3	4.33	E
Wooster, OH 2011	Gm04_48993297_T_G	4	C1	154.16	2.78	5.2	3.18	E
Knoxville, TN 2010-11 Wooster, OH 2011	Gm19_44937486_T_C	19	L	70.75	3.75	7.2	3.17	W
Knoxville, TN 2010-11 Wooster, OH 2011	Gm05_33176582_G_A	5	A1	33.77	3.44	7.8	2.56	W
Knoxville, TN 2010-11 Wooster, OH 2011	Gm02_47790307_C_T	2	D1b	150.38	2.56	5.7	3.26	E

- Six QTLs were identified using R/qtl on five chromosomes (2, 3, 4, 5 and 19)

GROUP A: ADDITIVE EFFECTS

SAS

ENVIRONMENT	MARKERS	CHR	MLG	LOC (cM)	R^2 (%)	ADDITIVE EFFECT	FAVORABLE ALLELE	P-VALUE
Knoxville, TN 2010	Gm19_44937486_T_C	19	L	76.71	8.17	5.75	W	<0.0001
Knoxville, TN 2010	Gm15_43797502_G_T	15	E	72.68	6.38	1.88	W	0.002
Knoxville, TN 2010	Gm02_47790307_C_T	2	D1b	121.66	6.04	3.39	E	0.0028
Knoxville, TN 2010	Gm09_6967374_C_T	9	K	15.94	4.64	0.88	E	0.0106
Wooster, OH 2011	Gm19_44955912_T_G	19	L	76.84	7.98	-4.22	W	<0.0001
Wooster, OH 2011	Gm10_47585270_T_G	10	O	108.89	5.35	2.27	E	0.0049
Wooster, OH 2011	Gm02_49126947_T_C	2	D1b	127.25	5.31	3.44	E	0.0051
Wooster, OH 2011	Gm01_1494600_C_T	1	D1a	5.52	4.73	2.44	E	0.009
Knoxville, TN 2010-11 Wooster, OH 2011	Gm19_44964042_C_T	19	L	76.91	8.12	3.21	W	<0.0001
Knoxville, TN 2010-11 Wooster, OH 2011	Gm18_8772679_T_C	18	D2	33.67	6.88	2.83	W	0.0002
Knoxville, TN 2010-11 Wooster, OH 2011	Gm11_5773052_G_A	11	B1	20.42	6.53	3.80	E	0.0018
Knoxville, TN 2010-11 Wooster, OH 2011	Gm13_27348409_A_G	13	F	150.28	6.07	4.13	E	0.0006
Knoxville, TN 2010-11 Wooster, OH 2011	Gm14_49107190_G_A	14	B2	102.52	5.97	6.14	W	0.003
Knoxville, TN 2010-11 Wooster, OH 2011	Gm03_47386481_A_C	3	N	120.71	5.67	5.81	E	0.004
Knoxville, TN 2010-11 Wooster, OH 2011	Gm02_49126947_T_C	2	D1b	127.25	5.07	5.82	E	0.0071

- Eleven QTLs using SAS on eleven chromosomes (2, 3, 9, 10, 11, 13, 14, 15 and 19)

COMPARING MARKER ASSISTED SELECTIONS TO PHENOTYPIC SELECTIONS

MAS Individual/Over Env.		PS Over Env.
Top 5%	⟷	Top 5%
Top 10%	⟷	Top 10%
Top 25%	⟷	Top 25%
Top 50%	⟷	Top 50%
Bottom 50%	⟷	Bottom 50%
Bottom 25%	⟷	Bottom 25%
Bottom 10%	⟷	Bottom 10%
Bottom 5%	⟷	Bottom 5%

