

Genome wide analysis of *Ga1-s* modifiers in maize

Preston Hurst^{1,2}, Zhikai Liang^{1,2}, Christine Smith^{1,2}, Melinda Yerka^{1,@}, Brandi Sigmon^{1,2}, Oscar Rodriguez^{1,+}, and James C Schnable^{1,2,*}

¹Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

²Center for Plant Science Innovation, University of Nebraska-Lincoln, Lincoln, Nebraska, USA

*Corresponding Author: schnable@unl.edu

+Present affiliation: Conagra Brands, Brookston, Indiana, USA

@Present affiliation: Department of Agriculture, Veterinary Rangeland Sciences, University of Nevada - Reno, Reno, Nevada, USA

ABSTRACT

A one way reproductive barrier exists between most popcorn varieties and dent corn varieties grown in the United States. This barrier is predominantly controlled by the *ga1* locus. Using data from a diverse population of popcorn accessions pollinated by a dent corn tester, we found that the non-reciprocal pollination barrier conferred by *ga1* is more complex than previously described. Individual accessions ranged from 0% to 100% compatible with dent corn pollen. Using conventional genotyping-by-sequencing data from 371 popcorn accessions carrying *Ga1-s*, seven significant modifiers of dent pollen compatibility were identified on five chromosomes. One locus may either be a nonfunctional *ga1* allele present within popcorn, or second necessary gene for the reproductive barrier in genetic linkage with *ga1*, while the other modifiers are clearly genetically unlinked. The existence of *ga1* modifiers segregating in a popcorn genetic background may indicate selective pressure to allow gene flow between populations, which should be incorporated into future models of the impact of genetic incompatibility loci on gene flow in natural and agricultural plant populations.

Introduction

Sexual incompatibility is a common phenomenon in the plant kingdom. Crossing barriers have been observed in several species, including tobacco^{1,2}, tomato³⁻⁵, and rice^{6,7} among others. Lateral barriers, such that members of one population are unable to produce offspring with members of another population, can reinforce speciation⁸ and genes that create these barriers can play a critical role in cases of sympatric speciation^{9,10}. Due to the preferential and selective nature of the barrier, a male and female action are both required to enable recognition and discrimination of self and non-self pollen¹¹. This type of incompatibility is seen in *Solanum*^{12,13}, as well as in some *Brassica* species¹⁴. It is possible that the action of both the male and female may be conferred through the actions of a protein encoded by a single gene, or the male and female functions of a crossing barrier may be provided by separate genes in tight genetic linkage, as has been shown for *tcb1*¹⁵.

In maize, the term gametophyte factor is employed to describe a locus where pollen grains carrying a particular allele exhibit increased fertilization success relative to pollen grains of the competing allele¹⁶. At least 10 such loci have been described for maize¹⁶. Of these *Ga1* was the first to be identified, with the first report of segregation distortion linked to the as yet unnamed *su1* locus on maize chromosome 4 published by Carl Correns in 1902¹⁷, the determination that *Ga1* was an independent genetic locus and not a second phenotype of the *su1* by R. A. Emerson in 1925¹⁸, and given the name *Ga1* by Mangelsdorf and Jones one year later¹⁹. *Ga1* creates a substantial unidirectional crossing barrier between plants carrying different alleles at the locus. Seed set when *Ga1-s* silks are pollinated with exclusively *ga1* pollen is consistently reported to be <5%^{20,21}, and in a competitive environment containing both *Ga1-s* and *ga1* pollen, estimates of the proportion of kernels successfully pollinated by *ga1* pollen grains range from 0.6-4%^{19,22,23}. The strength of the *Ga1* crossing barrier is notable relative to other crossing barriers characterized in maize. The reference allele of *Ga2* allows 10-25% fertilization of wild-type pollen in competitive pollination assays, while a more recently isolated allele of *Ga2* from teosinte provides somewhat stronger incompatibility with wild-type pollen²⁴. *Tcb1* provides near complete exclusion of *tcb1* pollen in a teosinte background, but exhibit attenuated effects when introgressed into maize backgrounds, suggesting the function of *Tcb1* depends on one or more modifier loci²⁵. The early identification of *Ga1*, combined with the relatively extreme incompatibility phenotype of *Ga1* relative to other gametophyte factor loci described in maize¹⁶ have made the locus a particular focus of genetic investigation over the last century.

Three functionally distinct alleles have been described at the *Ga1* locus. The predominant allele in US maize germplasm is *ga1*. *Ga1-s* is found in many popcorn lines and in some populations of teosinte, the wild progenitor of maize, creates the

unidirectional crossing barrier, as *Gal-s* silks are receptive to pollen carrying the *Gal-s* allele, but are largely non-receptive to pollen grains carrying the *gal* allele^{17,26}. A third allele *Gal-m*, conveys the male, but not female function of the *gal* locus. Thus, pollen carrying *Gal-m* can successfully pollinate *Gal-s* silks, but *Gal-m* silks can be successfully pollinated by *gal*, *Gal-m*, and *Gal-s* pollen²⁷. Data suggests the male function of *Gal-m* may be incomplete, as *Gal-s* pollen is still more successful than *Gal-m* pollen when pollinating *Gal-s* silks²⁸. More recent studies suggest that *Gal-m* may be the predominant allele in maize tropical landraces²⁹.

