Supplementary Note

1 The model reflecting the performance of homozygotes and heterozygote of one locus simulated according to Hill equation

The typical physiologihiscal theory had pointed that most of the mutants of genes encoding enzymes are dominant[1,2], which has been preliminarily verified in yeast and human[3]. Recent study also noticed that enzyme unsaturation caused by insufficient substrate may be the direct cause of dominance[4]. However, it is still not clear whether insufficient substrate background is the general mechanism of dominant or even overdominance, and the relationship between the sufficiency of background and the occurrence of dominant or overdominance is still not clearly elucidated in literature. In this study, we systematically simulated occurrence of additive, dominant to over-dominant inheritance of target receptor genes under different level of ligand background supply.

It is generally recognized that a ligand X binds to a receptor Y and reacts to produce a product is a common mechanism in biology. Dynamically, the number of molecule product of Y produced per unit time is a function of the concentration of ligand X on its active form X*:

$$\text{Production rate of Y} = f(X^*) \tag{1-1}$$

Typically, the input function of f(X) is a monotonic, S-shaped function. It is an increase function when X is an activator and a decrease one when X is a repressor[5]. The Hill input function for an activator is a curve that rises from zero and approaches a maximal saturated level:

$$f(X^*) = \frac{\beta X^{*n}}{K^n + X^{*n}} \qquad \text{Hill function for activator (1-2)}$$

The Hill function has three parameters, K, β and n.

Parameter K is termed as the activation coefficient, and has units of concentration. It defines the concentration of active X needed to significantly activate production. From the equation, we can see that half-maximal production is reached when X = K. The value of K related to the chemical affinity between ligand X and its receptor, as well as additional factors.

Parameter β is the maximal production level of Y. Maximal production is reached at high activator concentration, X>> K. Because at high concentration, X binds the receptor with high probability to generate more products per unit time.

Parameter n is known as Hill coefficient. It governs the steepness of the curve between two inflection points of the input function. Usually, it is moderately steep, with n = 1 – 4. The larger is n, the more step-like is the input function. Particularly, When n =1, hill function is equal to Michaelis Menten equation. As many functions in biology, the Hill function approaches a limiting value at high level of X, rather than increase indefinitely.

For a repressor, the Hill function is a decreasing S-shaped curve, whose shape depends on three similar parameters:

Hill input function for repressor (1-3)

$$f(x) = \frac{\beta}{1+\left(\frac{X^*}{K}\right)^n}$$

The production of Y is balanced by two process, degradation (destruction by specific proteins in the cell) and dilution (the reduction in concentration due to the increase of cell volume during growth). The degradation rate is a_{deg}, and the dilution rate is a_{dil}, giving a total degradation plus dilution rate (in units of 1/time) of

$$\alpha = \alpha_{deg} + \alpha_{dil} \quad (1\text{-}4)$$

The change in the concentration of Y due to the difference between its production and degradation plus dilution, as described by a dynamic equation:

$$dY/dt = f(X^*) - \alpha Y \quad (1\text{-}5)$$

At stead state, Y reaches a constant concentration Y_{st}, The steady-state concentration can be found by solving for dY/dt = 0. The steady-state concentration is:

$$Y_{st} = f(X^*)/\alpha \quad (1\text{-}6)$$

If reached its maximal level, we can also write as:

$$Y_{st} = \beta/\alpha \quad (1\text{-}7)$$

This makes sense: The higher is the production rate β, the higher will reach the steady-state concentration Y_{st}. The higher is the degradation/dilution rate α, the lower is Y_{st}.

Now let us consider one single locus with allele A and a, which are or code some kind of receptor and can be regulated by ligand X. The product of allele A is Y_1 at steady-state under concentration $[X_{11}^*]$ of active X (X_{11}^*), and that of a is Y_2 under concentration $[X_{22}^*]$. The production function of two alleles is expressed respectively as:

$$A: dY_1/dt = f([X_{11}^*]) - \alpha_1 Y_1 \quad (1\text{-}8a)$$

$$a: dY_2/dt = f([X_{22}^*]) - \alpha_2 Y_2 \quad (1\text{-}8b)$$

Where $\alpha_1 > 0$ and $\alpha_2 > 0$ are the relative degradation rate.

