1 **Genetic Components of Heterosis for Seedling Traits in an Elite Rice Hybrid**
2 **Analyzed Using an Immortalized F₂ Population**

3
4 Dan Zhu, Gang Zhou, Caiguo Xu, Qifa Zhang*

5
6 National Key Laboratory of Crop Genetic Improvement, National Center of Plant Gene
7 Research (Wuhan), Huazhong Agricultural University, Wuhan 430070, China

8
9 *Corresponding author (Fax: +86-27-8728 7092; Phone: +86-27-8728 2429; e-mail:
10 qifazh@mail.hzau.edu.cn)

11
12
13 **Manuscript information**

14 A total of 19 pages of the text, 3 figures, 7 tables and 5 supplementary tables.

15 **Word and character counts**

16 A total of 247 words in the abstract and 5317 characters in the paper.

17
18 **Abbreviations**

19 AA, additive by additive interaction; AD, additive by dominance interaction; DA,
20 dominance by additive interaction; DD, dominance by dominance interaction; IMF₂,
21 immortalized F₂ population; ND, negative dominance; NILs, near isogenic lines, NO,
22 negative overdominance; PD, positive dominance; PO, positive overdominance; RILs,
23 recombinant inbred lines; SNP, single nucleotide polymorphism.

Abstract

Utilization of heterosis has greatly contributed to rice productivity in China and many Asian countries. Superior hybrids usually show heterosis at two stages: canopy development at vegetative stage and panicle development at reproductive stage resulting in heterosis in yield. Although the genetic basis of heterosis in rice has been extensively investigated, all the previous studies focused on yield traits at maturity stage. In this study, we analyzed the genetic basis of heterosis at seedling stage making use of an “immortalized F₂” population composed of 105 hybrids produced by intercrossing recombinant inbred lines (RILs) from a cross between Zhenshan 97 and Minghui 63, the parents of Shanyou 63, which is an elite hybrid widely grown in China. Eight seedling traits, seedling height, tiller number, leaf number, root number, maximum root length, root dry weight, shoot dry weight and total dry weight, were investigated using hydroponic culture. We analyzed single-locus and digenic genetic effects at the whole genome level using an ultrahigh-density SNP bin map obtained by population re-sequencing. The analysis revealed large numbers of heterotic effects for seedling traits including dominance, overdominance and digenic dominance (epistasis) in both positive and negative directions. Overdominance effects were prevalent for all the traits, and digenic dominance effects also accounted for a large portion of the genetic effects. The results suggested that cumulative small advantages of the single-locus effects and two-locus interactions, most of which could not be detected statistically, could explain the genetic basis of seedling heterosis of the F₁ hybrid.

Keywords: Heterosis, seedling trait, immortalized F₂, overdominance, digenic dominance

Introduction

Utilization of heterosis has greatly contributed to rice productivity in China and many Asian countries as a result of widespread adoption of hybrid rice. Superior hybrids of rice usually show heterosis at two stages: canopy development at vegetative stage providing the biomass basis for yield, and panicle development at reproductive stage, which comprises a major component of yield. A number of previous studies focused on genetic investigation of heterosis in yield and yield component traits in rice evaluated at mature stage, which produced variable results regarding the genetic components and their relative importance to heterosis depending on the materials, the genetic designs of the experimental populations and the methods of analysis used in the studies (Xiao et al., 1995; Yu et al., 1997; Li et al., 2001; Luo et al., 2001, 2012; Hua et al., 2002, 2003; Li et al., 2008; Xin et al., 2011; Zhou et al., 2012). Heterosis has been shown to be a widespread phenomenon in many rice hybrids for various traits at seedling stage (Zhang et al., 2008; Ma et al., 2011; Zhai et al., 2013). However, there has not been reported study on the genetics of heterosis in rice at vegetative stage, although such studies have been reported in maize (Tollenaar et al., 2004; Hoecker et al., 2006; Frascaroli et al., 2007) and *Arabidopsis thaliana* (Barth et al., 2003; Kearsey et al., 2003; Meyer et al., 2004, 2010; Syed and Chen, 2005; Kusterer et al., 2007a, 2007b; Melchinger et al., 2007), revealing that heterosis for biomass and biomass-related traits in early seedling stage involves multiple genetic components including dominance, overdominance and epistasis.

Shanyou 63, a cross between Zhenshan 97 and Minghui 63, is an elite rice hybrid that has been widely cultivated in the last three decades in China. A series of studies have been conducted to learn the genetic basis of heterosis in yield and yield component traits of this hybrid including genetic analysis of an $F_{2:3}$ population (Yu et al., 1997) and an “immortalized F_2 ” (IMF2) population (Hua et al., 2002, 2003), which was derived from random crossing between recombinant inbred lines (RILs) (Xing et al., 2002). Xie et al. (2010) and Yu et al. (2011) genotyped the RILs by population sequencing, resulting in an ultrahigh-density linkage map composed of 1619 bins. Using this map, Zhou et al. (2012) assessed the relative contributions of the genetic

components to heterosis of grain yield per plant and yield component traits including tiller number per plant, grain number per panicle and grain weight. The results showed that single-locus dominance had relative small contributions in all the traits, while overdominance is predominant in heterosis of yield, grain number and grain weight. Dominance by dominance interaction expressed as digenic dominance was important in heterosis of tiller number per plant and grain weight, and also has roles in yield and grain number.

