



Overdominance is the major genetic basis of lint yield heterosis in interspecific hybrids between *G. hirsutum* and *G. barbadense*

Shuhua Tian^{1,2} · Xiaolong Xu³ · Xiefei Zhu¹ · Fang Wang³ · Xianliang Song³ · Tianzhen Zhang^{1,4}

Received: 8 January 2019 / Revised: 2 March 2019 / Accepted: 6 March 2019 / Published online: 22 March 2019
© The Genetics Society 2019

Abstract

The genetic basis of heterosis has not been resolved for approximately a century, although the role of loci with overdominant (ODO) effects has continued to be discussed by biologists. In the present investigation, a proposed model was studied in *Gossypium hirsutum* L. introgression lines (ILs) harbouring a segment of *G. barbadense*. These introgressions were confirmed by a single marker of *G. barbadense*. These ILs contained 396 quantitative trait loci (QTLs) for 11 yield and non-yield traits that were recorded in the field on homozygous and heterozygous plants for 5 years. After comparing the different types of QTLs between the yield group and the non-yield group, it was found that the yield group had significantly higher ODO QTL ratios. Moreover, 16 ODO QTLs identified for 5 yield-related traits were consistently detected during 5 cotton growing seasons (2010–2011 and 2013–2015): 6 of 7 for boll weight, 3 of 11 for seed-cotton yield per plant, 4 of 17 for boll number, 2 of 13 for lint yield per plant and 1 of 11 for lint percentage. Therefore, we propose that overdominance is the major genetic basis of lint yield heterosis in interspecific hybrids between *G. barbadense* and *G. hirsutum*. These findings have important implications in cotton breeding in that the boll weight can be improved by utilizing ODO QTLs via heterosis; thus, the stagnant yield barrier can be smashed to achieve sustainable increases in cotton production. Additionally, this concept can be translated to other field crops for improving their yield potential.

These authors contributed equally: Shuhua Tian, Xiaolong Xu

Supplementary information The online version of this article (<https://doi.org/10.1038/s41437-019-0211-5>) contains supplementary material, which is available to authorized users.

✉ Xianliang Song
songxl999@163.com

✉ Tianzhen Zhang
cotton@njau.edu.cn

¹ State Key Laboratory of Crop Genetics and Germplasm Enhancement, Cotton Hybrid R & D Engineering Center (the Ministry of Education), College of Agriculture, Nanjing Agricultural University, Nanjing 210095, China

² Shanghai Key Laboratory of Protected Horticultural Technology, Forestry and Fruit Tree Research Institute, Shanghai Academy of Agricultural Sciences (SAAS), Shanghai 201403, China

³ State Key Laboratory of Crop Biology/Agronomy College, Shandong Agricultural University, Taian 271018, China

⁴ Agronomy Department, College of Agriculture and Biotechnology, Zhejiang University, Zhejiang 310029, China

Introduction

Heterosis refers to the fact that hybrids derived from diverse inbred genotype varieties have a greater biomass, higher growth velocity and/or greater fertility than their parental genotypes (Chen 2010). However, the genetic basis for this phenomenon has remained obscure. To date, several hypotheses including dominance, overdominance, pseudo-overdominance and epistasis have been proposed for explaining the genetics behind heterosis. The dominance hypothesis emphasizes that the dominant genes, controlling the beneficial traits (Bruce 1910) shared from both parents, are concentrated in the first-generation hybrid. This hypothesis was successfully proven in a maize polymerization and improvement experiment (Richey and Sprague 1931). The dominance hypothesis is one of the theoretical bases for breeding high-yield hybrids in maize, rye and other cross-pollinated crops. However, there are still many questions that are difficult to answer based on this hypothesis. For example, the performance of hybrid offspring exceeding 20% or even 50% of homozygous dominant parents often occurs in maize cross-breeding, which cannot be addressed using the dominance hypothesis (Crow

1948), making this hypothesis unrealistic. To tackle this issue, a true overdominance (ODO) hypothesis was proposed to explain the increased performance of F_1 s owing to the interaction between heterozygous alleles, which is superior to the interaction between homozygous alleles at a locus (East 1936; Crow 1948). For example, a heterozygous genotype has been proven to be advantageous in sickle-cell anaemia (Allison 2009) and tomato (Krieger et al. 2010). Additionally, the pseudo-overdominance hypothesis (Jones 1917) was first proposed in 1917 as another type of ODO to explain heterosis. It was suggested that the linkage of loci containing dominant alleles in the repulsion phase is responsible for increased vigour in F_1 s. For example, a separate quantitative trait locus (QTL) for plant height (*qHT7.1*) was recognized as exhibiting repulsion linkage with the known auxin transporter *Dw3* gene between two inbreds, leading to heterosis in the hybrid (Li et al. 2015). The last hypothesis is that ‘epistasis’ explains that hybrid vigour resulting from the expression of a pair of alleles at a locus is inhibited by the other alleles present on different loci, i.e., interactions between genes on different loci (Yu et al. 1997).

In 1894, a seminal research report demonstrating the heterotic expression in an F_1 hybrid derived from a cross between *Gossypium hirsutum* and *Gossypium barbadense* was published (Mell 1894). Thereafter, multiple studies have demonstrated prominent heterosis in yield and its components, such as boll number and boll weight (Ball 1908; Loden and Richmond 1951; Meyer 1969; Davis 1979; Meredith and Bridge 1984; Meredith 1990; Zhang and Pan 1999). All these findings prepared a firm foundation in breeding interspecific hybrids, replacing the substantial acreage sown using conventional cultivars in several countries including India—which witnessed increased yield and improvements in fibre quality. Therefore, the development of such hybrids is an effective approach for enhancing sustainable yield and improved lint quality (Hua 1963; Zhang and Pan 1999).

We comprehensively describe the phenomenon of hybrid vigour in cotton using a set of 43 upland cotton (*G. hirsutum*) introgression lines (ILs) containing chromosome segments extracted from Pima cotton (*G. barbadense*). The presence of these chromosome segments was confirmed by individual DNA markers. The introgressed segment of each chromosome of *G. barbadense* replaced the homologous fragment of *G. hirsutum* line TM-1. The newly developed ILs are approximately isogenic lines, thus avoiding the entire epistatic interactions in the cotton genome, which aids in the accurate identification of QTLs conferring ODO effects. Additionally, our ILs were developed by interspecific crossing, thus maximizing genetic and phenotypic diversity in a wide range of traits, which also partly reduces

the systematic error of our experiment. Based on our findings, we propose that overdominance is the major genetic basis of lint yield heterosis in interspecific hybrids between *G. barbadense* and *G. hirsutum*.

