

GBS identification of genomic loci associated with polyembryony in *ig1* mutants



Sam Hokin, Matt Evans, Antony Chettoor & Vicki Tang
shokin@carnegiescience.edu



CARNEGIE
SCIENCE

Plant Biology

SUMMARY

The *LBD6* or *ASYMMETRIC LEAVES2* gene in maize is an important regulator in the transition from proliferation to differentiation in embryo sac development (Evans, 2007). The mutant, also known as *indeterminate gametophyte1* (*ig1*), produces embryo sacs which lack the ability to limit proliferation, leading to structural defects. These defects include embryo sacs with extra egg cells, extra central cells, and extra polar nuclei, which give rise to abnormal seed phenotypes post-fertilization. Extra egg cells in an embryo sac can be fertilized producing kernels with multiple viable embryos.

The *ig1-O* mutation results in different frequencies of multiple embryos, or 'twinning,' in different maize inbred backgrounds. Confocal microscopy confirms that Mo17 embryo sacs carrying the *ig1-O* allele have a significantly higher frequency of multiple egg cells than *ig1-O* embryo sacs in a B73 inbred background. The frequency of twinning is intermediate in the *ig1-O/+ Mo17/B73* hybrid compared to *ig1-O* inbred lines. This suggests that a genetic modifier(s) may be responsible for the difference in phenotypic frequency between these inbred lines.

Genotyping By Sequencing (GBS) data analysis and initial SSR marker fine mapping indicate broad enhancers of twinning on chromosomes 3 and 9, as well as narrow regions elsewhere.

Carrying the Mo17 allele appears to be sufficient to increase the frequency of twinning caused by the *ig1-O* mutation.

PHENOTYPES

The *ig1* mutant produces many seeds with twin embryos (as well as triplets and other abnormalities), especially on a Mo17 background.



Twin seedlings grow from an *ig1* mutant twin-embryo seed.