In this study, we investigated the genetic basis of heterosis in the hybrid Shanyou 63 at the seedling stage using the same IMF2 population. The analysis also revealed the prevalence of overdominance and dominance by dominance interactions in the genetic basis of heterosis in seedling traits.

Results

1. The genotypes of IMF2 population and heterosis of the seedling traits

The genotype of each cross in the IMF2 deduced from the RILs is presented in Fig. 1 and Table S1. Theoretically, the composition and frequencies of various genotypes of this population are similar to an F_2 population, allowing for detection of genetic components such as additive and dominance effects at single-locus level and interactions of multiple loci.

The measurements of seedling traits (Table 1) showed that there was statistically significant mid-parent heterosis and high-parent heterosis in all the seedling traits except seedling height, when the F_1 was compared to the two parents, Minghui 63 and Zhenshan 97, which is consistent to previous results (Ma et al., 2011). In particular, root number, root dry weight and shoot dry weight thus also total dry weight showed higher heterosis than other traits. Relative to Minghui 63, Zhenshan 97 was the high-value parent for all the traits except seedling height. This is different from the results observed at adult stage, in which Minghui 63 was the high-value parent for yield and yield component traits (Yu et al., 1997; Hua et al., 2002). Such complementary patterns of the parental performance and also the resulting heterosis in the hybrids during vegetative and reproductive stages provide important basis for the

superior performance of the hybrids relative to the parents throughout the life cycle of the rice plant. Transgressive segregation was observed in all the traits in both low and high ends and in both IMF2 and RIL populations (Table 1).

2. Genome-wide single-locus genetic effects

Additive and dominance effects of the bins were estimated for each trait making use of the ultrahigh-density bin map (Fig. 2 and Table S2). Both positive and negative dominance effects were prevalent across the genome in all the seedling traits. More bins showed positive dominance in leaf number, tiller number, shoot dry weight and total dry weight, while negative dominance effects were slightly more frequent in seedling height, maximum root length, root number, and root dry weight. Percentage of bins showing positive dominance (including overdominance) ranged from a high of 56.9% in leaf number to a low of 45.4% in seedling height (Table 2).

Degree of dominance effects was measured by the ratio of dominance to the absolute value of additive effects ($D/|a|$), as partial dominance (<1), full dominance ($=1$), and overdominance (>1), in both positive and negative directions. For all the seedling traits, the numbers of bins showing overdominance were slightly higher than dominance and partial dominance. The percentage of bins exhibiting positive overdominance ranged from 24.4% in seedling height to 32.9% in shoot dry weight. Negative overdominance was observed from 23.8% of the bins in leaf number to 33.0% of the bins in seedling height. Slightly more positive overdominance than negative overdominance effects were detected for all the seedling traits except seedling height (Table 2).

The total number of bins showing significant additive effects varied among the traits (Tables 2 and S3). Analysis of the bins by the *h*-test revealed that many of the bins showing partial or full dominance did not reach the levels for statistical significance. In particular, none of the bins with dominance effects reached the significance levels for seedling height, leaf number, tiller number or root dry weight (Tables 2 and S2). These results indicated that numbers of bins with significant additive and overdominance effects were predominant, while only a small number of

bins showed significant dominance effects for the seedling traits.

The bins with significant genetic effects were not randomly distributed in the genome (Fig. 3). We suspect that many adjacent bins showing significant effects might result from the same bins, which would lead to overestimation of the numbers and the effects of the bins. The merge approach used by Zhou et al. (2012) was employed to resolve the effects of adjacent heterotic bins, which involved dominance/overdominance effects. After the merging process, the numbers of bins showing negative overdominance were slightly higher than ones showing positive overdominance for seedling height, root number, root dry weight, shoot dry weight and total dry weight, while more bins showed positive than negative overdominance in tiller number, leaf number and maximum root length (Table 2).

3. Digenic interaction of the seedling traits in the entire genome

Two-way ANOVA analyses with permutation tests showed that significant interactions were widespread but varied in distribution for different traits. The numbers of significant interactions were larger for traits related to root morphology, such as root number, maximum root length and root dry weight. The number of interactions for tiller number, shoot dry weight and total dry weight was intermediate. Seedling height and leaf number showed the least numbers of two-locus interactions (Table 3).

We followed the approaches used in a previous study (Zhou et al., 2012) to reduce possible overestimation of the epistasis effects: reducing possible false positive effects by chance with permutations, and reducing redundancy resulting from adjacent two-locus combinations by the merging process. The number of significant interactions for each trait was greatly reduced (Table 3), and the epistatic interactions that remained might represent the minimum numbers of epistatic effects.

We partitioned each interaction into four terms through orthogonal contrast tests: additive (first locus) $\times$ additive (second locus) (AA), additive $\times$ dominance (AD), dominance $\times$ additive (DA), and dominance $\times$ dominance (DD), at the 0.05 significance level (Table S4). As observed in previous studies of similar analyses (Yu et al., 1997; Hua et al., 2002; Zhou et al., 2012), AA occurred at predominantly high

frequencies, DD interactions were the least frequent, AD and DA occurred with intermediate frequencies (Table 3).