Materials and methods

Plant material

Although the population of ILs originally consisted of 174 lines (Wang et al. 2012), 162 lines were available for conducting the present experiments. A set of hybrids between ILs and TM-1 (introgression line hybrids) were produced before performing our field test every year. Theoretically, the introgression line hybrids (ILHs) were heterozygous for the corresponding IL-substituted fragment. As a result, 162 sets of lines were developed in total, each set containing TM-1 and IL, as well as the corresponding hybrids, ILHs. In total, 162 sets of lines were planted at the Dezhou Cotton Breeding Station farm (Dezhou, China) in 2010 and 2011. Of the total, only 43 introgression lines were selected for conducting heterotic analysis. The same experiment, arranged in three randomized replications, was conducted at the Jiangpu Cotton Breeding Station (Nanjing, China) for 3 years (2013–2015). A total of 12 seedlings of each TM-1, homozygous IL and heterozygous ILH were transplanted in one plot. The planting density of these plants was one plant per square metre. Standard agronomic practices were applied from sowing to harvesting.

Phenotypic evaluation

Data of 17 traits were collected and analysed in the first 2 years (2010 and 2011). Of the 17 traits, 6 were yield related (lint yield per plant (LPP), g/plant; seed-cotton yield per plant (SPP), g/plant; boll number (BN), number; boll weight (BW), g; lint index (LI), g; lint percentage (LP), %), 6 were intermediate traits (fruit branches (FB), number; plant height (PH), cm; seed index (SI), g/100 seeds; short fibre index (SFI), %; first fruit branch position (FBP), number; fruiting-branch internodes length of first fruit branch (FBI), cm) and 5 were non-yield traits (fibre length (FL), mm; fibre strength (FS), cN/tex; micronaire (MIC); fibre uniformity ratio (FUR), %; fibre elongation (FE), %). Our objective was to determine the differences between the yield and non-yield groups. Therefore, we excluded the data of six intermediate traits (see Results for detailed classification) collected from the trial conducted for 3 years (2013–2015) at the Jiangpu Cotton Breeding Station, due to the heavy workload of the field investigation associated with data acquisition.

Statistical analysis and QTL mapping

Correlation analyses and analysis of variance were performed using R statistical software. The broad-sense heritability (h^2), expressed as $\sigma_G^2/\sigma_{G+E}^2$, was also computed for each trait. The genotype was denoted as a random effect factor, and the genetic variation was computed as a percentage of the total variation—comprising both the genetic and environmental variations.

With regard to QTL mapping, we used the previous classification methods reported for tomato by Semel et al. (2006). For example, a comparison between each IL or ILH was made (by *t*-test) with TM-1 and also to each other. A substantial difference in either of these from TM-1 (reference genotype) indicated the presence of a QTL with an introgression. In the present experiments, owing to a similar number of replications for the ILs and ILHs and a higher number of replications for TM-1, two levels of significance were performed for the identification of a QTL. The first-level comparison was performed to compare ILs and ILHs to TM-1 (a1), whereas the second-level comparison examined the differences between ILs and ILHs (a2). The analysis for QTL detection was conducted at two thresholds: a1 = 1% and a2 = 5% (significant analysis).

Qualitative classification of the mode of inheritance

Based on the classification method proposed by Semel et al. (2006), the phenotypic effect of a QTL was regarded as the effect of a significant line (IL or ILH) and is presented as a percentage of TM-1 (positive values indicate an increasing QTL, where the introgression is greater than TM-1, whereas negative values indicate a decreasing QTL). The specific classification method is such that if the IL and ILH contribute significantly in the same direction, then the higher value was used to represent the phenotypic effect of the QTL. Conversely, if a significant antagonistic effect of IL and ILH relative to TM-1 was found, the introgression was regarded as containing two QTLs; one QTL is increasing, whereas the other is decreasing. Based on decision boxes (Fig. 1), we determined the mode of inheritance of a QTL.

Quantitative classification of the mode of inheritance

To analyse heterosis more rigorously, a quantitative index of mode of inheritance was used to score each QTL that was similar to the analytical methods used in tomatoes (Semel et al. 2006). For example, if the IL significantly outperformed TM-1, then 100 mode-of-inheritance units were assigned to the phenotypic interval observed between the IL (located at 100 coordinates) and TM-1 (located at zero coordinate). Then, the quantitative index of a QTL was

determined based on the relative position of the ILH on the phenotypic interval between its parents (IL and TM-1). For example, if ILH and TM-1 were identical in phenotypic value, then it was presented as 0 (completely recessive); if the ILH phenotypic value was identical to the IL, it was identified as 100 (completely dominant); and if the phenotypic value of ILH was between the IL and TM-1, the index was confirmed as being a number between 0 and 100, reflecting the relative position of the ILH between IL and TM-1 (Fig. 1a–d). If the ILH value was beyond the range in the IL direction (Fig. 1e, f), the score was >100, and if the ILH was beyond the range in the TM-1 direction (Fig. 1g), the score was <0. Another probability was that the IL was less important than TM-1. In such circumstances, the ILH must be important compared with TM-1 (otherwise, it was not identified as a QTL). In this case, the quantitative index of mode of inheritance was computed as 100 plus the (absolute) phenotypic disparity (percentage of TM-1) between ILH and the higher parent (under the condition that ILH was highest) or the lower parent (under the condition that ILH was lowest) (Fig. 1f, h). In some extraordinary cases, IL and ILH were significantly opposite to TM-1 (one is on the higher side, whereas the other is on the lower side), in which case two QTLs were implied (Fig. 1i).

Pleiotropism-corrected QTL

The present research primarily revolves around the comparison of QTL mode of inheritance between yield and non-yield traits. Thus, there is a strong possibility that one QTL can be counted twice, especially for the correlated traits having pleiotropic effects. To counter this argument, we considered a single QTL for a genomic region impacting multiple related traits ($|r| > 0.5$) (Semel et al. 2006), as shown in Fig. 2.

Results

Heterosis in the interspecific ILs

To maximize the phenotypic and genetic diversity, data of various traits were collected entirely from interspecific ILs harbouring the majority of the *G. barbadense* genome in the TM-1 background. A total of 17 traits were compared in our research (Table S2). Among these, the heterotic effect was exhibited in lint yield per plant, seed-cotton yield per plant, boll number, boll weight and the lint index, where ILH mean values were distinctively higher than those of the ILs and TM-1. Other traits, such as lint percentage, fruit branches, plant height and the seed index, did not exhibit any heterotic effect, as the ILH average values were smaller than those of their parents. Among these traits, lint yield per