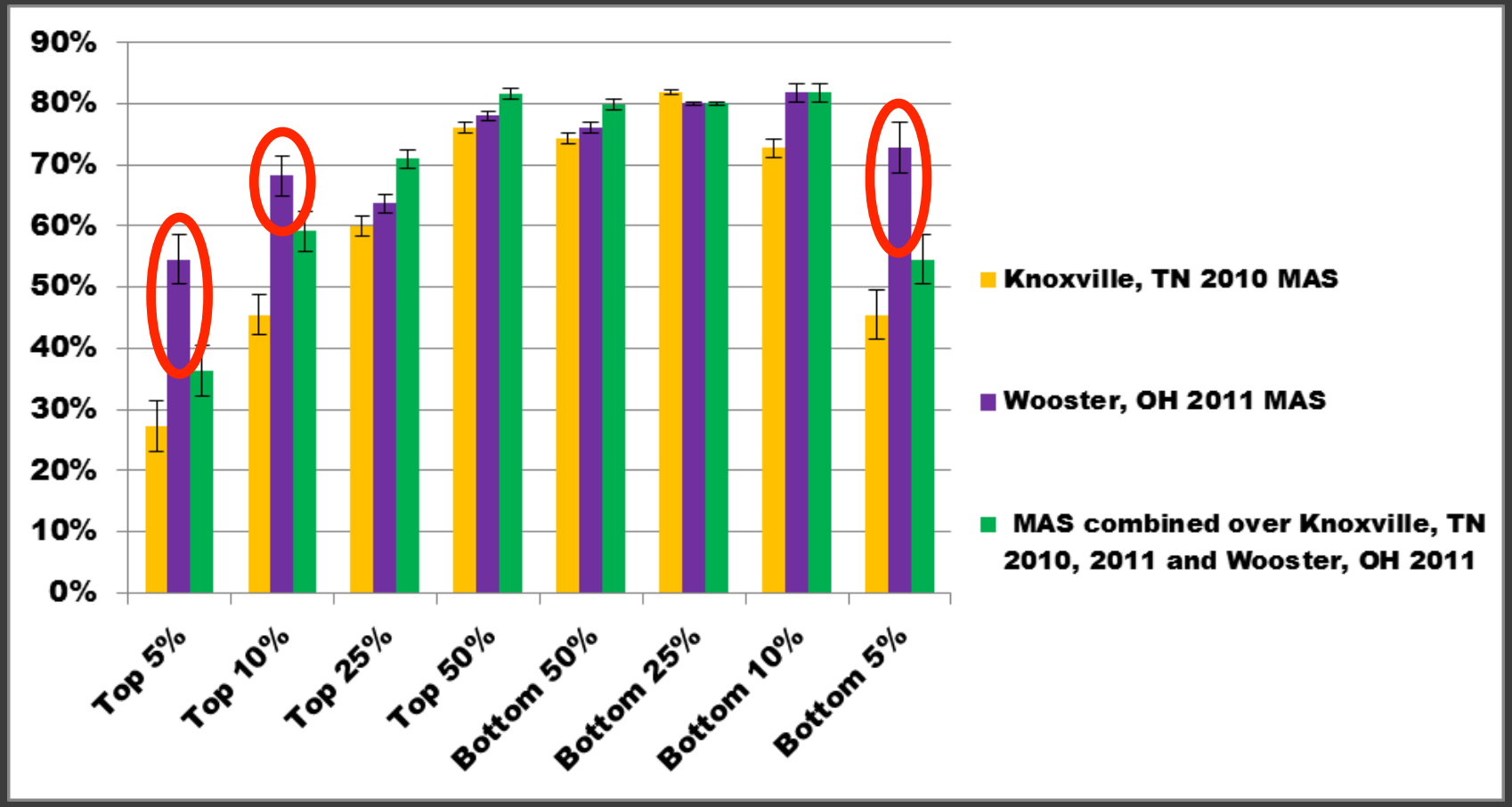
MARKER ASSISTED SELECTIONS						YIELD (kg ha^{-1})		
KNOXVILLE, TN 2010		WOOSTER, OH 2011		KNOXVILLE, TN 2010-11 WOOSTER, OH 2011		KNOXVILLE, TN 2010-11 WOOSTER, OH 2011		
LINE	RANK	LINE	RANK	LINE	RANK	LINE	YEILD	RANK
28	01	59	01	71	01	481	3319.2	01
45	02	62	02	90	02	833	3110.9	02
58	03	71	03	125	03	978	3003.4	03
90	04	86	04	144	04	689	2970.5	04
104	05	144	05	156	05	144	2969.8	05
106	06	225	06	211	06	463	2953.4	06
117	07	261	07	224	07	675	2875.7	07
120	08	337	08	260	08	578	2869.1	08
130	09	341	09	292	09	814	2828.7	09
134	10	344	10	344	10	756	2815.3	10
144	11	358	11	463	11	502	2808.5	11
146	12	428	12	481	12	292	2801.8	12
156	13	463	13	543	13	896	2801.8	13
203	14	481	14	583	14	632	2795.1	14
204	15	524	15	710	15	774	2795.1	15
211	16	592	16	751	16	637	2754.8	16
266	17	689	17	767	17	951	2748.1	17
291	18	737	18	814	18	668	2748.1	18
292	19	751	19	833	19	130	2727.9	19
358	20	756	20	896	20	454	2721.2	20
481	21	774	21	912	21	146	2714.5	21
487	22	814	22	951	22	751	2694.3	22

MARKER ASSISTED SELECTIONS						YIELD (kg ha^{-1})		
KNOXVILLE, TN 2010		WOOSTER, OH 2011		KNOXVILLE, TN 2010-11 WOOSTER, OH 2011		KNOXVILLE, TN 2010-11 WOOSTER, OH 2011		
LINE	RANK	LINE	RANK	LINE	RANK	LINE	YEILD	RANK
28	01	59	01	71	01	481	3319.2	01
45	02	62	02	90	02	833	3110.9	02
58	03	71	03	125	03	978	3003.4	03
90	04	86	04	144	04	689	2976.5	04
104	05	144	05	156	05	144	2969.8	05
106	06	224	06	211	06	463	2956.4	06
117	07	261	07	224	07	675	2879.7	07
120	08	337	08	260	08	578	2869.1	08
130	09	344	09	312	09	814	2838.7	09
134	10	344	10	344	10	756	2815.3	10
144	11	358	11	463	11	502	2808.5	11
146	12	428	12	481	12	292	2801.8	12
156	13	463	13	543	13	896	2801.8	13
203	14	481	14	583	14	632	2795.1	14
204	15	524	15	710	15	774	2795.1	15
211	16	592	16	751	16	637	2754.8	16
266	17	689	17	767	17	951	2748.1	17
291	18	737	18	814	18	668	2748.1	18
292	19	751	19	833	19	130	2727.9	19
358	20	756	20	896	20	454	2721.2	20
481	21	774	21	912	21	146	2714.5	21
487	22	814	22	951	22	751	2694.3	22

TOP 10% MAS

TOP 10% PS

GROUP A: TOP MAS RILS VS TOP YIELDING RILS



GROUP A: TOP MAS RILS VS TOP YIELDING RILS

- R/qtl
- 5 out of 11 RILs that were in the top yielding 5% were selected using MAS
- 7 out of 22 RILs that were in the top yielding 10% were selected using MAS