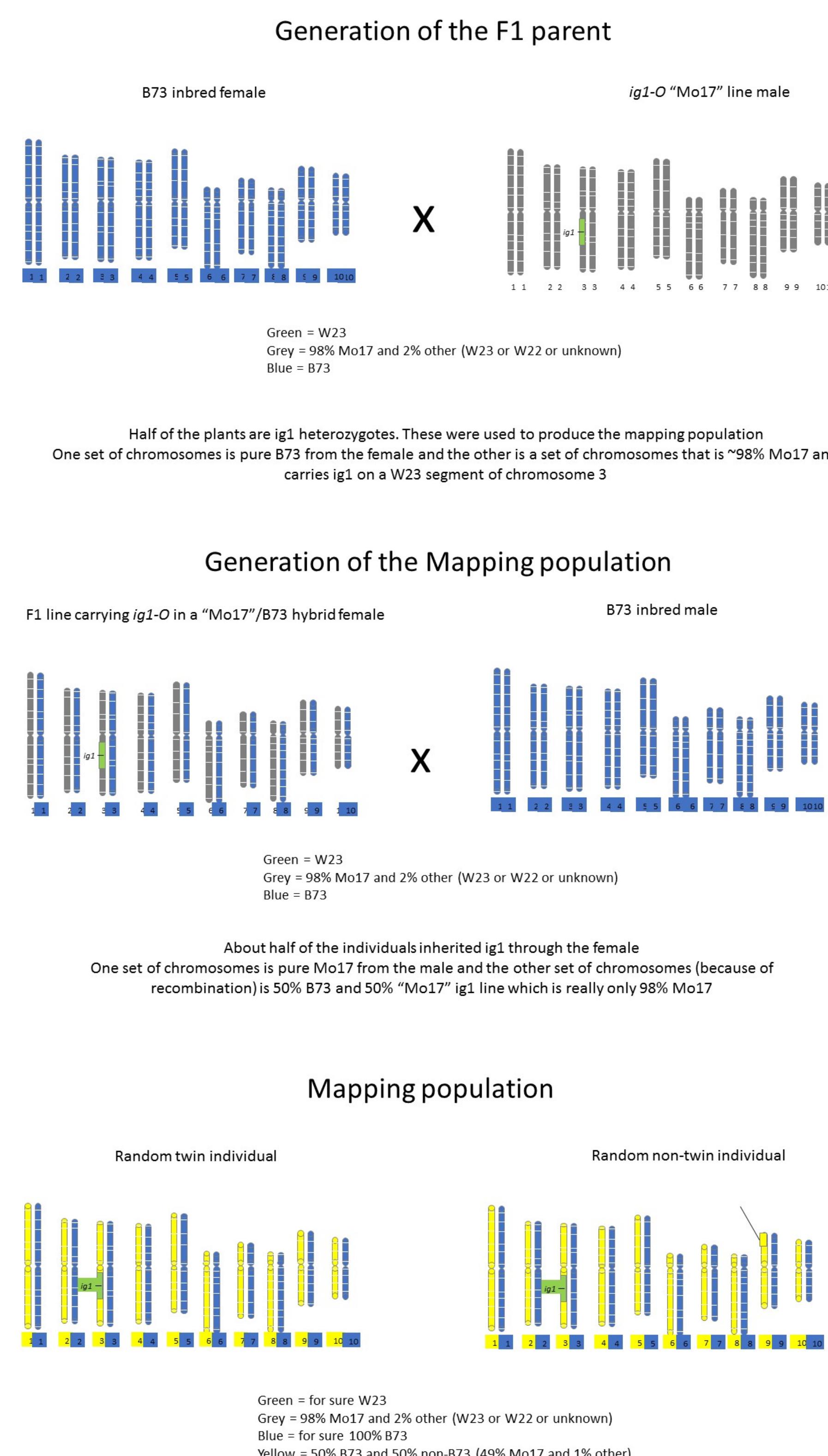
Rate of occurrence

The Mo17/B73 hybrid *ig1* mutants exhibit a twinning phenotype at a frequency intermediate between the frequency of twinning in individual inbred lines.

<i>ig1</i> inbred lines	# seeds	twins
B73	1372	0.5%
Mo17/B73 hybrid	5217	4.2%
Mo17	840	8.5%