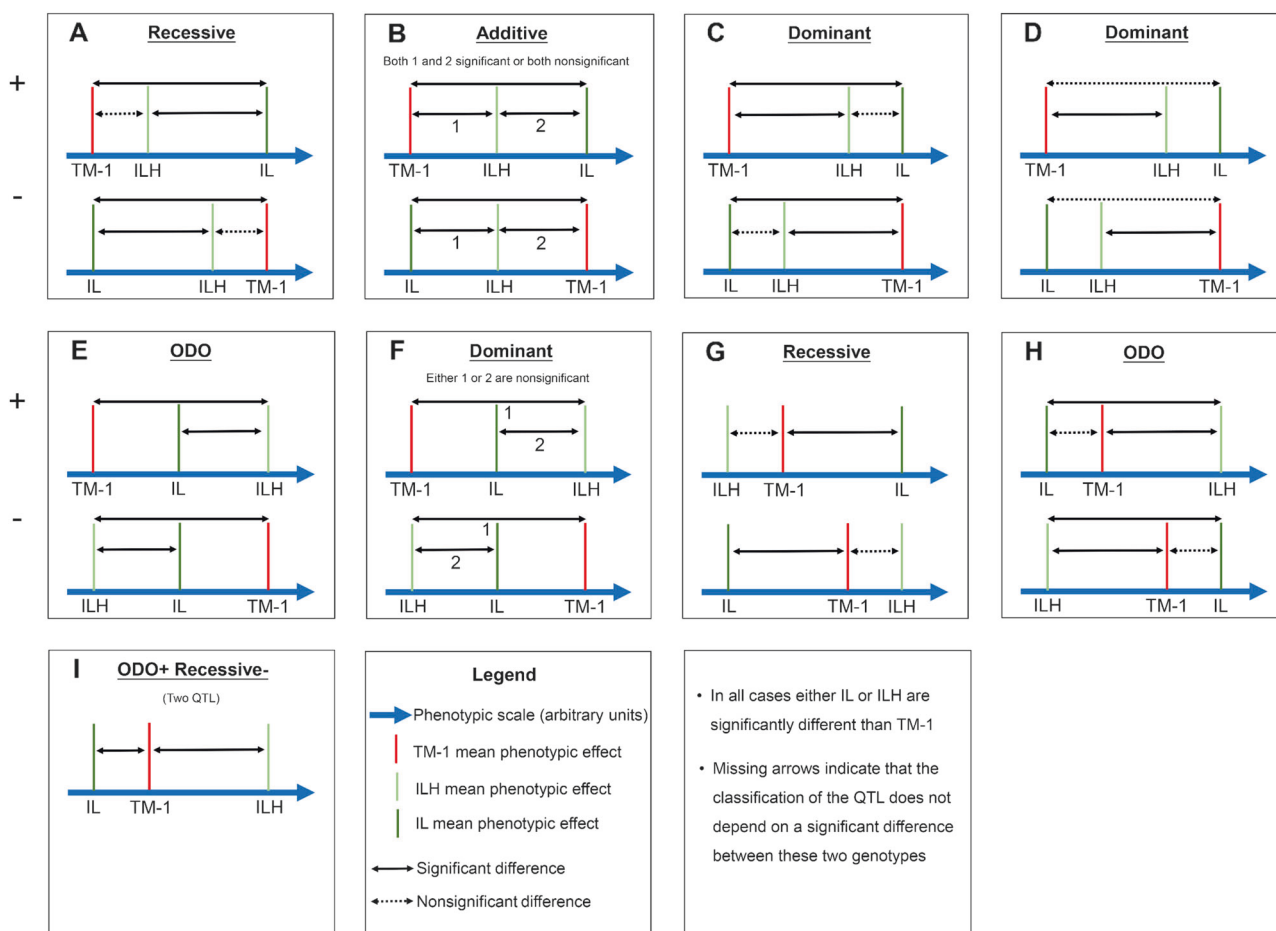


Fig. 1 Decision boxes for quantitative trait locus (QTL) mode of inheritance. Schematic illustration of the 9 (a–i) QTL types and their categorization of mode of inheritance indicating the relative influences (position on the x-axis) and their significance (arcs) for the introgression line (IL) and introgression line hybrid (ILH), as well as TM-1 plants. Solid lines are used to indicate significant differences, whereas dotted lines are used to indicate nonsignificant differences between

genotype pairs (refer to legend). A missing line represents that a QTL was divided into the specified categories regardless of the importance between the pair of genotypes. The upper scale on each square is used to represent an increase in the QTL (+), and the lower scale is used to represent a decrease in the QTL (–), except for square I, which represents two QTLs. It is worth noting that the IL or ILH is largely different from TM-1 in all of the above cases

plant exhibited the most pronounced heterotic effects. Of all the traits explored in the present investigation, only five yield-related traits exhibited heterotic effects. These data suggest that heterosis is ubiquitous in the interspecific ILs and is connected with yield-related phenotypes.

QTL mapping

To investigate whether heterotic phenotypes for yield traits and a particular mode-of-inheritance mechanism are related with each other at specific loci, QTL analysis of the entire phenotypic database was performed in such a manner that each IL and ILH was compared with the TM-1. A QTL in the introgressed part of the genome was considered if one of the lines exhibited a distinct effect. A statistical study, discussed in the Materials and methods, resolved 396 QTLs, conferring 11 yield and non-yield traits on average (Table

S3). Of these 396 QTLs, 214 represented the yield group and 182 belonged to the non-yield group. In addition, 6 traits studied in the first 2 years that were intermediate between yield- and non-yield-related traits resolved 164 QTLs (Table S3).

To thoroughly investigate this phenomenon, all of the QTLs were classified into the following categories of mode of inheritance (Fig. 1), i.e., additive, recessive, dominant or ODO. Thereafter, we compared the differences in QTL distribution between yield and non-yield groups. Surprisingly, the distribution data of the QTLs highlighted that the yield group included many more increasing ODO QTLs that were associated with recessive QTLs than in the non-yield group (Fig. 3). In this analysis and a subsequent analysis, an intermediate group of traits was between the yield and the non-yield groups in the distribution tendency of the mode-of-inheritance patterns of the QTL (refer to Table S4). In

