Original Paper Open Access

Genomic localization of the maize cross-incompatibility gene, *Gametophyte factor 1* (*ga1*)

Josie C Bloom, James B Holland*

USDA-ARS Plant Science Research Unit, Dep of Crop Science, Box 7620, North Carolina State University, Raleigh, NC 27695-7620 USA
*Corresponding author: E-mail: james_holland@ncsu.edu

Abstract

Gametophyte factors in maize (*Zea mays* L) mediate pollen-pistil interactions that confer preferential fertilization of some pistil genotypes by pollen carrying compatible alleles. Non-Mendelian segregation ratios caused by gametophyte factors have been observed in maize since the early 1920's. In this study, we mapped the *ga1* locus by analyzing segregation distortion and QTL for seed set in recombinant inbred lines from the cross B73 × Hp301. The dent corn parent of this cross, B73, carries *ga1*, whereas the popcorn parent Hp301 carries the *Ga1-s* allele conferring preferential fertilization and cross-sterility. High density genotyping-by-sequencing markers were used to delineate the position of *ga1* to a 2.6 Mbp region of the B73 reference sequence (AGP version 2). We also mapped *ga1* based on its inheritance in backcross between W22 (*ga1*) and a near-isogenic line carrying *Ga1-s* on a chromosome segment introgressed from a popcorn variety. The chromosome blocks introgressed from popcorn into W22 were identified with a high-density SNP assay. We created simple sequence repeat markers targeting sequences linked to *ga1* to track its inheritance in the W22 backcross populations. Recently, other researchers independently mapped *ga1* in crosses of Chinese dent and popcorn lines, and we compared our results to theirs with reference to the B73 reference sequence. All lines of evidence were consistent, with the intervals containing *ga1* among all studies overlapping in a 2.2 Mbp interval on chromosome 4 containing 13 predicted genes in the B73 reference sequence.

Keywords: gametophyte factor, pollination, compatibility

Introduction

Certain maize genotypes carry genes referred to as gametophyte factors that mediate pollen – pistil interactions and affect the success of fertilization (Nelson, 1994). In contrast to self-incompatibility systems common in outcrossing plant species (Barrett and Shore, 2008), the gametophyte factors of maize condition preferential fertilization by pollen carrying a gametophyte allele matching a pistil allele (Kermicle and Evans, 2005), thus increasing the potential for self-fertilization and assortative mating. Nelson (1994) suggested that 10 distinct gametophyte factors have been reported in maize. The first gametophyte factor reported in maize, *Gametophyte factor 1* (*ga1*), was identified by its effect on segregation distortion at the linked *sugary 1* locus in crosses between sweet (*ga1 ga1 su1 su1*) and pop (*Ga1 Ga1 Su1 Su1*) maize varieties when backcrosses were made to the popcorn as a female parent (Mangelsdorf and Jones, 1926; Nelson, 1994). On pistils of plants homozygous or heterozygous for a *Ga1* allele, *Ga1* pollen is favored over *ga1* pollen. A strong allele of *Ga1*, *Ga1-s*, confers nearly complete non-reciprocal cross-sterility (or "unilateral cross-incompatibility", Kermicle and Evans, 2010; Schwartz, 1950) with *ga1* pollen. In cultivated maize, most popcorn strains carry the *Ga1-s* allele at the *Gametophyte factor 1* locus, rendering them non-receptive to pollen from corn belt dent varieties, which carry the *ga1* allele (Nelson, 1952; Nelson, 1994). In contrast, the reciprocal cross is generally successful, as *ga1 ga1* pistils do not confer selective fertilization.

Gametophyte factors may play a role in restricting gene flow among maize populations and between maize and teosinte (Kermicle and Evans, 2010; Kermicle et al, 2006), potentially influencing the demography and evolution of the genus *Zea* (Ross-Ibarra et al, 2009; Van Heerwaarden et al, 2011). In addition, the use of gametophyte factors to prevent pollen contamination has been commonplace in popcorn production, but could be expanded to protect organic corn production fields from contamination by nearby varieties carrying transgenes, for example (Kermicle, 2001). *Ga1* was mapped by classical linkage methods to chromosome 4, approximately 23.2 cM from *Su1* (Mangelsdorf and Jones, 1926), but more precise localization of *ga1* would be useful for modern maize breeding and population genetics studies. The objective of this study was to map *ga1* relative to publicly available DNA markers. Recently, Zhang et al (2012) mapped *ga1* to a 1.5 cM region on chromosome 4S in crosses between Chinese pop and dent inbreds. The availability of the B73 reference genome sequence (Schnable et al, 2009) facilitates direct comparisons among independent mapping studies, and we com-

 RECEIVED 28/11/2011

pare results from two populations in our study and those of Zhang et al (2012) using this reference sequence to better resolve the genome interval containing *ga1*. Identifying the location of the *ga1* locus will aid marker-assisted selection of the gene and future efforts to clone the gene.

