

Dissecting repulsion linkage in the dwarfing gene *Dw3* region for sorghum plant height provides insights into heterosis

Xin Li^a, Xianran Li^a, Eyal Fridman^b, Tesfaye T. Tesso^c, and Jianming Yu^{a,1}

^aDepartment of Agronomy, Iowa State University, Ames, IA 50011; ^bInstitute of Plant Science, Agricultural Research Organization, The Volcani Center, Bet Dagan 50250, Israel; and ^cDepartment of Agronomy, Kansas State University, Manhattan, KS 66506

Edited by Ronald L. Phillips, University of Minnesota, St. Paul, MN, and approved August 21, 2015 (received for review May 31, 2015)

Heterosis is a main contributor to yield increase in many crop species. Different mechanisms have been proposed for heterosis: dominance, overdominance, epistasis, epigenetics, and protein metabolite changes. However, only limited examples of molecular dissection and validation of these mechanisms are available. Here, we present an example of discovery and validation of heterosis generated by a combination of repulsion linkage and dominance. Using a recombinant inbred line population, a separate quantitative trait locus (QTL) for plant height (*qHT7.1*) was identified near the genomic region harboring the known auxin transporter *Dw3* gene. With two loci having repulsion linkage between two inbreds, heterosis in the hybrid can appear as a single locus with an overdominance mode of inheritance (i.e., pseudo-overdominance). Individually, alleles conferring taller plant height exhibited complete dominance over alleles conferring shorter height. Detailed analyses of different height components demonstrated that *qHT7.1* affects both the upper and lower parts of the plant, whereas *Dw3* affects only the part below the flag leaf. Computer simulations show that repulsion linkage could influence QTL detection and estimation of effect in segregating populations. Guided by findings in linkage mapping, a genome-wide association study of plant height with a sorghum diversity panel pinpointed genomic regions underlying the trait variation, including *Dw1*, *Dw2*, *Dw3*, *Dw4*, and *qHT7.1*. Multilocus mixed model analysis confirmed the advantage of complex trait dissection using an integrated approach. Besides identifying a specific genetic example of heterosis, our research indicated that integrated molecular dissection of complex traits in different population types can enable plant breeders to fine tune the breeding process for crop production.

heterosis | repulsion linkage | pseudo-overdominance | plant height | genome-wide association studies

Heterosis, the better performance of a hybrid than either of its parents, can be traced back to Charles Darwin and has been of great interest to biologists. Hybrids derived from diverse inbred lines generally show higher growth rate and yield and better resistance to biotic and abiotic stresses (1). Approximately 15% of the yield increase in corn, starting in the early 1930s, comes from the introduction of hybrid corn (2). Hybrid rice has a 20–30% yield advantage over inbred rice cultivars and contributes significantly to food security (3).

Four hypotheses have been proposed for heterosis: dominance, overdominance, pseudo-overdominance, and epistasis. Advances in genomics shed new light on other theories for heterosis, including protein metabolism and epigenetic changes in hybrids (4–6). The dominance hypothesis assumes the better performance of F_1 results from accumulation of more dominant alleles in the F_1 genome than in either of its parents. These dominant alleles can mask the deleterious effects of recessive alleles. Conversely, inbreeding depression, the reduced fitness of progeny from crossing between related individuals, is caused predominately by exposing deleterious recessive alleles (7). The overdominance hypothesis argues heterozygous genotypes are better than either of the two homozygous genotype parents. Two specific examples were found in sickle-cell anemia (8) and tomato (9), where the advantage of heterozygous

genotype at a single locus was demonstrated. The pseudo-overdominance hypothesis was first proposed in 1917 as a possible explanation for heterosis (10). It assumes that many factors control a specific trait and that it is unlikely that one individual could have all the favorable alleles in its genome. If alleles with opposite effects at multiple loci are linked in repulsion phase in two parental inbreds, overdominance may be mistaken for a single locus unless detailed dissection is carried out in progenies. The epistasis hypothesis states that heterosis is caused by de novo interactions between different loci.

Research evidence from multiple crops supported all four hypotheses: dominance (11, 12), overdominance (9, 13, 14), pseudo-overdominance (15, 16), and epistasis (17–22). These results suggest that the genetic basis for heterosis can be trait and species dependent and may be explained by not a single unified theory but multiple theories together.

Plant height generally shows significant heterosis, making it an ideal trait to study and illustrate heterosis (6). Sorghum plant height shows 6–72% high-parent heterosis (23–28). In maize, high-parent heterosis of plant height can reach 32–120% (13, 29, 30). The high-parent heterosis of plant height can reach 10–20% in elite rice hybrids (12, 31, 32) and 22% in wheat hybrids (33).

In sorghum, previous research has identified four major loci controlling plant height, *Dw1–Dw4* (34). These four dwarfing loci were found to be unlinked and were phenotypically selected during conversion of sorghum into temperate-region production with machine harvesting. At each locus, the tall type allele shows complete dominance over the other allele. Among the four dwarfing loci, only *Dw3* has been molecularly characterized (Sb07g023730), encoding a P-glycoprotein auxin transporter (35). *Dw1* was mapped to

Significance

Heterosis, the better performance of hybrids over their parents, holds great economic and biological significance. Different theories have been proposed, but specific examples with detailed dissection are limited. If close linkage of alleles with opposite effects exists, the superiority of hybrids over inbreds is observed and may appear as overdominance at a single locus. We present a case of pseudo-overdominance generated by repulsion linkage between two quantitative trait loci in sorghum plant height. A combination of approaches were used: linkage mapping under a defined genetic background, genome-wide association study with a diversity panel, computer simulation, and designed crosses with selected genetic stocks. Our findings provide insights into heterosis and a tool box for plant breeders to develop ideal cultivars.

