Theor Appl Genet (2011) 123:667–679
DOI 10.1007/s00122-011-1617-5

ORIGINAL PAPER

Prevalence of segregation distortion in diploid alfalfa and its implications for genetics and breeding applications

Xuehui Li · Xiaojuan Wang · Yanling Wei · E. Charles Brummer

Received: 22 February 2011 / Accepted: 9 May 2011 / Published online: 31 May 2011

Abstract Segregation distortion (SD) is often observed in plant populations; its presence can affect mapping and breeding applications. To investigate the prevalence of SD in diploid alfalfa (*Medicago sativa* L.), we developed two unrelated segregating F_1 populations and one F_2 population. We genotyped all populations with SSR markers and assessed SD at each locus in each population. The three maps were syntenic and largely colinear with the *Medicago truncatula* genome sequence. We found genotypic SD for 24 and 34% of markers in the F_1 populations and 68% of markers in the F_2 population; distorted markers were identified on every linkage group. The smaller percentage of genotypic SD in the F_1 populations could be because they were non-inbred and/or due to non-fully informative markers. For the F_2 population, 60 of 90 mapped markers were distorted, and they clustered into eight segregation distortion regions (SDR). Most SDR identified in the F_1 populations were also identified in the F_2 population. Genotypic SD was primarily due to zygotic rather than allelic distortion, suggesting zygotic not gametic selection is the main cause of SD. On the F_2 linkage map, distorted markers in all SDR except two showed heterozygote excess. The severe SD in the F_2 population likely biased genetic distances among markers and possibly also marker ordering and could affect QTL mapping of agronomic traits. To reduce the effects of SD and non-fully informative markers, we suggest constructing linkage maps and conducting QTL mapping in advanced generation populations.

Communicated by J. Yu.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1617-5) contains supplementary material, which is available to authorized users.

X. Li · Y. Wei · E. C. Brummer (✉)
Forage Improvement Division, The Samuel Roberts Noble Foundation, Ardmore, OK 73401, USA
e-mail: ecbrummer@noble.org

X. Wang
School of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China

Introduction

Mendelian segregation of alleles at a locus is a fundamental tenet of classical genetics, enabling geneticists to predict the expression of simply inherited traits. Deviations from the Mendelian segregation expectations have been investigated since the early twentieth century (Mangelsdorf and Jones 1926; Rhoades 1942; Sandler et al. 1959), and the phenomenon was defined as segregation distortion (SD) (Sandler et al. 1959). SD of molecular markers has been observed in most genetic mapping studies across all the species, for example, maize (Lu et al. 2002), rice (Xu et al. 1997), tomato (Paterson et al. 1988), alfalfa (Brummer et al. 1993), and *Populus* (Yin et al. 2004). Distorted markers may bias the estimation of genetic distance between markers and may affect orders among markers (Lorieux et al. 1995a, b); consequently, it will affect the genetic mapping of phenotypic traits (Vogl and Xu 2000; Xu 2008).

Biologically, the cause of SD could result from selection among gametes and/or zygotes (Falconer and Mackay 1996). If an allele or alleles at a locus is detrimental to gamete or zygote fitness, then that locus will cause markers linked to it to deviate from the expected Mendelian segregation ratios (Zamir and Tadmor 1986). Thus, observing a cluster of markers showing SD suggests that the

chromosomal region may have one or more genes causing SD. The region of distorted markers is termed segregation distortion region (SDR, Lu et al. 2002) and the causative, underlying gene(s) segregation distortion locus (SDL, Vogl and Xu 2000). Based on marker clustering and skew direction, SDRs have been identified in maize (Lu et al. 2002), rice (Xu et al. 1997), grape (Riaz et al. 2008), and other species (Faris et al. 1998; McDaniel et al. 2007). A few studies have attempted to locate SDLs on genetic linkage maps (Hall and Willis 2005; Luo et al. 2005; Wang et al. 2005).

Diploid alfalfa includes two main subspecies, *M. sativa* subsp. *caerulea* and *M. sativa* subsp. *falcata*. Several genetic maps have been constructed for diploid alfalfa using partially inbred mapping populations (F_2 or backcross) derived from intra-subspecies and inter-subspecies crosses (Brummer et al. 1993; Echt et al. 1994; Kaló et al. 2000; Kiss et al. 1993). Serious SD (18–63%) was found in those studies, irrespective of taxonomic relationships between parental genotypes. Distorted markers were usually found clustered, as expected due to linkage, but interestingly, the SD was most commonly with heterozygote excess. A biological explanation is that alfalfa carries a high level of deleterious (and possibly lethal) recessive alleles due to its outcrossing and self-incompatible nature, and that selection against deleterious recessive alleles during inbreeding causes skewed ratios of marker genotypes in inbred generations. If the deleterious recessive alleles at different loci are linked in repulsion phase, then a preponderance of heterozygotes at linked marker loci would be observed among progeny.

In one non-inbred diploid alfalfa mapping population (F_1 population), only 9% of mapped marker loci showed significant SD, which supports the above explanation (Tavoletti et al. 1996). Another supporting evidence is that a relatively low SD (4–9% of mapped alleles) was observed in a tetraploid backcross mapping population, because deleterious recessive homozygotes have a smaller chance to be exposed and selected against in tetraploid compared to diploid populations (Brouwer and Osborn 1999). However, in non-inbred autotetraploid alfalfa mapping populations (F_1 populations), 35% of AFLP and 25% of SSR markers were significantly distorted in one (Julier et al. 2003) and about 32% of markers exhibited SD in another (Robins et al. 2007). Both biological and non-biological factors might contribute to the varied SD levels observed in those studies, such as different parental genotypes and differing genetic similarity between parents (Chetelat et al. 2000; Lu et al. 2002; Xu et al. 1997). Non-biological factors such as environment (especially conditions during gamete development or seed production), sample size, genotyping errors, and differing statistical methodologies may affect the estimation of SD in different experiments. Thus, analysis of SD in multiple mapping populations while controlling these non-biological factors may help to understand the genetic mechanisms of SD.

The objectives of this study were to test the hypotheses that (1) SD would be identified in diploid alfalfa mapping populations with different genetic backgrounds, (2) greater SD would be observed in a more inbred F_2 population compared to a less inbred F_1 population, both of which derived from the same parents, and (3) the distorted markers would be clustered and predominantly characterized by heterozygote excess in the F_2 population.

Materials and methods

Generation of mapping populations

We obtained four wild, diploid alfalfa plant introductions, PI464712 (*M. sativa* subsp. *caerulea* from Turkey), PI243225 (*M. sativa* subsp. *caerulea* from Iran), PI577551 (*M. sativa* subsp. *caerulea* from Canada), and PI631817 (*M. sativa* subsp. *falcata* from Russia), from the USDA-ARS National Plant Germplasm System. We hybridized a single genotype from two accessions (PI464712-4 × PI631817-1) to develop the inter-subspecific CF-F_1 population ("C" and "F" represent subsp. *caerulea* and *falcata*, respectively). One single plant from the F_1 population [(PI464712-4 × PI631817-1)-5] was self-pollinated to generate the CF-F_2 population. During the time we produced the F_2 population; we also produced additional F_1 individuals in the same greenhouse so that we could compare both populations using seeds that were produced under same environmental conditions.

For the CF-F_1 population, 191 of 240 seeds germinated, and 183 survived and were used for DNA extraction. For the CF-F_2 population, 161 of 237 seeds germinated, and DNA was extracted from 152 F_2 plants, 32 of which died subsequent to DNA extraction.

We also developed the intra-subspecies CC-F_1 population by hybridizing between a single genotype from each of the other two accessions (PI243225-A × PI577551-D). We obtained 190 CC-F_1 individuals from which DNA was extracted and used for genotyping.