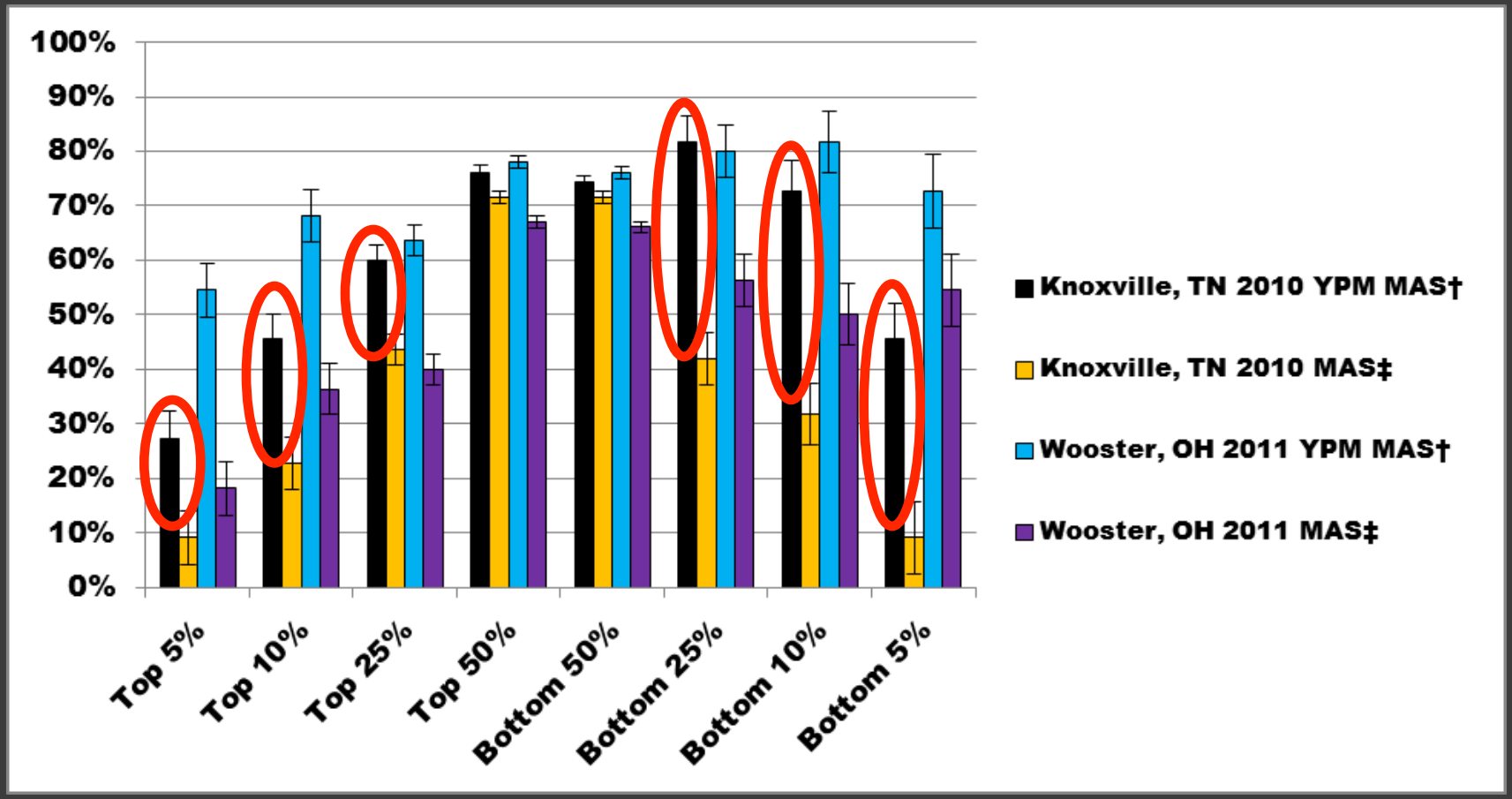
WOOSTER, OH 2011				
MAS		YIELD (kg ha^{-1})		
Line	Rank	Line	Yld	Rank
59	01	814	5227.4	01
62	02	292	5166.9	02
71	03	689	5160.2	03
86	04	559	4998.9	04
^{bb}**144**	05	978	4992.2	05
224	06	896	4918.3	06
261	07	481	4904.9	07
337	08	463	4857.8	08
341	09	144	4763.8	09
344	10	833	4710.0	10
358	11	146	4669.7	11
428	12	751	4642.8	12
^{bb}**463**	13	211	4636.1	13
^{bb}**481**	14	754	4575.6	14
524	15	148	4562.2	15
592	16	489	4562.2	16
^{bb}**689**	17	951	4562.2	17
737	18	767	4521.9	18
^{b}**751**	19	675	4521.9	19
756	20	774	4508.4	20
^{b}**774**	21	253	4508.4	21
^{bb}**814**	22	604	4501.7	22

GROUP A: EPISTATIC INTERACTIONS (R/QTL)