Author contributions: Xin Li and J.Y. designed research; Xin Li performed research; Xianran Li, E.F., and T.T.T. contributed new reagents/analytic tools; Xin Li analyzed data; and Xin Li, Xianran Li, and J.Y. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. Email: jmyu@iastate.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1509229112/-DCSupplemental.

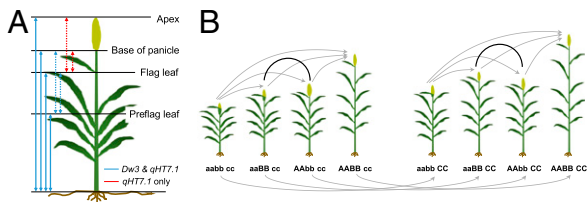


Fig. 1. Plant height measurements and crossing design. (A) Total plant height, height at the base of the panicle, height at the flag leaf, and preflag leaf height were measured directly (solid arrows). The flag leaf-to-apex interval, flag leaf-to-base of panicle interval, preflag leaf-to-flag leaf interval, and preflag leaf-to-base of panicle interval were calculated (dotted arrows). *qHT7.1* affects all eight plant height components, whereas the effect of *Dw3* is detected for six components. (B) Crossing design for hybrid development. Eight inbreds were selected by their genotypes at *qHT7.1* (AA or aa), *Dw3* (BB or bb), and *Dw1* (CC or cc) to make the crosses. The four on the left have the *dw1dw1* background, and the four on the right have the *Dw1Dw1* background. Two specific crosses (black lines) are expected to generate F₁s showing heterosis in plant height. For all other crosses (gray lines), the F₁ is expected to have a height similar to that of the second inbred parent indicated by the arrow.

chromosome 9 at ~57 Mb (36), and *Dw2* was mapped to chromosome 6 at ~42 Mb (37, 38).

In this study, we present an example of heterosis as a result of repulsion linkage for sorghum plant height through genetic mapping and designed cross (Fig. 1 A and B). We started by identifying an additional plant height quantitative trait locus (QTL) in a recombinant inbred line (RIL) population located 3 Mb away from the known *Dw3* gene on chromosome 7 (*qHT7.1*). A genome-wide association study (GWAS) of plant height using the sorghum association panel (39) further confirmed the discovery. Computer simulation was conducted to evaluate the influence of repulsion linkage on QTL detection in an F₂ population. The gene action at each QTL and heterosis were demonstrated by observing F₁ hybrids generated from crossing selected RILs with different combinations of QTL alleles. Finally, a close examination of allelic combinations of plant height QTL in diverse accessions revealed the potential of breeding with designed approaches.

Results

Linkage Mapping of Plant Height. An RIL population was developed from Tx430 and P898012. Plant height in this population was consistent across locations and years (Fig. S1), with the entry mean-based heritability of 0.96. P898012 is ~25 cm taller than Tx430, but plant height of RILs varied widely (74 ~236 cm), following a normal distribution (Fig. S2).

The genotyping-by-sequencing (GBS) procedure generated 8,960 SNP markers that are segregating in this population. A set of 1,756 selected markers was used to construct the linkage map. The total map length is 1,916 cM spanning 1,452 unique bins, leading to a mean marker bin density of 1.3 cM (Fig. S3 and Table S1).

Composite interval mapping with the genetic map defined by 1,756 SNPs detected three QTL on chromosomes 7 and 9, and single-marker scan with physical mapping of all 8,960 SNP markers supported these findings (Fig. S4). Two QTL on chromosome 7 are ~29 cM or 3 Mb apart and have opposite effects (Fig. 2 A and B and Table 1). The first QTL, located at ~91 cM, was designated *qHT7.1*, and the second QTL located at ~120 cM was mapped to the position of the known *Dw3* gene (35). The chromosome 9 QTL corresponds to the *Dw1* locus (36, 37). Next, we confirmed that the second QTL on chromosome 7 in the current study is *Dw3*, a P-glycoprotein responsible for polar transport of auxin in the stalk (35). The Tx430 allele has an 882-bp insertion in the fifth exon (Fig. S5), resulting in a nonfunctional protein.

Mapping of Plant Height Components. Both QTL on chromosome 7 affect four direct measurements: total plant height, base of panicle height, flag leaf height, and preflag leaf height. Both QTL affect the interval between preflag leaf and flag leaf. However, the *Dw3*

gene does not affect variation in the upper parts of the stem, namely the interval between flag leaf and the base of the panicle (peduncle length) or the interval between flag leaf and the top of the panicle (Fig. 2C and Fig. S4); it has a small effect on the interval between preflag leaf and base of the panicle (Fig. S4). *qHT7.1* has the largest phenotypic contribution to six of the eight plant height components and has the second largest phenotypic contribution to the remaining two traits (Table 1 and Fig. S4).

GWAS of Plant Height Across a Diversity Panel. Following the finding of two linked QTL with opposite effects on chromosome 7, we further explored the possibility of detecting this locus, *qHT7.1*, in the sorghum association panel (39), a much broader genetic background. The first genome scan with the linear mixed model detected the two known plant height QTL (*Dw1* on chromosome 9 and *Dw2* on chromosome 6) and another candidate QTL for *Dw4* on chromosome 4 (Fig. 3A). At this point, neither *Dw3* nor *qHT7.1* was significant. However, after *Dw1* and *Dw2* were incorporated as covariates in the second genome scan, both *Dw3* and *qHT7.1* were significantly associated with plant height. The association signal for *Dw3* is located at 58,390,034 bp on chromosome 7 ($P = 1.2 \times 10^{-9}$), and the most significant signal for *qHT7.1* is located at 55,229,509 bp ($P = 4.8 \times 10^{-8}$) (Fig. 3B). These two SNPs are right within the corresponding QTL peak regions of the linkage mapping, and they also are near the most significant SNPs found in the single-marker scan (Fig. S4). Based on a search of the literature, this QTL had not been detected in previous association mapping studies (36, 37, 40, 41), possibly because of the need for targeted analysis.