This suggests a model in which B73 and Mo17 differ for a single modifier with a major effect on twinning (the Mo17 allele of this modifier promoting twinning in *ig1* mutants).

GBS MAPPING POPULATION



GBS RESULTS

Mapping and SNP statistics

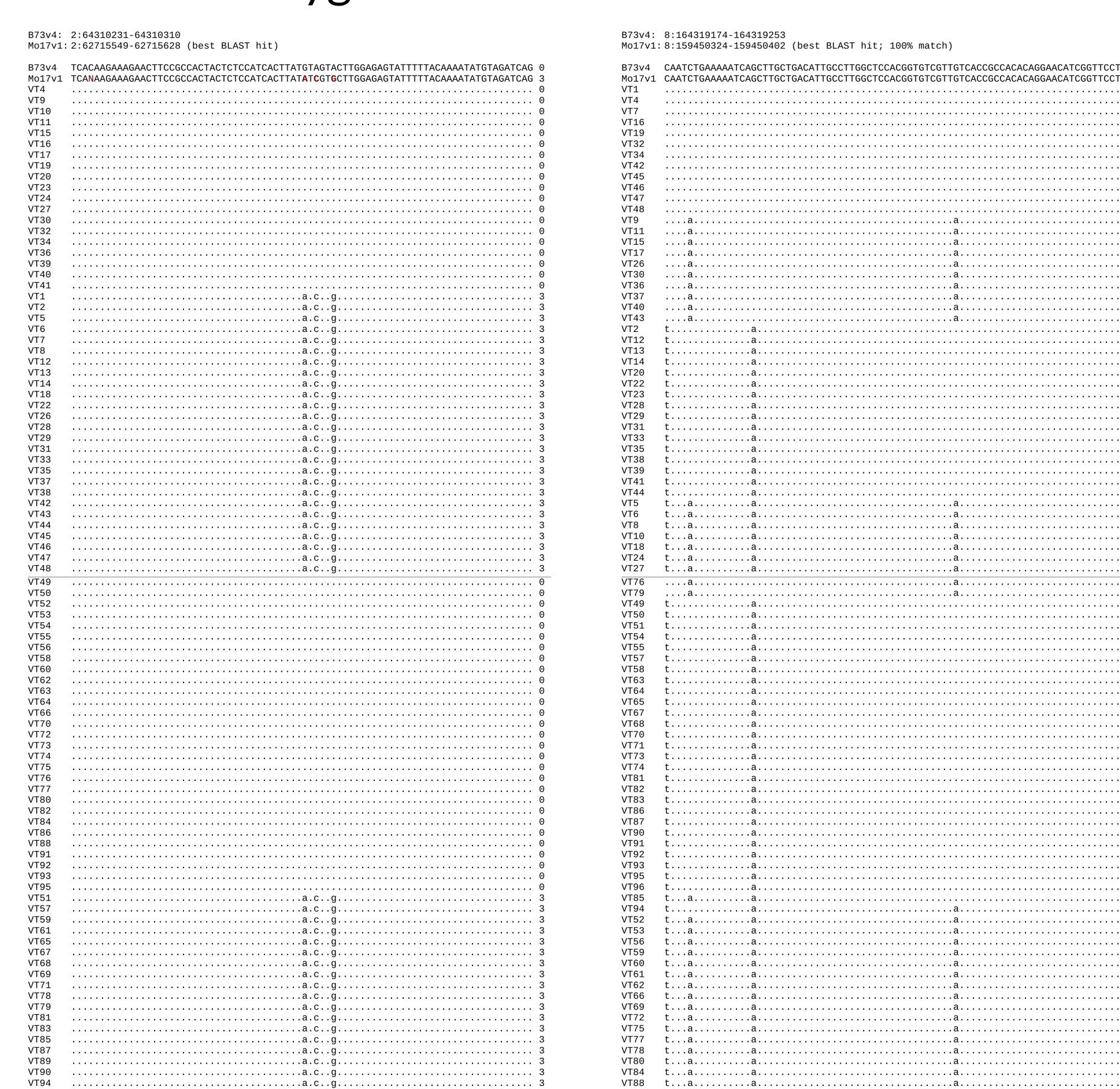
45 sequenced samples had non-twin phenotype, 48 had twin phenotype.