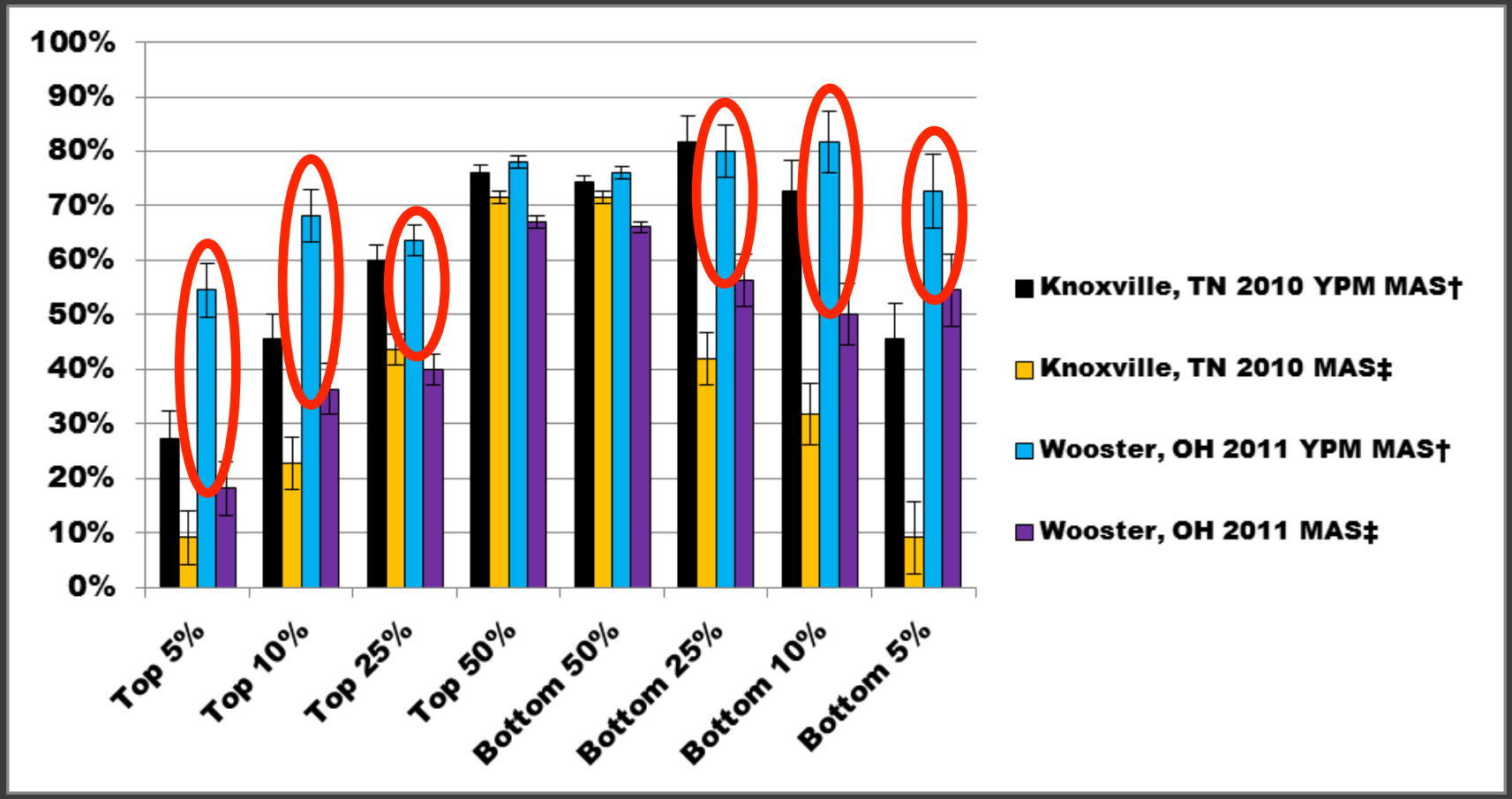
Given the Additive Effect of the QTL at Locus 1 and the Additive Effect at Locus 2

ENVIRONMENT	LOCUS 1	CHR	MLG	LOCUS 2	CHR	MLG	R^2 (%)	ADDITIVE X ADDITIVE EFFECT E	ADDITIVE X ADDITIVE EFFECT W
Knoxville, TN 2010	Gm19_44937486_T_C	19	L	GM15_10059948_T_C	15	E	3.12	5.80	3.01
				GM15_50338705_T_C	15	E	2.77	5.83	3.31
				GM20_41180602_G_A	20	I	3.01	5.72	3.10
Knoxville, TN 2010	Gm04_48782140_G_T	4	C1	GM06_45433980_G_A	6	C2	4.22	-0.46	3.09
				GM11_37065128_T_C	11	B1	4.20	-1.43	1.59
Wooster, OH 2011	Gm19_45198812_C_A	19	L	GM04_11182315_A_G	4	C1	3.54	0.19	5.91
				GM05_32908802_T_C	5	A1	5.14	-1.30	5.46
				GM13_28429921_T_C	13	F	3.68	-0.14	5.81
				GM20_12318232_A_G	20	I	3.52	5.18	-0.49
Wooster, OH 2011	Gm04_48993297_T_G	4	C1	GM06_49103970_C_T	6	C2	4.65	-0.65	5.77
				GM10_37618173_A_G	10	O	5.92	-2.44	4.68
				GM19_44478931_A_G	19	L	2.67	0.90	6.10
Knoxville, TN 2010-11 Wooster, OH 2011	Gm19_44937486_T_C	19	L	GM05_39611177_C_T	5	A1	1.94	4.83	7.09
				GM11_38762112_G_T	11	B1	1.78	4.65	6.70
				GM15_49657706_C_T	15	E	3.70	7.32	4.30
				GM19_42189531_T_C	19	L	1.66	9.48	5.19
Knoxville, TN 2010-11 Wooster, OH 2011	Gm05_33176582_G_A	5	A1	GM02_32518097_T_C	2	D1b	3.69	0.95	-1.62
				GM16_28901653_G_A	16	J	3.66	1.27	-1.24
				GM20_34223656_G_A	20	I	3.89	1.40	-1.32
Knoxville, TN 2010-11 Wooster, OH 2011	Gm02_47790307_C_T	2	D1b	GM02_46778366_G_A	2	D1b	4.42	-1.89	2.85
				GM04_29535808_A_G	4	C1	3.64	0.04	2.73
				GM18_48533018_G_A	18	D2	4.13	-0.03	2.88
				GM19_50486916_C_T	19	L	4.14	0.29	3.13