Materials and Methods

Plant materials

Two independent mapping populations were used to map *ga1*. The first population is a set of 192 recombinant inbred lines (RILs) derived from the cross of wild-type (*ga1 ga1*) dent corn inbred B73 as female to the popcorn inbred line, Hp301, which is homozygous for the *Ga1-s* allele. Each line was derived from a single F5 plant obtained from repeated self-fertilization with minimal selection from a unique F2 from this cross. These lines were developed as part of the maize nested association mapping (NAM) resource and have been genotyped at 1106 polymorphic single nucleotide polymorphism (SNP) markers (McMullen et al, 2009), of which 731 were polymorphic in this cross. Genotype data are publicly available at www.panzea.org. Seeds for this population are available from the USDA Maize Genetic Cooperation Stock Center (maizecoop.cropsci.uiuc.edu/).

The second population was derived from crosses between the wild type dent corn inbred W22 to a nearly-isogenic line homozygous for the *Ga1-s* allele. This *Ga1-s Su1* (W22) stock was developed by Dr. Jerry Kermicle (University of Wisconsin-Madison) by backcrossing the *Ga1-s* allele from the "White Cloud" hybrid popcorn variety five times into the W22 background (Kermicle et al, 2006). We crossed the nearly isogenic pair of lines W22 and *Ga1-s Su1* (W22) and backcrossed the F1 plants to W22 to form a BC1F1 population. BC1F1 plants were backcrossed to W22 to form a BC2F1 population.

Mapping Ga1-s in B73 × Hp301 RILs

The 192 B73 × Hp301 NAM RILs were planted in rows containing 15 – 25 plants each in an isolated field in the summer of 2008. Two rows of the wild-type corn belt dent F1 hybrid LH283 × LH287 were planted between every four rows of RILs in the field. Seeds of the wild-type pollinator were planted by hand in every row on five different dates over a period of five weeks to provide a consistent supply of wild-type *ga1* pollen during the time that RILs flowered. RILs were detasseled. At maturity, four ears from each RIL were scored for percent seed set relative to total number of cupules visible on the cob. Mean percentage seed set for each line was calculated.

QTL for mean seed set were mapped using R/qtl (Broman and Sen, 2009; Broman et al, 2003). Genome-wide significance thresholds were estimated with 1000 permutations for each analysis and support intervals were estimated as 95% Bayesian credible intervals.

Thirty-two B73 × Hp301 NAM RILs with recombination events on the short arm of 4S in the known general area of *Ga1* were selected for retesting in the summer of 2009 to confirm phenotype. Eight plants from each of these lines were crossed by hand as males onto the homozygous *Ga1-s* parent, Hp301 to test their pollen compatibility with *Ga1-s* silks. At maturity, each ear resulting from a hand pollination was scored for percentage seed set in 5% increments relative to the total potential seed set for ear. The mean percent seed set for each RIL was calculated.

Mapping Ga1-s in W22 backcross populations

W22 and *Ga1-s Su1* (W22) were screened for polymorphisms using one SSR marker per each of 85 genome "bins" (approximately 20 cM regions of the genome) and all publicly available SSR markers in bins 4.02 and 4.03 (19 markers at the time), representing the region to which *ga1* had been previously mapped by classical genetics. All DNA extractions were carried out using the ChargeSwitch® gDNA-plant kit (Invitrogen). The 14 µl polymerase chain reactions (PCR) were as follows: 3.5 µl DNA diluted 1:10, 1µl each forward and reverse primers at 0.05 µg/µl, 0.05 µl Taq DNA polymerase, and 8.45 µl 1x PCR reaction buffer (containing 10x PCR buffer,15 mM $MgCl_2$, 10 mM dNTPs, and 5 M betaine). PCR products were separated on a 4% super fine resolution agarose gel.

New SSR markers were designed by searching www.maizesequence.org for repeat motifs in an 8 cM region suspected to harbor the *ga1* locus based on classical linkage map information. The Gramene (www.gramene.org) simple sequence repeat identification tool (Temnykh et al, 2001) was used to identify SSRs in the available sequence, and Oligo Perfect software (Invitrogen) was used to design primer pairs to amplify SSRs. One hundred-forty new markers were developed and tested on the parents and four markers with reliable amplification and easily scored

Table 1 - Newly developed simple sequence repeat (SSR) markers used to map *ga1* in the W22*3(*Ga1-s Su1* (W22)) BC2F1 population.

Marker	BAC	Repeat sequence	Forward primer	Reverse primer
NC1041	AC188008	(TA) 10	CAAGGTTGTCCAAAGAACCAA	AAGCTGGTGGAGACTGGAGA
NC1104	AC208641	(TAC) 5	CAGGCTGCACAAGGTGATAA	GGGAACAGGTACCGAGAAAA
NC1139	AC184846	(CT) 24	TGTTGGGTTTCACTCCTTGA	ATCACAACATGTGCCTGCAT
NC1140	AC191791	(GA) 26	AAAGCTTCAATAGGATCCA	TGTTGGGTTTCACTCCTTGA

polymorphisms were identified (Table 1).

In the summer of 2007, 700 BC1F1 seeds from this cross were planted and individually staked. Cold temperatures and soil crusting soon after planting inhibited germination, and only 291 plants emerged. Highly variable emergence times resulted in substantial variation in flowering time; thus we were able to backcross only 62 BC1F1 plants to form 62 BC2F1 families.