Then, the product of two alleles at steady-state is respectively:

$$A: Y_1 = f(X_1^*)/\alpha_1 = \frac{\beta_1}{\alpha_1}\frac{[X_{11}^*]^{n_1}}{K_{11}^{n_1} + [X_{11}^*]^{n_1}} \quad (1\text{-}9a)$$

$$a: Y_2 = f(X_2^*)/\alpha_2 = \frac{\beta_2}{\alpha_2}\frac{[X_{22}^*]^{n_2}}{K_{22}^{n_2} + [X_{22}^*]^{n_2}} \quad (1\text{-}9b)$$

With $\mu_j = \frac{\beta_j}{\alpha_j}$, then Equations 1-9a and 1-9c are transformed into:

$$A: Y_1 = \mu_1\frac{[X_{11}^*]^{n_1}}{K_{11}^{n_1} + [X_{11}^*]^{n_1}} \quad (1\text{-}10a)$$

$$a: Y_2 = \mu_1\frac{[X_{22}^*]^{n_2}}{K_{22}^{n_2} + [X_{22}^*]^{n_2}} \quad (1\text{-}10b)$$

Regarding the relationship between products of homozygotes (AA and aa) and heterozygote (Aa) of the locus, we consider three scenarios. The general assumption for three scenarios is that: (1) the ligand background concentration keeps constant among two homozygotes (representing the parents) and heterozygote (representing the F_1 hybrid); so if two alleles share the same kind of ligand background, the ligand background concentration in homozygotes and heterozygote will be $[X_{11}^*]$ or $[X_{22}^*]$, with $[X_{11}^*] = [X_{22}^*]$; if two alleles have their respective ligand backgrounds, two homozygotes and the heterozygote will maintain the same concentration of both ligands, $[X_{11}^*]$ and $[X_{22}^*]$; (2) the ligand background can be equally and randomly allocated to two alleles in the homozygote, and the reaction of two alleles in heterozygote is independent[6] and the ligand background will be allocated to two different alleles under the rule as defined in different scenarios; (3) there is a basal product m in two homozygotes and heterozygote when there is no ligand.

Firstly, we consider the three scenarios under the situation that the ligand works as an activator.

Scenario 1: null allele vs one functional allele of one polymorphic site under one ligand background (**Supplementary Fig 32 and Supplementary Fig 33a-b**). The ligand background concentration in two homozygotes and heterozygote will be $2[X_{11}^*] = 2[X_{22}^*] = 2[X^*]$. The product of AA, aa and Aa at steady state will be:

$$AA: Y_{11} = m + \mu_1\frac{[X^*]^{n_1}}{K^{n_1} + [X^*]^{n_1}} + \mu_1\frac{[X^*]^{n_1}}{K^{n_1} + [X^*]^{n_1}} \quad (1\text{-}11a)$$

$$aa: Y_{22} = m + 0 \quad (1\text{-}11b)$$

$$Aa: Y_{12} = m + \mu_1 \frac{(2[X^*])^{n_1}}{K_{11}^{n_1} + (2[X^*])^{n_1}} \quad (1\text{-}11c)$$

Scenario 2: two alleles of one polymorphic site under two independent backgrounds, that is, two alleles of one polymorphic site of the receptor can be bound by two respective and independent ligands as the backgrounds of the receptor (**Supplementary Fig 34 and Supplementary Fig 36**). The ligand background concentration in two homozygotes and heterozygote will be $2[X_{11}^*]$ and $2[X_{22}^*]$, but X_{11}^* can only be allocated to allele A and X_{22}^* to allele a. The product of AA, aa and Aa at steady state will be:

$$AA: Y_{11} = m + \mu_1 \frac{[X_{11}^*]^{n_1}}{K_{11}^{n_1} + [X_{11}^*]^{n_1}} + \mu_1 \frac{[X_{11}^*]^{n_1}}{K_{11}^{n_1} + [X_{11}^*]^{n_1}} \quad (1\text{-}12a)$$

$$aa: Y_{22} = m + \mu_2 \frac{[X_{22}^*]^{n_2}}{K_{22}^{n_2} + [X_{22}^*]^{n_2}} + \mu_1 \frac{[X_{22}^*]^{n_2}}{K_{22}^{n_2} + [X_{22}^*]^{n_2}} \quad (1\text{-}12b)$$

$$Aa: Y_{12} = m + \mu_1 \frac{(2[X_{11}^*])^{n_1}}{K_{11}^{n_1} + (2[X_{11}^*])^{n_1}} + \mu_1 \frac{(2[X_{22}^*])^{n_2}}{K_{22}^{n_2} + (2[X_{22}^*])^{n_2}} \quad (1\text{-}12c)$$

Scenario 3: two alleles of one polymorphic site with shared background, that is, two alleles of one polymorphic site of the receptor can be bound by the same ligand as the background of the receptor (**Supplementary Fig 38**). The ligand background concentration in two homozygotes and heterozygote will be $2[X_{11}^*] = 2[X_{22}^*] = 2[X^*]$. If the ligand background X* was equally allocated to each of the two alleles in heterozygote as the simulation previously reported[6,7], the product of AA, aa and Aa at steady state will be:

$$AA: Y_{11} = m + \mu_1 \frac{[X^*]^{n_1}}{K_{11}^{n1} + [X^*]^{n_1}} + \mu_1 \frac{[X^*]^{n_1}}{K_{11}^{n1} + [X^*]^{n_1}} \quad (1\text{-}10a)$$

$$aa: Y_{22} = m + \mu_2 \frac{[X^*]^{n_2}}{K_{22}^{n_2} + [X^*]^{n_2}} + \mu_2 \frac{[X^*]^{n_2}}{K_{22}^{n_2} + [X^*]^{n_2}} \quad (1\text{-}10b)$$

$$Aa: Y_{12} = m + \mu_1 \frac{[X^*]^{n_1}}{K_{11}^{n_1} + [X^*]^{n_1}} + \mu_2 \frac{[X^*]^{n_2}}{K_{22}^{n_2} + [X^*]^{n_2}} \quad \text{(Equal allocation)} \quad (1\text{-}10c)$$