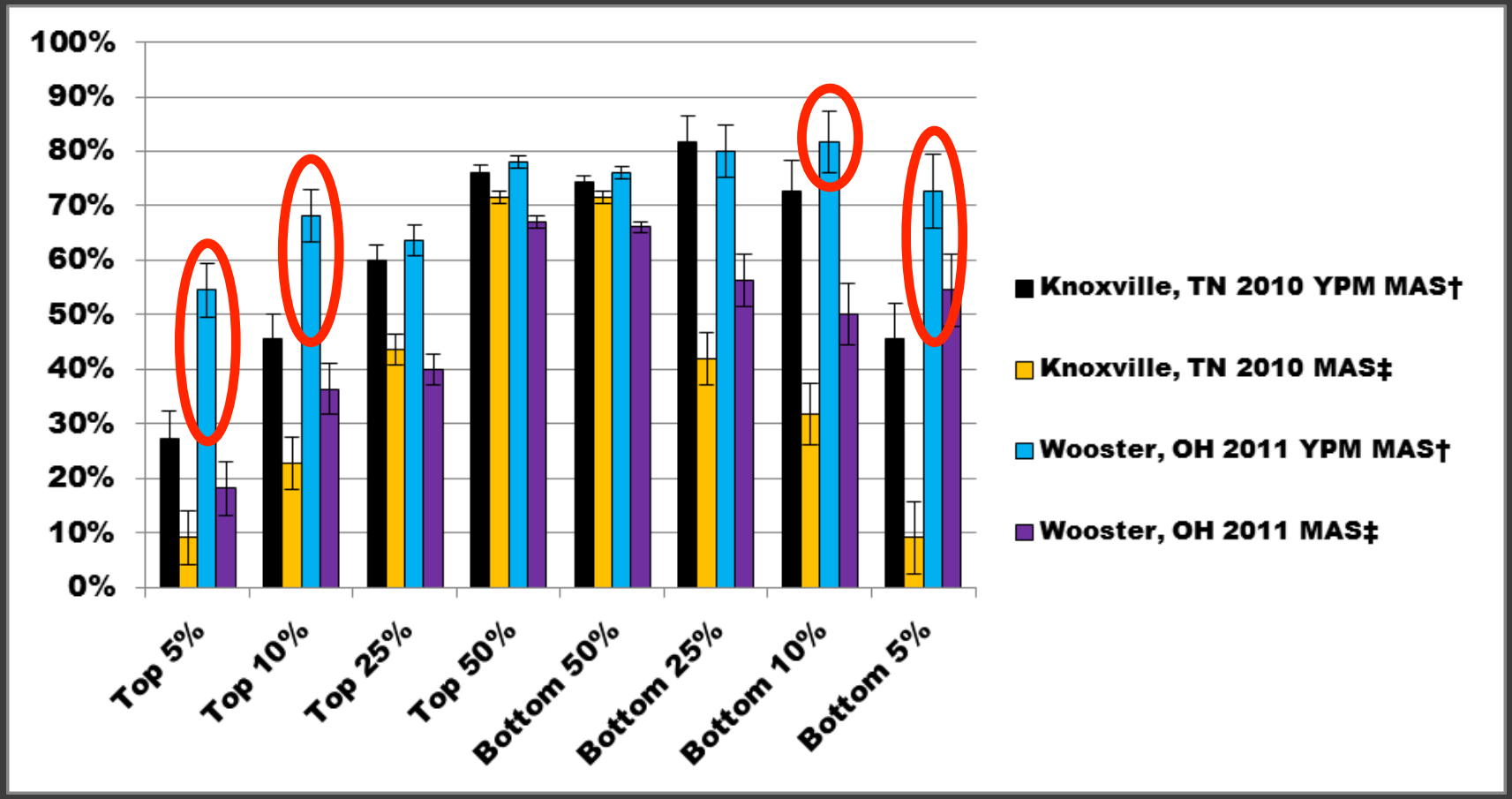
GROUP A: TOP MAS RILS VS TOP YIELDING RILS COMPARED TO USING THE YPM



GROUP A: TOP MAS RILS VS TOP YIELDING RILS COMPARED TO USING THE YPM



GROUP A: TOP MAS RILS VS TOP YIELDING RILS COMPARED TO USING THE YPM



YPM USING ADD. AND ADD. X ADD. WOOSTER, OH 2011 DATA

- R/qtl

- 9 out of 11 RILs in the top yielding 5% were selected using MAS

- 15 out of 22 RILs in the top yielding 10% were selected using MAS

YPM WOOSTER, OH 2011		YIELD (kg ha^{-1}) WOOSTER, OH 2011	
LINE	RANK	LINE	YIELD
689	01	**bb814**	5227.4
481	02	**bb292**	5166.9
951	03	**bb689**	5160.2
463	04	559	4998.9
144	05	**bb978**	4992.2
774	06	**bb896**	4918.3
814	07	**bb481**	4904.9
978	08	**bb463**	4857.8
292	09	**bb144**	4763.8
337	10	**bb833**	4710.0
211	11	146	4669.7
751	12	**b751**	4642.8
896	13	**b211**	4636.1
487	14	754	4575.6