Encouraged by these findings, we conducted the third genome scan with the multilocus mixed model (MLMM) method (42), which sequentially adds significant QTL as covariates in the model before testing all other SNPs. This third scan identified the same SNP at 55,229,509 bp as the fourth most significant marker, following *Dw1* on chromosome 9, *Dw2* on chromosome 6, and *Dw3* on chromosome 7 (Fig. 3C and Table 2). Although 80% of the accessions in this diverse panel have the allele with the positive

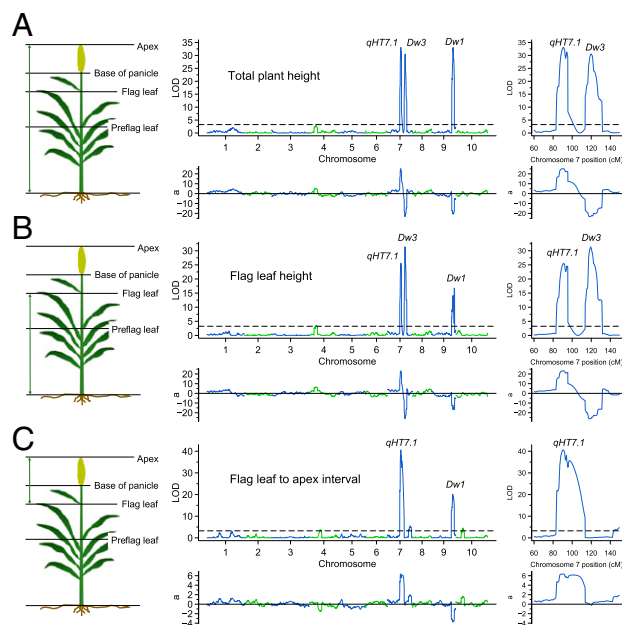


Fig. 2. Linkage mapping of plant height. (A–C, Left) Diagrams defining measurement points. (Center and Right) Results from composite interval mapping. (Center, Upper) The logarithm of the odds (LOD) score profile with the permutation threshold indicated by the horizontal line. (Center, Lower) The additive effect (a) with the Tx430 allele as the reference. (Right, Upper) The LOD score profile for enlarged chromosome 7 region. (Right, Lower) The additive effect for enlarged chromosome 7 region.

Table 1. Three QTL identified through linkage mapping for three plant height traits

Trait	<i>qHT7.1</i> , Chr7: ~91 cM	<i>Dw3</i> , Chr7: ~120 cM	<i>Dw1</i> , Chr9: ~146 cM
Total plant height	25.1 (34.7%)	-22.2 (27.1%)	-20.4 (30.2%)
Flag leaf height	24.1 (28.8%)	-26.1 (31.5%)	-20.0 (24.9%)
Flag leaf-to-apex interval	6.3 (45.4%)	Not significant	-3.8 (16.8%)

The additive effect of each QTL is shown in centimeters for the Tx430 allele. Numbers in parentheses are the phenotypic variation explained by each QTL.

effect at the *qHT7.1* locus, 57% of accessions have the haplotypes with opposite effects at the *qHT7.1* and *Dw3* loci (Table S2).

Gene Actions of Each Plant Height QTL and Heterosis. Because only homozygous inbreds were involved in linkage mapping and GWAS, the dominance effect of each QTL cannot be estimated. To examine dominance effects and heterosis, we decided to make hybrid crosses between inbreds. This crossing scheme can confirm our expectation that heterosis in plant height should be observed in hybrids derived from a pair of inbreds in which the two QTL on chromosome 7 are in repulsion linkage (Fig. 1B). In one cross involving repulsion linkage between *qHT7.1* and *Dw3* under the *dw1dw1* background (aaBB cc × AA bb cc), the F₁ hybrid was 44 cm taller than the tall parent, equivalent to 35% high-parent heterosis (Fig. 4A and Table S3). In another cross involving repulsion linkage between *qHT7.1* and *Dw3* under the *Dw1Dw1* background (aaBB CC × AA bb CC), the F₁ hybrid was 30 cm taller than the tall parent, equivalent to 19% high-parent heterosis (Fig. 4B and Table S3). Other plant height components, except for the flag leaf-to-apex interval, showed 18–50% high-parent heterosis. The flag leaf-to-apex interval of both hybrids is similar to that of the tall parents, as expected, because only *qHT7.1* was detected for the upper part of the plant (Fig. 2), and the tall parents are homozygous dominant at *qHT7.1*. Together, these findings provide evidence that the observed heterosis is caused by the repulsion linkage on chromosome 7 rather than other uncharacterized loci. With crosses involving different alleles of *qHT7.1* under different backgrounds, the tall allele shows complete dominance over the short allele (Fig. 4C and D and Table S4). Similarly, *Dw1* and *Dw3* showed nearly complete dominance on plant height (Table S4), agreeing with previous research (34).

Simulation Appraisal of the Challenges in Dissecting QTL with Repulsion Linkage. To understand the influence of repulsion linkage on mapping power and effect estimation, we conducted a series of simulations with an F₂ population. We intended to answer two questions: (i) Can the two QTL be clearly separated instead of detecting one QTL spanning the whole region? (ii) What would the ratio of dominant to additive effect (d/a) of the detected QTL be? The default value of d/a was one because complete dominance was simulated at each locus. Results showed that heritability and distance between two linked QTL are most important, followed by distance between markers and population size. When genetic variance increased (i.e., heritability increased from 0.1 to 0.3), the probability of separating the two QTL increased from 22% to 54% (Fig. 5A and D). The power to separate the two QTL also increased by changing the population size from 200 to 500 (Fig. S6). The estimated d/a was higher when only one QTL was detected than when two separate QTL were detected, particularly when the distance between two QTL was less than 20 cM (Fig. 5B, C, E, and F). The simulation results showed that the lowest estimate of d/a was 1.06 if the distance between two QTL was less than 50 cM: The dominant effect was overestimated and the additive effect was underestimated because of repulsion linkage of the two QTL (pseudo-overdominance). Increasing the distance between the two QTL increased the probability of separating the two QTL and led to more accurate estimates of d/a. Increasing marker density, that is, reducing the distance between markers, also led to better estimates of d/a but had no significant effect on the ability to separate the two QTL.