SSR genotyping

Tissue from young leaves of greenhouse grown plants was freeze-dried, ground, and used for DNA extraction with the CTAB method (Doyle and Doyle 1990). SSR markers used in previous alfalfa genetic mapping projects (Julier et al. 2003; Robins et al. 2007; Sledge et al. 2005) as well as additional markers obtained from the Noble Foundation, Ardmore, OK were used for mapping (Supplementary

Table 1). Primers were synthesized by Integrated DNA Technologies (IDT, http://www.idtdna.com), with 18 nucleotides of M13 universal primer sequence added onto the 5′ end of the forward primer (Schuelke 2000). The M13 universal primer sequence labeled with blue (6-FAM), green (HEX), or yellow (NED) fluorescent tags were synthesized by Applied Biosystems (http://www.appliedbiosystems.com). The PCR recipes and ingredients were exactly the same as Sledge et al. (2005). The PCR program used was either the same as Julier et al. (2003) 4 min at 94°C, followed by 35 cycles with 30 s at 94°C, 1 min at 55°C and 1 min at 72°C, plus a final elongation step of 7 min at 72°C (program A in Supplementary Table 1) or as follows: 95°C for 2 min, followed by 30 cycles with 30 s at 95°C, 45 s at 60°C, and 45 s at 72°C, plus 10 cycles with 30 s at 95°C, 45 s at 53°C, 45 s at 72°C, and finalized with an elongation step of 7 min at 72°C (program B in Supplementary Table 1). PCR products from four to eight SSR markers were diluted ten times and pooled for each individual, mixed with 0.2 µl of GeneScan-500 ROX size standard (ABI catalogue # 401734) and analyzed on an ABI 3730 DNA analyzer. The data files from the sequencer were analyzed using the Genemarker software (http://www.softgenetics.com), verified by visual inspection. All markers were scored as co-dominant loci.

Map construction

In total 105 SSR markers were scored and used to construct the genetic linkage map for the CC-F_1 population, 108 markers for the CF-F_1 population, and 103 markers for the CF-F_2 population. Ninety-one markers were common between CF-F_1 and CF-F_2 populations; 42 markers were common between CF-F_1 and CC-F_1 population. For each mapping population, a linkage map was constructed using JOINMAP 4.0. Markers were grouped with a LOD $\geq$ 3.0, and map distances were calculated with the Haldane function. Primer sequences were used to search *M. truncatula* pseudomolecules produced from the *M. truncatula* euchromatic genome sequence, build 3.0 with BLAST at the website http://medicago.org/genome/cvit_blast.php. If both forward and reverse primers hit the same BAC location with a predicted amplicon size similar to our observed fragment size, we used the position with the smaller number on the alignment between the BAC sequence and primer as the physical location of the corresponding marker.

Segregation distortion analysis

F_1 population

The parents used to construct the populations were not inbred (alfalfa suffers from severe inbreeding depression precluding the development of inbred lines). As a consequence, each parent could contain up to two alleles per locus, and the alleles in one parent could be different from those in the other parent. Therefore, seven genotypic segregation patterns are possible for our F_1 populations (Table 1). For each marker, a χ^2 test was used to test genotypic SD as the deviation of the observed genotypic distribution from the expected Mendelian genotypic ratio (Table 1). To test if the genotypic SD was caused by gametic or zygotic selection, an allelic and a zygotic SD test were conducted for each marker. For the allelic SD test, a χ^2 test was used to test the deviation of the observed allelic distribution from the expected allelic ratio of 1:1 for each heterozygous parent (Table 1). This test is the same as that used previously in the two-way pseudo-testcross for evaluating SD in F_1 full-sib progenies (Grattapaglia and Sederoff 1994; Tavoletti et al. 1996). If the parental allele transmission could not be determined (e.g., $A_1A_2 \times A_1A_2$), then we did an overall χ^2 test for allelic SD. For the zygotic SD test, we first estimated allele frequencies based on the observed genotype frequencies in F_1 population (Table 1). Then we used a χ^2 test to test the deviation of the observed genotypic distribution from the genotypic ratio that was expected given allele frequency estimates. When three or more markers with genotypic SD were clustered together, we considered the group of markers as a candidate SDR if the markers in the group showed similar allelic and/or genotypic distortion patterns and had similar allelic and zygotic SD tests.

F_2 population

In the F_2 population, the expected genotypic segregation ratio of A_1A_1:A_1A_2:A_2A_2 is 1:2:1. A χ^2 test was used to identify the markers with genotypic SD. We defined significance at $P < 0.01$. To test if the genotypic SD was caused by gametic or zygotic selection, a modified method of two successive χ^2 tests (Lorieux et al. 1995b) was applied for each marker. Assuming a locus with two alleles, "A_1" and "A_2" in the F_2 mapping population, p represents the frequency of allele "A_1" and q represents the frequency of allele "A_2" The maximum likelihood estimator (MLE) of p is $\hat{p} = \frac{n_{A_1A_1} + 1/2 n_{A_1A_2}}{n}$ and that of q is $\hat{q} = \frac{n_{A_2A_2} + 1/2 n_{A_1A_2}}{n}$, in which n is the total number of genotypes evaluated and $n_{A_1A_1}$, $n_{A_1A_2}$, and $n_{A_2A_2}$ are the observed numbers in each of the three genotypic classes (viz., A_1A_1, A_1A_2, A_2A_2). Allelic SD was tested as a deviation from the expected allelic ratio for A_1:A_2 of 1:1. Zygotic SD was tested as deviation of the observed genotypic ratio from the expected genotypic ratio given the estimated allele frequency (A_1A_1:A_1A_2:$A_2A_2 = n\hat{p}^2 : 2n\hat{p}\hat{q} : n\hat{q}^2$).

Table 1 Expected allelic and genotypic segregation patterns in an F_1 population derived from diploid, non-inbred parents