146	15	148	4562.2
854	16	**b489**	4562.2
489	17	**b951**	4562.2
675	18	767	4521.9
86	19	**b675**	4521.9
833	20	**b774**	4508.4
72	21	253	4508.4
454	22	604	4501.7

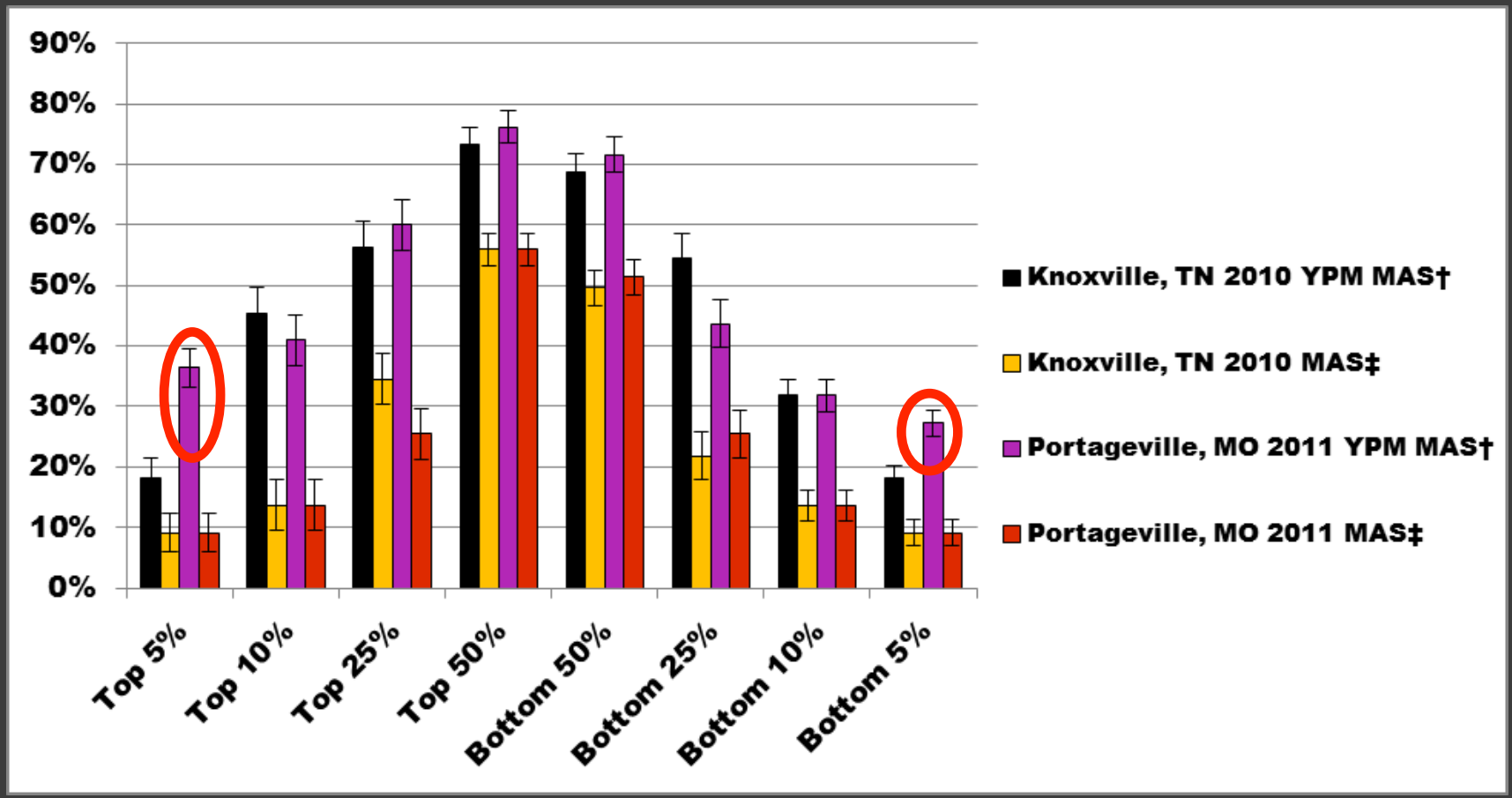
GROUP B: AGRONOMIC TRAITS

- Belleville, IL had an average yield (3434 kg ha^{-1}) that was significantly higher than the average yield in Knoxville, TN in 2010 (2327 kg ha^{-1}) and 2011 (1835 kg ha^{-1}).
- The yield in Knoxville, TN in 2010 averaged 67% and in 2011 averaged 53% of the yield in Belleville, IL in 2011

GROUP C: AGRONOMIC TRAITS

- Portageville, MO had an average yield (3808 kg ha^{-1}) that was significantly higher than the average yield in Knoxville, TN in 2010 (2188 kg ha^{-1}) and 2011 (1914 kg ha^{-1}).
- The yield in Knoxville, TN in 2010 averaged 57% and in 2011 averaged 50% of the yield in Portageville, MO in 2011