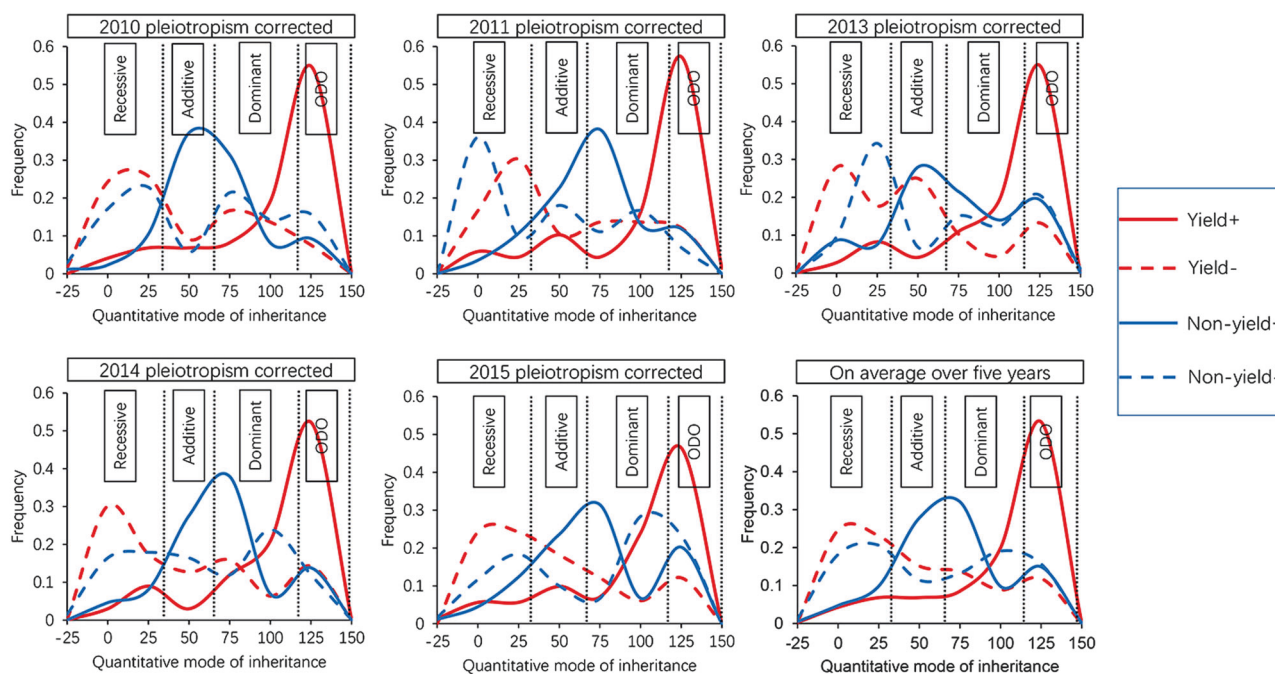


Fig. 2 Distribution of quantitative mode-of-inheritance quantitative trait loci (QTLs) for the yield and non-yield groups over 5 years. To correct for pleiotropism, a genomic area that influences multiple related traits ($|r| > 0.5$) and has an identical inheritance mode was regarded to be a single QTL. The features that were influenced by a specific area were regarded as vertices of the line in the graph, and the related features were linked by a line. In this hypothetical graph, each isolated cluster was regarded as a single QTL, and this method of pleiotropism correction relies on the threshold of the correlation coefficient (r). For example, the criterion that $r = 1$ is identical to classifying each feature as a QTL, regardless of the link between them, as the hypothetical graph features no connections (because the correlation cannot be > 1); therefore, each trait can be regarded as an isolated

cluster. Based on the relative position of the introgression line hybrids (ILHs) on the phenotypic interval between the two parents, which are the introgression lines (ILs) and TM-1, we can determine the quantitative index of the QTL. The region from -25 to $+150$ of the quantitative index of a QTL is represented by the x -axis, whereas the frequency of these indices is represented by the y -axis. The four curves indicate the distribution of increasing or decreasing QTLs for features in yield or non-yield groups. In other words, the quantitative index of QTLs identified as recessive is negative or positive up to approximately 33 (the first third in the interval of IL–TM-1). The indices of the additive QTL are relatively lower than 33 and higher than 66 (the second third), and those of the dominant are higher than 66 and may be greater than 100, whereas the ODO QTL must be higher than 100

our analyses, we only concentrated on the differences observed between the yield and non-yield groups, ignoring the intermediate traits.

Pleiotropism-corrected QTL

A correlation study was conducted for all the traits to correct for the bias generated by the pleiotropic effect of each QTL (Table S1). For example, the phenotype of seed-cotton yield per plant was found to be closely associated with boll number. A total of 40 QTLs for seed-cotton yield per plant and 43 QTLs for boll number were identified in 2010; 39 of these were co-mapped to the identical ILs, suggesting pleiotropism for both traits. Therefore, QTL data for all the traits were required to correct the redundancy. Thus, we counted the number of QTLs conferring highly related traits with greater similar inheritance (where r is higher than 0.5 or r is lower than -0.5) as a single IL QTL (refer to the Materials and methods). This type of pleiotropism-corrected procedure has been found worthwhile for making

comparisons between groups of traits as opposed to assuming there is no pleiotropism, thus avoiding the redundant QTLs (refer to Table S5 for a list of all pleiotropism-corrected QTLs).

It was found that even after applying the pleiotropism-correction method, some remarkable discrepancies for mechanisms of mode of inheritance among the non-redundant QTLs were observed between groups (refer to Table 1). This was much more similar to the results calculated without considering pleiotropism, i.e., the yield traits depicted a remarkable increase in ODO QTLs. For example, 31 of 138 (22.5%) QTLs in the yield group were increasing ODO QTLs, whereas in the non-yield group, only 1 of 156 (0.9%) QTLs was an increasing ODO QTL. Additionally, a rise and decline in a similar number of QTLs in non-yield traits was observed in the mode-of-inheritance category (except additive), whereas the yield group exhibited a tendency towards a recessive decrease and ODO/dominant increase in QTLs. We therefore deployed a more stringent statistical analysis that considered complete

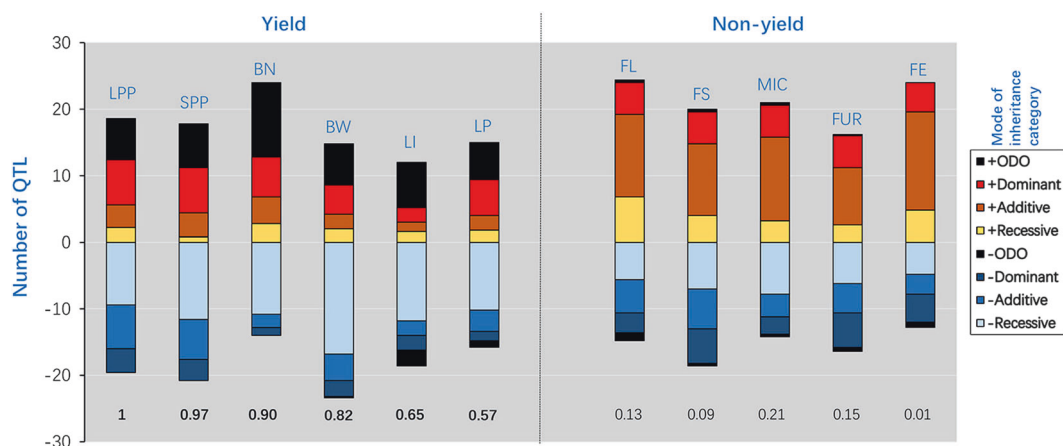


Fig. 3 Distribution of quantitative trait locus (QTL) mode of inheritance for cotton traits on average over 5 years. The QTL number for a specific feature is represented by a vertical bar, the colour of which is based on the inheritance mode category. The number of increasing QTLs is represented by those bars above the zero line, whereas the number of decreasing QTLs relative to TM-1 is represented by the negative bars. The correlation value for each feature to lint yield for each plant is shown at the bottom, and those features with a correlation

value higher than 0.5 are presented in bold type. LPP lint yield per plant, SPP seed-cotton yield per plant, BN boll number, BW boll weight, LI lint index, LP lint percentage, FB fruit branches, PH plant height, SI seed index, SFI short fibre index, FBP first fruit branch position, FBI fruiting-branch internode length of first fruit branch, FL fibre length, FS fibre strength, MIC micronaire, FUR fibre uniformity ratio, FE fibre elongation