No. of alleles	Parental matings	Expected genotypic ratio	Estimated allele frequencies[a]	Expected allelic ratio in the heterozygous parent(s)		Expected genotypic ratio given the estimated allele frequencies
				Maternal parent	Paternal parent	
2	$A_1A_2 \times A_1A_1$	A_1A_2:A_1A_1 = 1:1	$\hat{p}_{A_1} = n_{A_1A_1}/n$ $\hat{p}_{A_2} = n_{A_1A_2}/n$	A_1:A_2 = 1:1		A_1A_2:$A_1A_1 = n\hat{p}_{A_1}\hat{p}_{A_2} : n\hat{p}_{A_1}^2$
2	$A_1A_1 \times A_1A_2$	A_1A_1:A_1A_2 = 1:1	$\hat{p}_{A_1} = n_{A_1A_1}/n$ $\hat{p}_{A_2} = n_{A_1A_2}/n$		A_1:A_2 = 1:1	A_1A_1:$A_1A_2 = n\hat{p}_{A_1}^2 : n\hat{p}_{A_1}\hat{p}_{A_2}$
2	$A_1A_2 \times A_1A_2$	A_1A_1:A_1A_2:A_2A_2 = 1:2:1	$\hat{p}_{A_1} = (n_{A_1A_1} + \frac{1}{2}n_{A_1A_2})/n$ $\hat{p}_{A_2} = (n_{A_2A_2} + \frac{1}{2}n_{A_1A_2})/n$	A_1:A_2 = 1:1	A_1:A_2 = 1:1	A_1A_1:A_1A_2:$A_2A_2 = n\hat{p}_{A_1}^2 : 2n\hat{p}_{A_1}\hat{p}_{A_2} : n\hat{p}_{A_2}^2$
3	$A_1A_2 \times A_3A_3$	A_1A_3:A_2A_3 = 1:1	$\hat{p}_{A_1} = n_{A_1A_3}/n$ $\hat{p}_{A_2} = n_{A_2A_3}/n$	A_1:A_2 = 1:1		A_1A_3:$A_2A_3 = n\hat{p}_{A_1}\hat{p}_{A_3} : n\hat{p}_{A_2}\hat{p}_{A_3}$
3	$A_1A_1 \times A_2A_3$	A_1A_2:A_1A_3 = 1:1	$\hat{p}_{A_2} = n_{A_1A_2}/n$ $\hat{p}_{A_3} = n_{A_1A_3}/n$		A_2:A_3 = 1:1	A_1A_2:$A_1A_3 = n\hat{p}_{A_1}\hat{p}_{A_2} : n\hat{p}_{A_1}\hat{p}_{A_3}$
3	$A_1A_2 \times A_1A_3$	A_1A_1:A_1A_3:A_1A_2:A_2A_3 = 1:1:1:1	$\hat{p}_{A_{1m}} = (n_{A_1A_1} + n_{A_1A_3})/n$ $\hat{p}_{A_2} = (n_{A_1A_2} + n_{A_2A_3})/n$ $\hat{p}_{A_{1p}} = (n_{A_1A_1} + n_{A_1A_2})/n$ $\hat{p}_{A_3} = (n_{A_1A_3} + n_{A_2A_3})/n$	A_1:A_2 = 1:1	A_1:A_3 = 1:1	A_1A_1:A_1A_3:A_1A_2:$A_2A_3 = n\hat{p}_{A_{1m}}\hat{p}_{A_{1p}} : n\hat{p}_{A_{1m}}\hat{p}_{A_3} : n\hat{p}_{A_{1p}}\hat{p}_{A_2} : n\hat{p}_{A_2}\hat{p}_{A_3}$
4	$A_1A_2 \times A_3A_4$	A_1A_3:A_1A_4:A_2A_3:A_2A_4 = 1:1:1:1	$\hat{p}_{A_1} = (n_{A_1A_3} + n_{A_1A_4})/n$ $\hat{p}_{A_2} = (n_{A_2A_3} + n_{A_2A_4})/n$ $\hat{p}_{A_3} = (n_{A_1A_2} + n_{A_2A_3})/n$ $\hat{p}_{A_4} = (n_{A_1A_4} + n_{A_2A_4})/n$	A_1:A_2 = 1:1	A_3:A_4 = 1:1	A_1A_3:A_1A_4:A_2A_3:$A_2A_4 = n\hat{p}_{A_1}\hat{p}_{A_3} : n\hat{p}_{A_1}\hat{p}_{A_4} : n\hat{p}_{A_2}\hat{p}_{A_3} : n\hat{p}_{A_2}\hat{p}_{A_4}$

[a] $\hat{p}_{A_1}, \hat{p}_{A_2}, \hat{p}_{A_3}$, and $\hat{p}_{A_4}$ represent estimated frequency of alleles A_1, A_2, A_3, and A_4, respectively; $\hat{p}_{A_{1m}}$ represents estimated frequency of allele A_1 from maternal parent; $\hat{p}_{A_{1p}}$ represents estimated frequency of allele A_1 from paternal parent; n is the total number of genotypes; $n_{A_1A_1}$, $n_{A_1A_2}$, $n_{A_2A_2}$, $n_{A_1A_3}$, $n_{A_1A_4}$, $n_{A_2A_3}$, and $n_{A_2A_4}$ are the observed numbers in each genotypic class (viz., A_1A_1, A_1A_2, A_2A_2, A_1A_3, A_1A_4, A_2A_3, A_2A_4, respectively)

The χ^2 test for allelic SD is as follows:

$$\chi_1^2 = \frac{(2n\hat{p} - n)^2 + (2n\hat{q} - n)^2}{n} \quad (1)$$

The χ^2 test for zygotic SD is as follows:

$$\chi_1^2 = \frac{(n_{A_1A_1} - n\hat{p}^2)^2}{n\hat{p}^2} + \frac{(n_{A_1A_2} - 2n\hat{p}\hat{q})^2}{2n\hat{p}\hat{q}} + \frac{(n_{A_2A_2} - n\hat{q}^2)^2}{n\hat{q}^2} \quad (2)$$

As with F_1 populations, when three or more markers with significant genotypic SD were clustered together, we considered the group of markers as a candidate SDR if the markers in the group showed similar distortion patterns. We denoted the most distorted marker in each SDR as the putative SDL. Based on the relative genotypic frequency at that marker, we estimated the relative viability of each genotype assuming there were no gametic effects. For a locus with two alleles "A_1" and "A_2"(A_1 is dominant allele with higher observed frequency, which could be paternal or maternal grandparent allele), the relative viability of A_1A_1, A_1A_2, and A_2A_2 are 1, $1 - hs$, and $1 - s$, where s is the selection coefficient and h is the degree of dominance. The estimation of h and s was done as described by Luo et al. (2005). A degree of dominance of 0.5 suggests additive effects at the SDL, between 0.5 and 1 suggests recessiveness, between 0 and 0.5 suggests dominance, less than zero suggests over-dominance, and greater than one suggests under-dominance.

Epistasis

To identify loci with epistatic interactions for viability, we conducted a contingency test for each pair of markers using Fisher's exact test. The test of epistatic interactions was only applied to the F_2 population because of the complicated marker segregation patterns in the F_1 populations. Pairs of markers that were linked on one chromosome generally showed non-independence because of linkage disequilibrium. Therefore, we focused on marker pairs from different chromosomes. When one or more contiguous markers from one chromosome showed non-independent segregation with three or more contiguous markers from another chromosome at $P < 0.01$, we considered the two loci to be involved in an epistatic interaction.

Results

Genetic linkage maps

The two F_1 populations, each with 99 mapped markers, had maps of similar total length at just under 550 cM (Table 2). The F_2 population had 90 mapped markers and a shorter map partially due to a number of smaller linkage groups that did not link together (Table 2; Fig. 1). In all cases, markers were spaced on average about 5 cM apart throughout the genome. Comparing the orders of common markers, the genetic maps of the three mapping populations showed a high level of overall synteny (Fig. 1). The alfalfa genetic maps also showed high levels of synteny to the *M. truncatula* physical map (Fig. 1).

Segregation distortion levels

Between one-quarter and one-third of mapped markers had genotypic SD ($P < 0.01$) in the two F_1 populations (CC-F_1 = 34%; CF-F_1 = 24%); 68% were distorted in CF-F_2 (Table 2; Fig. 1; Supplementary Tables 2, 3, 4). In both F_1 populations, all markers with genotypic SD showed allelic SD, but none of them showed zygotic SD (Supplementary Tables 2, 3). In CF-F_2, 26 of 90 mapped markers showed allelic SD and 40 showed zygotic SD ($P < 0.01$). In a number of cases, however, markers on the F_1 populations were not fully informative. For example, if one parent was homozygous for a marker locus, then distortion in this parent would not be identified from the progeny ratios. Therefore, the distortion levels in the F_1 populations are likely underestimated. We have denoted markers that do not show SD but that also are not fully informative in green in Fig. 1.