Each family was planted in the field nursery in summer 2009 paired with an adjacent row of *Ga1-s Su1* (W22) stock. Eight BC2F1 plants from within each family were crossed by hand as males to the *Ga1-s Su1* (W22) stock. Families segregating for *Ga1-s ga1* and *ga1 ga1* genotypes are expected to segregate for the capacity to effectively pollinate *Ga1-s Su1* (W22), whereas families uniformly *ga1 ga1* are expected to be incapable of producing good seed set when used as pollinators of the *Ga1-s Su1* (W22) stock. At maturity, each crossed ear was harvested and scored for percentage of seed set in 5% increments relative to the total potential seed set for ear. Genotypes at the ga1 locus were inferred based on the observed seed set, where seed set lower than 10% was considered to be the result of pollination by a *ga1 ga1* plant. In our experience, scanty seed set (between 0 and 10%) can sometimes occur in such crosses due to contamination.

A separate sample of five seeds of each BC2F1 family was used for genotypic analysis of each family. Coleoptile tissue of each plant was collected separately into labeled plates for DNA extraction according to the ChargeSwitch® gDNAplant kit (Invitrogen) instructions. PCR reactions were carried out as described above using four newly developed SSR primers (Table 1) and one previously published SSR primer (nc004), which amplifies part of the *adh2* locus. Markers were ordered by hand relative to *ga1* to minimize double recombination events, recombination frequencies among markers were estimated by directly counting recombinant families, and map distances were calculated using Kosambi's map function. Physical sequence locations of markers were identified on the B73 AGP version 2 reference se-

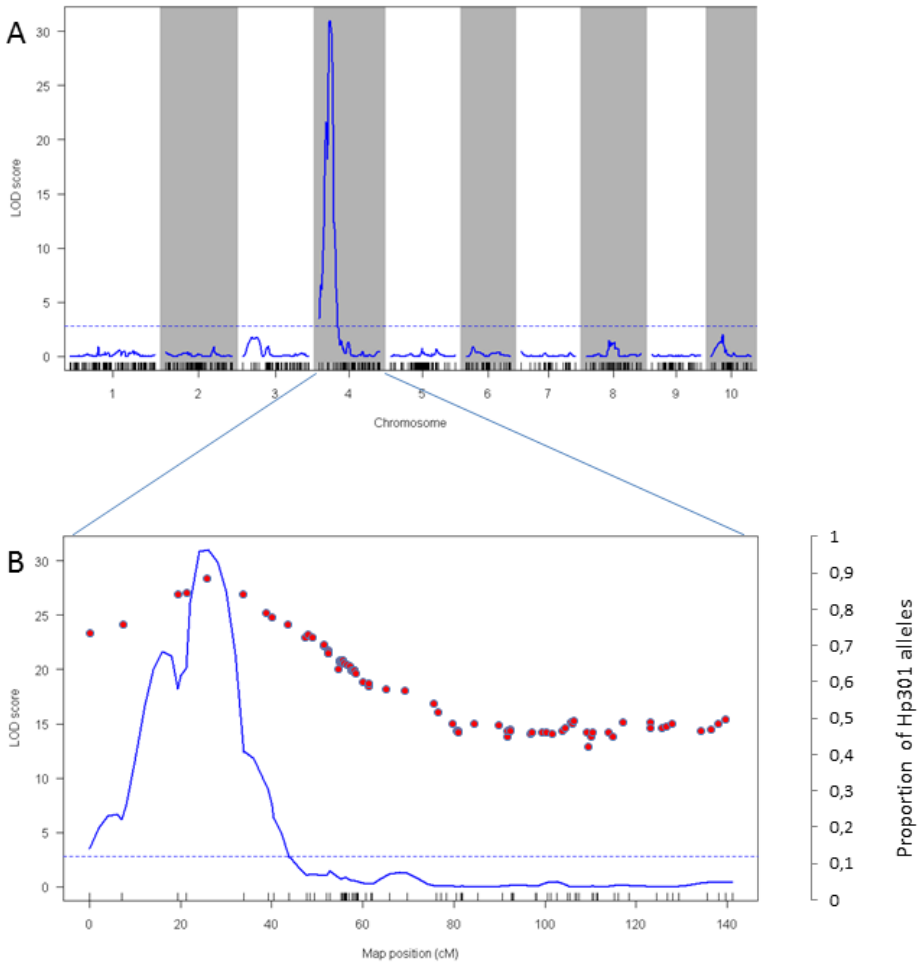

Figure 1 - Genome scan in B73 × Hp301 RILs. (A) Genome scan for QTL affecting seed set of B73 × Hp301 RILs pollinated by ga1ga1 pollen in an isolation block. LOD scores for QTL test are plotted against genome positions. (B) Close-up view of genome scan results for chromosome 4, including frequency of Hp301 allele at each marker plotted with red circles.

quence by BLAST analysis at www.maizesequence.org to permit comparisons across maps.

High density SNP assays of diverse maize lines

A single plant each of W22 and *Ga1-s Su1* (W22) were genotyped at high density with the Illumina MaizeSNP50 Genotyping BeadChip (www.illumina.com) containing 56,110 validated SNP probes derived from the B73 sequence. This information allowed us to define the regions of sequence that differ between the nearly isogenic stocks. Stretches of three or more consecutive polymorphic SNPs or two or more polymorphic SNPs within a 10 kb region were considered to represent genomic segments introgressed from the donor parent into the NIL.