Mapping *Gal* proved challenging. Early maps based on phenotypic markers placed the locus on chr4 between *del* and *sul* but at a substantial distance, perhaps 43 centimorgans (cM) from *del* and 33 cM from *sul*²². Zhang and colleagues employed molecular markers to constrain the location of *Gal* to a 1.5 cM interval on chromosome 4²⁰. Employing additional data from the B73 x Hp301 NAM subpopulation and a backcross population carrying *Gal-s*, Bloom and Holland constrained the location of *gal* to a location containing 13 predicted genes in the B73 reference genome²¹. Lauter et al.(2017) provided evidence PME3, a pectin methylesterase, is responsible for the female action of the locus³⁰. Zhang et al.(2018) identified a PME gene as responsible for the male action of *Gal-s*³¹. These studies suggest that both the male and female actions of unilateral dent sterility are related to the deesterification of pectin. Recently another pectin methylesterase *Pertunda* has also been identified as responsible for the female action of the *Tcb1* crossing barrier³², although pollen carrying the *Tcb1-s* and *gal* alleles is compatible with *Gal-s* silks, and vice versa.

Mapping receptivity to *gal* pollen in a popcorn by dent corn recombinant inbred line population (B73 x Hp301) identified only a single locus corresponding to *Gal*²¹. Fine mapping estimated that locus to be slightly larger than 1MB³³. In contrast to the result from the B73 x Hp301 population, apparent modifiers of *gal* function have been reported in the popcorn variety Supergold³⁴, from which many of the inbreds from one the three major heterotic groups in commercial hybrid popcorn production are derived³⁵. Here we employed data from *gal* pollination tests of a diverse panel of maize popcorns, including representatives of the Supergold, Amber Pearl, and South American popcorn heterotic groups to assess the degree of variation present for *Gal* and whether specific modifiers segregating in popcorn could be identified.

Methods

Scoring Dent Sterility

A diverse panel of 311 popcorn lines planted in an isolation plot at the University of Nebraska Experimental Station at Mead, NE during the 2017 mid-western United States growing season were assessed for quantitative compatibility with *gal* pollen. Each genotype was represented by 20 plants planted in a 17.5 foot long row, with a 30 inch alley between each range and each row. The perimeter of the field was planted with two rows of a purple *gal* line in a dent corn background. Within the field, every fifth row was planted with the same purple dent *gal* line. Lines of the diversity panel were detasseled prior to anthesis and female flowering dates were recorded for each accession to confirm exposure of the silks to freshly shed pollen from the purple dent *gal* line. Each line was scored for incompatibility with *gal* pollen at maturity. Ears with significant numbers of yellow or white kernels, indicating pollen contamination from a source either than the intended *gal* pollen donor were disregarded. (Figure S1)

Genotyping

327 commercial inbred popcorn lines from the ConAgra Foods popcorn breeding program and 44 public popcorn lines from USDA were genotyped using conventional genotyping-by-sequencing (GBS)³⁶. DNA samples were submitted to the Cornell Institute for Genomic Diversity for GBS library construction and sequencing. Samples were sequenced in four distinct barcoded libraries. A total of 827M raw Illumina reads were generated. Reads were aligned to the B73 reference genome (RefGen V3) using bwa aln³⁷. 238,772 SNPs were called using the TASSEL-GBS pipeline v3.0³⁸ with the MAF (minor allele frequency) filtering threshold set to 0.05 and mnF (minimum inbreeding coefficient defined as 1-Ho/He) set to 0.8. Missing SNP calls were imputed using Beagle version 4.1³⁹. After imputation, SNPs with a minor allele frequency below 0.01 were removed from the dataset. The final dataset consisted of 238,772 SNPs scored in 311 maize popcorn lines for which both genotypic and phenotypic data was available.

Genome Wide Association Analysis and Interpretation

Association analyses were performed in parallel using the mixed linear model (MLM) implemented in GEMMA with a significance threshold of 0.01⁴⁰ and the FarmCPU algorithm (a multi-locus mixed model (MLMM)) implemented in MVP (<https://github.com/XiaoleiLiuBio/rMVP>) at a Bonferroni corrected significance threshold of 0.01⁴¹.

The first 5 principal components of population structure, as determined by TASSEL⁴², were including in both analysis to control for population structure effects⁴³. Kinship was calculated using MVP and the same kinship matrix was employed in both analyses. (Figure S2) Structural and functional gene annotations for version 3 of the B73 reference genome were

obtained from Phytozome v12.1. All genes within 10kb of the SNP were considered potential candidate genes. All linkage disequilibrium calculations were performed in TASSEL, version 5⁴².

Results

Significant variation in the degree of *gal* fertilized seed set was observed among the accessions of this popcorn population (Figure S1). Half of accessions showed >5% seed set with purple kernels indicating successful pollination by *gal* pollen. Five lines showed near complete compatibility with *gal* pollen. This observation is notably different from reports which assessed *Gal-s* introgressed into dent corn backgrounds where *Gal-s* consistently produces a high degree (>95%) of sterility when fertilized with *gal* pollen^{20,21}.

A genome-wide association study was conducted to test whether specific loci outside the *Gal-S* locus were associated with variation in the effectiveness of the *Gal-s* allele's female function. Using genotype calls for 238,772 segregating SNPs genotyped using conventional genotyping by sequencing, seven statistically significant trait associated SNPs were identified on five chromosomes using the FarmCPU algorithm: 2, 3, 4, 5 and 10 (Figures 1, S2). Three of these trait associated SNPs (Chr5 bp:99960879, Chr4 bp:7020954, and Chr2 bp:72824565) were also supported by MLM based GWAS (Figures 1, S3). The trait associated SNP identified on chromosome 4 is located at 7,020,954 bp on the B73_RefGenV3 reference genome, corresponding to approximately 7.8MB in B73_RefGenV4. This position is in close proximity to two genes identified at the *Gal* locus^{30,31}.