As our simulation indicated, the locus will always appear to be additive under the situation of equal allocation ($X^* = (X_{11}^* + X_{22}^*)/2$). We need consider the situations that the background X* was allocated to two alleles in an asymmetric way due to different affinities of two alleles with the ligand background. If the ligand background X* was allocated to each of the two alleles in heterozygote in proportion to their respective affinities, the product of Aa at steady state will be:

$$\text{Aa: } Y_{12} = m + \mu_1 \frac{(2[X^*]K_{22}/(K_{11}+K_{22}))^{n_1}}{K_{11}^{n_1}+(2[X^*]K_{22}/(K_{11}+K_{22}))^{n_1}} + \mu_2 \frac{(2[X^*]K_{11}/(K_{11}+K_{22}))^{n_2}}{K_{22}^{n_2}+(2[X^*]K_{11}/(K_{11}+K_{22}))^{n_2}} \quad \text{(Asymmetric allocation)}$$

(1-10d)

We also proposed an optimal strategy to maximize the output of the heterozygote. Let $S_1+S_2 = 2[X^*]$, S_1 and S_2 represent the ligand concentration allocated to allele A and a in heterozygote, respectively, when the product of heterozygote Y_{12} is maximized at the ligand concentration 2[X*] (**Supplementary Fig 39-40**). The product of Aa at steady state will be:

$$\text{Aa: } Y_{12} = m + \max\left(\mu_1 \frac{S_1^{n_1}}{K_{11}^{n_1}+S_1^{n_1}} + \mu_2 \frac{S_2^{n_2}}{K_{22}^{n_2}+S_2^{n_2}}\right) \quad \text{(Maximized allocation)} \quad (1\text{-}10e)$$

Secondly, we consider the three scenarios under the situation that the ligand works as a repressor.

Regarding Scenario 1, null allele vs one functional allele of one polymorphic site under one ligand background (**Supplementary Fig 33c-d**). The product of AA, aa and Aa at steady state for negative regulation will be:

$$\text{AA: } Y_{11} = m - \mu_1\left(1-\frac{K^{n_1}}{K^{n_1}+[X^*]^{n_1}}\right) - \mu_1\left(1-\frac{K^{n_1}}{K^{n_1}+[X^*]^{n_1}}\right) \quad (1\text{-}11a)$$

$$\text{aa: } Y_{22} = m - 0 \quad (1\text{-}11b)$$

$$\text{Aa: } Y_{12} = m - \mu_1\left(1-\frac{K^{n_1}}{K^{n_1}+(2[X^*])^{n_1}}\right) \quad (1\text{-}11c)$$

Regarding Scenario 2, two alleles of one polymorphic site under two independent backgrounds (**Supplementary Fig 37**). The product of AA, aa and Aa at steady state for negative regulation will be:

$$\text{AA: } Y_{11} = m - \mu_1\left(1-\frac{K_{11}^{n_1}}{K_{11}^{n_1}+[X_{11}^*]^{n_1}}\right) - \mu_1\left(1-\frac{K_{11}^{n_1}}{K_{11}^{n_1}+[X_{11}^*]^{n_1}}\right) \quad (1\text{-}12a)$$

$$\text{aa: } Y_{22} = m - \mu_2\left(1-\frac{K_{22}^{n_2}}{K_{22}^{n_2}+[X_{22}^*]^{n_2}}\right) - \mu_2\left(1-\frac{K_{22}^{n_2}}{K_{22}^{n_2}+[X_{22}^*]^{n_2}}\right) \quad (1\text{-}12b)$$

$$\text{Aa: } Y_{12} = m - \mu_1\left(1-\frac{K_{11}^{n_1}}{K_{11}^{n_1}+(2[X_{11}^*])^{n_1}}\right) - \mu_2\left(1-\frac{(K_{22})^{n_2}}{K_{22}^{n_2}+(2[X_{22}^*])^{n_2}}\right) \quad (1\text{-}12c)$$