93 samples	mean	s.d.	max	min
mapped reads	3,633,114	676,410	5,204,340	1,956,272
# SNPs	85,118	16,892	132,620	54,132
avg. depth	8.79	1.06	11.02	6.04
# > 1x	2.86%	0.24	3.61%	2.27%
# > 4x	1.62%	0.15	1.96%	1.10%
Het rate	0.026%	0.005	0.039%	0.017%

The GBS sequencing was done by Novogene. The reads were mapped against the B73v4 genome. Positions with at least four reads qualified for SNP calls.

Example non-segregated and segregated loci

Example regions in which roughly half of the samples have heterozygous SNP calls.



There is no segregation of twins vs. non-twins in this region. The three SNPs match the corresponding Mo17v1 best-BLAST-hit region.

Twin samples segregate significantly from non-twin samples in this region, with more twins carrying three SNPs, which don't match Mo17v1.

Fisher's exact test

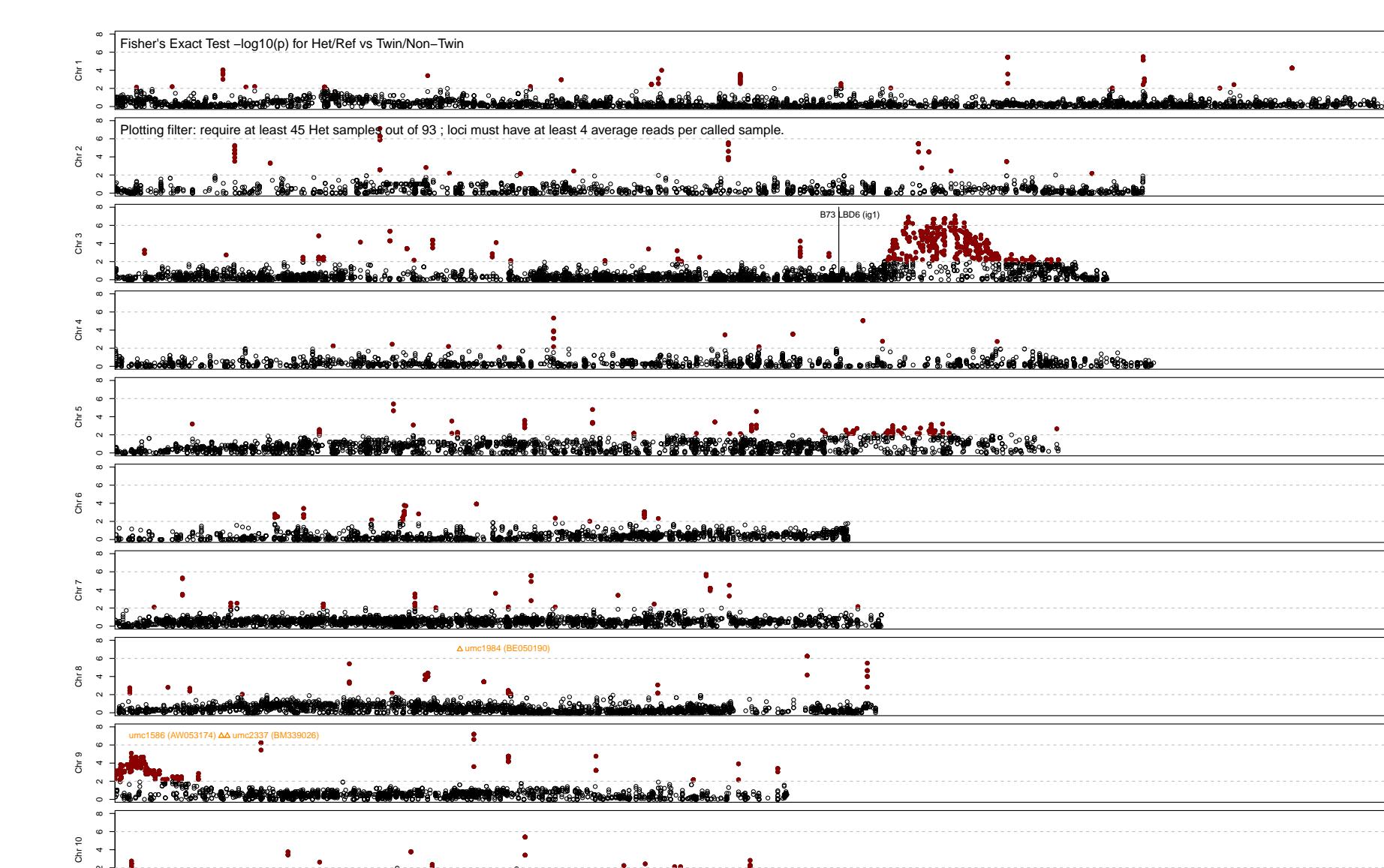
Each SNP location on the genome provides a classic 2 X 2 contingency table, twin/non-twin vs. het/no call, appropriate for Fisher's Exact Test. Fisher's Exact Test was therefore applied to every location, generating a *p*-value at every position. For example:

2:28422677	reads	no call	het call
non-twin	417	27	18
twin	773	7	41
total	1190	34	59

Fisher's exact test on calls at this locus gives $p = 5.5 \times 10^{-6}$. The number of called samples, 59, and the average reads per called sample, $1190/59=20.2$, pass the filter used for the whole-genome Manhattan plot below.

Regions of segregation

Fisher's exact test was applied to all calls in the GBS study for the null hypothesis that the number of Het calls does not differ between the twin and non-twin groups. Loci were plotted if they had at least 4 reads per called sample and had at least 45 Het calls out of 93 samples.



The position of the *LBD6* (*ig1*) gene on Chr3 is indicated, as well as the locations of SSR markers that exhibit segregation between twins and non-twins. Red indicates significance at $p < 0.01$. Dashed lines are drawn at $p = 10^{-2}$ and 10^{-6} .

SSR MARKERS

For comparison to the GBS results, SSR markers on Chr8 and Chr9 were applied to 73 DNA samples of *ig1* twins from different plants and progenitor wild-type inbred lines to identify the parental alleles for each locus. Of the primers ordered and tested, those for **umc1984** on Chr8, and **umc1586** and **umc2337** on Chr9 were polymorphic between B73 and Mo17 and showed the clearest bands.

On Chr8, **umc1984** showed a significantly higher number of heterozygous samples than expected ($p < 0.05$, chi-squared test), while on Chr9, both **umc1586** and **umc2337** showed a significantly higher number of heterozygous samples than expected ($p < 0.01$, and $p < 0.05$, respectively).

More fine mapping will need to be done to narrow down the modifying regions and to identify candidate genes for these enhancers and determine the differences between the B73 and Mo17 alleles.

CONCLUSIONS

References

M. M. Evans. The indeterminate gametophyte1 gene of maize encodes a LOB domain protein required for embryo Sac and leaf development. *Plant Cell*, 19(1):46–62, Jan 2007.

This research was funded by National Science Foundation grant #XXXXXXX.