CC-F_1 population

We identified five candidate SDRs that included 33 of the 36 distorted markers in the CC-F_1 population (red-filled

Table 2 Number of markers scored and mapped in each of the three alfalfa populations, number of markers showing genotypic segregation distortion, map length, and average map distance between markers

Population	No. of markers scored/mapped	No. of markers with genotypic SD/no. of mapped	Map length (cM)	No. of linkage groups	Average map density (cM)
CC-F_1	105/99	36/36	528	8	5.3
CF-F_1	108/99	26/22	547	10	5.5
CF-F_2	103/90	70/60	391	13	4.3

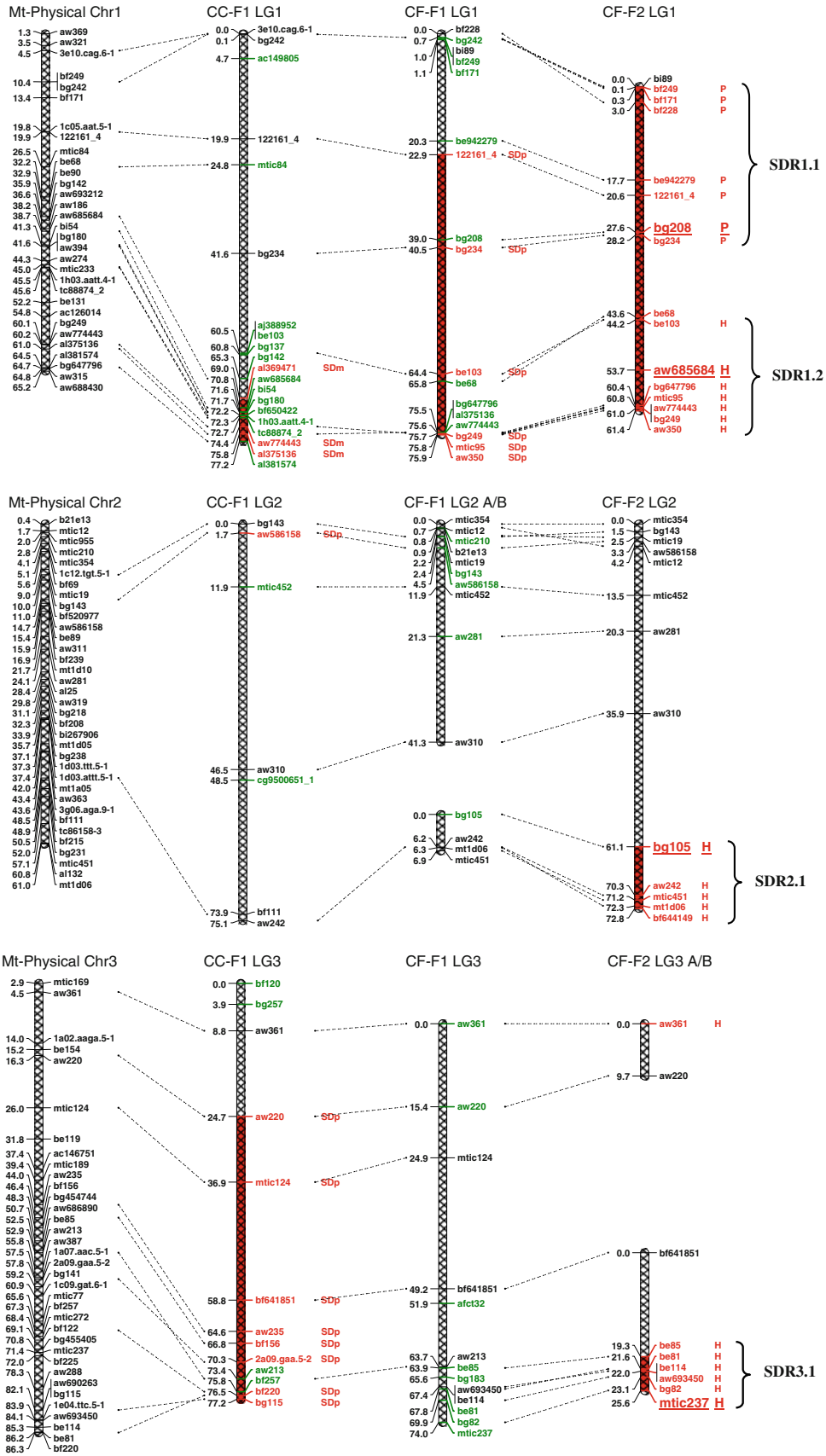


Fig. 1 Physical map of *M. truncatula* (Build 3.0) and genetic linkage maps for three alfalfa mapping populations based on SSR markers. The physical locations indicated on the maps are all in the scale of $\times 10^6$ base pairs. Markers in *red* exhibited genotypic segregation distortion at $P < 0.01$. Markers in *green* on F_1 linkage maps did not show distortion but were not fully informative. Regions of linkage groups considered to be candidate segregation distortion regions (SDR) are denoted by *red fill*. The labels "SDm", "SDp", and "SDmp" on F_1 linkage maps indicate that the distorted marker showed allelic SD for maternal parent alleles, paternal parent alleles, or both maternal and paternal parent alleles, respectively. "M", "P", and "H" were labeled for the distorted markers on the F_2 linkage map that skewed toward homozygotes with maternal grandparent allele, homozygotes with paternal grandparent allele, or heterozygotes, respectively. The segregation distortion regions identified in the F_2 population are designated by the *bracket to the right of the linkage groups*; the most distorted marker in each SDR is *underlined* and *in larger font size* and is considered to be the closest marker to the segregation distortion locus (color figure online)