Allelic Combinations for Designed Breeding Schemes. To explore the breeding implications of our findings further, we classified the 307 sorghum accessions used in the GWAS into 15 groups based on the genotypes at *Dw1*, *Dw2*, *Dw3*, and *qHT7.1*, the four most significant loci (Fig. 6). Interestingly, one particular genotype (homozygous dominant at *Dw2* and *Dw3* and homozygous recessive at *Dw1* and *qHT7.1*) does not exist in the sorghum association panel. Of the 307 accessions, 240 (78%) have either two or three recessive loci. A dominant allele of *qHT7.1* consistently increased plant height in different backgrounds of three other loci (Fig. 6, Bottom Row), as indicated by the higher median plant height in the even-numbered groups than in the odd-numbered groups. Variation within each group is expected because of effects at other loci. A hybrid generated by crossing two parents from different groups will show heterosis if the hybrid has more loci with dominant alleles in its genome than either of its parents, as in the current study (Fig. 4A and B). Therefore specific crosses can be made to exploit heterosis from dominance complementation. On the other hand, the expected plant height of different hybrids can be obtained so that only certain crosses are made to meet certain production restrictions on height parameters.

Discussion

Research into heterosis using both classical and modern genetics shows that different mechanisms are favored for different traits in different species, so finding a single unified theory to explain heterosis may be too challenging (4, 6, 43, 44). However, the dominance hypothesis has support from studies in maize and rice

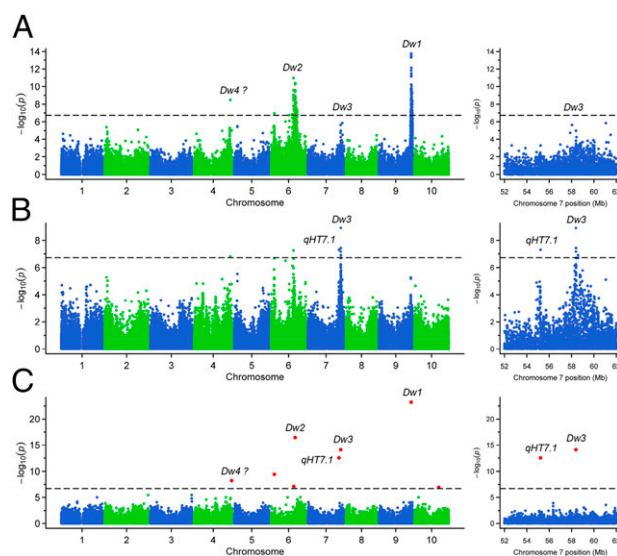


Fig. 3. Genome-wide association mapping of plant height. (A) Initial genome scan. (B) Second genome scan with a model including *Dw1* and *Dw2* as covariates. (C) Genome scan using the MLMM. (Left) Genome-wide results. (Right) Enlarged chromosome 7 region. The horizontal line in each subsection is the Bonferroni-corrected significance threshold. The positions of *Dw1*, *Dw2*, *Dw3*, *qHT7.1*, and the possible position for *Dw4* are indicated. The red points in C are the significant SNPs (entering the final model as covariates) identified by MLMM.

Table 2. Significant SNPs (covariates) detected by the MLM in GWAS

SNP	Locus	Effect	MAF	P value
S9_57236778	<i>Dw1</i>	-0.44	0.36	5.9×10^{-24}
S6_42744899	<i>Dw2</i>	0.36	0.36	3.4×10^{-17}
S7_58390034	<i>Dw3</i>	-0.29	0.43	7.8×10^{-15}
S7_55229509	<i>qHT7.1</i>	0.43	0.20	2.9×10^{-13}
S6_5702897	—	-0.86	0.01	3.7×10^{-10}
S4_66720692	<i>Dw4 ?</i>	-0.31	0.15	5.7×10^{-9}
S6_40343762	—	-0.46	0.04	7.5×10^{-8}
S10_43486183	—	0.71	0.01	1.1×10^{-7}

Effect estimates are for the Tx430 allele and are in SD units because the data used for association mapping were standardized. Chromosome number and SNP position are given after "S" in the first column. MAF, minor allele frequency.

(11, 12, 19, 45–47). Many QTL detected for maize heterosis are located near the centromere. The low recombination frequency around the centromere may conserve repulsion linkage, possibly explaining the observed overdominance (13). The estimate of dominance effect decreased as intercrossing of maize populations advanced, indirectly supporting the pseudo-overdominance hypothesis (48, 49). Recent findings of excessive residual heterozygosity around maize pericentromeric regions also were speculated to be evidence of the dominance hypothesis for heterosis (50).

In an earlier study of heterosis of maize yield, one QTL showing the overdominance effect in the initial analysis was later separated into two QTL 19 cM apart with dominant effects but in repulsion phase, resulting in pseudo-overdominance (15, 16). In our study, the two QTL are located on the long arm of chromosome 7, away from the centromere region. Although they are only 3 Mb apart, there is adequate recombination (29 cM) that allowed us to detect both loci. Having the repulsion linkage captured in two mapping parents facilitated the discovery of this example of heterosis in plant height generated by pseudo-overdominance.