4. Pleiotropic effects of single-locus and digenic dominance

A total of sixteen bins were observed to show significant dominance effects on five or more traits (Table 4) likely due to pleiotropic effects, of which nine bins showed positive effects and seven bins showed negative effects. Most of the bins exhibited similar gene actions on the traits (i.e., PO or NO), except bin233 that showed negative overdominance (or dominance) for all traits. Similarly, bin1394 showed positive overdominance for all traits except shoot dry weight and total dry weight (Table 4).

We also found four pairs of bins showing significant DD interactions that simultaneously impacted several traits (Table 5), suggesting that DD interactions might also have pleiotropic effects especially on biomass and biomass-related traits.

5. Contributions of the genetic components to heterosis of the hybrid

As argued by Zhou et al. (2012), only the effects involving heterozygosity would be relevant to heterosis of hybrid F_1 , since all the loci that are polymorphic between the parents would be heterozygous in the F_1 . This included dominance (including partial, full and overdominance) at the single-locus level, and DD interactions between loci in the present analysis. We estimated the relative contributions of these genetic components to the seedling heterosis of F_1 hybrid for each trait, calculated as the summation of the significant effects for each trait over the bins across the genome divided by the trait mean (Table 6). The single-locus effect was not included in the calculation if it was involved in a DD interaction.

The estimated total relative heterotic genetic effects for root number and leaf number, 4.55 and 1.60, respectively, were larger than those of other traits (Table 6). In general overdominance effects, both positive and negative, were predominant for all the traits, while dominance effects were rarely detected. Digenic dominance effects, also both positive and negative, were also large, which suggested that cumulative small advantages of two-locus interactions also accrued to the heterosis of the F_1 hybrid. The

widespread negative overdominance and digenic dominance indicated that heterozygosity was not necessarily always favorable for the expression of the trait even in highly heterotic hybrids. In summary, our results indicated that overdominance and digenic dominance involving large numbers of loci play important roles in rice seedling heterosis.

We also noted that the relative magnitudes of these estimates (Table 6) did not seem to correspond well with the observed mid-parent heterosis for the traits (Table 2). Two reasons may account for such discrepancies. First, cancelation between positive and negative effects of the dominance, overdominance and digenic dominance of each trait resulted in small positive net effect on the trait. And second the majority of the dominance effects that were widespread in the genome did not reach the thresholds for statistical significance thus were not included in the final calculation. Both of the cases pose great challenges for characterizing and quantification of the genetic components underlying heterosis.

Discussion

We performed a comprehensive dissection of genetic effects of heterosis for seedling traits of an elite rice hybrid using an IFM2 population derived from RILs from a cross between the parents of the best rice hybrid in China. The results revealed various heterotic effects for seedling traits in the whole genome including the dominance, overdominance and epistasis in both positive and negative directions. One important observation was a preponderance of evidence for the role of overdominance and epistasis in heterosis for all the seedling traits, with partial dominance seemingly playing a less important role. In *A. thaliana*, overdominance and epistasis were also found to be important in the manifestation of heterosis for biomass and biomass-related traits (Kusterer et al., 2007b; Melchinger et al., 2007; Meyer et al., 2010). In maize, a study showed that the average degree of dominance for seedling weight was >1 , which indicated likely involvement of overdominance (Frascaroli et al., 2007). The results of Semel et al. (2006) in a tomato study showed that most of the reproductive traits related to seed and fruit yield exhibited overdominance, while nonreproductive

traits like morphological characteristics had lower levels of overdominance. Our results indicate that overdominance effects were also prevalent in non-reproductive traits. It should be noted however that the apparent prevalence of overdominance relative to dominance may represent a statistical artifact of the *h*-test. This is because the majority of the partial dominance effects did not reach the thresholds for statistical significance although they were undoubtedly prevalent in the genome, and made considerable contributions to the genetic basis of heterosis, which should not be neglected. Zhou et al. (2012) discussed in detail the perplex between the apparent prevalence of overdominance in reproductive (yield) traits in IMF2 derived from the same RIL population and the rarity of overdominance in crosses of near isogenic lines (NILs) and cloned genes. Here we also observed the prevalence of overdominance in the seedling traits. The same puzzle may also remain although both NILs and cloned genes are rare for seedling traits in order to make similar comparison. We compared the bins with significant pleiotropic effects identified here with the results of Zhou et al. (2012) for heterosis for yield and yield component traits, not many bins were found overlapping between the two analyses (Table 7).

Another noticeable point is that the relative phenotypic values of the two parents Zhenshan 97 and Minghui 63. Zhenshan 97 is the higher parent for all the seedling traits except seedling height, which indicated a faster growth rate and larger biomass accumulation in Zhenshan 97 at the seedling stage. The lack of heterosis in seedling height could be explained by the results that a larger number of negative overdominance loci were detected than ones showing positive effects. Overall these results were in agreement with the previous studies that Zhenshan 97 was the high-value parent for seedling traits (Cui et al., 2002, 2008; Xu et al., 2004). In contrast, Minghui 63 is the high-value parent for yield and yield component traits as well as a large array of traits observed at the maturity stage (Yu et al., 1997; Hua et al., 2002; Xing et al., 2002). This has very important implication to the strong heterosis of the hybrid Shanyou 63: The genetic contribution of Zhenshan 97 plays an important role for the fast early growth of the hybrid for establishing the canopy, while the large stature of the plant and sizes of the yield traits of Minghui 63 provide the basis for the

yield level. In combination, heterosis at the two stages makes Shanyou 63 a successful hybrid.