Fig. 1 continued

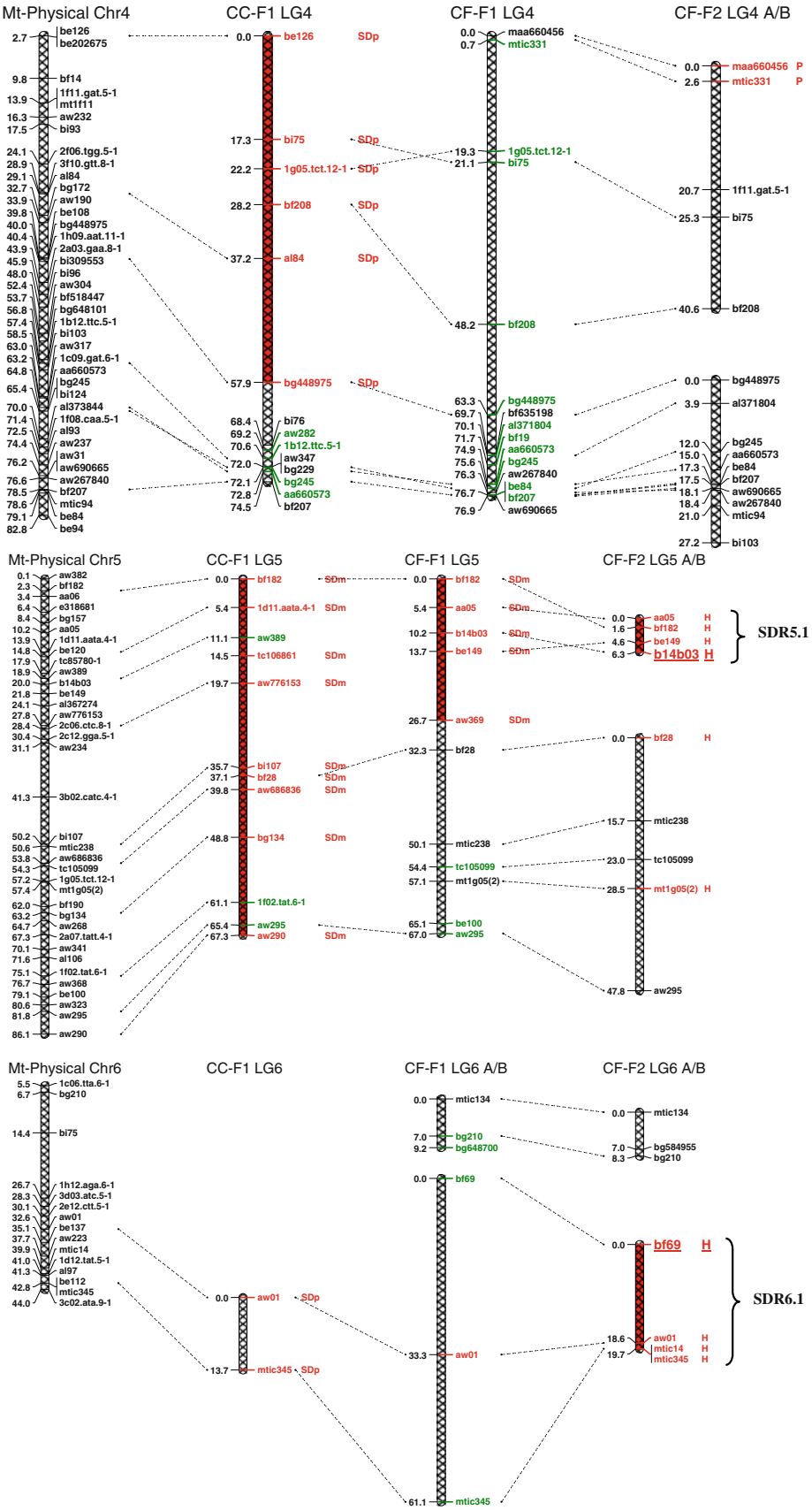
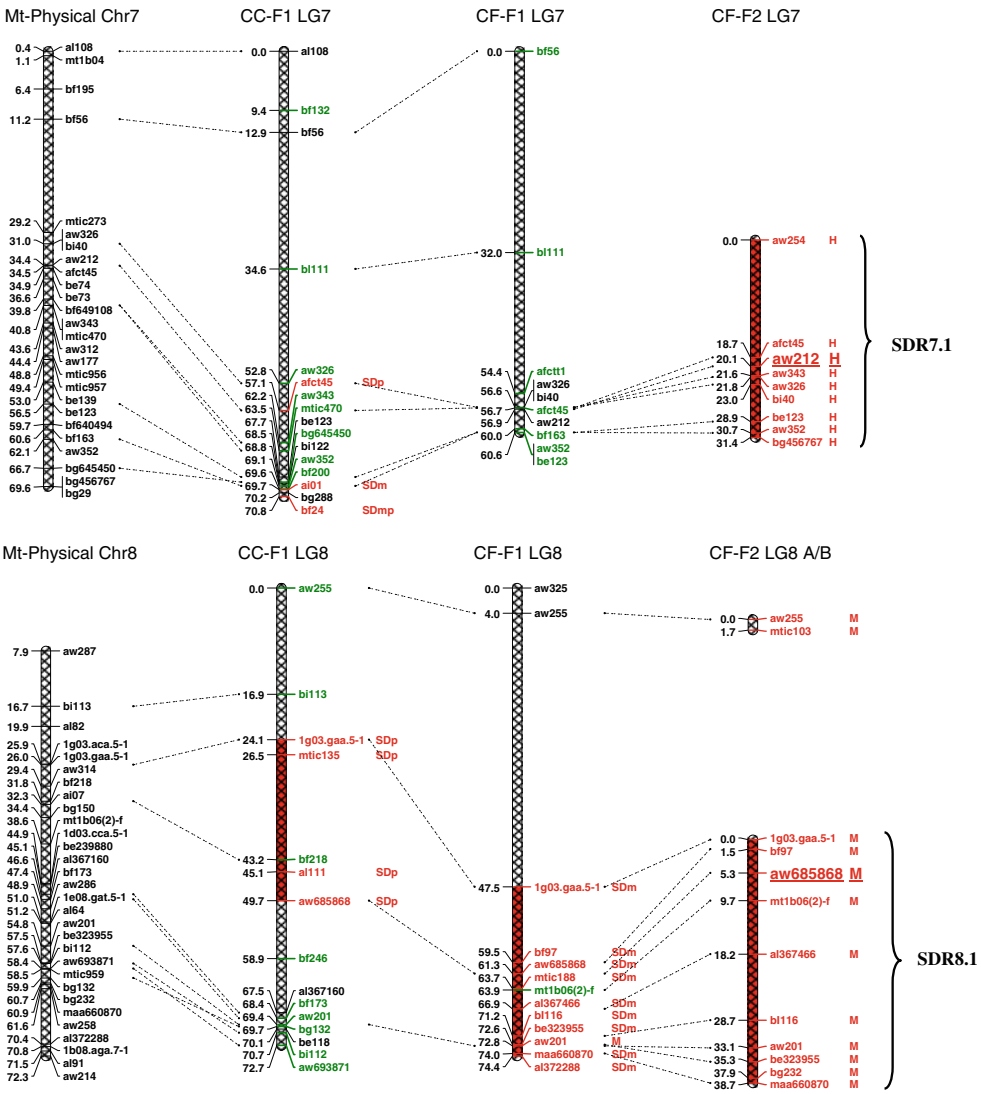


Fig. 1 continued



linkage group segments in Fig. 1); they were located on LG1, LG3, LG4, LG5, and LG8. Three distorted markers al369471, aw774443, and al375136 on LG1 showed allelic SD for maternal alleles (Fig. 1; Supplementary Table 2). The maternal parent was not polymorphic for a group of six non-distorted markers between al369471 and aw774443, which could explain the lack of SD for those markers (Supplementary Table 2). Therefore, we considered the region between marker al369471 and al375136 as a candidate SDR. Except for the SDR on LG4, all SDRs contained one or more non-fully informative markers, which could have masked distortion that was actually present (Fig. 1; Supplementary Table 2). The markers within SDRs on LG1 and LG5 showed allelic SD for the maternal parental alleles; and the markers within the other three SDRs showed allelic SD for the paternal parental alleles (Fig. 1; Supplementary Table 2).

CF-F$_1$ population

We identified three candidate SDRs in the CF-F$_1$ population that included 18 of the 22 markers with genotypic SD (red-filled linkage group segments in Fig. 1); these were located on LG1, LG5, and LG8. The markers within the SDR on LG1 showed allelic SD for the paternal alleles, and the markers within SDR on LG5 and LG8 showed allelic SD for the maternal alleles (Fig. 1; Supplementary Table 3).

CF-F$_2$ population

Because all markers in the CF-F$_2$ population were fully informative and because the F$_2$ population is more inbred than the F$_1$ populations, we identified more SD in this population than in the F$_1$ populations. Of the 60 markers with genotypic SD, 53 were placed into seven groups with three or

Table 3 Segregation distortion loci (SDLs) within segregation distortion regions (SDRs) in the CF-F_2 population, their degree of dominance (*h*) and the relative viability of the three marker genotypes (A_1A_1, A_1A_2, and A_2A_2)

Marker	SDR	*h*	$A_1A_1^a$ 1	A_1A_2 $1 - hs$	A_2A_2 $1 - s$
bg208	SDR1.1	0.31	1 (P)	0.78	0.29
aw685684	SDR1.2	−7.67	1 (P)	1.85	0.89
bg105	SDR2.1	−2.70	1 (P)	3.13	0.21
mtic237	SDR3.1	−1.53	1 (M)	1.84	0.45
b14b03	SDR5.1	−3.88	1 (M)	2.29	0.67
bf69	SDR6.1	−2.17	1 (P)	3.05	0.05
aw212	SDR7.1	−18.33	1 (M)	7.88	0.63
aw685868	SDR8.1	0.83	1 (M)	0.18	0.01

[a] A_1 the dominant allele with the higher observed frequency, *P* the paternal grandparental allele is the dominant allele, *M* the maternal grandparent allele is the dominant allele for a specific SDL

more consecutive distorted markers (Fig. 1). From the seven groups, we identified eight SDRs based on the skew directions of alleles and genotypes and the similarity of allelic and zygotic SD tests (Fig. 1; Supplementary Table 4). In addition to the eight SDRs we defined, two distorted markers on LG4A could be a potential SDR, and two markers on LG8A could be an additional SDR or part of SDR8.1, but we do not consider them as candidate SDRs in this study.