Two high density SNP assays of diverse maize inbred lines were evaluated in an attempt to identify SNPs associated with *Ga1-s* alleles in diverse maize germplasm. A set of 126 lines composed of 30 diverse public maize inbreds (including the founders of the NAM population) and 96 inbred lines developed by private companies with expired PVP certificates were assayed with the Illumina Maize SNP50 Genotyping BeadChip. To our knowledge, Hp301 is the only inbred that carries *Ga1-s* in this set. We searched for SNPs unique to Hp301 and within the genomic interval on chromosome 4S defined by the mapping experiments in this set of markers. An independent survey of 281 diverse maize lines comprising a maize association panel (Flint-Garcia et al, 2005) was conducted with genotyping by sequencing (GBS) (Elshire et al, 2011). We analyzed raw GBS scores to identify markers within the *Ga1-s* interval defined by the mapping studies that were unique to popcorns carrying *Ga1-s*, which include inbred lines 4722, SA24, and the Iowa dent sterile lines (I29, IDS28, IDS69, and IDS91), but exclude the Supergold (Sg) popcorn lines (Ashman, 1975).

High density GBS marker profiles have recently been released for the entire set of NAM RILs (data available at www.panzea.org/lit/data_sets.html). We used GBS genotypes for all B73 × Hp301 RILs to estimate recombination frequencies between GBS markers in the region defined by NAM map SNPs encompassing the QTL region for seed set in the 2008 field experiment. We compared GBS marker genotypes to the mean seed set phenotypes from the 2008 and 2009 field experiments of the B73 × Hp301 RIL to identify the set of GBS markers that cosegregated with the *ga1* phenotype.

We identified predicted genes within the intervals identified by genetic mapping by searching the maize filtered gene set (ZmB73_5b_FGS, available at www.maizesequence.org). Predicted functions of genes were obtained from annotations available for the filtered gene set at www.maizesequence.org and by manual BLAST analysis (Altschul et al, 1990) of each gene's predicted protein sequence against the NCBI nonredundant protein database (blast.ncbi.nlm.nih.gov/Blast.cgi).

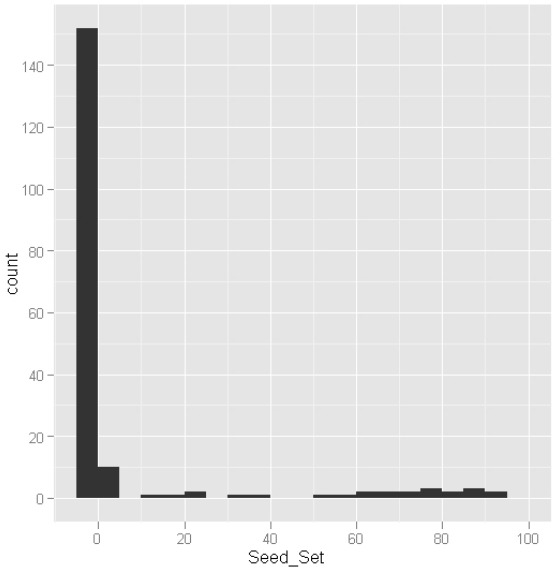

Figure 2 - Histogram of mean seed set scores for B73 x Hp301 RIL population pollinated by a dent corn (*ga1 ga1*) pollinator in an isolation block.

Results and Discussion

Mapping Ga1-s in the B73 × Hp301 NAM RILs

Segregation distortion was observed in the *ga1* region in the B73 × Hp301 NAM family, but not in the other 24 NAM families (McMullen et al, 2009). SNP markers between 0 and 58 cM on the B73 × Hp301 chromosome 4 linkage map exhibited significant ($p < 0.01$) excess of Hp301 alleles (Figure 1). The proportion of Hp301 alleles at each locus in this region ranged from 63% to 88%. This maximum value is about what is expected at the *ga1* locus itself in F5-derived lines if *ga1* pollen is 12% effective at fertilizing on *Ga1-s/ga1* silks, similar to a value of 20% effectiveness previously suggested by Mangelsorf and Jones (1926). Markers showing the most severe segregation distortion (> 78% Hp301 alleles) were located from 19.4 to 33.9 cM on the map, delimiting the position of *ga1* to this region, with the maximum segregation distortion observed at marker PZA00975.1 located at 25.9 cM (Figure 1).

We mapped QTL for percent seed set when B73 × Hp301 RIL silks were pollinated with wild type pollen in an isolation block. Mirroring the high frequency of Hp301 alleles at markers in the *ga1* region, we observed that 88% of the RILs had the Hp301 "dent sterile" phenotype, with less than 5% of the potential seed set realized when pollen was exclusively ga1 type (Figure 2). A single QTL was detected above the permutation-based genome-wide $\alpha = 0.05$ significance threshold, and the most likely QTL position was 26.1 cM on the linkage map (Figure 1). The 95% Bayes credible interval for the QTL was 2 cM, from 24.1 to 26.1 cM on the linkage map, inside of the 13 cM interval flanked by SNP markers PZA00436.7 and PZA01122.1. These marker positions define an inter-