The precise placement of the *Gal* locus in the B73 reference genome remains somewhat ambiguous. Multiple groups who have attempted to map the *Gal* region have only been able to narrow it down to a large interval^{20,21,31,33}, likely as a result of structural differences between the popcorn derived *Gal-s* and dent corn derived *gal* haplotypes at this locus. Consistent with this explanation, three large gaps containing no SNP markers of size 429kb, 826kb, and 454kb were observed within the estimated position of the *Gal-s* locus, in comparison to 1 SNP every 12kb in the whole dataset (Figure 3). Here we employ the interval reported by Zhang et al which placed the causal gene between 8.52MB and 10.23MB on chromosome 4 in B73 RefGenV4. After translation, the equivalent is approximately 7.66Mb to 9.35Mb on chromosome 4 in B73 RefGenV3³¹. The range of the gaps found in the current study span from 7.623Mb to 9.358Mb.

The distribution of phenotypic scores for individuals carrying the minor allele at the chr4 locus are also consistent with the minor allele of this trait associated SNP marking a *gal* allele, either from introgression of dent germplasm during popcorn breeding and improvement, or already present at low frequencies in popcorn germplasm. However, based on patterns of linkage disequilibrium in the region from 6.7Mb to 10.5Mb on chromosome 4 (Figure 3) it appears that the trait associated SNP Chr4 bp:7020954 is in not in linkage with the region from 8.53Mb to 10.23Mb in the B73 reference genome determined to correspond to the popcorn region carrying the pectin methylesterases responsible for the *gal* reproductive barrier³¹. In addition, Chr4 bp:7020954 exhibits significant LD with adjacent markers in the genome, indicating the marker is likely not misplaced, whether as a result of a misassembly in the B73 reference genome or structural variation between popcorn and dent corn (Figure 3).

The minor alleles of the trait associated SNPs identified on chromosomes 2 and 5 were also associated with an increase in successful fertilization by *gal* pollen (i.e. decreased effectiveness of the reproductive barrier conveyed by *Gal-s*). The minor alleles of trait associated SNPs on chromosomes 3 and 10 are associated with a decrease in successful fertilization by *gal* pollen (i.e. increased function of *Gal-s* as a reproductive barrier (Figure 2). Three trait associated SNPs (Chr5 bp:192080965, Chr4 bp:7020954, Chr10 bp:117590335) are within larger intervals identified by composite interval mapping⁴⁴. Composite interval mapping also identified a signal on chromosome 2 that, while not sufficiently strong to be statistically significant on its own, overlaps with and lends additional confidence to the two trait associated SNPs (Chr2 bp:70339953 and Chr2 bp:72824565) identified on that chromosome⁴⁴. With two exceptions, Chr2 bp:72824565 and Chr5 bp:192080965, each trait associated SNP was within 10KB of one or more annotated maize gene models (Table ??). However, caution should be used in interpreting these gene lists as significant presence absence variation is present in maize^{45,46} and genes involved in reproductive barriers in popcorn may be absent from the B73 maize reference genome as was, indeed, the case for both *Gal-s* and *Tcb1-s*³⁰⁻³².

Discussion

The results presented above suggest the genetic basis of non-reciprocal cross sterility conferred at the *Gal* locus is more complex than previously thought. Anecdotal, popcorn breeders have long noticed a complex mode of inheritance³⁴, and at least one previous study identified mapping intervals which appeared to act as modifiers of *Gal-s*⁴⁴. The existence of these modifiers pose challenges to efforts to employ the *Gal-s* allele in commercial breeding programs.

One trait associated SNP, Chr4 bp:7020954, is in close proximity to the *gal* region, and may indeed represent a *gal* allele, either from introgression of dent germplasm during popcorn breeding and improvement, or already present at low frequencies in popcorn germplasm. However linkage disequilibrium analysis was not consistent with this interpretation (Figure 3). One alternative explanation is that Chr4 bp:7020954 marks a second gene in tight linkage with the PME genes already identified as essential to *Gal-s* function which modifies the function of this crossing barrier. GRMZM2G157241, located 8kb from

Chr4 bp:7020954, is homologous to characterized calcium binding proteins. It has been proposed that transfer of calcium ions formed from the action of pectin methylesterase enzymes plays a role in the Gal locus³¹. As pectin is deesterified by PME, free Ca²⁺ becomes crosslinked with pectin and the cell wall of the pollen tube becomes stiff⁴⁷. However further research will be needed to determine whether Chr4 bp:7020954 indeed marks a tightly linked modifier of *Gal-s* or, despite the present linkage disequilibrium evidence, marks a knockout of at least the female function of *Gal-s* equivalent to either *Gal-m* or *gal*.

The existence of modifiers for *Gal-s* function also has implications for specialty maize breeding and production efforts. One recent proposal has been to introgress the *Gal-s* allele into maize refuge lines and grow in a mixed bag refuge system with Bt maize⁴⁸. As the Bt *gal* pollen would be unable to pollinate the ears of the non-Bt *Gal-s* refuge plants, in theory this approach would increase the amount of non-Bt grain available to insect pests, decreasing the selective pressure for Bt resistance in pest species. *Gal-s* modifiers present a challenge and an opportunity, as *gal* pollen is likely to dramatically outnumber *Gal-s* pollen on the silks of refuge plants within Bt fields, and this any increase or decrease in *Gal-s* efficiency is likely to produce substantial shifts in the effectiveness of this refuge strategy. Some organic producers also utilize the *Gal-s* allele as an additional approach to avoiding any detectable GM traits from pollen drift which might threaten organic certification⁴⁹. Here again, knowledge of loci that increase or decrease the effectiveness of the crossing barrier could have significant economic impact.

If allelic differences impact the degree to which populations inter-mate, it will surely have an impact on allele frequencies in other areas of the genome and impact the amount of gene flow from one population to another. If speciation genes are subject to modification by other loci, one would expect that all loci involved would have an impact on the history and future of the organism. It is possible that selective sweeps exist around these loci, that may explain some of the phenotypic variation between wild and domesticated genotypes, as well as different classes of maize such as dent, flint and popcorn.