Materials and methods

1. Plant materials and population design

The IMF2 population consisting of 105 crosses was derived by pairwise hybridization of 210 RILs from a cross between Zhenshan 97 and Minghui 63 (Xing et al., 2002). In constructing the IMF2 population, the 210 RILs were randomly divided into two groups, and each RIL in the first group was paired with a RIL in the second group (Table S5). Because of very different flowering time, the RILs were sown at intervals of 7 days, and transplanted in pairs according to the mating design. The crosses were made in two consecutive growing seasons in Wuhan (summer, 2010) and Lingshui (winter, 2011) in China.

For phenotyping, the IMF2 population, RILs, the parental lines Zhenshan 97 and Minghui 63, and the F_1 were grown hydroponically. Approximately 150 seeds of each line were soaked in tap water for two days, and allowed to germinate for one day. The germinated seeds were sown in large plastic boxes filled with sand. Ten days after sowing in the sand bed, 40 healthy seedlings of each line were selected for transplanting into concrete frames in eight large pools each with a volume of 3540 L of nutrient solution, prepared according to the protocol (Yoshida et al., 1976). Fresh nutrient solution was added every week with pH adjusted, and completely refreshed after two weeks. There were 80 concrete frames in a pool with 14 holes in each frame, so that the seedlings were spaced 10 cm apart within each frame and 15 cm between seedlings in different frames. The seedling was anchored in the frame by wrapping with sponge.

The lay out of the seedlings followed a randomized complete block design with two replications. Each replication consisted of 106 plots with four rows of 10 seedlings each, two rows for each hybrid and one row for each of the respective parents (RILs). True hybrid seedlings of each cross were identified with seven pairs of InDel markers after two weeks of hydroponic culture. After 30 days of hydroponic growth (so that the

total growth period of seedling is 40 days, including 10 days of sand culture), five healthy seedlings of each genotype in each plot were individually measured for the following traits: seedling height, tiller number, leaf number, root number, maximum root length, shoot dry weight, root dry weight and total dry weight. All the traits were measured in one day, except shoot and root dry weight that were measured after oven drying at 70 °C for three days. The average of the five seedlings was used as the score of the trait in each plot. Trait measurements averaged over the two replications were used as the raw data in the analyses.

2. Genotyping

The analysis was based on the ultrahigh-density bin map consisting of 1619 bins (Xie et al., 2010; Yu et al., 2011). The genotype of each cross in the IMF2 population was deduced from the genotypes of their RIL parents (Table S1).

3. Data analyses

3.1 Analysis of single-locus effects

For each trait of the IMF2 population, the additive effect of each bin was calculated as half of the difference between the means of the two homozygotes, and the dominance effect was the difference between the value of the heterozygote and the means of the two homozygotes. The bins with significant effects of any kind among three genotypes were identified by one-way ANOVA, with a threshold of F value of 3.084 ($P = 0.05$), followed by 1000 permutation tests. Bins showing significant additive effects were identified by t -test between the two homozygotes of each locus. Significant ($P \leq 0.05$) dominance effect of each bin was identified by an h test (Huang et al., 2006),

$$h = [F_1 - (P_1 + P_2)/2] / \sqrt{V_{F_1}/n_{F_1} + \frac{1}{4}(V_{P_1}/n_{P_1} + V_{P_2}/n_{P_2})}$$
, where F_1 is the mean of the heterozygote, and P_1 and P_2 are, respectively, the means of the two homozygotes. n_{F_1} , n_{P_1} and n_{P_2} are the numbers of individuals that were heterozygous and homozygous for the two parental genotypes respectively. V_{F_1} , V_{P_1} and V_{P_2} are the variances of phenotypic data of the three genotypes. This test is essentially similar to

the *t* test.

Similar genetic effects detected in adjacent bins likely resulted from the same factor and thus were grouped into a cluster using a merging process. In doing so, the most significant bin of each group was chosen as the core locus, and genetic effects of the bins next to the core locus was calculated sequentially using the information from the recombinants between the two bins. The adjacent bin and the core locus were regarded as one cluster if the genetic effects of the adjacent bin were not significant, or otherwise they belonged to different clusters. The dominance value from the bin with lowest *P*-value of each cluster was taken for the cluster (Zhou et al., 2012).