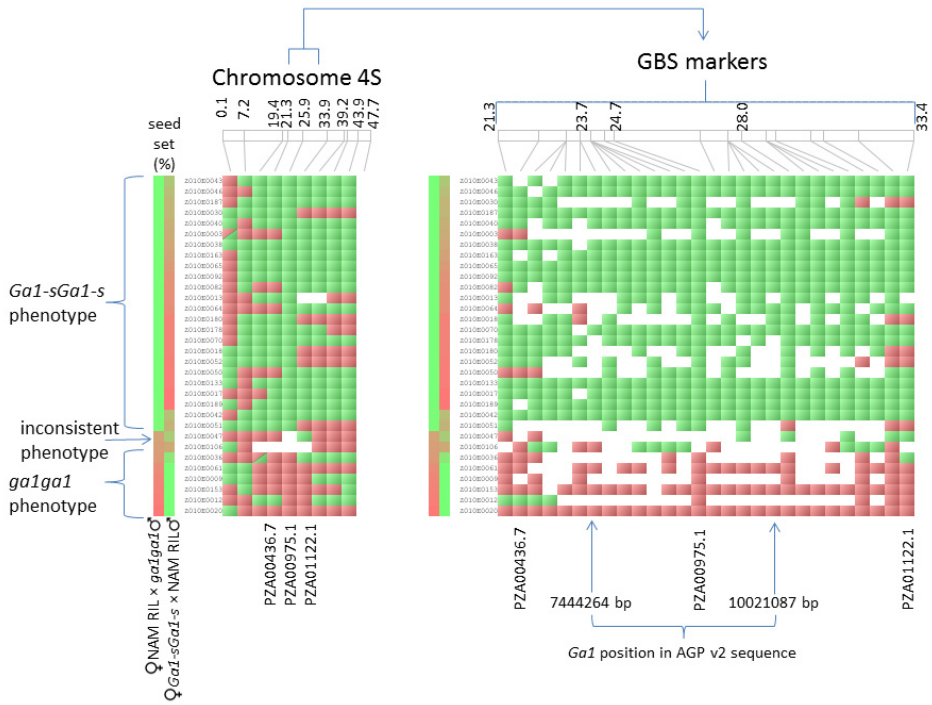

Figure 3 - Genotypes and phenotypes of B73 × Hp301 RILs with recombinations near *ga1*. Pollen compatibility phenotypes and SNP genotypes for 32 B73 × Hp301 RILs with recombinations on chromosome 4S. Seed set phenotypes are indicated by heat map on left side, with colors representing the percent seed set, from zero (green) to dark red (100%). Left hand phenotype heat map indicates seed set of RILs when pollinated by *ga1 ga1* male; right hand phenotype heat map indicates seed set of RILs when used as males to pollinate a *Ga1-s Ga1-s* female. RILs are sorted by phenotypes so that lines near the top have the *Ga1-s Ga1-s* phenotype, whereas those near the bottom have the ga1 ga1 phenotype. Leftmost set of graphical genotypes indicate alleles at NAM map SNPs on chromosome 4S; rightmost set of graphical genotypes indicate alleles of same RILs at three NAM SNPs defining the ga1 QTL interval along with informative and non-redundant GBS markers inside the interval. Genetic map positions are given at top of graphical genotypes; physical sequence positions are given at bottom. Genetic positions for the GBS markers were estimated from recombinations observed among all 192 B73 × Hp301 RILs; the total distance across the interval estimated in this way was similar to but not exactly identical to the original NAM SNP marker map of the population (a difference of 0.5 cM was observed).

val between 6408214 and 12609493 bp on the chromosome 4 AGP version 2 reference sequence.

To check the reliability of seed set scores, 32 RIL silks exhibiting recombinations in the putative *ga1* region were tested for their capacity to effectively set seed as pollinators of Hp301, the *Ga1-s Ga1-s* parent. RILs carrying *Ga1-s* should be compatible with Hp301 and capable of producing good seed set, whereas RILs homozygous for *ga1* should be incompatible and incapable of producing good seed set. Of the 32 RILs tested, 30 exhibited phenotypes consistent with their previous evaluations as female parents pollinated by the wild type tester in the previous year (Figure 3). This result indicates the general reliability of scoring seed set in the two tests, but suggests that a low frequency of phenotypic classification errors was observed, presumably due to pollen contamination in the test of male ga1 function or due to poor pollination success because of environmental factors or timing of flowering in the test of female ga1 function. Alternatively, at least one RIL that was successfully fertilized by wild type pollen and whose pollen was able to fertilize Hp301 (Figure 3) could represent a recombination event that separates the male gametophytic function of *Ga1-s* from its female sporophytic function, resulting in a *Ga1-m* allele (Kermicle et al, 2006). Results from the two years of testing of NAM RILs with recombinations near the QTL peak clearly demonstrate that marker PZA00975.1 (25.9 cM) has the most consistent cosegregation with the Ga1-s phenotype, and that flanking markers PZA00436.7 (21.3 cM) and PZA01122.1 (33.9 cM) delineate the position of *ga1* (Figure 3).