Regarding Scenario 3, two alleles of one polymorphic site with shared background (**Supplementary Fig 41-42**). The product of AA, aa and Aa at steady state for negative regulation will be:

$$\text{AA: } Y_{11} = m - \mu_1\left(1-\frac{K_{11}^{n_1}}{K_{11}^{n_1}+[X^*]^{n_1}}\right) - \mu_1\left(1-\frac{K_{11}^{n_1}}{K_{11}^{n_1}+[X^*]^{n_1}}\right) \quad (1\text{-}13a)$$

$$\text{aa: } Y_{22} = m - \mu_2\left(1-\frac{K_{22}^{n_2}}{K_{22}^{n_2}+[X^*]^{n_2}}\right) - \mu_2\left(1-\frac{K_{22}^{n_2}}{K_{22}^{n_2}+[X^*]^{n_2}}\right) \quad (1\text{-}13b)$$

$$Aa:\ Y_{12} = m - \mu_1\left(1-\frac{K_{11}^{n_1}}{K_{11}^{n_1}+[X^*]^{n_1}}\right)-\mu_2\left(1-\frac{K_{22}^{n_2}}{K_{22}^{n_2}+[X^*]^{n_2}}\right) \quad \text{(Equal allocation)} \quad (1\text{-}13c)$$

Or,

$$Aa:\ Y_{12} = m - \mu_1\left(1-\frac{K_{11}^{n_1}}{K_{11}^{n_1}+(2[X^*]K_{22}/(K_{11}+K_{22}))^{n_1}}\right)-\mu_2\left(1-\frac{K_{22}^{n_2}}{K_{22}^{n_2}+(2[X^*]K_{11}/(K_{11}+K_{22}))^{n_2}}\right)$$

(Asymmetric allocation) (1-13d)

Or,

$$Aa:\ Y_{12} = m - \max\left(\mu_1\left(1-\frac{K_{11}^{n_1}}{K_{11}^{n_1}+S_1^{n_1}}\right)+\mu_2\left(1-\frac{K_{22}^{n_2}}{K_{22}^{n_2}+S_2^{n_2}}\right)\right) \quad \text{(Maximized allocation)} \quad (1\text{-}13e)$$

Where $S_1+S_2 = 2[X^*]$.

According the simulated values of Y_{11},Y_{12} and Y_{22}, we calculated the degree of dominance (d/a) for the locus as:

$$d/a=(Y_{12}-(Y_{11}+Y_{22})/2|Y_{22}-Y_{11}|) \quad (1\text{-}14)$$

2 The model reflecting the performance of homozygotes and heterozygote of one locus simulated according to trimer ABA assembly

The balance between genes involved in a biological complex is one important hypothesis about heterosis. The typical example for gene balance was reported by Balazs and coleagues[8]. Their studies indicated that mutation of the subunit in a trimer ABA complex can result in imbalance and thus is harmful, which might impact gene imbalance on dominance. However, these studies did not consider the effects from the counterpart background. Thus, we simulated the effects of complex background on dominance of one single polymorphic locus that codes A or B.

In the system of trimer ABA complex, A and B are monomers, AB is the bridge dimer without active function, the trimer ABA is the functional entity. The reaction among monomers, dimer and trimer could be illustrated by the following chemical formula:

$$A + B <=> AB \quad (2\text{-}1)$$

$$AB + A <=> ABA \quad (2\text{-}2)$$

For simplicity, we consider a pseudo equilibrium state, that is: A and B were input once in an enclosed environment and no degradation was considered; after a period of time, a chemical equilibrium state will be achieved. Set S_A and S_B as the initial input concentration of A and B, k_{AB} as the association

rate from left to right in formula (2-1), m_{AB} as the dissociation rate from right to left of formula (2-1), k_{ABA} and m_{ABA} as the association rate from left to right and the dissociation rate from right to left in formula (2-2). And let [A], [B], [AB] and [ABA] represent the concentration of A, B, AB and ABA at equilibrium state. So we have:

$$S_A = [A] + [AB] + 2[ABA]$$

$$S_B = [B] + [AB] + [ABA]$$

$$k_{AB} \times [A][B] = m_{AB} \times [AB]$$

$$[AB] = k_{AB}/m_{AB} \times [A][B]$$

$$k_{ABA} \times [AB][A] = m_{ABA} \times [ABA]$$

$$[ABA] = k_{ABA}/ m_{ABA} \times [AB][A]$$

$$[ABA] = k_{ABA}/ m_{ABA} \times k_{AB}/m_{AB}[A][B][A]$$

We define association coefficient by the ratio of association to dissociation for two steps as:

$$K_1 = k_{AB}/m_{AB}$$

$$K_2 = k_{ABA}/m_{ABA}$$

then, we derived that

$$[AB] = K_1 \times [A][B]$$

$$[ABA] = K_1 \times K_2 \times [A][B][A]$$

According to the formula of stoichiometry balance, we can have:

$$S_A = [A] + K_1 \times [A][B] + 2K_1 \times K_2 \times [A][B][A]$$

$$S_B = [B] + K_1 \times [A][B] + K_1 \times K_2 \times [A][B][A]$$

Through the above equation, we get

$$[B] = S_B/(1 + K_1 \times [A] + K_1 \times K_2 \times [A]^2)$$

Then we introduced [B] into S_A, we get,

$$S_A = [A] + S_B \times K_1 \times [A]/(1 + K_1 \times [A] + K_1 \times K_2 \times [A]^2) +$$

$$2\, S_B \times K_1 \times K_2 \times [A]^2/(1 + K_1 \times [A] + K_1 \times K_2 \times [A]^2)$$