Author contributions: BS, JCS, MY and OR conceived the study; CS, OR, and ZL generated the data, PH and ZL analyzed the results, JCS and PH wrote the paper

Acknowledgements: Supported by research funding from ConAgra Brands to JCS and OR. The authors thank Aaron Lorenz and Amritpal Singh for their assistance in popcorn genotyping.

References

1. de Nettancourt, D. Incompatibility in angiosperms. *Sex. plant reproduction* **10**, 185–199 (1997).
2. Lee, C. B., Page, L. E., McClure, B. A. & Holtsford, T. P. Post-pollination hybridization barriers in nicotiana section alatae. *Sex. Plant Reproduction* **21**, 183 (2008).
3. Liedl, B., McCormick, S. & Mutschler, M. Unilateral incongruity in crosses involving lycopersicon pennellii and l. esculentum is distinct from self-incompatibility in expression, timing and location. *Sex. plant reproduction* **9**, 299–308 (1996).
4. Moyle, L. C. & Graham, E. B. Genetics of hybrid incompatibility between lycopersicon esculentum and l. hirsutum. *Genetics* **169**, 355–373 (2005).
5. Moyle, L. C. & Nakazato, T. Comparative genetics of hybrid incompatibility: sterility in two solanum species crosses. *Genetics* (2008).
6. Matsubara, K., Sano, Y. *et al.* A gene block causing cross-incompatibility hidden in wild and cultivated rice. *Genetics* **165**, 343–352 (2003).
7. Chen, J. *et al.* A triallelic system of *s5* is a major regulator of the reproductive barrier and compatibility of indica-japonica hybrids in rice. *Proc. Natl. Acad. Sci.* **105**, 11436–11441 (2008).
8. De Nettancourt, D. Self-and cross-incompatibility systems. In *Plant Breeding*, 203–212 (Springer, 1993).
9. Nosil, P. & Schluter, D. The genes underlying the process of speciation. *Trends Ecol. & Evol.* **26**, 160–167 (2011).
10. Wu, C.-I. & Ting, C.-T. Genes and speciation. *Nat. Rev. Genet.* **5**, 114 (2004).
11. De Nettancourt, D. *Incompatibility and incongruity in wild and cultivated plants*, vol. 3 (Springer Science & Business Media, 2001).
12. Covey, P. A. *et al.* Multiple features that distinguish unilateral incongruity and self-incompatibility in the tomato clade. *The Plant J.* **64**, 367–378 (2010).
13. Tovar-Méndez, A. *et al.* Restoring pistil-side self-incompatibility factors recapitulates an interspecific reproductive barrier between tomato species. *The Plant J.* **77**, 727–736 (2014).
14. Takada, Y. *et al.* Genetic analysis of novel intra-species unilateral incompatibility in brassica rapa (syn. campestris) l. *Sex. plant reproduction* **17**, 211–217 (2005).

15. Lu, Y., Kermicle, J. L. & Evans, M. M. Genetic and cellular analysis of cross-incompatibility in *zea mays*. *Plant reproduction* **27**, 19–29 (2014).
16. Nelson, O. E. The gametophyte factors of maize. In *The maize handbook*, 496–503 (Springer, 1994).
17. Correns, C. E. Scheinbare ausnahmen von der mendel'schen spaltungsregel für bastarde apparent exceptions to the mendelian cleavage rule for hybrids. *Berichte der deutschen botanischen gesellschaft* **20**, 159–172 (1902).
18. Emerson, R. A possible case of selective fertilization in maize hybrids. *Anat. Rec* **29**, 136 (1925).
19. Mangelsdorf, P. C. & Jones, D. F. The expression of mendelian factors in the gametophyte of maize. *Genetics* **11**, 423 (1926).
20. Zhang, H. *et al.* Genetic analysis and fine mapping of the *ga1-s* gene region conferring cross-incompatibility in maize. *Theor. applied genetics* **124**, 459–465 (2012).
21. Bloom, J. C. & Holland, J. B. Genomic localization of the maize cross-incompatibility gene, gametophyte factor 1 (*ga1*). *Maydica* **56** (2012).
22. Emerson, R. Relation of the differential fertilization genes, *ga ga*, to certain other genes of the *su-tu* linkage group of maize. *Genetics* **19**, 137 (1934).
23. Demerec, M. Cross sterility in maize. *Zeitschrift für Induktive Abstammungs-und Vererbungslehre* **50**, 281–291 (1929).
24. Kermicle, J. L. & Evans, M. M. S. The *zea mays* sexual compatibility gene *ga2*: Naturally occurring alleles, their distribution, and role in reproductive isolation. *J. Hered.* **101**, 737–749, DOI: [10.1093/jhered/esq090](https://doi.org/10.1093/jhered/esq090) (2010). [/oup/backfile/content_public/journal/jhered/101/6/10.1093_jhered_esq090/2/esq090.pdf](https://oup/backfile/content_public/journal/jhered/101/6/10.1093_jhered_esq090/2/esq090.pdf).
25. Evans, M. & Kermicle, J. Teosinte crossing barrier1, a locus governing hybridization of teosinte with maize. *Theor. Appl. Genet.* **103**, 259–265 (2001).
26. Schwartz, D. The analysis of a case of cross-sterility in maize. *Proc. Natl. Acad. Sci. United States Am.* **36**, 719 (1950).
27. JIMÉNEZ T, J. R. & NELSON, O. E. A new fourth chromosome gametophyte locus in maize. *J. Hered.* **56**, 259–263 (1965).
28. Kermicle, J. L. & Evans, M. M. Pollen–pistil barriers to crossing in maize and teosinte result from incongruity rather than active rejection. *Sex. Plant Reproduction* **18**, 187–194 (2005).
29. Kermicle, J., Taba, S. & Evans, M. The gametophyte-1 locus and reproductive isolation among *zea mays* subspecies. *Maydica* (2006).
30. Lauter, A. N. M., Muszynski, M. G., Huffman, R. D. & Scott, M. P. A pectin methylesterase *zmpme3* is expressed in gametophyte factor1-s (*ga1-s*) silks and maps to that locus in maize (*zea mays* l.). *Front. Plant Sci.* **8** (2017).
31. Zhang, Z. *et al.* A pectin methylesterase gene at the maize *ga1* locus confers male function in unilateral cross-incompatibility. *Nat. communications* **9**, 3678 (2018).
32. Lu, Y., Hokin, S. A., Kermicle, J. L. & Evans, M. M. S. A silk-expressed pectin methylesterase confers cross-incompatibility between wild and domesticated strains of *zea mays*. *bioRxiv* DOI: [10.1101/529032](https://doi.org/10.1101/529032) (2019). <https://www.biorxiv.org/content/early/2019/01/23/529032.full.pdf>.
33. Liu, X. *et al.* Fine mapping of the maize cross-incompatibility locus gametophytic factor 1 (*ga1*) using a homogeneous population. *Crop. Sci.* **54**, 873–881 (2014).
34. Ashman, R. Modification of cross-sterility in maize. *J. heredity* **66**, 5–9 (1975).
35. Kantety, R. V., Zeng, X., Bennetzen, J. L. & Zehr, B. E. Assessment of genetic diversity in dent and popcorn (*zea mays* l.) inbred lines using inter-simple sequence repeat (issr) amplification. *Mol. breeding* **1**, 365–373 (1995).
36. Elshire, R. J. *et al.* A robust, simple genotyping-by-sequencing (gbs) approach for high diversity species. *PloS one* **6**, e19379 (2011).
37. Li, H. & Durbin, R. Fast and accurate long-read alignment with burrows–wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
38. Glaubitz, J. C. *et al.* Tassel-gbs: a high capacity genotyping by sequencing analysis pipeline. *PloS one* **9**, e90346 (2014).
39. Browning, B. L. & Browning, S. R. Documentation for beagle 2.1. (2007).
40. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. genetics* **44**, 821 (2012).