The entire NAM population has also been genotyped with GBS markers, permitting us to further refine the interval of the *ga1* locus by more precisely identifying the positions of recombinations within the 12 cM interval around the mapped gene that coincides with the segregation distortion peak. We identified 89 informative SNP markers inside this interval with reliably imputed genotype scores (i.e., no double crossovers within very small physical distances). By comparing ga1 phenotypes to recombination positions within this region, we defined the position to a

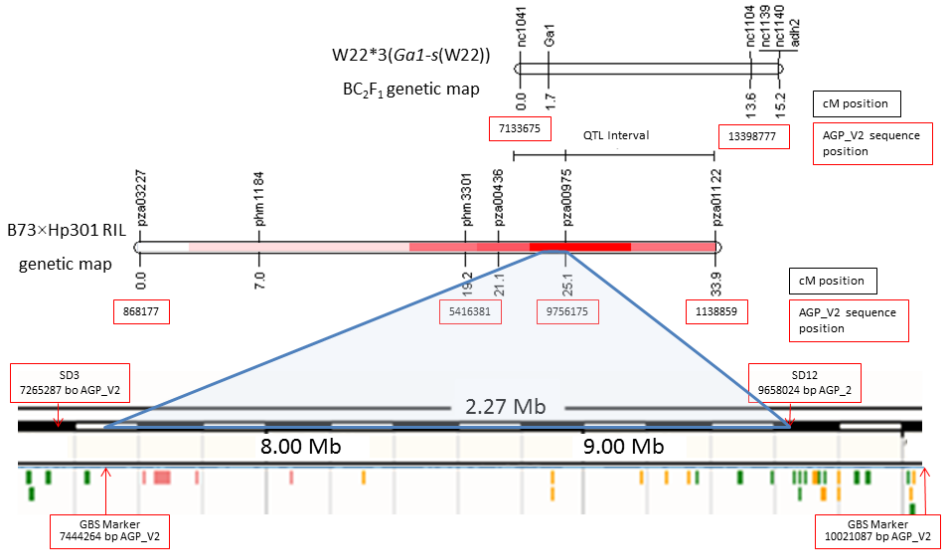

Figure 4 - Genetic and physical maps of ga1 region. Genetic and physical maps of ga1 region in W22 × Ga1-s Su1 (W22) backcross families, B73 × Hp301 RILs, and SDGa25-derived mapping families (Zhang et al, 2012). Overlapping regions were identified by reference to the B73 AGP version 2 reference sequence, and genes in this region were identified in the maize filtered gene set. SD3 and SD12 refer to SSR markers mapped by Zhang et al (2012). Intensity of red color on B73 × Hp301 RIL linkage map indicates frequency of Hp301 allele, with darker red indicating higher frequency of Hp301 allele. Colored rectangles below the physical sequence indicate predicted genes.

2.6 Mbp interval between SNP markers at AGP V2 positions 7444264 bp and 10021087 bp on the AGP v2 reference sequence (Figures 3 and 4).

Mapping Ga1-s in the W22 × Ga1-s Su1 (W22) backcross population

Five of 19 SSR markers located in chromosome bins 4.02 and 4.03 surveyed were polymorphic between W22 and *Ga1-s Su1* (W22). In addition, SSR polymorphisms between these two lines were detected in bins 1.04 and 3.03 but not in any of the other 84 bins surveyed. BAC sequences in the region defined by the five polymorphic SSR markers were screened to identify additional simple sequence repeats, with the objective of creating new PCR-based markers for this region. We identified 140 SSR motifs in the region and created primer pairs for each. Four of these new SSR markers amplified reliably and produced clear size polymorphisms between the parents that could be identified using separation on agarose gels.

High-density SNP genotyping with more than 50,000 SNPs revealed 14 genomic regions distributed on chromosomes 1, 3, 4, 5, 8, and 10 that were different between W22 and *Ga1-s Su1* (W22). Combined, these regions represented about 76.7 Mb, or about 3% of the approximately 2.5 Gb B73 reference sequence. These results demonstrated that *Ga1-s Su1* (W22) is nearly-isogenic to W22, with only a small amount of sequence difference outside the *ga1* locus region. We identified two regions on chromosome 4S where W22 and *Ga1-s Su1* (W22) had a high frequency of non-identical-in-state SNP genotypes: 5980646 to 11081179 bp and 12050310 to 14113339 bp on the B73 AGP version 2 reference sequence.

As expected for a gametophyte factor, the *Ga1-s* allele did not affect preferential fertilization when backcrossed into a *ga1 ga1* background. We observed 45% of the BC2F1 families to be segregating for *Ga1-s ga1* and *ga1 ga1* genotypes, not significantly different from the expectation of 50%. Limited marker density and sample size precluded high resolution mapping of *ga1* in the backcross of *Ga1-s Su1* (W22) to its near-isogenic parental stock, W22. Nevertheless, *ga1* was localized to a 13.6 cM interval defined by two newly developed SSR markers, nc1041 (7133675 bp AGP v2) and nc1104 (13398777 bp AGP v2) that is consistent with the non-identical in state region between the parents of the population and also with the interval defined by QTL mapping in the B73 × Hp301 RIL population (Figure 4).