Table 1 Qualitative mode-of-inheritance distribution

Mode of inheritance	Yield, <i>n</i> (%)	Non-yield, <i>n</i> (%)	<i>P</i> χ^2
+Overdominant (ODO)	31 (22.5)	1 (0.9)	2.02E-09
-ODO	4 (2.6)	2 (1.5)	0.33
+Dominant	18 (13.2)	21 (13.4)	0.92
-Dominant	10 (7.4)	19 (12)	0.16
+Additive	13 (9.7)	46 (29.2)	1.81E-05
-Additive	16 (11.6)	19 (11.9)	0.88
+Recessive	9 (6.7)	20 (12.7)	0.07
-Recessive	36 (26.2)	29 (18.4)	0.12
Total	138 (100)	156 (100)	

The qualitative distribution of inheritance mode indicates the mean number of quantitative trait loci (QTLs) over a 5-year period, divided into yield and non-yield groups (the percentage of QTLs in this classification of all the QTLs of that group is represented by the number in parentheses). The symbols that precede the inheritance mode represent an increasing (+) or decreasing (-) QTL relative to TM-1. The yield and non-yield groups are statistically compared in each inheritance mode with the help of the χ^2 test by dividing the QTLs into those that belong to this inheritance mode and those that do not, with one degree of freedom

pleiotropism (Table S6). With this method, only one QTL was counted for the QTL affecting multiple traits belonging to the same group, expressing a much more identical mode of inheritance, regardless of the correlation among traits (Semel et al. 2006). Despite applying highly stringent assumptions, the ODO of the yield traits was 25 times higher than that of the non-yield traits, as shown in Table S7.

Mode of inheritance on a quantitative scale

To achieve precision related to the mode of inheritance of each QTL, a large dataset of quantitative measurements was explored for adding a quantitative index to each QTL's mode of inheritance (refer to the Materials and methods). The frequency distribution of the mode-of-inheritance index for QTLs in the yield and non-yield groups is presented in Fig. 2. Corresponding to the distribution of the QTL quantitative mode of inheritance for increasing yield traits, the 'yield' curve attained the highest value in the ODO domain, suggesting that many QTLs follow this mode of inheritance. Inversely, the majority of the QTLs for the non-yield group and for the decreasing yield phenotypes were found in a recessive-additive-dominant domain. Thus, comparisons for the mode of inheritance of four curved distribution lines are presented in a quantitative manner in Fig. 2. The increasing yield QTL is also shown to be overrepresented, beginning in the dominant domain and increasing sharply towards the ODO domain. This propensity is supported by the results obtained conducting the qualitative analysis (Fig. 3; Table 1). Heterosis may be partially fractionated into small genomic areas that transmit dominance in a heterozygous state (ODO QTL). These genomic regions cumulatively confer a genome-wide effect.

ODO QTL for yield-related traits

To provide a good reference for high-yield breeding, overdominant quantitative trait loci detected in yield-related

Table 2 Overdominant (ODO) quantitative trait loci (QTLs) for yield-related traits

Trait	ODO QTL	Chr.	Marker	2010	2011	2013	2014	2015	Mean effect	Mean quantitative mode of inheritance	
LPP	<i>oLPP-A3-1</i>	A3	BNL2443		√				3.6	103.6	
	<i>oLPP-A4-1</i>	A4	NAU3469	√					7.9	104.1	
	<i>oLPP-A4-2</i>	A4	Y1315	√					5.7	100.7	
	<i>oLPP-A9-1</i>	A9	NAU3888	√	√	√		√	7.0	103.2	
	<i>oLPP-A9-2</i>	A9	NAU3414	√	√	√	√	√	6.1	103.0	
	<i>oLPP-A13-1</i>	A13	NAU3017	√					7.6	100.7	
	<i>oLPP-D1-2</i>	D1	NAU3543		√				2.8	102.8	
	<i>oLPP-D13-1</i>	D13	CIR221	√	√				3.2	127.0	
	<i>oLPP-D5-1</i>	D5	NAU2708				√		5.0	202.0	
	<i>oLPP-D10-1</i>	D10	NAU6562	√	√	√	√		6.9	103.2	
	<i>oLPP-D8-1</i>	D8	NAU2292	√					3.6	100.8	
	<i>oLPP-D6-1</i>	D6	NAU3532	√	√	√	√	√	3.6	103.6	
	<i>oLPP-D6-2</i>	D6	NAU6269	√	√	√		√	4.3	103.9	
	SPP	<i>oSPP-A2-1</i>	A2	HAU2583	√	√	√	√	√	21.1	110.2
<i>oSPP-A3-1</i>		A3	BNL2443	√	√	√	√	√	20.1	109.3	
<i>oSPP-A3-2</i>		A3	NAU3479		√				19.3	103.6	
<i>oSPP-A9-1</i>		A9	NAU3888		√				10.7	101.1	
<i>oSPP-A12-2</i>		A12	NAU3519	√	√		√	√	14.9	106.7	
<i>oSPP-D2-1</i>		D2	JESPR156	√	√	√	√	√	13.2	112.8	
<i>oSPP-D1-2</i>		D1	NAU3543	√		√	√	√	10.1	109.3	
<i>oSPP-D7-1</i>		D7	NAU3594	√					10.8	100.8	
<i>oSPP-D13-1</i>		D13	CIR221		√	√			14.6	114.6	
<i>oSPP-D9-1</i>		D9	CIR286		√				12.7	101.8	
BN	<i>oBN-A1-1</i>	A1	NAU3254		√	√		√	3.3	100.5	
	<i>oBN-A3-1</i>	A3	BNL2443	√	√	√		√	3.9	101.9	
	<i>oBN-A5-1</i>	A5	BNL3992	√	√	√	√	√	4.1	101.7	
	<i>oBN-A8-1</i>	A8	NAU3793	√	√	√	√	√	4.3	102.1	
	<i>oBN-A9-1</i>	A9	NAU3888		√	√		√	2.1	101.1	
	<i>oBN-A9-2</i>	A9	NAU3414	√			√		3.9	101.8	
	<i>oBN-A12-2</i>	A12	NAU3519		√	√		√	3.9	102.1	
	<i>oBN-D1-1</i>	D1	NAU5302			√			0.7	152.3	
	<i>oBN-D3-2</i>	D3	NAU3016	√			√		3.7	102.2	
	<i>oBN-D13-1</i>	D13	CIR221	√	√	√	√	√	3.5	217.4	
	<i>oBN-D13-2</i>	D13	NAU3223	√			√	√	3.7	101.4	
	<i>oBN-D10-2</i>	D10	NAU3916		√	√		√	2.8	101.6	
	<i>oBN-D11-2</i>	D11	NAU3373	√			√		3.8	102.1	
	<i>oBN-D6-1</i>	D6	NAU3532		√	√	√	√	3.3	100.3	
	<i>oBN-D6-2</i>	D6	NAU6269			√		√	4.9	100.4	
	<i>oBN-D12-1</i>	D12	NAU4925	√	√	√	√	√	6.8	104.0	
	<i>oBN-D12-2</i>	D12	NAU3032	√	√		√	√	2.4	102.2	
	BW	<i>oBW-A9-2</i>	A9	NAU3414	√	√	√	√	√	1.0	100.5
		<i>oBW-A11-2</i>	A11	NAU2152		√				0.5	100.1
		<i>oBW-D7-1</i>	D7	NAU3594	√	√	√	√	√	1.0	100.5
<i>oBW-D13-2</i>		D13	NAU3223	√	√	√	√	√	0.8	100.4	
<i>oBW-D6-1</i>		D6	NAU3532	√	√	√	√	√	0.5	100.5	
<i>oBW-D12-1</i>	D12	NAU4925	√	√	√	√	√	1.0	100.5		