GROUP C: TOP MAS RILS VS TOP YIELDING RILS COMPARED TO USING THE YPM



YPM USING ADD. AND ADD. X ADD. KNOXVILLE, TN 2010 DATA

- R/qtl

- 8 out of 11 RILs that were in the top yielding 5% were selected using MAS

- 14 out of 22 RILs that were in the top yielding 10% were selected using MAS

YPM		YIELD (kg ha^{-1})	
KNOXVILLE, TN 2010		KNOXVILLE, TN 2011	
LINE	RANK	LINE	YIELD
671	01	^{aa}199	38.8
932	02	^{aa}938	38.5
265	03	^{aa}378	38.1
378	04	^{aa}448	37.9
469	05	^{aa}450	37.8
760	06	849	37.8
426	07	^{aa}426	37.7
198	08	^{aa}63	37.5
523	09	263	37.1
448	10	183	36.8
382	11	^{aa}78	36.6
620	12	460	36.6
938	13	764	36.5
466	14	^{a}867	36.4
377	15	^{a}932	36.2
553	16	^{a}523	36.2
867	17	^{a}198	36.1
63	18	612	36.1
898	19	359	36.0
450	20	^{a}620	35.9
1006	21	430	35.8
199	22	^{a}382	35.7

YPM USING ADD. AND ADD. X ADD. KNOXVILLE, TN 2010 DATA

- R/qtl

- 2 out of 11 RILs that were in the top yielding 5% were selected using MAS

- 6 out of 22 RILs that were in the top yielding 10% were selected using MAS

YPM		YIELD (kg ha^{-1})	
KNOXVILLE, TN 2010		PORTAGEVILLE, MO 2011	
LINE	RANK	LINE	YIELD
671	01	**bb213**	5301.3
932	02	352	4911.6
265	03	263	4763.8
378	04	607	4710.0
469	05	**bb450**	4696.6
760	06	680	4649.5
426	07	36	4602.5
198	08	966	4602.5
523	09	908	4595.8
448	10	505	4589.1
382	11	141	4582.4
620	12	**b760**	4555.5
938	13	165	4508.4
466	14	320	4481.6
377	15	**b1006**	4474.9
553	16	**b867**	4468.1
867	17	311	4461.4
63	18	572	4461.4
898	19	596	4441.3
450	20	**b378**	4421.1
1006	21	963	4407.7
199	22	270	4387.5

YPM USING ADD. AND ADD. X ADD. KNOXVILLE, TN 2010, 2011 AND PORTAGEVILLE, MO 2011 DATA

- R/qtl

- 6 out of 11 RILs that were in the top yielding 5% were selected using MAS

- 9 out of 22 RILs that were in the top yielding 10% were selected using MAS

YPM		YIELD (kg ha^{-1})	
KNOXVILLE, TN 2010-11 PORTAGEVILLE, MO 2011		KNOXVILLE, TN 2011	
LINE	RANK	LINE	YIELD
263	01	bb199	2608.7
867	02	bb938	2583.5
213	03	bb378	2561.6
932	04	448	2548.2
612	05	bb450	2539.8
760	06	849	2536.4
450	07	426	2529.7
505	08	63	2521.3
938	09	bb263	2491.1
165	10	183	2470.9
633	11	bb78	2460.8
378	12	460	2460.8
121	13	764	2450.8
78	14	b867	2447.4
786	15	b932	2430.6
553	16	523	2430.6
956	17	198	2425.6
607	18	b612	2423.9
803	19	359	2418.8
898	20	620	2410.4
199	21	430	2407.1
680	22	382	2395.3

YPM USING ADD. AND ADD. X ADD. KNOXVILLE, TN 2010, 2011 AND PORTAGEVILLE, MO 2011 DATA

- R/qtl

- 7 out of 11 RILs that were in the top yielding 5% were selected using MAS

- 11 out of 22 RILs that were in the top yielding 5% were selected using MAS

YPM	
KNOXVILLE, TN 2010-11 PORTAGEVILLE, MO 2011	
LINE	RANK
263	01
867	02
213	03
932	04
612	05
760	06
450	07
505	08
938	09
165	10
633	11
378	12
121	13
78	14
786	15
553	16
956	17
607	18
803	19
898	20
199	21
680	22

YIELD (kg ha^{-1})	
PORTAGEVILLE, MO 2011	
LINE	YIELD
cc**213**	5301.3
352	4911.6
cc**263**	4763.8
cc**607**	4710.0
cc**450**	4696.6
cc**680**	4649.5
36	4602.5
966	4602.5
cc**908**	4595.8
cc**505**	4589.1
141	4582.4
c**760**	4555.5
c**165**	4508.4
320	4481.6
1006	4474.9
c**867**	4468.1
311	4461.4
572	4461.4
596	4441.3
c**378**	4421.1
963	4407.7
270	4387.5

GROUP D: AGRONOMIC TRAITS

- Plymouth, NC had an average yield (2191 kg ha^{-1}) that was not significantly higher than the average yield in Knoxville, TN in 2010 (2354 kg ha^{-1}) and 2011 (1720 kg ha^{-1}).
- Group D was the only group in which each environment had significantly similar yields.