3.2 Analysis of digenic interactions

Digenic interactions for seedling traits in the whole genome were analyzed employing two-way ANOVA. The whole genome was searched for the effects of digenic interactions for each trait, using all possible two-locus (bin) combinations in the IMF2 population. Only two-locus data sets with all the nine genotypes containing three or more crosses in the IMF2 population were included in the calculation. The calculation was based on unweighted cell means (Snedecor and Cochran, 1980) and the sums of squares were multiplied by the harmonic means of the cell sizes to form the test criteria. Significant digenic interactions (two bins) were identified at $P \leq 0.001$ followed by 10,000 permutations. To eliminate the possibility that the digenic interactions involving adjacent bins might result from the same interaction, we followed the approach of Zhou et al. (2012) to merge such adjacent interactions. Finally, all the significant digenic interactions were partitioned into components: additive by additive effect (AA), additive by dominance effect (AD), dominance by additive effect (DA) and dominance by dominance effects (DD) using an orthogonal contrast test (Cockerham, 1954) at $P \leq 0.05$. Digenic dominance (DD) effects were measured as the advantage of double heterozygote relative to the means of the parental homozygotes.

Acknowledgments

346 This research was supported by grants from the National Natural Science Foundation
347 of China (No. 31330039), National 863 Project (No. 2014AA10A604) and the Bill &
348 Melinda Gates Foundation.

References

- Barth, S., Busimi, A.K., Friedrich Utz, H., Melchinger, A.E., 2003. Heterosis for biomass yield and related traits in five hybrids of *Arabidopsis thaliana* L. Heynh. Heredity 91, 36-42.
- Cui, H., Peng, B., Xing, Z., Xu, G., Yu, B., Zhang, Q., 2002. Molecular dissection of seedling-vigor and associated physiological traits in rice. Theor. Appl. Genet. 105, 745-753.
- Cui, K., Huang, J., Xing, Y., Yu, S., Xu, C., Peng, S., 2008. Mapping QTLs for seedling characteristics under different water supply conditions in rice (*Oryza sativa*). Physiol. Plant. 132, 53-68.
- Frascaroli, E., Cane, M.A., Landi, P., Pea, G., Gianfranceschi, L., Villa, M., Morgante, M., Pe, M.E., 2007. Classical genetic and quantitative trait loci analyses of heterosis in a maize hybrid between two elite inbred lines. Genetics 176, 625-644.
- Hoecker, N., Keller, B., Piepho, H.P., Hochholdinger, F., 2006. Manifestation of heterosis during early maize (*Zea mays* L.) root development. Theor. Appl. Genet. 112, 421-429.
- Hua, J., Xing, Y., Wu, W., Xu, C., Sun, X., Yu, S., Zhang, Q., 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, 2574-2579.
- Hua, J.P., Xing, Y.Z., Xu, C.G., Sun, X.L., Yu, S.B., Zhang, Q.F., 2002. Genetic dissection of an elite rice hybrid revealed that heterozygotes are not always advantageous for performance. Genetics 162, 1885-1895.
- Huang, Y., Zhang, L., Zhang, J., Yuan, D., Xu, C., Li, X., Zhou, D., Wang, S., Zhang, Q., 2006. Heterosis and polymorphisms of gene expression in an elite rice hybrid as revealed by a microarray analysis of 9198 unique ESTs. Plant Mol. Biol. 62, 579-591.
- Kearsey, M.J., Pooni, H.S., Syed, N.H., 2003. Genetics of quantitative traits in *Arabidopsis thaliana*. Heredity 91, 456-464.
- Kusterer, B., Muminovic, J., Utz, H.F., Piepho, H.P., Barth, S., Heckenberger, M.,

- Meyer, R.C., Altmann, T., Melchinger, A.E., 2007a. Analysis of a triple testcross design with recombinant inbred lines reveals a significant role of epistasis in heterosis for biomass-related traits in *Arabidopsis*. *Genetics* 175, 2009-2017.
- Kusterer, B., Piepho, H.P., Utz, H.F., Schon, C.C., Muminovic, J., Meyer, R.C., Altman, T., Melchinger, A.E., 2007b. Heterosis for biomass-related traits in *Arabidopsis* investigated by quantitative trait loci analysis of the triple testcross design with recombinant inbred lines. *Genetics* 177, 1839-1850.
- Li, L.Z., Lu, K.Y., Chen, Z.M., Mu, T.M., Hu, Z.L., Li, X.Q., 2008. Dominance, overdominance and epistasis condition the heterosis in two heterotic rice hybrids. *Genetics* 180, 1725-1742.
- Li, Z.K., Luo, L.J., Mei, H.W., Wang, D.L., Shu, Q.Y., Tabien, R., Zhong, D.B., Ying, C.S., Stansel, J.W., Khush, G.S., Paterson, A.H., 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.
- Luo, L.J., Li, Z.K., Mei, H.W., Shu, Q.Y., Tabien, R., Zhong, D.B., Ying, C.S., Stansel, J.W., Khush, G.S., Paterson, A.H., 2001. Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics* 158, 1755-1771.
- Luo, X.J., Xin, X.Y., Yang, J.S., 2012. Genetic-basis analysis of heterotic loci in Dongxiang common wild rice (*Oryza rufipogon* Griff.). *Genet. Res.* 94, 57-61.
- Ma, Q., Hedden, P., Zhang, Q., 2011. Heterosis in rice seedlings: its relationship to gibberellin content and expression of gibberellin metabolism and signaling genes. *Plant Physiol.* 156: 1905-1920.
- Melchinger, A.E., Piepho, H.P., Utz, H.F., Muminovic, J., Wegenast, T., Torjek, O., Altmann, T., Kusterer, B., 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, 1827-1837.
- Meyer, R.C., Kusterer, B., Lisec, J., Steinfath, M., Becher, M., Scharr, H., Melchinger, A.E., Selbig, J., Schurr, U., Willmitzer, L., Altmann, T., 2010. QTL analysis of early stage heterosis for biomass in *Arabidopsis*. *Theor. Appl. Genet.* 120,

227-237.