Table 2 (continued)

Trait	ODO QTL	Chr.	Marker	2010	2011	2013	2014	2015	Mean effect	Mean quantitative mode of inheritance
LI	<i>oBW-D12-2</i>	D12	NAU3032	√	√	√	√	√	1.0	100.5
	<i>oLI-A1-1</i>	A1	NAU3254	√		√		√	1.7	101.6
	<i>oLI-A3-2</i>	A3	NAU3479			√			1.2	100.1
	<i>oLI-A8-1</i>	A8	NAU3793	√		√		√	1.5	100.8
	<i>oLI-A8-2</i>	A8	JESPR232			√	√		1.2	100.2
	<i>oLI-D2-1</i>	D2	JESPR156		√				1.7	100.5
	<i>oLI-D1-1</i>	D1	NAU5302	√	√	√	√		1.2	101.0
	<i>oLI-D3-2</i>	D3	NAU3016		√				0.0	100.0
	<i>oLI-D13-1</i>	D13	CIR221	√		√	√	√	1.2	165.0
	<i>oLI-D11-2</i>	D11	NAU3373	√		√	√	√	2.5	101.3
	<i>oLI-D4-1</i>	D4	NAU3781		√				0.8	100.5
	<i>oLI-D8-1</i>	D8	NAU2292	√		√	√	√	2.1	101.0
	<i>oLI-D6-1</i>	D6	NAU3532		√				1.4	101.3
	<i>oLI-D6-2</i>	D6	NAU6269	√			√		1.0	101.0
<i>oLI-D12-1</i>	D12	NAU4925	√	√	√			2.3	101.3	
LP	<i>oLP-A2-1</i>	A2	HAU2583	√	√				4.8	102.5
	<i>oLP-A8-1</i>	A8	NAU3793	√	√	√			4.9	102.8
	<i>oLP-A9-1</i>	A9	NAU3888				√		4.6	100.6
	<i>oLP-A10-1</i>	A10	NAU5323		√		√	√	3.7	103.4
	<i>oLP-A11-1</i>	A11	NAU3341	√	√	√		√	5.4	102.7
	<i>oLP-D1-1</i>	D1	NAU5302				√		2.3	102.3
	<i>oLP-D10-2</i>	D10	NAU3916		√				4.9	104.9
	<i>oLP-D11-1</i>	D11	TMG06			√		√	3.8	103.8
	<i>oLP-D11-2</i>	D11	NAU3373	√	√	√	√	√	5.4	102.6
	<i>oLP-D12-1</i>	D12	NAU4925	√	√				8.0	105.0
	<i>oLP-D12-2</i>	D12	NAU3032	√	√	√		√	6.4	103.1

The above data are counted under the situation with no pleiotropism (Table S3). The symbol '√' means detected in this year. LPP lint yield per plant, SPP seed-cotton yield per plant, BN boll number, BW boll weight, LI lint index, LP lint percentage

traits during the 5 years of our study are summarized in Table 2.

Lint yield per plant and seed-cotton yield per plant

In total, 13 and 11 ODO QTLs were identified during at least 1 year of the study for LPP and SPP traits, respectively. For LPP, ODO QTLs were found on 10 different chromosomes (chr. nos. A3, A4, A9, A13, D1, D5, D6, D8, D10 and D13). Of the 13 QTLs, two (15.4%; *oLPP-A9-2* and *oLPP-D6-1*) were detected in all 5 years, three in 4 years, one in 2 years and seven in 1 year. For SPP, QTLs were located on chromosomes A2, A3, A9, A12, D1, D2, D6, D7, D9 and D13. Of the 11 QTLs identified during at least 1 year of the study, three (27.3%; *oSPP-A2-1*, *oSPP-A3-1* and *oSPP-D2-1*) were found in all 5 years, three in 4 years, one in 2 years and four in 1 year. Therefore, SPP had a higher proportion of stable ODO QTLs than that of LPP.

Boll number and boll weight

The distribution of ODO QTLs for boll number was significantly different than that of boll weight. Only 7 ODO QTLs were found to be associated with boll weight compared with 17 for boll number. The boll weight QTLs were located on five chromosomes (chr. nos. A9, A11, D6, D7, D12 and D13) vs. 13 chromosomes (chr. nos. A1, A3, A5, A8, A9, A12, D1, D3, D6, D10, D11, D12 and D13) for boll number. Specifically, of the 7n ODO QTLs, 6 ODO QTLs (*oBW-A9-2*, *oBW-D6-1*, *oBW-D7-1*, *oBW-D12-1*, *oBW-D12-2* and *oBW-D13-2*) were identified in all 5 years, which is much higher than the 4 of 17 (23.5%) ODO QTLs (*oBN-A5-1*, *oBN-A8-1*, *oBN-D12-1* and *oBN-D13-1*) found for boll number. Therefore, the ODO QTLs related to boll weight had a more stable performance over the 5-year study period.

Lint index and lint percentage

The ODO QTLs for the lint index and lint percentage exhibited ordinary performance. No ODO QTL for the lint index was found to be consistent over the 5-year period; only four were identified in four of the years, three in 3 years, two in 2 years and five in 1 year. The 14 ODO QTLs for the lint index resided on 12 chromosomes (chr. nos. A1, A3, A8, D1, D2, D3, D4, D6, D8, D11, D12 and D13). Similarly, a total of 11 ODO QTLs for lint percentage were anchored on 9 chromosomes (chr. nos. A2, A8, A9, A10, A11, D1, D10, D11 and D12). Among these, only one ODO QTL (*oLP-D11-2*) was consistently detected in each one of the 5 years, two in 4 years, two in 3 years, three in 2 years and three in 1 year.