Computer simulation is a powerful tool to test hypotheses when real testing conditions are not readily available (51). It allowed us to evaluate the influences of repulsion linkage on genetic mapping and effect estimation. Our simulation results showed that if two QTL in repulsion linkage are close to one another (<10 cM), they may not be separated, and the overdominance effect could be observed. Clearly, it would be desirable to identify other cases of pseudo-overdominance in which two loci have much closer genetic distance than we found in the current case. Additional case studies at the individual loci level for other quantitative traits, including biomass and grain yield, would generate empirical evidence and feedback for the theoretical research.

Direct comparison of results and leverage of information in linkage mapping, the GWAS, and designed crosses were possible in our study because of the large number of SNPs with known genome positions. The three QTL (*Dw1*, *Dw3*, and *qHT7.1*) detected in linkage mapping have approximately equal effect sizes. In the GWAS with diverse accessions, the *Dw1* gene was the most significant QTL because it has both the largest effect size and also a relatively high minor allele frequency. The QTL at ~120 cM on chromosome 7, which corresponds to the cloned *Dw3* gene, was detected by both linkage mapping and association mapping in many previous studies (36, 37, 52). Only one linkage-mapping study suggested two QTL in this chromosome region (53).

Examining results across genome scans with different methods enriches our understanding of trait genetic architecture. In a GWAS, *Dw3* and *qHT7.1* can be detected only after *Dw1* and *Dw2* are included in the linear mixed model. This result agrees with general expectation that specifying major loci in the GWAS model can improve power of detecting other loci. Likewise, the MLM method, which sequentially adds detected loci to the scanning models (42), supported the findings in *Dw3* and *qHT7.1* and identified additional

loci. Interestingly, both *Dw3* and *qHT7.1* were detected with the appropriate models. The low linkage disequilibrium between these two loci may be caused by haplotype frequencies captured in breeding materials through intentional selection based on phenotypes (Table S2).

Because *Dw3* is responsible for auxin transport in the stem from top to bottom, the lower part of the plant is expected to be influenced more than the upper part of the plant. The upper part will have enough auxin for elongation regardless of the existence of the auxin transporter as long as the auxin production system is functional. *qHT7.1*, on the other hand, affects all height component traits in this study (both direct measurements and indirect calculations), meaning that this QTL may be involved in a different pathway regulating plant height. The gibberellic acid (GA) and brassinosteroid (BR) pathways are key players regulating plant height in many species (54, 55). We speculate the gene underlying *qHT7.1* may be involved in these pathways.

Previous studies have showed that the four major genes controlling plant height in sorghum, *Dw1*–*Dw4*, are not linked, and the tall phenotype shows complete dominance (34). The crossing design illustrated in Fig. 1B allowed us to evaluate the gene action under mostly defined genetic backgrounds. For each cross, the two inbred lines were selected based on their similarity at other positions in the genome to minimize background effects on the target QTL. Although this approach cannot eliminate residual background effects completely, the hybrid crosses made based on these three major loci allowed us to examine the heterosis and dominance effects from these target QTL.

The gene region harboring *Dw3* and *qHT7.1* is an introgression region from the sorghum conversion program, together with the *Ma1* and *Dw2* gene region on chromosome 6 and the *Dw1* gene region on chromosome 9 (56). However, selection was conducted based on phenotypic observations. Genetic analysis and our specific research indicate that we have a large space for further optimizing plant height in sorghum. Repulsion linkage is also the reason for linkage drag in plant breeding. It poses a challenge for introgression of desirable chromosome segments into another background (16, 57). Dissecting the chromosome region into different QTL can help identify the desirable allelic combination in breeding populations. Fine tuning plant height can lead to sorghum hybrids that are high biomass and high lodging-resistant at the same time. Our findings of the QTL *qHT7.1*, its

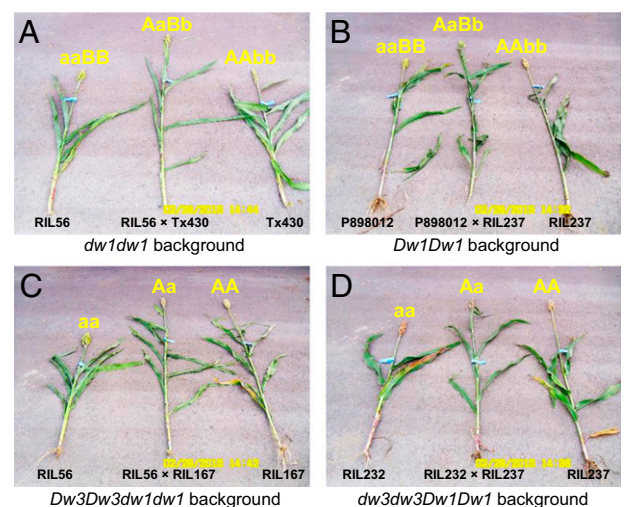


Fig. 4. Plant height of parental lines and corresponding F₁ hybrids. (A and B) Heterosis caused by repulsion linkage of *qHT7.1* and *Dw3*, under different backgrounds of *Dw1*. (C and D) The tall allele of *qHT7.1* shows complete dominance over the short allele of *qHT7.1* under different backgrounds of *Dw3* and *Dw1*. In each picture, the seed parent is on the left, and the pollen parent is on the right.