The markers in SDR1.2, SDR2.1, SDR3.1, SDR5.1, and SDR7.1 showed zygotic SD but not allelic SD, and all the markers exhibited excess heterozygosity (Fig. 1; Supplementary Table 4). The markers within SDR6.1 showed both allelic SD and zygotic SD (Supplementary Table 4). The degree of dominance (h value) among these putative SDLs varied from −1.53 to −18.33 (Table 3), which suggested over-dominant zygotic selection. Markers within SDR1.1 and SDR8.1 showed allelic SD but not zygotic SD (Supplementary Table 4). For the markers with allelic SD, the markers within SDR1.1 and 6.1 skewed toward the paternal grandparent allele; the markers within SDR8.1 skewed toward the maternal grandparent allele (Fig. 1; Supplementary Table 4).

Comparison among the populations based on common markers indicated that all three candidate SDRs identified on the CF-F_1 linkage map were also identified in the CF-F_2 linkage map (Fig. 1). Because the two populations were derived from the same parents, the SDRs identified in both populations are likely caused by the same underlying SDLs. The large SDR on LG1 identified on the CF-F_1 linkage map probably contains two SDLs based on comparison with the two SDRs in the CF-F_2 population. The CC-F_1 population appears to have an SDR on LG4 that is not present in the other populations (although two distorted loci are located in the same region in CF-F_2) and the possibility of two separate SDL in the large SDRs on LG3

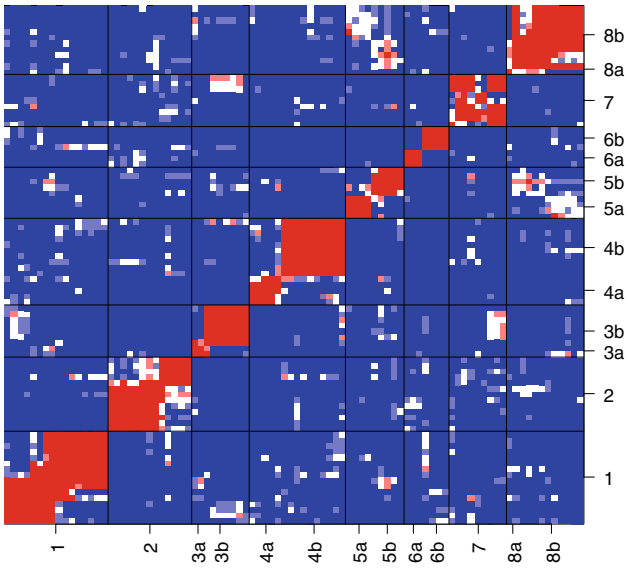


Fig. 2 Genomewide map of pairwise genotype associations, measured by *P* value. The 13 linkage groups in the CF_F_2 linkage map we identified are shown along the *x*- and *y*-axes. The *P* value for the significance of the test for pairwise interaction for each pair of loci is color coded as follows: *red* $P < 0.001$, *light red* $P < 0.01$, *white* $P < 0.05$, *light blue* $P < 0.1$, and *blue* $P \leq 1$ (color figure online)

and LG5. Otherwise, the CC-F_1 SDRs are also identified in the CF-F_2 population.

Epistatic interaction contributed to segregation distortion

Epistatic interactions contributing to viability were identified between two pairs of chromosomal regions in the CF-F_2 population ($P < 0.01$; Fig. 2). The first pair of regions included five markers on LG3 (between 19.3 and 23.1 cM) and three markers on LG7 (between 28.9 and 31.4 cM). All were distorted with an excess of heterozygotes and deficit of paternal grandparent homozygotes. Interestingly, the paternal grandparent homozygous genotype for the five markers within SDR3.1 was only observed when the three markers on SDR7.1 were heterozygous, which suggested a possible deleterious/lethal epistatic interaction. The other pair of regions included three markers on LG5B (between 15.7 and 28.5 cM) and two markers LG8B (between 0 and 1.5 cM). No paternal grandparent homozygotes for the two markers on LG8B were observed.

Discussion

Genetic linkage maps

A high level of synteny has been found between genetic maps in alfalfa and *M. truncatula* (Choi et al. 2004). Our

results confirm the findings on a broad scale although the relatively long interval between markers on our maps (~5 cM on average) does not enable us to detect small rearrangements or deletions between the species. A reciprocal translocation has been identified between chromosome 4 and 8 of *M. truncatula* reference accession A17 (Kamphuis et al. 2007), but Build 3.0 of the *M. truncatula* genome sequence, which we used for our comparisons, did not identify that translocation, and hence our linkage groups had similar orders as the *M. truncatula* chromosomes 4 and 8. Further investigations with more dense genetic maps and physical sequences of alfalfa are needed to fully understand the genome relationships between these *Medicago* species.

We also noted similar marker orders between the two F_1 genetic linkage maps, one derived from hybridization of two genetically distinct taxa (*caerulea* and *falcata*) (Sakiroglu et al. 2010) and one an intra-subspecies cross of *caerulea* genotypes. Taxa that are genetically distant often produce hybrids that have abnormal meiosis and/or large chromosome rearrangements, leading to substantial loss of viability and/or fertility (Lynch and Walsh 1998). However, normal meiosis (Sprague 1959) and high fertility (Li and Brummer 2009) have been observed for inter-subspecies crosses between *M. sativa* subsp. *caerulea* or its tetraploid counterpart subsp. *sativa* and diploid or tetraploid *M. sativa* subsp. *falcata*. Taken together, these results suggest that the genomes of the two alfalfa taxa have few if any differences in organization. We did identify some minor differences in marker order or spacing across small regions of linkage groups, but these are most likely caused by map distance estimates that have high standard errors due to the relatively small population sizes, resulting in few informative recombinants among tightly linked markers.

Segregation distortion in diploid alfalfa

Why fewer markers with SD were found in F_1 populations

Compared to the F_2 population, the two F_1 populations had only half as many markers with genotypic SD. One possible explanation for the disparity is that some markers were not fully informative in the F_1 population, and hence, SD may have occurred but could not be detected at these loci (marked in green in Fig. 1). In some instances, fully informative markers linked to these partially informative markers showed genotypic SD in the F_1 population and allelic SD for one parent. Thus, these partially informative markers likely mask SD. The second explanation for the disparity is inbreeding depression caused by recessive alleles that are not exposed as homozygotes in an F_1 population but that could be observed in the F_2 population. For example, the distorted markers in SDR2.1, SDR3.1, and SDR7.1 on the CF-F_2 linkage map were not distorted in CF-F_1, probably for this reason (Fig. 1; Supplementary Table 4).

The genetic mechanism of segregation distortion

Similar to previous studies on diploid, partially inbred alfalfa mapping populations (Brummer et al. 1993; Kaló et al. 2000), we found a large percentage of markers (68%) showing SD in our F_2 population. Distorted markers were mostly clustered, and markers within each cluster had a same skew direction, which indicated that the distorted markers were caused by biological factors, such as an SDL, rather than by experimental error (mis-scoring, etc.). Across the three populations, at least one region on every chromosome had an SDL affecting gamete or zygote fitness. The observed distorted genotypic ratios could result from selection at one or more of the following life stages: (1) gametophyte development, (2) fertilization, (3) seed development, (4) seed germination, and/or (5) plant growth prior to tissue collection for DNA extraction (Fishman et al. 2001; Zamir and Tadmor 1986). In this study, 161 of 231 seeds germinated for the CF-F_2 population. Of 161 seedlings, 41 seedlings were very weak and died before transplanting to larger pots. DNA was successfully isolated from 152 seedlings, including 32 of the 41 seedlings that died early after germination. We did not find obvious differences in the genotypic segregation ratios of the 120 surviving plants from that of the 32 plants that died early, except for the distorted markers within SDR3.1 (Fig. 1; Supplementary Table 4) (Some distorted markers among the 120 plant sample did not show significant SD in the 32 plant samples probably due to low testing power because of the small sample size in the latter group.). This indicated that selection at most SDL probably happened before seedling growth began. However, for those markers in SDR3.1 on LG3B, about 80% of the 32 plants that died early were homozygous for the maternal grandparental allele, which indicated that the SDL within SDR3.1 could be related to early seedling growth.