IDENTIFIED QTL

- Based on CIM 23 yield QTL were identified
- 21 additional QTL were detected using single factor ANOVA
- QTLs explained 4.5% to 11.9% of the phenotypic variation for yield
- QTLs were identified on all 20 chromosomes
- Five of the 44 QTLs have not been previously reported
- QTL analysis was conducted separately for each group, in each individual environment and combined over environments, with each program

CONCLUSION

- Some top yielding lines might be missed by MAS unless the prediction equation uses data from the targeted environment
- MAS from one year can successfully identify some of the top yielding lines in subsequent years and distant environments
- This leads to credibility for future MAS studies in soybean
- Hopefully, this study along with previous studies will provide further insight into what QTL and tools are available for soybean yield improvement by MAS

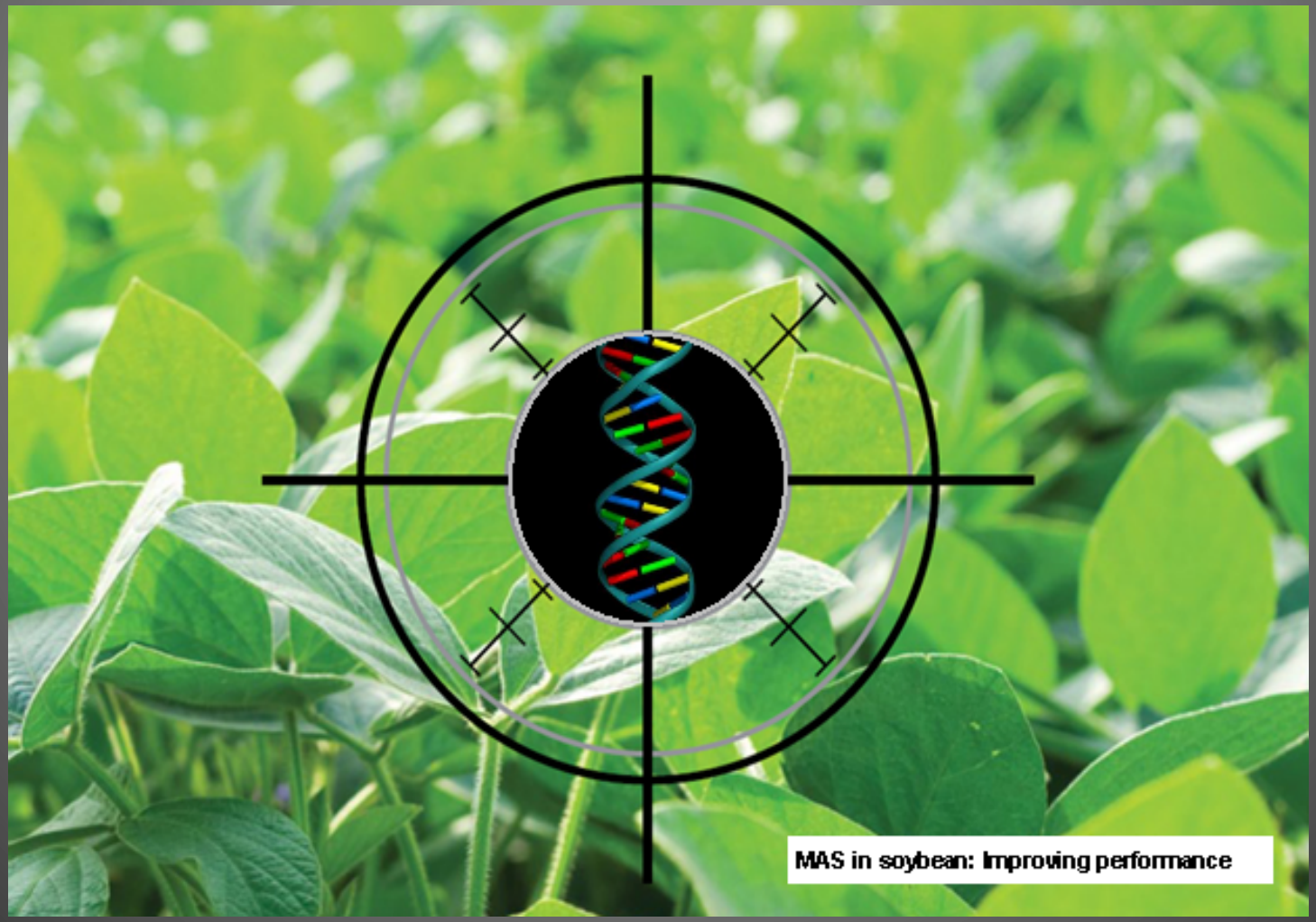
ACKNOWLEDGEMENTS

- My committee members: Dr. Pantalone, Dr. Allen, Dr. Kopsell and Dr. Saxton
- The research staff at ETREC
- The Bean Team at UT: Jeneen Abrams, Jeffrey Boehm, Deborah Ellis, Beth Meyers, Chris Smallwood and Nicole Tacey
- Collaborators: Wooster, OH, Belleville, IL, Portageville, MO and Plymouth, NC
- Support provided by the Tennessee Soybean Promotion Board and the United Soybean Board
- David Hyten and Perry Cregan – Soybean Genomics and Improvement Laboratory Beltsville, MD

USB FUNDED

TENNESSEE SOYBEAN PROMOTION BOARD

Plant Sciences UT



ANY QUESTIONS ??

Image adapted from: http://www.croplife.com/article/10901/insects-weave-a-tangled-web and http://www.mitochondrialdnatesting.com/nuclear-dna.html