Meyer, R.C., Torjek, O., Becher, M., Altmann, T., 2004. Heterosis of biomass production in *Arabidopsis*. Establishment during early development. Plant Physiol. 134, 1813-1823.

Semel, Y., Nissenbaum, J., Menda, N., Zinder, M., Krieger, U., Issman, N., Pleban, T., Lippman, Z., Gur, A., Zamir, D., 2006. Overdominant quantitative trait loci for yield and fitness in tomato. Proc. Natl. Acad. Sci. USA 103, 12981-12986.

Snedecor, G.W., Cochran, W.G., 1980. Statistical Methods. Ames, Iowa State College Press.

Syed, N.H., Chen, Z.J., 2005. Molecular marker genotypes, heterozygosity and genetic interactions explain heterosis in *Arabidopsis thaliana*. Heredity 94, 295-304.

Tollenaar, M., Ahmadzadeh, A., Lee, E.A., 2004. Physiological basis of heterosis for grain yield in maize. Crop Sci. 44, 2086-2094.

Xiao, J., Li, J., Yuan, L., Tanksley, S.D., 1995. Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. Genetics 140, 745-754.

Xie, W., Feng, Q., Yu, H., Huang, X., Zhao, Q., Xing, Y., Yu, S., Han, B., Zhang, Q., 2010. Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing. Proc. Natl. Acad. Sci. USA 107, 10578-10583.

Xin, X.Y., Wang, W.X., Yang, J.S., Luo, X.J., 2011. Genetic analysis of heterotic loci detected in a cross between *indica* and *japonica* rice (*Oryza sativa* L.). Breed. Sci. 61, 380-388.

Xing, Y.Z., Tan, Y.F., Hua, J.P., Sun, X.L., Xu, C.G., Zhang, Q., 2002. Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. Theor. Appl. Genet. 105, 248-257.

Xu, C.G., Li, X.Q., Xue, Y., Huang, Y.W., Gao, J., Xing, Y.Z., 2004. Comparison of quantitative trait loci controlling seedling characteristics at two seedling stages using rice recombinant inbred lines. Theor. Appl. Genet. 109, 640-647.

Yoshida, S., Forno, D.A., Cock, J., 1976. Laboratory manual for physiological studies

439 of rice, 3 ed, International Rice Research Institute, Manila, Philippines.

440 Yu, H., Xie, W., Wang, J., Xing, Y., Xu, C., Li, X., Xiao, J., Zhang, Q., 2011. Gains in

441 QTL detection using an ultra-high density SNP map based on population

442 sequencing relative to traditional RFLP/SSR markers. PLoS One 6, e17595.

443 Yu, S.B., Li, J.X., Xu, C.G., Tan, Y.F., Gao, Y.J., Li, X.H., Zhang, Q., Saghai Maroof,

444 M.A., 1997. Importance of epistasis as the genetic basis of heterosis in an elite

445 rice hybrid. Proc. Natl. Acad. Sci. USA 94, 9226-9231.

446 Zhai, R., Feng, Y., Wang, H., Zhan, X., Shen, X., Wu, W., Zhang, Y., Chen, D., Dai, G.,

447 Yang, Z., Cao, L., Cheng, S., 2013. Transcriptome analysis of rice root heterosis

448 by RNA-Seq. BMC Genomics 14, 19.

449 Zhang, H.Y., He, H., Chen, L.B., Li, L., Liang, M.Z., Wang, X.F., Liu, X.G., He, G.M.,

450 Chen, R.S., Ma, L.G., Deng, X.W., 2008. A genome-wide transcription analysis

451 reveals a close correlation of promoter INDEL polymorphism and heterotic gene

452 expression in rice hybrids. Mol. Plant 1, 720-731.

453 Zhou, G., Chen, Y., Yao, W., Zhang, C., Xie, W., Hua, J., Xing, Y., Xiao, J., Zhang, Q.,

454 2012. Genetic composition of yield heterosis in an elite rice hybrid. Proc. Natl.

455 Acad. Sci. USA 109, 15847-15852.

Figure legends

Fig. 1. Genotypes of the 1619 bins for the 105 crosses in the immortalized F₂ population.

The horizontal axis is the physical position of each bin, and vertical axis is the serial number of each cross. Red indicates Zhenshan 97 genotype; blue indicates Minghui 63 genotype; green indicates heterozygous genotype.

Fig. 2. Genome-wide additive and dominance effects for the seedling traits.

The horizontal axis is bin position on the 12 chromosomes, the green dashed line indicates the end of each chromosome. The vertical axis shows the values of the additive and dominance effects. Blue line represents additive effect, and red line represents dominance effect.

Fig. 3. Dominance and overdominance effects for seedling traits at whole genome level.