41. Liu, X., Huang, M., Fan, B., Buckler, E. S. & Zhang, Z. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS genetics* **12**, e1005767 (2016).
42. Bradbury, P. J. *et al.* Tassel: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
43. Lipka, A. E. *et al.* Gapit: genome association and prediction integrated tool. *Bioinformatics* **28**, 2397–2399 (2012).
44. Shrestha, V. The search for modifiers of the maize gametophyte factor *gal-s* and quantitative trait polymorphisms emerging from doubled-haploid maize lines. (2016).
45. Springer, N. M. *et al.* Maize inbreds exhibit high levels of copy number variation (cnv) and presence/absence variation (pav) in genome content. *PLoS genetics* **5**, e1000734 (2009).
46. Swanson-Wagner, R. A. *et al.* Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. *Genome research* (2010).
47. Catoire, L., Pierron, M., Morvan, C., du Penhoat, C. H. & Goldberg, R. Investigation of the action patterns of pectin-methylesterase isoforms through kinetic analyses and nmr spectroscopy implications in cell wall expansion. *J. Biol. Chem.* **273**, 33150–33156 (1998).
48. Albertsen, M. C., Higgins, L. S. & Trimnell, M. Enhancing the effectiveness of blended refuge (2018). US Patent App. 15/551,822.
49. Jones, Z. G., Goodman, M. M. & Krakowsky, M. D. Identification of resistance to the *gal-m* gametophyte factor in maize. *Euphytica* **206**, 785–791 (2015).

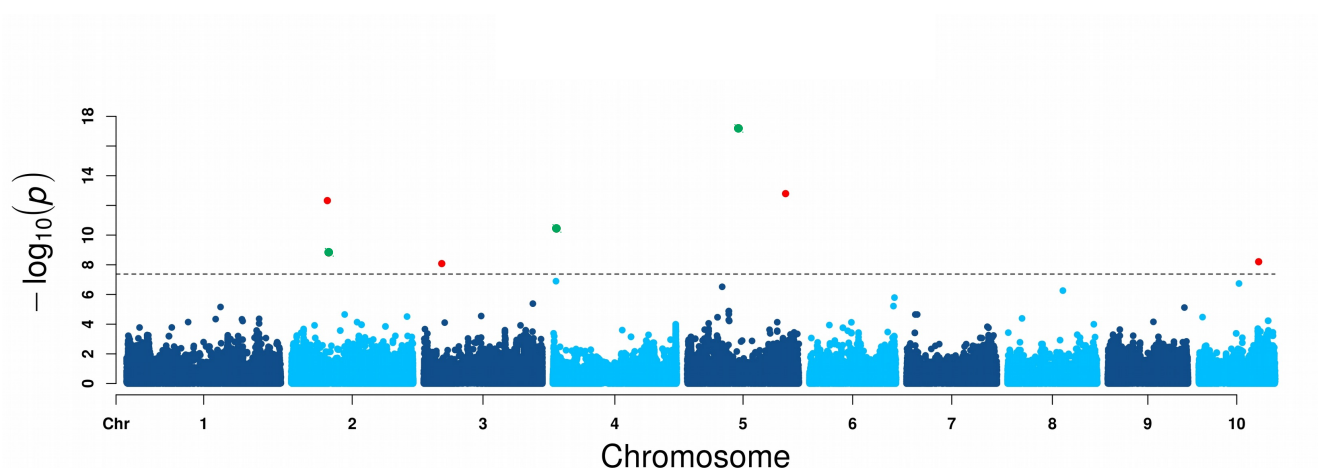


Figure 1. Associations between 238,772 genetic markers and degree of receptivity to *gal* pollen using data from 311 popcorn lines. Each point indicates a single SNP, and position on the y-axis indicates the statistical significance of the association between that marker and sterility/fertility as determined using FarmCPU. Dashed line indicates a threshold of $p=4.18 \times 10^{-8}$ (bonferroni corrected threshold based on the number of SNPs tested and an original threshold of $p = 0.01$). For SNPs above the bonferroni corrected threshold, larger green points were also statistically significant in a parallel MLM GWAS analysis using the same genetic and phenotypic data.