We set a target function f([A]),

$$f([A]) = (S_A - ([A] + S_B \times K_1 \times [A]/(1 + K_1 \times [A] + K_1 \times K_2 \times [A]^2) +$$

$$2\, S_B \times K_1 \times K_2 \times [A]^2/(1 + K_1 \times [A] + K_1 \times K_2 \times [A]^2))^2$$

Among of the function of f([A]), only [A] is the unknown parameters, S_A, S_B, K_1 and K_2 were all the predefined data, thus the value of [A] that minimizes f([A]) is the solution of the concentration of A at

equilibrium state. Once the concentration of A at equilibrium state is obtained, the concentration of B, AB, and ABA at equilibrium state could be easily calculated according to the above equations.

For given S_A, S_B, K_1 and K_2, we solve the equation by using the optimize function in R and get the concentration of A, B, AB and ABA at the equilibrium state, the solutions of parent and F_1 was follow the same equations described above (**Supplementary Fig 63**).

We simulated two scenarios as following:

Scenario1, keep the input concentration of B fixed and constant among two homozygotes and the heterozygote of A, and A was coded by one polymorphic locus (**Supplementary Fig 63b-c**): S_A ranges from 0 to 20 nmol/L, with S_B = 2.5 nmol/L, K_1 = 1, K_2 = 100. The simulated data for the genotype of AA, aa and Aa as follow:

AA: $S_{A(AA)}$ = 0 - 20 nmol/L, $S_{B(AA)}$ = 2.5 nmol/L

aa: $S_{A(aa)}$ = 0 - 20 nmol/L, $S_{B(aa)}$ = 2.5 nmol/L

Aa: $S_{A(Aa)} = (S_{A(AA)} + S_{A(aa)})/2$, $S_{B(Aa)}$ = 2.5 nmol/L

Scenario2, keep the input concentration of A fixed and constant among two homozygotes and the heterozygote of B, and B was coded by one polymorphic locus (**Supplementary Fig 63d-e**): S_B ranges from 0 to 20 nmol/L, with S_A = 5 nmol/L, K_1 = 1, K_2 = 100. The simulated data for the genotype of AA, aa and Aa as follow:

BB: $S_{A(BB)}$ = 5 nmol/L, $S_{B(BB)}$ =0 - 20 nmol/L

bb: $S_{A(bb)}$ = 5 nmol/L, $S_{B(bb)}$ = 0 - 20 nmol/L

Bb: $S_{A(Bb)}$ =5 nmol/L, $S_{B(Bb)} = (S_{A(BB)} + S_{A(bb)})/2$

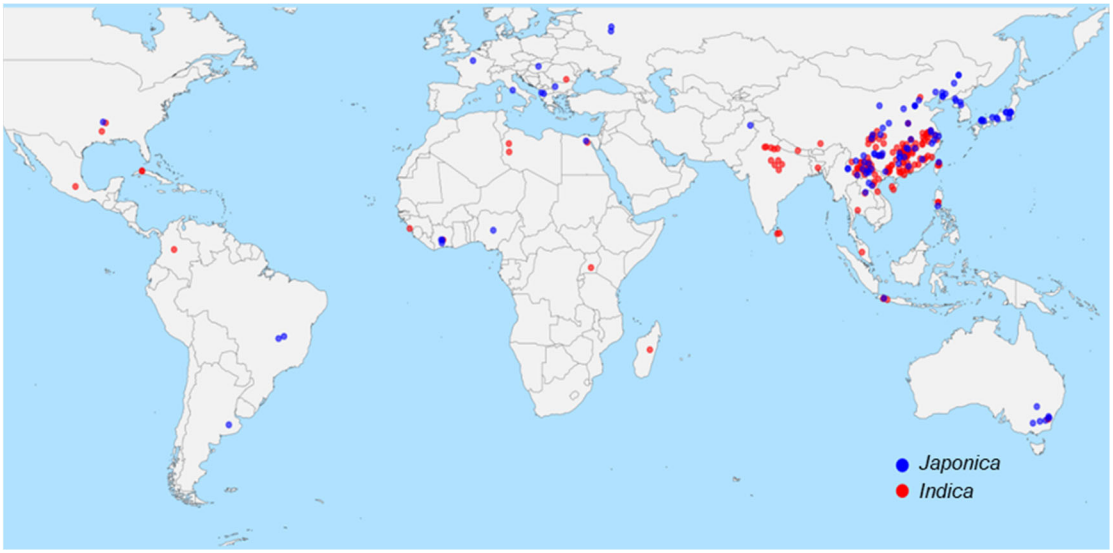
Same as the above simulation according to Hill function, according the simulated values of Y_{11},Y_{12} and Y_{22}, we calculated the degree of dominance (d/a) for the locus as:

$$\mathrm{d/a} = (Y_{12} - (Y_{11} + Y_{22})/2)/|Y_{22} - Y_{11}|$$

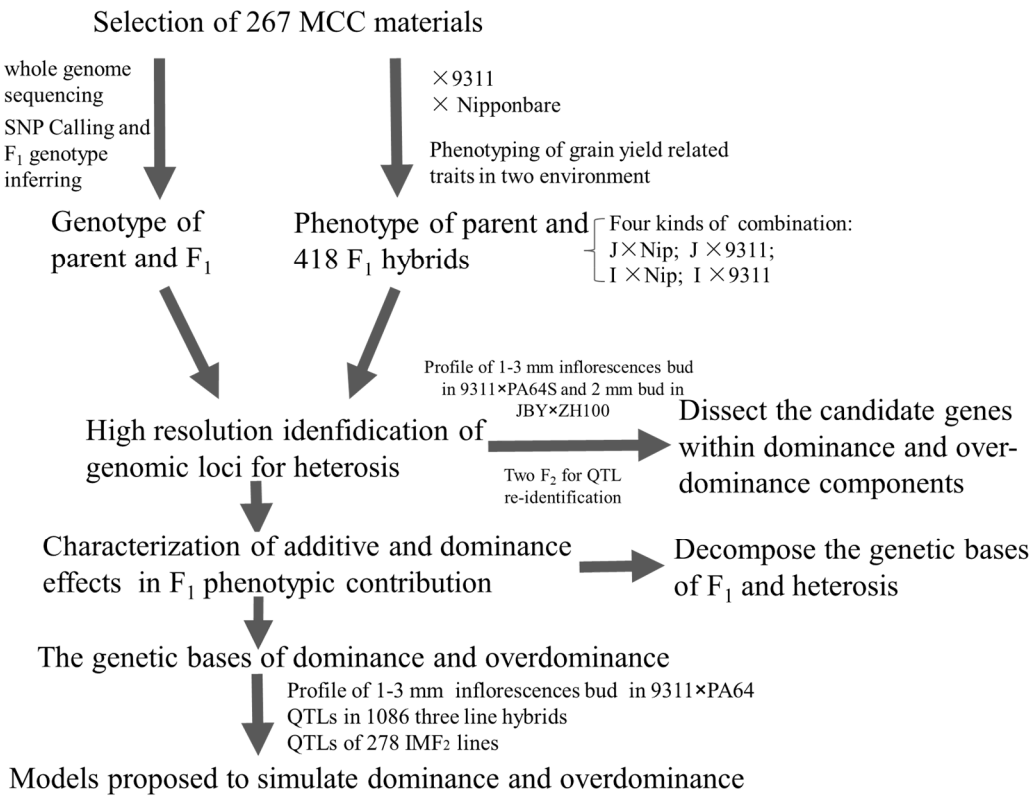
1. Wright, S. Physiological and evolutionary theories of dominance. *American Naturalist* **68**, 24-53 (1934).
2. Kacser, H. & Burns, J.A. The molecular basis of dominance. *Genetics* **97**, 639-66 (1981).
3. Kondrashov, F.A. & Koonin, E.V. A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications. *Trends Genet* **20**, 287-90 (2004).
4. Bürger, R. & Bagheri, H. Dominance and Its Evolution. *Encyclopedia of Ecology* (2008).

5. Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*, (2006).
6. Omholt, S.W., Plahte, E., Oyehaug, L. & Xiang, K. Gene regulatory networks generating the phenomena of additivity, dominance and epistasis. *Genetics* **155**, 969-80 (2000).
7. Veitia, R.A. A generalized model of gene dosage and dominant negative effects in macromolecular complexes. *FASEB J* **24**, 994-1002 (2010).
8. Papp, B., Pal, C. & Hurst, L.D. Dosage sensitivity and the evolution of gene families in yeast. *Nature* **424**, 194-7 (2003).