SNP alleles associated with Ga1-s in diverse maize lines

Of the more than 50,000 SNP markers included in the Illumina Maize SNP50 Genotyping BeadChip, 151 SNPs map within the physical interval defined to carry *ga1* in the mapping studies. Hp301 did not carry a unique allele at any of these loci compared to the 126 other diverse *ga1* maize lines included in the survey. A denser sample of 1198 SNP markers within this region was available for the 282 diverse maize lines in the GBS data set. We identified two SNPs at which all of the *Ga1-s* popcorn inbred lines carried an allele that matches the Hp301 alleles, but for which all other inbreds, including *ga1* popcorn inbreds, had a different allele. These SNPs were at

AGP v2 positions 10217183 (inside predicted gene GRMZM2G068398) and 10485210 (inside predicted gene GRMZM2G008507) on chromosome 4. These two genes are related by sequence and are highly homologous to sucrose-phosphate synthase genes from wheat and other plants. The association between these SNPs and *Ga1-s* is likely due to population history, however, since all of the *Ga1-s* type popcorns in this sample are derived from the South American popcorn group. Furthermore, a set of tropical-derived dent lines developed at North Carolina State University included in this set (NC296, NC296A, NC302, NC336, NC338, NC346, NC348, NC352, and NC354) carry a gametophyte factor that appears to be compatible with *Ga1-s* (Major Goodman, pers. communication), but do not share the Hp301 allele at these two SNPs. This could indicate that the Ga1-type NC dent lines carry a distinct allele of *ga1* compared to the popcorn lines, or more likely, this indicates that these two SNPs are in linkage disequilibrium with the *ga1* locus, but neither SNP is causative of the Ga1-s phenotype. These two SNPs are also just outside of the *ga1* interval defined by the linkage mapping studies, further suggesting that their association with *Ga1-s* is due to population structure rather than causality.

Comparison to ga1 position mapped by Zhang et al. (2012)

Zhang et al (2012) recently mapped *ga1* with high resolution in four families derived from backcrossing the *Ga1-s* donor line SDGa25 to different dent lines, using the *ga1 ga1* parents as females to prevent segregation distortion around *ga1*. They developed new SSR markers, some of which cosegregated with *ga1* in more than 1400 BC1F1 plants. Recombination analysis resolved the interval containing *ga1* to a 1.5 cM genetic interval defined by their newly created SSR markers SD3 and SD12 (Zhang et al, 2012). We identified the positions of these two markers on the B73 AGP version 2 reference sequence as 7265287 and 9658024 bp, respectively (Figure 4). The total size of this interval, 2.4 Mbp, is slightly larger than indicated by Zhang et al (2012), but still slightly smaller than the 2.6 Mbp interval defined in the B73 × Hp301 population. Importantly, the interval defined in their study and the present study are consistent and the two intervals overlap in the region between 7444264 and 9658024. The original interval defined by Zhang et al (2012) contained 15 predicted genes in the filtered predicted maize gene set (www.maizesequence.org). By combining their information with ours, we have refined the interval to one containing only 13 predicted genes (Table 2). Most of these genes have no obvious homology to known gene classes or protein domains, and those with identified homologies do not have an obvious relation to the ga1 cross-incompatibility phenotype (Table 2). One of the genes, GRMZM2G135056, encodes a nodulin-like protein, which is likely to contain transmembrane domains (Gollhofer et al, 2011), potentially permitting it to play a role in pollen-style interactions. Another gene, GRMZM2G039983, has homology to WDL1 of Arabidopsis, which regulates anisotropic growth of cells (Yuen et al, 2003), and perhaps could be involved in pollen tube growth.

Non-colinearity and insertion/deletion variation for whole genes appears common among maize inbreds (Fu and Dooner, 2002; Lai et al, 2005; Morgante et al, 2005). Thus, the difference between *Ga1-s* and *ga1* alleles may be caused by the presence or absence of a gene, such that the gene does not exist in B73. In this case, searching the predicted gene set of the B73 reference sequence is not an appropriate strategy. In the absence of a large insert sequence library for Hp301 or *Ga1-s Su1* (W22) lines, high resolution genetic mapping of the gene must proceed to an interval sufficiently small to permit its amplification by PCR from the line carrying the gene. Alternatively, the presence/absence of sequences generated by GBS of Hp301 and absent from the B73 reference

Table 2 - Predicted genes in the interval overlapping regions containing *ga1* across several populations and studies.

	AGP v2 sequence positions (bp)		
Gene ID	Transcript start	Transcript end	Conserved domains
GRMZM2G012821	7616846	7618466	F-box domain cyclin-like
GRMZM2G424553	7653177	7691914	kinesin motor domain
GRMZM2G135056	7780877	7782970	nodulin-like
GRMZM2G181073	8078275	8079905	
GRMZM2G029496	8305887	8308705	
GRMZM5G835418	8899536	8900563	
AC196002.2_FG002	8901387	8901950	
AC201986.3_FG002	9183034	9183546	
GRMZM2G702344	9259652	9260731	
GRMZM5G817995	9325329	9325631	
GRMZM2G419836	9351020	9354236	thioredoxin-like fold
GRMZM2G027021	9485207	9494351	GTP-binding protein hflX
GRMZM2G039983	9589010	9592389	Xklp2 targeting protein, WDL1

sequence could be mapped in the RIL population. Hp301-unique sequences that cosegregate with *Ga1-s* could represent the gene. Lastly, the gene could be identified by transposon tagging by knockout of the dominant *Ga1-s* gene in the Hp301, *Ga1-s Su1* (W22), or other background.