SNP	Significance	Significant in MLM	Gene(s)	Chr	Start	Stop	Function	Mid-Point from SNP
Chr5 bp:9960879	6.27E-18	Yes	GRMZM2G103937	5	9960879	9966861	ZINC FINGER FYVE domain containing protein	2991
Chr2 bp:70339953	4.74E-13	No	AC235541.1_FG002	2	70334195	70338285	DNAJ homolog subfamily C member	3713
			GRMZM5G826658	2	70340627	70344760	Uncharacterized nodulin-like protein	2740.5
Chr4 bp:7020954	3.68E-11	Yes	GRMZM2G157241	4	7012345	7013228	Calcium-binding protein CML30-related	8167.5
			GRMZM2G157247	4	7020204	7022982		639
			GRMZM2G702335	4	7027907	7028428		7213.5
Chr10 bp:117590335	6.22E-09	No	GRMZM2G072174	10	117593705	117595286	PSBP domain-containing protein 2, chloroplastic	4160.5
			GRMZM2G072322	10	117590821	117592686	Late embryogenesis abundant domain-containing protein-related	1418.5
			GRMZM2G072337	10	117586271	117590284	K15865 - Threonylcarbamoyladenosine tRNA methyltransferase CDKAL1	2057.5
			GRMZM2G072513	10	117576572	117580676	Domain of unknown function (DUF3475)	11711
Chr3 bp:35495570	8.29E-09	No	GRMZM2G180065	3	35495835	35503247	Non-specific serine/threonine protein kinase, Threonine-specific protein kinase	4171
			GRMZM2G180071	3	35492943	35495573	Protein tyrosine kinase, Di-glucose binding within ER (Malectin)	1112
Chr5 bp:192080965	1.61E-13	No	N/A					
Chr2 bp:70339953	4.74E-13	Yes	N/A					

Table 1. Table 1: All genes within 10KB of significant SNPs.

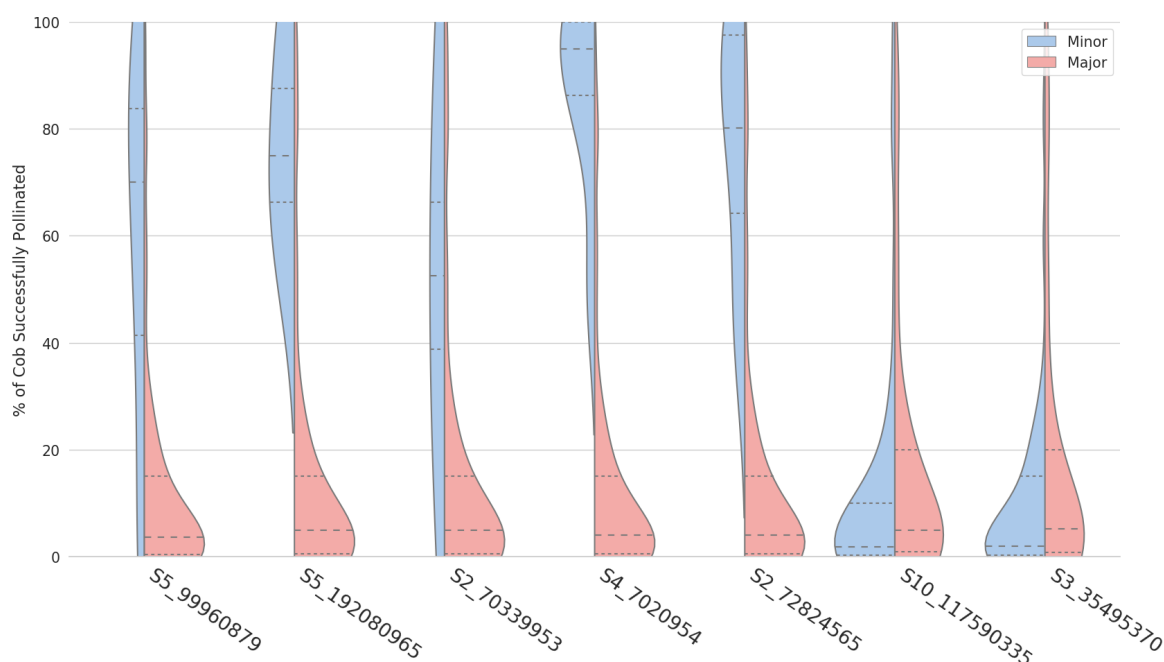


Figure 2. Distribution of % receptivity to *gal* pollen for individuals carrying either the minor or major alleles of each of the seven significantly trait associated SNPs identified in the FarmCPU GWAS analysis. Trait associated SNPs are ordered from most significant p-value (left) to least significant p-value (right). Lines with longer dashes indicate the median trait value for each subpopulation. Lines with shorter dashes indicate the 25th and 75th percentiles of the trait distribution for each subpopulation.