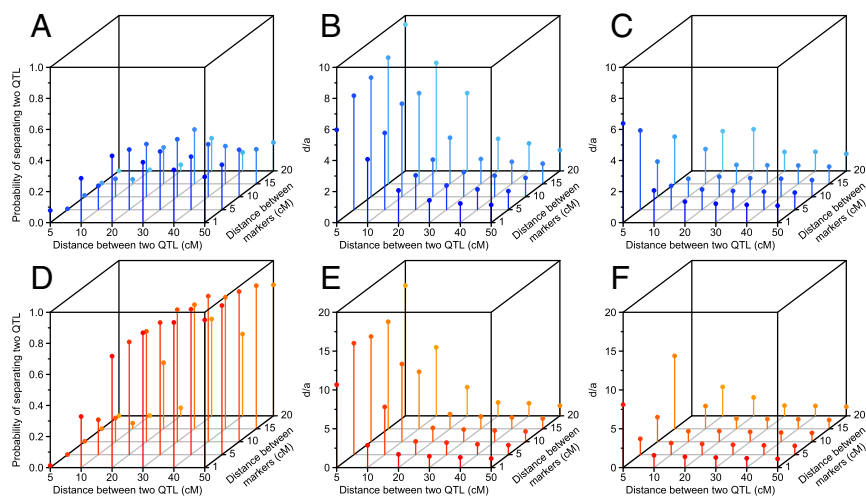


Fig. 5. Effect of repulsion linkage on QTL detection in a simulated F_2 population with 200 individuals. (A–C) Heritability of 0.1. (D–F) Heritability of 0.3. (B and E) The d/a ratios when one QTL was detected. (C and F) The d/a ratios when two QTL were detected. For all d/a ratios, each point is the median value from 5,000 runs, and missing points indicate that those scenarios were not observed.

haplotype with *Dw3*, and heterosis and gene action demonstration provide additional tools for breeders to manipulate plant height.

Materials and Methods

Mapping Population and Experimental Design. The 250 RILs used in this study are derived by single-seed descent from a cross between Tx430 and P898012. Tx430 is a grain sorghum inbred commonly used as pollinator in hybrid production. The other parent, P898012, is a grain sorghum inbred with good drought-tolerance traits (58). The field experiments were carried out in five environments (location-year combinations): Guayanilla, Puerto Rico (2011 and 2012); Manhattan, KS (2011 and 2012); and Ames, IA (2013). The population was planted in a randomized complete block (RCB) design with two replications in each environment.

Phenotypic Measurement. Total plant height was measured as the distance between ground and the top of the panicle. In the summer 2013 growing season, three other measurements (base of panicle height, flag leaf height, and preflag leaf height) were collected on the RIL population as complementary information of total plant height (Fig. 1A). By treating genotype as a random effect, best linear unbiased prediction (BLUP) values and variance components were obtained with PROC MIXED in SAS 9.3 software (SAS Institute). The entry mean-based heritability was calculated from variance components estimates.

Genotyping and Marker Screening. Genomic DNA was extracted from each RIL using the standard cetyltrimethylammonium bromide (CTAB) method (59). The concentrations of the DNA samples were quantified using Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen). The GBS procedure was carried out according to a standard protocol (60) and identified 10,380 initial SNPs segregating within the RIL population. A set of 8,960 SNPs was used for genetic mapping after SNPs with genotyping errors or misalignments were removed.

Linkage Map Construction and Genetic Mapping. To take full advantage of the high-density markers, we first conducted a single-marker scan using 8,960 SNPs with their physical positions in the genome. The analysis was conducted in R (R Development Core Team). By setting a missing data threshold of 10%, we selected 1,756 SNPs to build the genetic map using program MSTMap (61). Composite interval mapping using 1,756 SNPs was carried out in Windows QTL Cartographer 2.5 (62).

GWAS with Diverse Accessions. Association mapping was conducted with the sorghum association panel (39). The phenotypic data were the published plant height measurements (36). Based on the overlap of the 378 accessions with the published GBS data (37), a set of 307 accessions with 265K SNP markers was used for the GWAS. Missing data in the 265K SNPs were imputed with Beagle 4.0 (63). The analysis was carried out in the Genomic Association and Prediction Integrated Tool (64). The first three principle components were used as covariates for population structure in the compressed mixed linear models (65). After the initial analysis, a second genome scan was carried out with a mixed model

that included the most significant SNPs under *Dw1* and *Dw2* as additional covariates. A third genome scan was also carried out using the multilocus mixed model (42), which selects the most significant SNP from the stepwise regression to enter into the model sequentially as a covariate.

Computer Simulations. To assess the influence of repulsion linkage on genetic mapping in the F_2 population, a computer simulation experiment was conducted (51). Four factors were considered: heritability, distance between markers, distance between two QTL, and population size. Heritability (the genetic contribution of two QTL to phenotypic variance) was set at 0.1 or 0.3, distance between markers was set at 1, 5, 10, 15, or 20 cM, the distance between two QTL was set at 5, 10, 20, 30, 40, or 50 cM, and population size was set at 200 or 500 individuals. The two linked QTL were assumed to have equal effect sizes, and gene action at each locus was complete dominance. With 120 different combinations of the four factors, each combination was repeated 5,000 times. Genotypic and phenotypic data were generated by the UNIX version of the QTL cartographer software v1.17. The composite interval mapping method was used for linkage mapping. Three summary statistics were calculated: the proportion of simulation runs with two linked QTL detected separately, the d/a ratio when only one QTL was detected, and the d/a ratio when two QTL were detected. The median values were presented for the d/a ratios.