The horizontal axis is bin position on the 12 chromosomes, the green dashed line indicates the end of each chromosome. The vertical axis indicates value of dominance effect of each bin. Red columns represent bins showing overdominance, orange columns indicate bins showing dominant effects, and the black columns show the effects that does not reach statistical significance.

Fig 1

Figure 2

Figure 3

1 **Tables**

2 Table 1. Trait measurements of the immortalized F₂ population, RIL population and the parent heterosis of F₁ hybrid.

Trait	Immortalized F ₂		RIL		Zhenshan	Minghui	Shanyou	Midparent
	Range	Mean ± σ	Range	Mean ± σ	97	63	63	heterosis (%)
Seedling height (cm)	38.9–61.3	49.8±3.8	39.8–66.0	48.6± 4.4	48.9	54.4**	52.2	1.1
Leaf number	25.2–61.7	44.8±7.2	20.3–60.1	40.2±7.2	50.9	46.3**	55.9	15.0**
Tiller number	5.6–15.2	11.5±2.0	5.8–14.9	10.4±2.0	12.6	12.1*	13.7	9.9**
Maximum root length (cm)	17.0–25.8	20.9±1.6	12.8–24.3	20.5±1.9	19.5	17.8**	20.2	7.9*
Root number	44.4–131.9	71.5±16.1	36.0–115.7	67.6±14.4	78.0	68.2**	95.6	23.5**
Root dry weight (g)	0.338–0.760	0.501±0.085	0.175–0.650	0.474±0.075	0.555	0.477**	0.647	20.3**
Shoot dry weight (g)	1.552–3.930	2.712±0.532	0.978–3.690	2.455±0.478	2.650	2.438**	3.395	25.1**
Total dry weight (g)	1.918–4.601	3.213±0.606	1.152–4.268	2.929±0.534	3.205	2.915**	4.042	24.2**

3 * and ** indicate significant differences at $P \leq 0.05$ and ≤ 0.01 , respectively.

4 Table 2. Numbers of bins showing additive, dominant and overdominant effects detected in the whole genome, with *h*-test and merging process
5 for the seedling traits.

Trait	Additive	Positive overdominant			Positive dominant			Negative overdominant			Negative dominant		
	<i>t</i> -test ^a	Whole genome	<i>h</i> -test ^b	After merging	Whole genome	<i>h</i> -test ^b	After merging	Whole genome	<i>h</i> -test ^b	After merging	Whole genome	<i>h</i> -test ^b	After merging
Seedling height (cm)	203	396	61	25	339	0	0	534	114	33	350	0	0
Leaf number	247	528	138	49	393	0	0	385	81	22	313	0	0
Tiller number	247	518	130	49	341	0	0	459	78	36	301	0	0
Maximum root length (cm)	350	432	125	47	364	1	1	401	87	42	422	16	5
Root number	177	499	120	41	302	0	0	444	111	47	374	2	1
Root dry weight (g)	204	426	72	23	349	0	0	466	81	35	378	0	0
Shoot dry weight (g)	247	533	93	28	330	1	1	410	76	34	346	5	5
Total dry weight (g)	244	518	95	29	345	1	1	432	88	40	324	5	3

6 a and b indicate significant differences at $P \leq 0.05$.

7 Table 3. Significant digenic interactions of the seedling traits ($P \leq 0.001$).

Trait	Seedling height (cm)	Leaf number	Tiller number	Maximum root length (cm)	Root number	Root dry weight (g)	Shoot dry weight (g)	Total dry weight (g)
Whole genome level	960	805	1515	2407	2856	1985	1217	1263
Permutation test	404	238	694	1189	1389	990	584	627
After merging	52	53	66	80	89	83	68	69
AA	43	47	51	68	66	64	50	53
AD	16	28	32	31	45	40	27	29
DA	25	14	28	39	44	40	30	31
DD	18	12	23	17	21	23	15	16

8 The digenic interactions were identified by searching a total of 910,018 possible two-locus combinations at the whole genome level, and a
9 randomization test with 10,000 permutations, followed by merging process and orthogonal contrast test. AA, additive by additive interaction; AD,
10 additive by dominance interaction; DA, dominance by additive interaction; DD, dominance by dominance interaction.