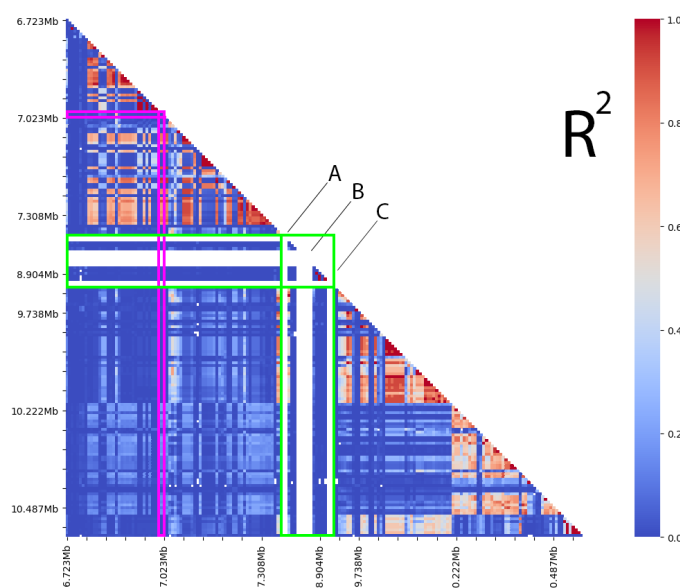


Figure 3. Linkage disequilibrium among SNP markers between 6.7Mb to 10.5Mb on chromosome 4. Each row/column represents a single marker. The purple outline marks the trait associated SNP S_7020954. The green outline marks the interval for *Gal-s* reported by Zhang et al. 2018, translated from B73 RefGenV4 to B73 RefGenV3. Three large gaps without any SNP markers identified in this population, are indicated by the white gaps – not drawn to scale – labeled A, B, and C. A, is a gap without any markers from 7.623Mb to 8.052Mb; B, is a gap without any markers from 8.078Mb to 8.904Mb; C, is a gap without any markers from 8.904Mb to 9.358Mb.

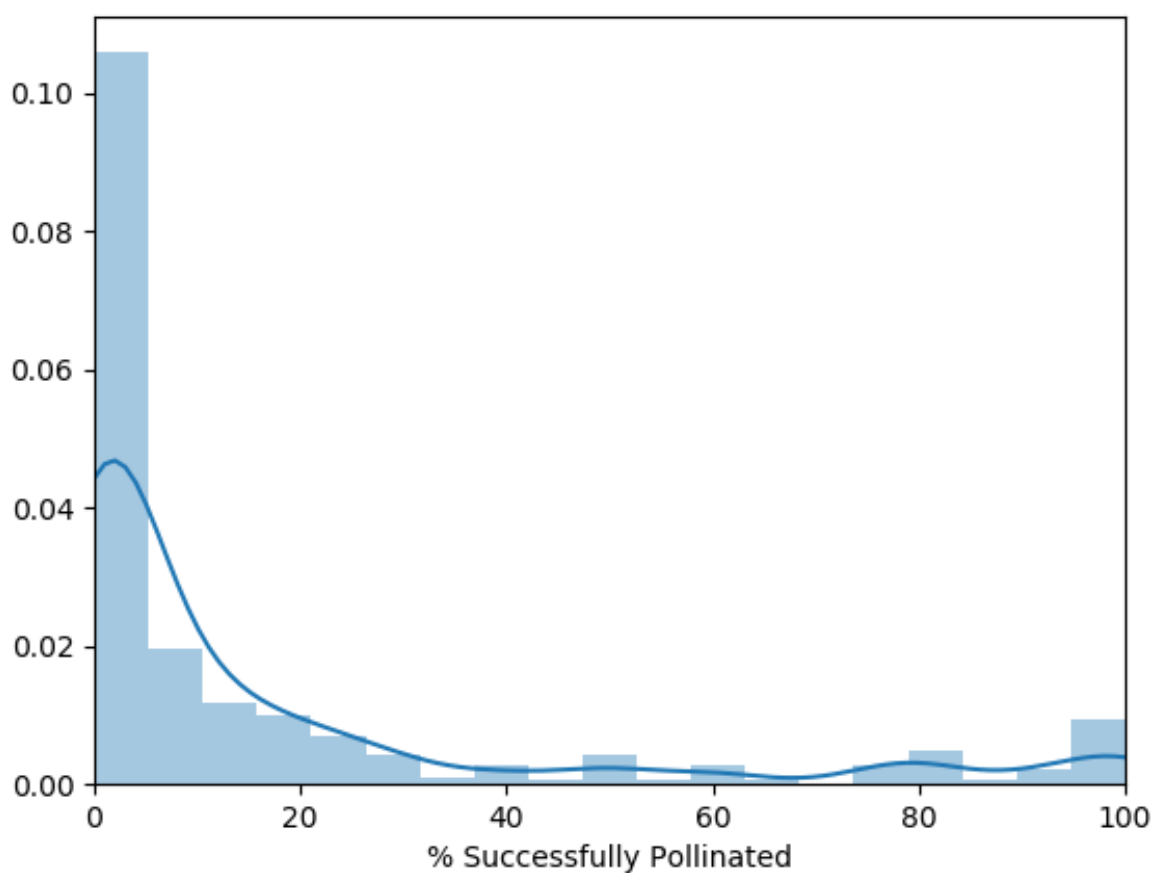


Figure S1. Distribution of percent compatibility with *gal* pollen among the 311 popcorn lines employed in this study. **Let's remove the tick marks from the bottom of the histogram -James**