Further genetic analysis should lead to the identification of the *ga1* gene itself. The identification of its genomic position relative to publicly available DNA markers reported in this study can be immediately useful for marker-assisted breeding applications to introduce *Ga1-s* into maize populations that will be aided by a barrier to pollination by dent maize. This could include popcorns (where it is already deployed), sweet corns, or organic maize varieties that require isolation from transgenic maize pollen. The role of *ga1* in the population genetics and evolution of maize and teosinte and its potentially unique mechanism of action in the fertilization process should be amenable to study once the gene has been identified.

Acknowledgements

We thank Dr. Jerry Kermicle for the gift of *Ga1-s su1* (W22) stock seeds. Research supported by USDA-Agricultural Research Service and USA National Science Foundation (DBI–0321467 and IOS–0820619).

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, 1990. Basic Local Alignment Search Tool. J Mol Biol 215: 403-410

Ashman RB, 1975. Modification of cross-sterility in maize. J Hered 66: 5-9

Barrett SCH, Shore JS, 2008. New insights on heterostyly: Comparative biology, ecology and genetics, pp. 3-32. In: Self-Incompatibility in Flowering Plants – Evolution, Diversity, and Mechanisms. V. Franklin-Tong E ed. Springer-Verlag, Berlin

Broman K, Sen Ś, 2009. A guide to QTL mapping with R/qtl. Springer, Dordrecht

Broman K, Wu H, Sen Ś, Churchill G, 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics. 19: 889-890

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE, 2011. A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. PLoS One 6

Flint-Garcia, SA, Thuillet AC, Yu J, Pressoir G, Romero SM, Mitchell SE, Doebley J, Kresovich S, Goodman MM, Buckler ES, 2005. Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J 44: 1054-1064

Fu H, Dooner HK, 2002. Intraspecific violation of genetic colinearity and its implications in maize. Proc Natl Acad Sci USA 99: 9573-9578

Gollhofer J, Schläwicke C, Jungnick N, Schmidt W, Buckhout TJ, 2011. Members of a small family of nodulin-like genes are regulated under iron deficiency in roots of Arabidopsis thaliana. Plant physiology and biochemistry. 49: 557-564

Kermicle JL, 2001. Genetic barriers that restrict hybridization in corn and teosinte. 56th Annual Corn and Sorghum Seed Research Conference. 56: 17-23

Kermicle JL, Evans MMS, 2005. Pollen-pistil barriers to crossing in maize and teosinte result from incongruity rather than active rejection. Sex Plant Reprod 18: 187-194

Kermicle JL, Evans MMS, 2010. The Zea mays Sexual Compatibility Gene *ga2*: Naturally Occurring Alleles, Their Distribution, and Role in Reproductive Isolation. J Hered 101: 737-749

Kermicle JL, Taba S, Evans MMS, 2006. The *gametophyte-1* locus and reproductive isolation among *Zea mays* subspecies. Maydica 51: 219-225

Lai J, Li Y, Messing J, Dooner HK, 2005. Gene movement by Helitron transposons contributes to the haplotype variability of maize. Proc Natl Acad Sci USA 102: 9068-73

Mangelsdorf PC, Jones DF, 1926. The expression of Mendelian factors in the gametophyte of maize. Genetics. 11: 423-455

McMullen MD, Kresovich S, Sanchez Villeda H, Bradbury P, Li H, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C, Brown P, Browne C, Eller M, Guill K, Harjes C, Kroon D, Lepak N, Mitchell SE, Peterson B, Pressoir G, Romero S, Oropeza M Rosas, Salvo S, Yates H, Hanson M, Jones E, Smith S, Glaubitz JC, Goodman M, Ware D, Holland JB, Buckler ES, 2009. Genetic properties of the maize Nested Association Mapping population. Science 325: 737-740

Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, Rafalski A, 2005. Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. Nat Genet 37: 997-1002

Nelson OE, 1952. Non-reciprocal cross-sterility in maize. Genetics 37: 101-124

Nelson OE, 1994. The gametophyte factors of maize, pp. 496-503. In: The Maize Handbook. Freeling M, Walbot V eds. Springer-Verlag, New York

Ross-Ibarra J, Tenaillon M, Gaut BS, 2009. Historical divergence and gene flow in the genus *Zea*. Genetics 181: 1399-1413

Schnable, PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E,

Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, L McMahan, Van Buren P, Vaughn MW et al, 2009. The B73 maize genome: Complexity, diversity, and dynamics. Science 326: 1112-1115

Schwartz D, 1950. The analysis of a case of cross-sterility in maize. Proc Natl Acad Sci USA 36: 719-724

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S, 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L): Frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11: 1441-1452

Van Heerwaarden J, Doebley J, Briggs WH, Glaubitz JC, Goodman MM, Sanchez Gonzalez JJ, Ross-Ibarra J, 2011 Genetic signals of origin, spread, and introgression in a large sample of maize landraces. Proc Natl Acad Sci USA 108: 1088-1092

Yuen CYL, Pearlman RS, Silo-suh L, Hilson P, Carroll KL, Masson PH, 2003. WVD2 and WDL1 modulate helical organ growth and anisotropic cell expansion in Arabidopsis. Plant Physiology 131: 493-506

Zhang H, Liu X, Zhang Y, Jiang C, Cui D, Liu H, Li D, Wang L, Chen T, Ning L, Ma X, Chen H, 2012. Genetic analysis and fine mapping of the *Ga1-S* gene region conferring cross-incompatibility in maize. Theor Appl Genet 124: 459-65