Discussion

Overdominance is responsible for yield heterosis in interspecific hybrid cottons

Heterosis in hybrids is often associated with the genetic diversity present in their parents. For example, high genetic divergence in parent genotypes was found to be associated with the heterosis in first-generation hybrids of corn (Joshi et al. 2001). However, it is worth noting that other studies reported that the correlation between marker diversity and yield is marginally low or non-existent. Fluctuation in the occurrence of such a type of correlation largely depends on the type of crop and the nature of the germplasm used, at least partially (Semel et al. 2006). Because many segregating loci interact throughout a genome, it is difficult to separate the heterosis into components using the traditional F_2 , backcross, or recombinant inbred populations (refer to Li et al. 2001; Luo et al. 2001). To minimize the impact of epistasis, many traits were phenotyped in the ILs, carrying single chromosome segments derived from the extra-long staple species, *G. barbadense*. Heterosis was fractionated into defined genomic regions using the relatively isogenic population. In total, 396 QTLs for 11 yield-related and yield-unrelated traits, studied for 5 years, were mapped. These findings resolved the disparities found in the mode of inheritance between increasing QTLs for yield traits and reducing yield QTLs and QTLs for non-yield traits. We substantially eliminated the genome-wide epistasis in our population. The other major component of heterosis in this research is ODO as revealed by the detection of many ODO QTLs.

Theoretically, ODO can be achieved only if a single gene is in a heterozygous state. Despite the fact that many ODO QTLs have been reported in several genetic mapping experiments (Zhuang et al. 2000; Shang et al. 2016), only

few reports have verified the model of pseudo-overdominance responsible for heterosis (Jones 1917). This is the explicit imitation of the dominance model due to the interrelation of the involved mutations. As a result, there is no example supporting the overdominance of a single gene (Lippmand and Zamir 2007; Charlesworth and Willis 2009). However, in non-crop plants (Rédei 1962) and animals (Wallace 1957), some reports have indicated that overdominance may be caused by heterozygosity for single gene mutations (Muller and Falk 1961). To determine which mechanism leads to ODO phenotypes, yield and non-yield groups were compared, and it was revealed that both the groups contained many decreasing recessive and increasing dominant QTLs (Table 1). Therefore, if ODO phenotypes can result from pseudo-overdominance, suggesting that ODO phenotypes are caused by the linkage of decreasing recessive and increasing dominant QTLs in repulsion, then a similar representation of ODO QTLs for both yield and non-yield groups should be expected. In the present study, only one ODO QTL for non-yield traits was identified. These findings suggest that pseudo-overdominance resulting from randomly linked genes has not been a major contributor to the heterosis of ILH hybrids. Conversely, it may be responsible for ODO being exhibited by a single locus.

ODO in field crops

Our speculation that ODO is associated with yield QTLs in cotton interspecific ILs was supported by a similar trend reported in various groups of traits in corn (Wang et al. 2016), tomato (Semel et al. 2006) and rice (Zhou et al. 2012; Goff and Zhang 2013; Tao et al. 2016). For example, Lu et al. (2003) analysed four traits in two corn backcross populations. A total of 24 of 28 QTLs (86%) conferred ODO for yield according to the absolute value of d/a . The ODO level for the three other non-yield traits was much less, *i.e.*, only two ODO QTLs out of 16 for grain moisture, one out of eight for stalk lodging and four out of 11 for plant height were detected. The analyses of quantitative genetic and phenotypic effects were performed by Semel et al. (2006) on a population of ILs of tomato, who identified 841 QTLs for 35 diverse traits measured on homozygous and heterozygous crop plants. The ILs exhibited strong reproductive fitness, characterized by the common ubiquity of ODO QTLs, and almost no common QTL was found for non-reproductive traits. In the present study, a large number of traits were investigated and compared, which also made the comparison between reproductive traits and non-reproductive traits more accurate. However, we investigated and compared some important yield and non-yield traits in cotton. To compensate for the errors caused by the relatively small number of traits under

investigation, we repeated the experiments for 5 years to obtain valid comparison results. Finally, we reached a similar conclusion as was demonstrated in tomato (Semel et al. 2006): ILs have strong yield ability, which are characterized by the common ubiquity of ODO QTLs, and almost no common QTL exists for the non-yield traits. Therefore, we are of the opinion that ODO is ubiquitous in field crops, although some inconsistent findings exist due to fluctuations found in populations or environments (Li et al. 2008).

ODO in cotton breeding evolution

We suggest that ODO QTLs associated with yield traits played a major role in cotton domestication by early farmers and later through planned breeding practices. Yield-related traits, a product of the farming economy, placed more importance on breeding than fibre quality and other agronomic traits. During the domestication of cotton, initially, high-yielding plants were selected that paved the way for strong directional selection (Mell 1894) vs. stabilizing natural selection for non-yield traits during the course of evolution. Typically, stabilizing selection supports those individuals that exhibit phenotypic values near or at the mean, whereas plants exhibiting extreme phenotypes are at a selective disadvantage.

ODO QTLs contribute to yield improvement

G. hirsutum accounts for over 95% of the annual global cotton production and is characterized by its high-yield potential (Wang et al. 2012), which is difficult to improve solely by adopting commonly used intraspecific hybridization procedures. Fortunately, the barriers of stagnating yield can be overcome by exploiting the heterosis of the interspecific hybridization between *G. hirsutum* and *G. barbadense*. In our study, we found that an increase in yield was often accompanied by a large number of ODO QTLs. Several ODO QTLs were detected for the six yield-related traits investigated, some of which were consistently detected over the 5-year study period (Table 2), for example, the boll weight trait demonstrated consistent performance for 5 years under fluctuating environments. These findings have important implications in cotton breeding, whereby the boll weight can be easily improved by utilizing ODO QTLs via heterosis.

The ODO QTLs identified in the present study are consistent with some yield QTLs that have previously been reported. Taking lint yield as an example, *oLPP-A4-1* corresponds to *F_{2.3}-qLY-c4-2* on chr. A4 (Yu et al. 2011); *oLPP-A13-1* corresponds to *qLY-07A-c13-1* on chr. A13 (Yu et al. 2013); and *oLPP-D8-1* corresponds to *qLY-D8-1* on chr. D8 (Shen et al. 2007). It may be inferred that there is

great potential of these stable yield QTLs contributing to sustainable increases in cotton lint yield. These findings can augment the efforts of cotton breeders aimed at improving cotton productivity beyond the year 2030. Similarly, the same concept may be translated to other field crops for improving yield potential.

Data archiving

This article does not report new empirical data or software.