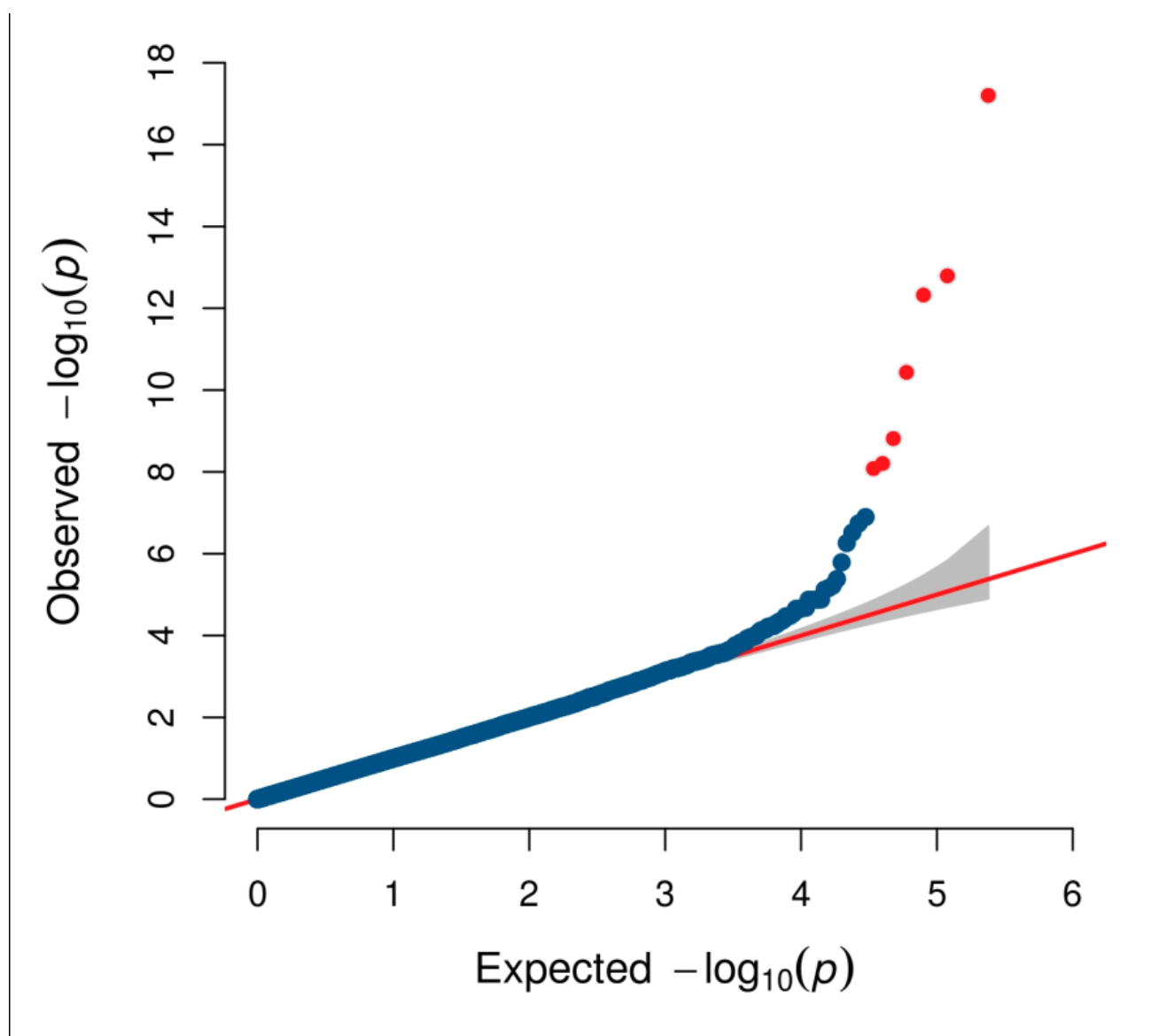


Figure S2. Overrepresentation of significant p-values in the FarmCPU analysis relative to the distribution expected from random data. Points in red indicate p-values which considered to be statistically significant in this analysis (based on a bonferroni corrected p-value threshold of 0.01)

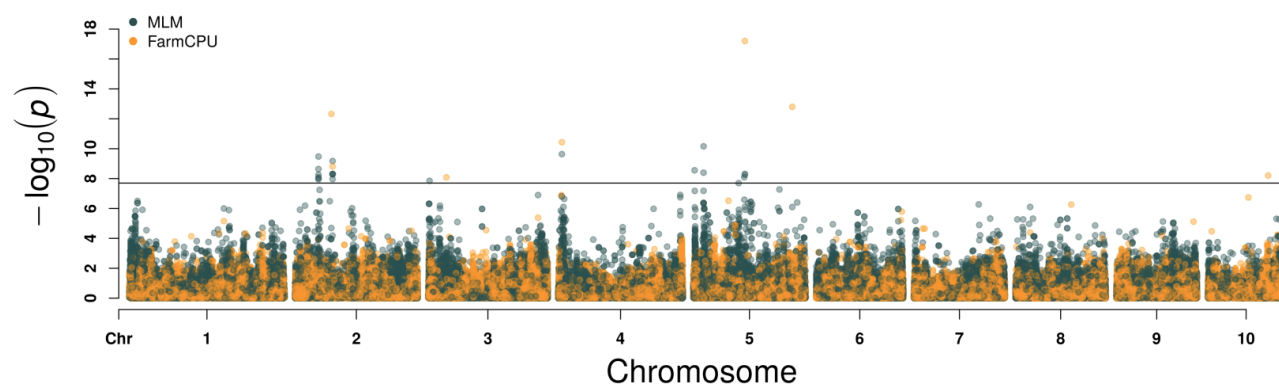


Figure S3. Combined results from both MLM and FarmCPU GWAS mapping. Each SNP is indicated by two separate points one indicating the p-value for the association between that SNP and variation in *gal* pollen compatibility in the MLM GWAS analysis (green) and a second indicating the p-value for the association between that SNP and variation in *gal* pollen compatibility in the FarmCPU GWAS analysis (green)