Hybrid Development. To verify the hypothesis that heterosis of sorghum plant height can be explained by repulsion linkage, crosses were made during

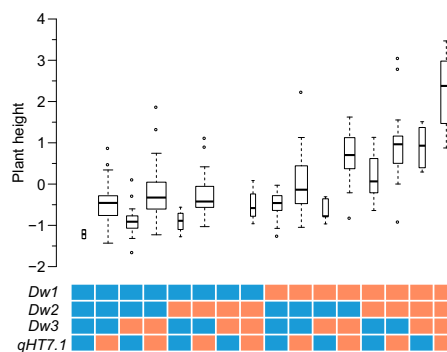


Fig. 6. Standardized plant height for the 16 genotype combinations in the sorghum association panel ordered by their genotypes at *Dw1*, *Dw2*, *Dw3*, and *qHT7.1*. The width of each box in the boxplot is proportional to the square root of the number of accessions in each group. Orange rectangles represent homozygous dominant genotypes, and blue rectangles represent homozygous recessive genotypes. No accessions were found in the seventh genotype combination.

summer 2014 at Ames, IA (Fig. 1B). Eight RILs were selected based on their genotypes at the three QTL detected in this study (*qHT7.1*, *Dw3*, and *Dw1*). For each cross, the two inbred lines were selected based on their genotype at the target plant height QTL and their similarity at other positions in the genome to minimize background effects. In addition, two inbreds were chosen to have similar flowering times to facilitate the crossing. The first two sets of six crosses focused on the two QTL on chromosome 7 with the QTL on chromosome 9 fixed as homozygous dominant or homozygous recessive. The remaining four

crosses were made to examine the effect of the chromosome 9 QTL. The hybrids together with their parents were grown at Santa Isabel, Puerto Rico in 2014.

ACKNOWLEDGMENTS. This work was supported by Binational Agricultural Research and Development Program IS-4256-12, by Agriculture and Food Research Initiative Competitive Grant 2011-03587 from the US Department of Agriculture National Institute of Food and Agriculture, by the Kansas State University Center for Sorghum Improvement, and by the Iowa State University Raymond F. Baker Center for Plant Breeding.