Acknowledgements This work was supported by National Key R & D Program for Crop Breeding project (2016YFD0101403), the Natural Science Foundation of Jiangsu Province (BK20150671), the Major Projects for Transgenic Breeding of China (2017ZX08005-004-006), the System of Modern Agriculture Industrial Technology (SDAIT-03-03/05) and the Natural Science Foundation (ZR2017MC057) of Shandong Province.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Allison AC (2009) Genetic control of resistance to human malaria. *Curr Opin Immunol* 21:499–505
- Balls WL (1908) Mendelian studies of Egyptian cotton. *J Agric Sci* 2:346–379
- Bruce AB (1910) The Mendelian theory of heredity and the augmentation of vigor. *Science* 32:627–628
- Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. *Nat Rev Genet* 10:783
- Chen ZJ (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci* 15:57–71
- Crow JF (1948) Alternative hypotheses of hybrid vigor. *Genetics* 33:477–487
- Davis DD (1979) Hybrid cotton: specific problems and potentials. *Adv Agron* 30:129–157
- East EM (1936) Heterosis. *Genetics* 21:375–397
- Goff SA, Zhang Q (2013) Heterosis in elite hybrid rice: speculation on the genetic and biochemical mechanisms. *Curr Opin Plant Biol* 16:221–227
- Hua XN (1963) Studies on heterosis exploitation of F₁ hybrids between *G. barbadense* and *G. hirsutum* L. *Acta Agron Sin* 2:1–27
- Jones DF (1917) Dominance of linked factors as a means of accounting for heterosis. *Proc Natl Acad Sci USA* 2:466–479
- Joshi SP, Bhavne SG, Chowdari KV, Apte GS, Dhonukshe BL, Lalitha K et al. (2001) Use of DNA markers in prediction of hybrid performance and heterosis for a three-line hybrid system in rice. *Biochem Genet* 39:179–200
- Krieger U, Lippman ZB, Zamir D (2010) The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nat Genet* 42:459–463
- Li L, Lu K, Chen Z, Mu T, Hu Z, Li X (2008) Dominance, overdominance and epistasis condition the heterosis in two heterotic rice hybrids. *Genetics* 180:1725–1742

- Li X, Li X, Fridman E, Tesso TT, Yu J (2015) Dissecting repulsion linkage in the dwarfing gene Dw3 region for sorghum plant height provides insights into heterosis. *Proc Natl Acad Sci USA* 112:11823–11828
- Li ZK, Luo LJ, Mei HW, Wang DL, Shu QY, Tabien R et al. (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* 158:1737–1753
- Lippman ZB, Zamir D (2007) Heterosis: revisiting the magic. *Trends Genet* 23:60–66
- Loden HD, Richmond TR (1951) Hybrid vigor in cotton: cytogenetic aspects and practical applications. *Econ Bot* 5:387–408
- Lu H, Romero-Severson J, Bernardo R (2003) Genetic basis of heterosis explored by simple sequence repeat markers in a random-mated maize population. *Theor Appl Genet* 107:494–502
- Luo LJ, Li ZK, Mei HW, Shu QY, Tabien R, Zhong DB et al. (2001) Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield. *Compon Genet* 158:1755–1771
- Mell PH (1894) Experiments in crossing for the purpose of improving the cotton fiber. *Ala Agri Exp Sta Bull* 56:18–20
- Meredith WR, Bridge RR (1984) Genetic contributions to yield changes in upland cotton. In: Fehr WR (ed) *Genetic contributions to yield gains of five major crop plants*. Crop Science Society of America, Madison, pp 75–87
- Meredith WRJ (1990) Yield and fiber-quality potential for second-generation cotton hybrids. *Pak J Biol Sci* 6:1045–1048
- Meyer VG (1969) Some effects of genes, cytoplasm, and environment on male sterility of cotton (*Gossypium*). *Crop Sci* 9:237–242
- Muller HJ, Falk R (1961) Are induced mutations in *Drosophila* overdominant? II. Experimental results. *Genetics* 46:727–735
- Rédei GP (1962) Single locus heterosis. *Mol Genet Genom* 93:164–170
- Richey FD, Sprague GF (1931) Experiments on hybrid vigor and convergent improvement in corn. *Tech Bull* 267:18–21
- Semel Y, Nissenbaum J, Menda N, Zinder M, Krieger U, Issman N et al. (2006) Overdominant quantitative trait loci for yield and fitness in tomato. *Proc Natl Acad Sci USA* 103:12981–12986
- Shang L, Liang Q, Wang Y, Zhao Y, Wang K, Hua J (2016) Epistasis together with partial dominance, over-dominance and QTL by environment interactions contribute to yield heterosis in upland cotton. *Theor Appl Genet* 129:1429–1446
- Shen X, Guo W, Lu Q, Zhu X, Yuan Y, Zhang T (2007) Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. *Euphytica* 155:371–380
- Tao Y, Zhu J, Xu J, Wang L, Gu H, Zhou R et al. (2016) Exploitation of heterosis loci for yield and yield components in rice using chromosome segment substitution lines. *Sci Rep* 6:36802
- Wallace B (1957) The effect of heterozygosity for new mutations on viability in *Drosophila*: a preliminary report. *Proc Natl Acad Sci USA* 43:404–407
- Wang H, Zhang X, Yang H, Liu X, Li H, Yuan L et al. (2016) Identification of heterotic loci associated with grain yield and its components using two CSSL test populations in maize. *Sci Rep* 6:38205
- Wang P, Zhu Y, Song X, Cao Z, Ding Y, Liu B et al. (2012) Inheritance of long staple fiber quality traits of *Gossypium barbadense* in *G. hirsutum* background using CSILs. *Theor Appl Genet* 124:1415–1428
- Yu SB, Li JX, Xu CG, Tan YF, Gao YJ, Li Xh et al. (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc Natl Acad Sci USA* 94:9226–9231
- Yu J, Yu S, Gore M, Wu M, Zhai H, Li X et al. (2011) Identification of quantitative trait loci across interspecific F₂, F_{2,3} and testcross populations for agronomic and fiber traits in tetraploid cotton. *Euphytica* 191:375–389
- Yu J, Zhang K, Li S, Yu S, Zhai H, Wu M et al. (2013) Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* × *Gossypium barbadense* backcross inbred line population. *Theor Appl Genet* 126:275–287
- Zhang TZ, Pan JJ (1999) Hybrid seed production in cotton. In: Basra AS (ed) *Heterosis and hybrid seed production in agronomic crops*. Food Products Press, New York, pp 149–175
- Zhou G, Chen Y, Yao W, Zhang C, Xie W, Hua J et al. (2012) Genetic composition of yield heterosis in an elite rice hybrid. *Proc Natl Acad Sci USA* 109:15847–15852
- Zhuang J, Fan Y, Wu J, Xia Y, Zheng K (2000) Identification of over-dominance QTL in hybrid rice combinations. *Hereditas* 22:205–208