The relatively low incidence of allelic SD compared to zygotic SD in the F_2 mapping population suggested that zygotic selection is more common than gametic selection. Therefore, the observed distortion of markers in this study is most likely attributed to selection among zygotes between fertilization and seed germination. The markers within SDR1.1 and SDR8.1 showed allelic SD but not zygotic SD, which suggested gametic selection (Lorieux et al. 1995b). However, the degree of dominance of the putative SDLs on SDR1.1 and SDR8.1 were 0.31 and 0.83, respectively (Table 3). These suggest that partially dominant zygotic selection for SDR1.1 and partially recessive zygotic selection for SDR8.1 could also explain the

observed distorted genotypic ratio and in turn cause the estimated distorted allelic ratio. Similarly, genotypic SD of markers identified in F_1 populations could be caused by zygotic selection although none of the distorted markers showed zygotic SD.

Except for SDR1.1 and SDR8.1, distorted markers in all other SDRs in CF-F_2 showed heterozygote excess. SD with heterozygote excess is commonly found in diploid alfalfa. Thirty of 51 distorted markers skewed toward excess heterozygotes in a diploid *M. sativa* segregating population (Brummer et al. 1993). Distorted marker regions on all eight chromosomes except for chromosome 4 had heterozygote excess in a *M. sativa* subsp. *falcata* × *M. sativa* subsp. *caerulea* population (Kaló et al. 2000). Heterozygote excess has also been observed in other species (Faris et al. 1998; Fishman et al. 2001; Kesseli et al. 1994; Lu et al. 2002; Prince et al. 1993). However, none of those species showed distortions toward heterozygotes as extreme as that we observed for some loci in alfalfa [e.g., markers on LG7 in the CF-F_2 population (Supplementary table 4), where over 95% of some loci were heterozygotes].

The genetic cause of the higher fitness of heterozygotes has been long debated in the heterosis literature (Lynch and Walsh 1998). The most obvious cause is overdominance, or heterozygote advantage; in our case, over-dominant zygotic selection appears predominant in alfalfa. However, classical and molecular genetic studies on heterosis and inbreeding depression suggest that what appears to be overdominance is actually due to complementary linked partially to completely dominant alleles, resulting in pseudo-overdominance (Charlesworth and Charlesworth 1999; Charlesworth and Willis 2009). Complementary gene action appears to be the major cause of biomass heterosis in alfalfa (Bingham et al. 1994) and would seem to be a likely explanation for the heterozygote excess SD in alfalfa. In conifers, heterozygote excess was attributed to pseudo-overdominance of multiple lethal factors although gametic selection between maternal and paternal gametes was also possible (Williams et al. 2001, 2003). Heterozygote excess was not only observed in inter-subspecies F_2 populations in this study and a previous study (Kaló et al. 2000) but also in an intra-subspecies F_2 population (Brummer et al. 1993). Thus, the observed heterozygote excess likely is not due to specific allelic interactions between the two subspecies.

Lethal epistatic interactions between unlinked loci (Dobzhansky-Muller incompatibilities) could contribute to SD and can be identified by testing pair-wise associations of marker genotypes (McDaniel et al. 2007; Payseur and Hoekstra 2005). We found two pairs of interacting chromosomal regions in CF-F_2 population that might contribute to SD. In addition to the two pairs of regions, isolated pairs of markers from different linkage groups showed non-independent distribution (Fig. 2). Potential additional regions involved in epistatic interactions may not have been detected because of low map saturation in this study.

Effect of SD on the estimation of genetic distances and orders among markers

The estimation of genetic distance between distorted markers could be biased (Bailey 1949; Lorieux et al. 1995a, b). The distance could be underestimated if the distorted markers skewed to a same direction and overestimated if skewed to different directions (Lorieux et al. 1995a; Zhu et al. 2007a, b). The bias level depends on the distortion level and real genetic distance between the markers (Lorieux et al. 1995a). The standard error of the estimated genetic distance between a pair of markers is also clearly increased when markers show SD (Lorieux et al. 1995a, b). In this study, most distorted markers were clustered with a same skew direction, and thus, the estimated genetic distances among the markers within each SDR were likely underestimated. Epistatic interactions between two SDLs on the same linkage group could bias the estimation of genetic distance (Zhu and Zhang 2007). Epistatic effects could either increase or decrease the actual genetic distance, depending on the mode of epistasis (e.g., additive by additive, dominance by dominance, etc.). To our knowledge, no model describing the effects of epistasis between SDL on different linkage groups that affect genetic distance has been published; hence, we do not have a clear idea of what effect the epistatic interaction between SDLs in our mapping population would have on our map.

The precise ordering of markers depends on precise estimation of pair-wise genetic distances (Lorieux et al. 1995a). Some differences in marker orders were observed among the three genetic linkage maps and the *M. truncatula* physical map, which might be caused by the biased estimations of distances between distorted markers. However, the different marker orders were only observed in small regions, which is more likely caused by the low confidence level of our estimates due to a limited sample size and/or rare informative recombinants among tightly linked markers than by SD. The increased standard errors of genetic distance estimates between distorted markers exacerbate the problem.

Several approaches have been investigated to adjust genetic distance between distorted markers considering SDL effects (Mitchell-Olds 1995; Vogl and Xu 2000; Zhu et al. 2007a), none of which has been integrated into a standard mapping program, especially one for a full-sib family mapping population. An alternative approach to more accurately locate SDL in the alfalfa genome is to construct genetic maps in advanced generations to break

linkages between SDL and surrounding markers. These experiments are currently under way.

Implication for QTL mapping and breeding in alfalfa

The F_1 populations had a lower level of SD than the F_2 population, but use of these populations in genetic mapping experiments may miss potential QTL because fewer markers were fully informative. The SDR on LG1 of the CC-F_1 map was based on three fully informative markers (Supplementary Table 2); the other 11 markers did not segregate in the female parent and did not exhibit SD. A favorable QTL allele from the female parent located in this region may not be identified efficiently because its surrounding markers are not polymorphic for the female parent. In the present study, only 48 out of 105 SSR markers were fully informative for the CC-F_1 population and 48 out of 108 for CF-F_1 population. Partially inbred F_2 populations have high levels of SD, which bias the estimation of genetic distance and possibly also marker order, also affecting QTL mapping. However, unlike F_1 populations, all markers in an F_2 population are fully informative.

Theoretically, the entire chromosome could exhibit distortion caused by just a few SDLs (Xu 2008). Potential SDLs were identified on all of the eight chromosomes in alfalfa from this study. This indicates that segregation of a substantial number of genes could be affected, including genes controlling important agronomic traits. When SDL and QTL alleles for agronomic traits are linked in repulsion, the selection or introgression of agronomic traits could be negatively affected. Therefore, locating SDL and understanding their genetic mechanisms can facilitate selection and breeding for the agronomic traits. A denser genetic linkage map and populations resulting from additional recombination will be helpful to identify precise SDL locations and investigate their mechanisms of action. Finally, the nature of SDL is not clear. Certainly some of them may be lethal factors that prevent seed development, but given their effects on fitness, they could also be loci directly involved in biomass (fitness) and/or heterosis for biomass. Future mapping will help elucidate these questions.