11 Table 4. Possible pleiotropic effects of the bins on seedling traits.

Locus	Leaf number		Tiller number		Root number		Root dry weight		Shoot dry weight		Total dry weight		Degree of dominance [§]
	A	D	A	D	A	D	A	D	A	D	A	D	
bin35	-0.32	-3.13	-0.07	-0.77	0.16	-2.31	0.01	-0.04	-0.02	-0.25	-0.01	-0.29	NO
bin148	0.18	2.79	0.03	0.73	-0.95	6.56	0.01	0.04	0.03	0.31	0.03	0.34	PO
bin154	0.02	2.81	-0.04	0.69	-0.71	7.39	0.01	0.04	0.01	0.33	0.01	0.37	PO
bin162	-0.82	2.40	-0.31	0.67	-2.25	5.79	-0.01	0.03	-0.07	0.29	-0.07	0.33	PO
bin233	1.66	-2.67	NA	NA	6.19	-5.21	0.01	-0.03	0.16	-0.21	0.17	-0.24	ND/NO
bin258	-0.91	-3.57	-0.21	-0.82	NA	NA	-0.01	-0.03	-0.04	-0.22	-0.05	-0.25	NO
bin485	NA	NA	0.35	-0.76	4.38	-5.94	0.01	-0.04	0.11	-0.19	0.13	-0.24	NO
bin518	0.73	2.44	0.35	0.75	5.72	5.85	0.03	0.03	NA	NA	0.11	0.19	PO
bin1032	-0.60	-3.48	-0.23	-1.16	-2.54	-8.83	-0.01	-0.04	-0.06	-0.22	-0.07	-0.26	NO
bin1094	-1.07	2.93	-0.41	0.72	NA	NA	0.01	0.03	0.07	0.19	0.06	0.22	PO
bin1272	NA	NA	-0.22	-0.83	2.09	-7.05	0.01	-0.03	0.01	-0.21	0.01	-0.25	NO
bin1314	2.01	-2.57	NA	NA	0.91	-7.17	0.02	-0.03	0.11	-0.22	0.13	-0.25	NO
bin1394	-2.02	2.80	-0.64	0.70	-4.21	6.94	-0.02	0.03	-0.21	0.19	-0.23	0.22	PD/PO
bin1456	0.63	2.77	0.14	1.08	-1.11	8.52	0.01	0.04	0.01	0.22	0.01	0.25	PO
bin1601	0.24	2.73	-0.21	0.72	2.07	7.33	0.02	0.04	0.04	0.24	0.06	0.28	PO
bin1604	0.23	2.54	-0.21	0.69	1.77	6.19	0.02	0.04	0.03	0.21	0.05	0.24	PO

12 A, additive effect. A positive value indicates that the allele from Zhenshan 97 is in the direction of increasing the phenotypic value, while a negative value
13 indicates that the allele from Minghui 63 increases the phenotypic value. D, dominance effect. A positive value indicates that the heterozygote has higher
14 phenotypic value than the average of two homozygotes, while a negative value indicates that the heterozygote has lower value than the average of two
15 homozygotes. NA indicates an nonsignificant effect. §, degree of dominance calculated as $D/|a|$. PO, positive overdominance; PD, positive dominance; NO,
16 negative overdominance; ND, negative dominance.

17 Table 5. Possible pleiotropic effects of DD interactions on the seedling traits.

Locus 1	Locus 2	Leaf number		Tiller number		Root number		Root dry weight		Shoot dry weight		Total dry weight	
		DD	<i>P</i> -value	DD	<i>P</i> -value	DD	<i>P</i> -value	DD	<i>P</i> -value	DD	<i>P</i> -value	DD	<i>P</i> -value
bin652	bin1032	-11.13	9.92E-05	-3.46	5.87E-06	NA	NA	-0.07	2.68E-02	-0.68	1.34E-03	-0.76	1.64E-03
bin1040	bin1060	-9.52	4.88E-04	-2.31	2.41E-03	NA	NA	-0.11	1.11E-03	-0.11	1.11E-03	-0.71	2.74E-03
bin659	bin1612	12.58	6.18E-06	2.98	1.12E-04	20.94	1.17E-03	0.13	1.30E-04	0.95	2.78E-06	1.08	2.80E-06
bin256	bin1035	NA	NA	NA	NA	16.89	5.87E-03	0.08	1.83E-02	0.55	9.83E-03	0.62	9.80E-03

18 DD, dominance by dominance interaction. *P*-value is from the orthogonal contrast test at the significance level of $P \leq 0.05$. NA indicates an nonsignificant
19 effect.

20

21 Table 6. Relative amounts of single-locus dominance, overdominance, and digenic
 22 dominance effects identified in the entire genome in the immortalized F₂ population.
 23 The values were summed over the bins divided by the mean of each trait.

Trait	PO	PD	NO	ND	PDD	NDD	Sum
Seedling height	0.710	0	-0.892	0	0.945	-0.415	0.348
Leaf number	2.905	0	-1.094	0	1.086	-1.302	1.596
Tiller number	3.197	0	-1.963	0	1.797	-2.871	0.160
Maximum root length	1.475	0.025	-1.026	-0.137	0.629	-0.786	0.179
Root number	3.862	0	-3.510	-0.071	4.483	-0.208	4.554
Root dry weight	1.482	0	-2.102	0	2.618	-1.442	0.556
Shoot dry weight	2.331	0.069	-2.351	-0.325	1.906	-1.273	0.358
Total dry weight	2.265	0.069	-2.562	-0.182	2.176	-1.266	0.500

24 The meanings of PO, PD, NO, and ND abbreviations are the same as in Table 4. PDD,
 25 positive digenic dominance; NDD, negative digenic dominance; Sum, the total effects of all
 26 the genetic components.

27 Table 7. Bins showing pleiotropic effects on seedling traits that overlap with yield
 28 and yield component traits in Zhou et al. (2012).

Trait	Yield	Grain per panicle	Tiller per plant	Grain weight
bin233	NA	NA	NO	NO
bin1032	PO	NA	PO	NA
bin1093	PO	PO	NA	NA
bin1272	PO	PO	NA	NA
bin1314	NA	PO	NA	NA
bin1456	PO	PO	NO	NA

29 PO, positive overdominance; NO, negative overdominance. NA indicates an nonsignificant
 30 effect.