- Lippman ZB, Zamir D (2007) Heterosis: Revisiting the magic. *Trends Genet* 23(2):60–66.
- Duvick DN (2001) Biotechnology in the 1930s: The development of hybrid maize. *Nat Rev Genet* 2(1):69–74.
- Lin SC, Yuan LP (1980) Hybrid rice breeding in China. Innovative approaches to rice breeding. Selected Papers from the 1979 International Rice Research Conference (International Rice Research Institute, Manila, Philippines) pp 35–51.
- Chen ZJ (2013) Genomic and epigenetic insights into the molecular bases of heterosis. *Nat Rev Genet* 14(7):471–482.
- Goff SA (2011) A unifying theory for general multigenic heterosis: Energy efficiency, protein metabolism, and implications for molecular breeding. *New Phytol* 189(4):923–937.
- Schnable PS, Springer NM (2013) Progress toward understanding heterosis in crop plants. *Annu Rev Plant Biol* 64:71–88.
- Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. *Nat Rev Genet* 10(11):783–796.
- Allison AC (2009) Genetic control of resistance to human malaria. *Curr Opin Immunol* 21(5):499–505.
- Krieger U, Lippman ZB, Zamir D (2010) The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat Genet* 42(5):459–463.
- Jones DF (1917) Dominance of Linked Factors as a Means of Accounting for Heterosis. *Genetics* 2(5):466–479.
- García AA, Wang S, Melchinger AE, Zeng Z-B (2008) Quantitative trait loci mapping and the genetic basis of heterosis in maize and rice. *Genetics* 180(3):1707–1724.
- Xiao J, Li J, Yuan L, Tanksley SD (1995) Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics* 140(2):745–754.
- Larièpe A, et al. (2012) The genetic basis of heterosis: Multiparental quantitative trait loci mapping reveals contrasted levels of apparent overdominance among traits of agronomical interest in maize (*Zea mays* L.). *Genetics* 190(2):795–811.
- Zhou G, et al. (2012) Genetic composition of yield heterosis in an elite rice hybrid. *Proc Natl Acad Sci USA* 109(39):15847–15852.
- Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* 132(3):823–839.
- Graham GI, Wolff DW, Stuber CW (1997) Characterization of a yield quantitative trait locus on chromosome five of maize by fine mapping. *Crop Sci* 37(5):1601–1610.
- Li ZK, et al. (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* 158(4):1737–1753.
- Yu SB, et al. (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc Natl Acad Sci USA* 94(17):9226–9231.
- Tang J, et al. (2010) Dissection of the genetic basis of heterosis in an elite maize hybrid by QTL mapping in an immortalized F2 population. *Theor Appl Genet* 120(2):333–340.
- Melchinger AE, et al. (2007) Genetic basis of heterosis for growth-related traits in Arabidopsis investigated by testcross progenies of near-isogenic lines reveals a significant role of epistasis. *Genetics* 177(3):1827–1837.
- Hua J, et al. (2003) Single-locus heterotic effects and dominance by dominance interactions can adequately explain the genetic basis of heterosis in an elite rice hybrid. *Proc Natl Acad Sci USA* 100(5):2574–2579.
- Kusterer B, et al. (2007) Heterosis for biomass-related traits in Arabidopsis investigated by quantitative trait loci analysis of the triple testcross design with recombinant inbred lines. *Genetics* 177(3):1839–1850.
- Liang GH, Dayton AD, Reddy CR (1972) Heterosis, inbreeding depression, and heritability estimates in a systematic series of grain-sorghum genotypes. *Crop Sci* 12(4):409–411.
- Pfeiffer TW, Bitzer MJ, Toy JJ, Pedersen JF (2010) Heterosis in sweet sorghum and selection of a new sweet sorghum hybrid for use in syrup production in Appalachia. *Crop Sci* 50(5):1788–1794.
- Ben-Israel I, Kilian B, Nida H, Fridman E (2012) Heterotic trait locus (HTL) mapping identifies intra-locus interactions that underlie reproductive hybrid vigor in Sorghum bicolor. *PLoS One* 7(6):e38993.
- Packer DJ, Rooney WL (2014) High-parent heterosis for biomass yield in photoperiod-sensitive sorghum hybrids. *Field Crops Res* 167:153–158.
- Quinby JR (1963) Manifestations of hybrid vigor in sorghum. *Crop Sci* 3(4):288–291.
- Sinha SK, Khanna R (1975) Physiological, biochemical, and genetic basis of heterosis. *Adv Agron* 27:123–174.
- Uzarowska A, et al. (2007) Comparative expression profiling in meristems of inbred-hybrid triplets of maize based on morphological investigations of heterosis for plant height. *Plant Mol Biol* 63(1):21–34.
- Flint-García SA, Buckler ES, Tiffin P, Ersoz E, Springer NM (2009) Heterosis is prevalent for multiple traits in diverse maize germplasm. *PLoS One* 4(10):e7433.
- Shen G, Zhan W, Chen H, Xing Y (2014) Dominance and epistasis are the main contributors to heterosis for plant height in rice. *Plant Sci* 215–216:11–18.
- Xiao J, Li J, Yuan L, McCouch SR, Tanksley SD (1996) Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theor Appl Genet* 92(6):637–643.
- Zhang Y, Ni Z, Yao Y, Nie X, Sun Q (2007) Gibberellins and heterosis of plant height in wheat (*Triticum aestivum* L.). *BMC Genet* 8:40.
- Quinby JR, Karper RE (1954) inheritance of height in sorghum. *Agron J* 46(5):211–216.
- Multani DS, et al. (2003) Loss of an MDR transporter in compact stalks of maize br2 and sorghum dw3 mutants. *Science* 302(5642):81–84.
- Brown PJ, Rooney WL, Franks C, Kresovich S (2008) Efficient mapping of plant height quantitative trait loci in a sorghum association population with introgressed dwarfing genes. *Genetics* 180(1):629–637.
- Morris GP, et al. (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc Natl Acad Sci USA* 110(2):453–458.
- Klein RR, et al. (2008) The effect of tropical sorghum conversion and inbred development on genome diversity as revealed by high-resolution genotyping. *Crop Sci* (Supplement 1):S-12–S-26.
- Casa AM, et al. (2008) Community resources and strategies for association mapping in sorghum. *Crop Sci* 48(1):30–40.
- Murray SC, Rooney WL, Hamblin MT, Mitchell SE, Kresovich S (2009) Sweet sorghum genetic diversity and association mapping for brix and height. *Plant Genome* 2(1):48–62.
- Upadhyaya HD, Wang Y-H, Gowda CLL, Sharma S (2013) Association mapping of maturity and plant height using SNP markers with the sorghum mini core collection. *Theor Appl Genet* 126(8):2003–2015.
- Segura V, et al. (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat Genet* 44(7):825–830.
- Birchler JA, Yao H, Chudalayandi S, Vaiman D, Veitia RA (2010) Heterosis. *Plant Cell* 22(7):2105–2112.
- Kaeppler S (2012) Heterosis: Many genes, many mechanisms - end the search for an undiscovered unifying theory. *ISRN Botany* 2012:682824.
- Li L, et al. (2008) Dominance, overdominance and epistasis condition the heterosis in two heterotic rice hybrids. *Genetics* 180(3):1725–1742.
- Russell WA, Eberhart SA, Vega UA (1973) Recurrent selection for specific combining ability for yield in 2 maize populations. *Crop Sci* 13(2):257–261.
- Huang X, et al. (2015) Genomic analysis of hybrid rice varieties reveals numerous superior alleles that contribute to heterosis. *Nat Commun* 6:6258.
- Gardner CO, Lonquist JH (1959) Linkage and the degree of dominance of genes controlling quantitative characters in maize. *Agron J* 51(9):524–528.
- Moll RH, Lindsey MF, Robinson HF (1964) Estimates of genetic variances and level of dominance in maize. *Genetics* 49(3):411–423.
- McMullen MD, et al. (2009) Genetic properties of the maize nested association mapping population. *Science* 325(5941):737–740.
- Li X, Zhu C, Wang J, Yu J (2012) Computer simulation in plant breeding. *Adv Agron* 116:219–264.
- Higgins RH, Thurber CS, Assaranurak I, Brown PJ (2014) Multiparental mapping of plant height and flowering time QTL in partially isogenic sorghum families. *G3-Genes Genomes Genetics* 4(9):1593–1602.
- Srinivas G, et al. (2009) Identification of quantitative trait loci for agronomically important traits and their association with genic-microsatellite markers in sorghum. *Theor Appl Genet* 118(8):1439–1454.
- Salas Fernandez MG, Becraft PW, Yin Y, Lübberstedt T (2009) From dwarves to giants? Plant height manipulation for biomass yield. *Trends Plant Sci* 14(8):454–461.
- Davies PJ (2004) *Plant Hormones: Biosynthesis, Signal Transduction, Action!* (Springer, Dordrecht) Ed. 3.
- Thurber CS, Ma JM, Higgins RH, Brown PJ (2013) Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. *Genome Biol* 14(6):R68.
- Zamir D (2001) Improving plant breeding with exotic genetic libraries. *Nat Rev Genet* 2(12):983–989.
- Yu J, Hamblin M, Tuiistra M (2013) *Association Genetics Strategies and Resources. Genomics of the Saccharinae, Plant Genetics and Genomics: Crops and Models*, ed Paterson AH (Springer, New York), Vol 11, pp 187–203.
- Doyle JJ (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15.
- Elshire RJ, et al. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6(5):e19379.
- Wu Y, Bhat PR, Close TJ, Lonardi S (2008) Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genet* 4(10):e1000212.
- Wang S, Basten CJ, Zeng Z-B (2012) *Windows QTL Cartographer 2.5* (Department of Statistics, North Carolina State University, Raleigh, NC).
- Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 81(5):1084–1097.
- Lipka AE, et al. (2012) GAPIT: Genome association and prediction integrated tool. *Bioinformatics* 28(18):2397–2399.
- Zhang Z, et al. (2010) Mixed linear model approach adapted for genome-wide association studies. *Nat Genet* 42(4):355–360.