Acknowledgment This research was funded by the USDA-DOE Plant Feedstock Genomics for Bioenergy program, award # 2009-65504-05809 to ECB.

References

Bailey NTJ (1949) The estimation of linkage with differential viability, II and III. Heredity 3:220–228

Bingham ET, Groose RW, Woodfield DR, Kidwell KK (1994) Complementary gene interaction in alfalfa is greater in autotetraploids than diploid. Crop Sci 34:823–829

Brouwer DJ, Osborn TC (1999) A molecular marker linkage map of tetraploid alfalfa (*Medicago sativa* L.). Theor Appl Genet 99:1194–1200

Brummer EC, Bouton JH, Kochert G (1993) Development of an RFLP map in diploid alfalfa. Theor Appl Genet 86:329–332

Charlesworth B, Charlesworth D (1999) The genetic basis of inbreeding depression. Genet Res 74:329–340

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nat Rev 10:783–796

Chetelat RT, Meglic V, Cisneros P (2000) A genetic map of tomato based on BC(1) *Lycopersicon esculentum* × *Solanum lycopersicoides* reveals overall synteny but suppressed recombination between these homeologous genomes. Genetics 154:857–867

Choi HK, Kim D, Uhm T, Limpens E, Lim H, Mun JH, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T, Kiss GB, Cook DR (2004) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. Genetics 166:1463–1502

Doyle JJ, Doyle JI (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Echt CS, Kidwell KK, Knapp SJ, Osborn TC, Mccoy TJ (1994) Linkage mapping in diploid alfalfa (*Medicago sativa*). Genome 37:61–71

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Longman, Essex

Faris JD, Laddomada B, Gill BS (1998) Molecular mapping of segregation distortion loci in *Aegilops tauschii*. Genetics 149:319–327

Fishman L, Kelly AJ, Morgan E, Willis JH (2001) A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. Genetics 159:1701–1716

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137:1121–1137

Hall MC, Willis JH (2005) Transmission ratio distortion in intraspecific hybrids of *Mimulus guttatus*: implications for genomic divergence. Genetics 170:375–386

Julier B, Flajoulot S, Barre P, Cardinet G, Santoni S, Huguet T, Huyghe C (2003) Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biol 3:9

Kaló P, Endre G, Zimányi L, Csanádi G, Kiss GB (2000) Construction of an improved linkage map of diploid alfalfa (*Medicago sativa*). Theor Appl Genet 100:641–657

Kamphuis LG, Williams AH, D'Souza NK, Pfaff T, Ellwood SR, Groves EJ, Singh KB, Oliver RP, Lichtenzveig J (2007) The *Medicago truncatula* reference accession A17 has an aberrant chromosomal configuration. New Phytol 174:299–303

Kesseli RV, Paran I, Michelmore RW (1994) Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. Genetics 136:1435–1446

Kiss GB, Csanadi G, Kalman K, Kalo P, Okresz L (1993) Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. Mol Gen Genet 238:129–137

Li XH, Brummer EC (2009) Inbreeding depression for fertility and biomass in advanced generations of inter- and intrasub specific hybrids of tetraploid alfalfa. Crop Sci 49:13–19

Lorieux M, Perrier X, Goffinet B, Lanaud C, DGd León (1995a) Maximum-likelihood models for mapping genetic markers showing segregation distortion. 1. Backcross populations. Theor Appl Genet 90:73–80

Lorieux M, Perrier X, Goffinet B, Lanaud C, DGd León (1995b) Maximum-likelihood models for mapping genetic markers showing segregation distortion. 2. F2 populations. Theor Appl Genet 90:81–89

Lu H, Romero-Severson J, Bernardo R (2002) Chromosomal regions associated with segregation distortion in maize. Theor Appl Genet 105:622–628

Luo L, Zhang YM, Xu S (2005) A quantitative genetics model for viability selection. Heredity 94:347–355

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA

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

McDaniel SF, Willis JH, Shaw AJ (2007) A linkage map reveals a complex basis for segregation distortion in an interpopulation cross in the moss *Ceratodon purpureus*. Genetics 176:2489–2500

Mitchell-Olds T (1995) Interval mapping of viability loci causing heterosis in *Arabidopsis*. Genetics 140:1105–1109

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726

Payseur BA, Hoekstra HE (2005) Signatures of reproductive isolation in patterns of single nucleotide diversity across inbred strains of mice. Genetics 171:1905–1916

Prince JP, Pochard E, Tanksley SD (1993) Construction of a molecular linkage map of pepper and a comparison of synteny with tomato. Genome 36:404–417

Rhoades MM (1942) Preferential segregation in maize. Genetics 27:395–407

Riaz S, Tenscher AC, Rubin J, Graziani R, Pao SS, Walker MA (2008) Fine-scale genetic mapping of two Pierce's disease resistance loci and a major segregation distortion region on chromosome 14 of grape. Theor Appl Genet 117:671–681

Robins JG, Luth D, Campbell IA, Bauchan GR, He CL, Viands DR, Hansen JL, Brummer EC (2007) Genetic mapping of biomass production in tetraploid alfalfa. Crop Sci 47:1–10

Sakiroglu M, Doyle JJ, Charles Brummer E (2010) Inferring population structure and genetic diversity of broad range of wild diploid alfalfa (*Medicago sativa* L.) accessions using SSR markers. Theor Appl Genet 121:403–415

Sandler L, Hiraizumi Y, Sandler I (1959) Meiotic drive in natural populations of *Drosophila melanogaster*. I. the cytogenetic basis of segregation-distortion. Genetics 44:233–250

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Sledge MK, Ray IM, Jiang G (2005) An expressed sequence tag SSR map of tetraploid alfalfa (*Medicago sativa* L.). Theor Appl Genet 111:980–992

Sprague EW (1959) Cytological and fertility relationships of *Medicago sativa*, *Medicago falcata*, and *Medicago gaetula*. Agron J 51:249–252

Tavoletti S, Veronesi F, Osborn TC (1996) RFLP linkage map of an alfalfa meiotic mutant based on an F-1 population. J Hered 87:167–170

Vogl C, Xu S (2000) Multipoint mapping of viability and segregation distorting loci using molecular markers. Genetics 155:1439–1447

Wang C, Zhu C, Zhai H, Wan J (2005) Mapping segregation distortion loci and quantitative trait loci for spikelet sterility in rice (*Oryza sativa* L.). Genet Res 86:97–106

Williams CG, Zhou Y, Hall SE (2001) A chromosomal region promoting outcrossing in a conifer. Genetics 159:1283–1289

Williams CG, Auckland LD, Reynolds MM, Leach KA (2003) Overdominant lethals as part of the conifer embryo lethal system. Heredity 91:584–592

Xu S (2008) Quantitative trait locus mapping can benefit from segregation distortion. Genetics 180:2201–2208

Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid and recombinant inbred populations in rice (*Oryza sativa* L.). Mol Gen Genet 253:535–545

Yin TM, DiFazio SP, Gunter LE, Riemenschneider D, Tuskan GA (2004) Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. Theor Appl Genet 109:451–463

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot Gaz 147:355–358

Zhu CS, Zhang YM (2007) An EM algorithm for mapping segregation distortion loci. BMC Genet 8:82

Zhu C, Wang C, Zhang YM (2007a) Modeling segregation distortion for viability selection. I. Reconstruction of linkage maps with distorted markers. Theor Appl Genet 114:295–305

Zhu CS, Wang FH, Wang JF, Li GJ, Zhang HS, Zhang YM (2007b) Reconstruction of linkage maps in the distorted segregation populations of backcross, doubled haploid and recombinant inbred lines. Chin Sci Bull 52:1648–1653