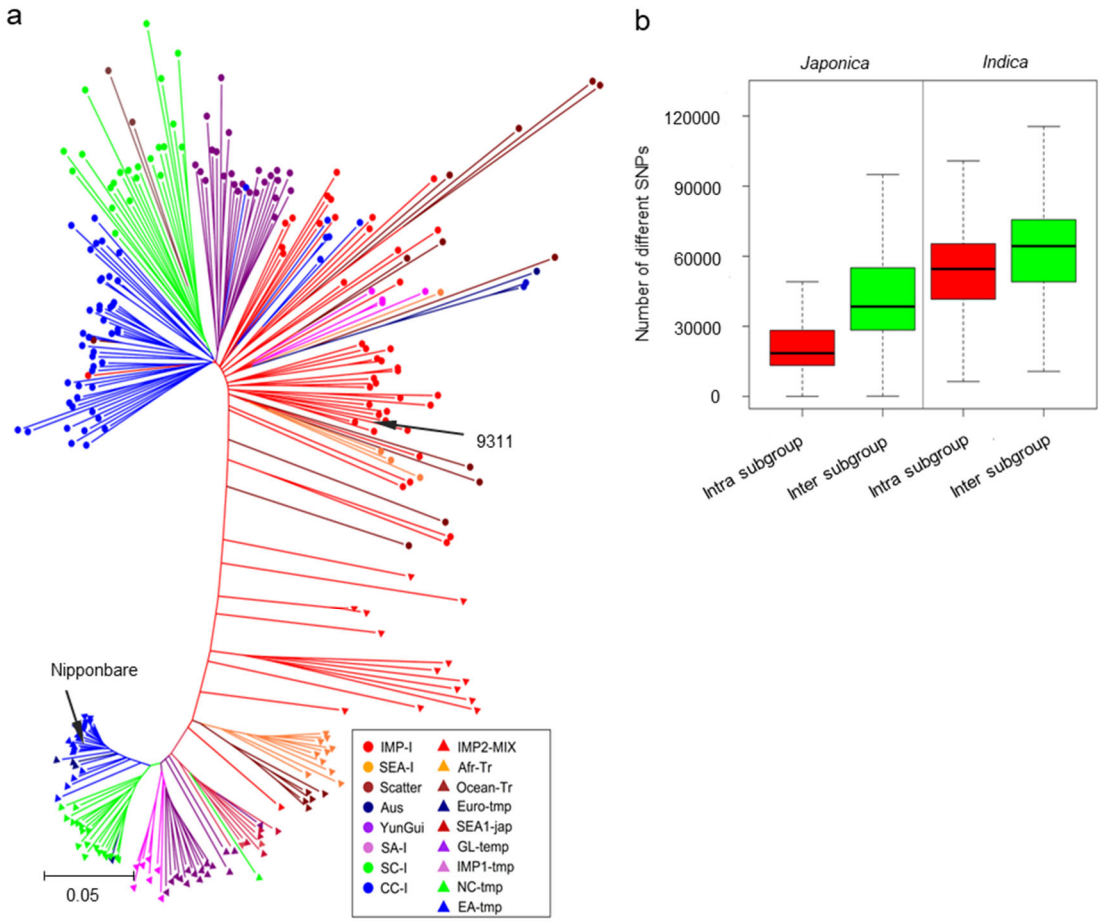
Supplementary Figures



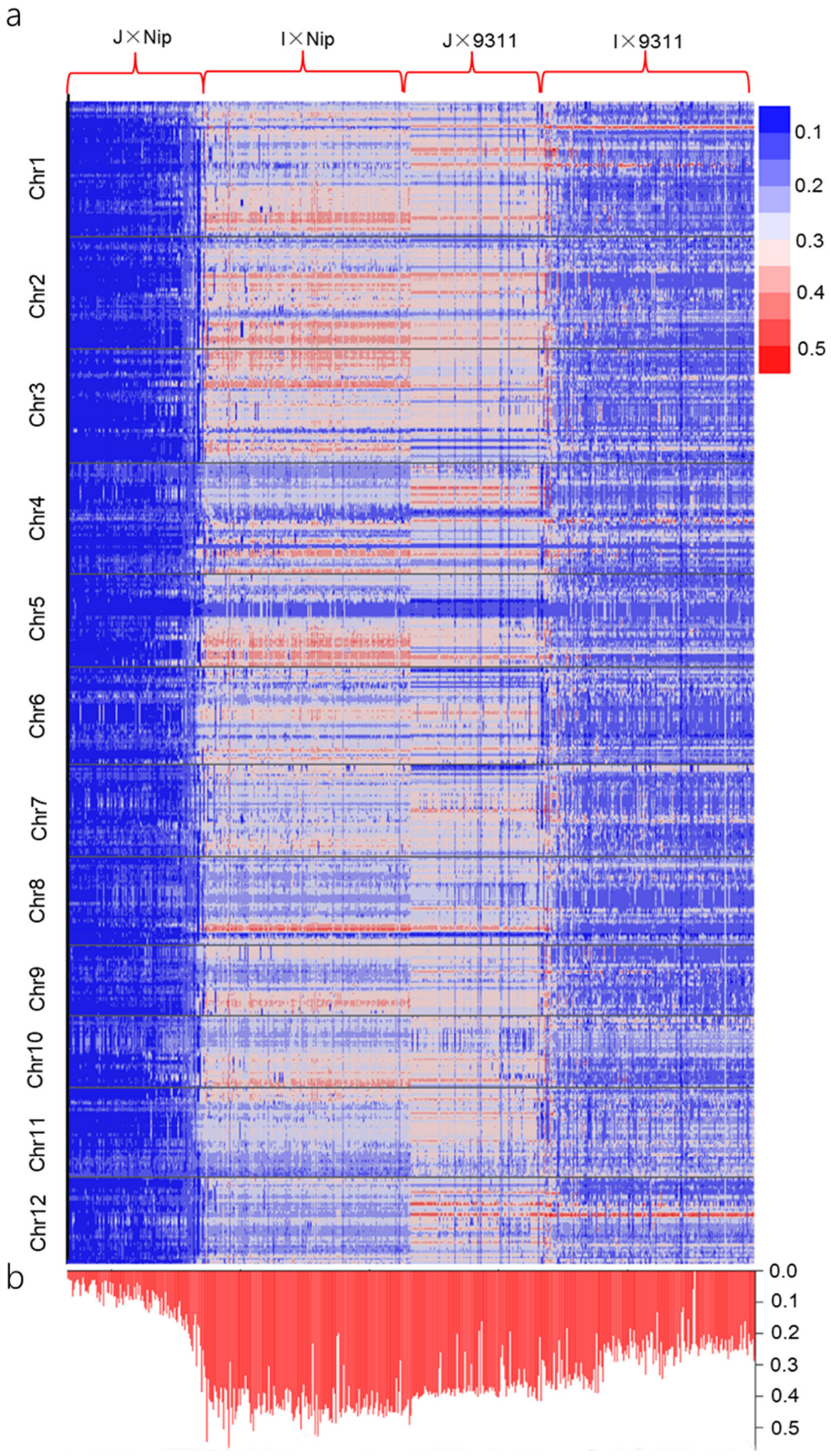
Supplementary Figure 1 Geographical distribution of 267 rice varieties. The red dots represent *Indica* varieties, the blue dots represent *Japonica* varieties.



Supplementary Figure 2 The experimental design and analysis procedure used in this study.

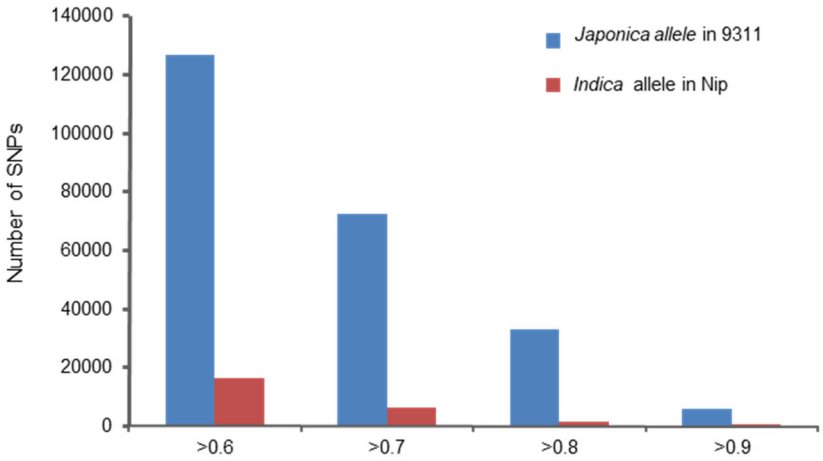


Supplementary Figure 3 Genetic structure of the 267 rice accessions. (a) The neighbor-joining tree of 267 rice accessions constructed from simple matching distance of 1.3 million SNPs. There were 8 and 9 subpopulations were identified in *Indica* and *Japonica* subspecies respectively. *Indica* sub-population including*:* IMP-I, Improved *Indica*; SEA-I, South east Asian *Indica;* Scatter, landraces which contains the admixed fragment and several inter *Indica-Japonica* type materials; Aus, the aus sub population; YunGui, Yunnan and Guizhou high altitude *Indica* from China; SA-I, South asian *Indica*; SC-I, South China *Indica*; CC-I, Center China *Indica. Japonica* sub-population including: IMP2-MIX, improved *Japonica* with admixed genomic fragment; Afr-Tr, Africa tropical *Japonica.* Ocean-Tr, Oceanica tropical *Japonica*. Euro-tmp, European temperate *Japonica;* SEA1-jap, Yunnan and Guizhou high altitude *Japonica* lines that subjected to Southeast Asian subtropical *Japonica*; GL-tmp, Yunnan and Guizhou high altitude Glutinous *Japonica*; IMP1-tmp, improved temperate *Japonica* lines; NC-tmp, north china temperate *Japonica;* EA-tmp from China, east asian temperate *Japonica*. (b) The average number of SNPs that different between pairwise individual of intra and inter sub-population in *Japonica* and *Indica.*

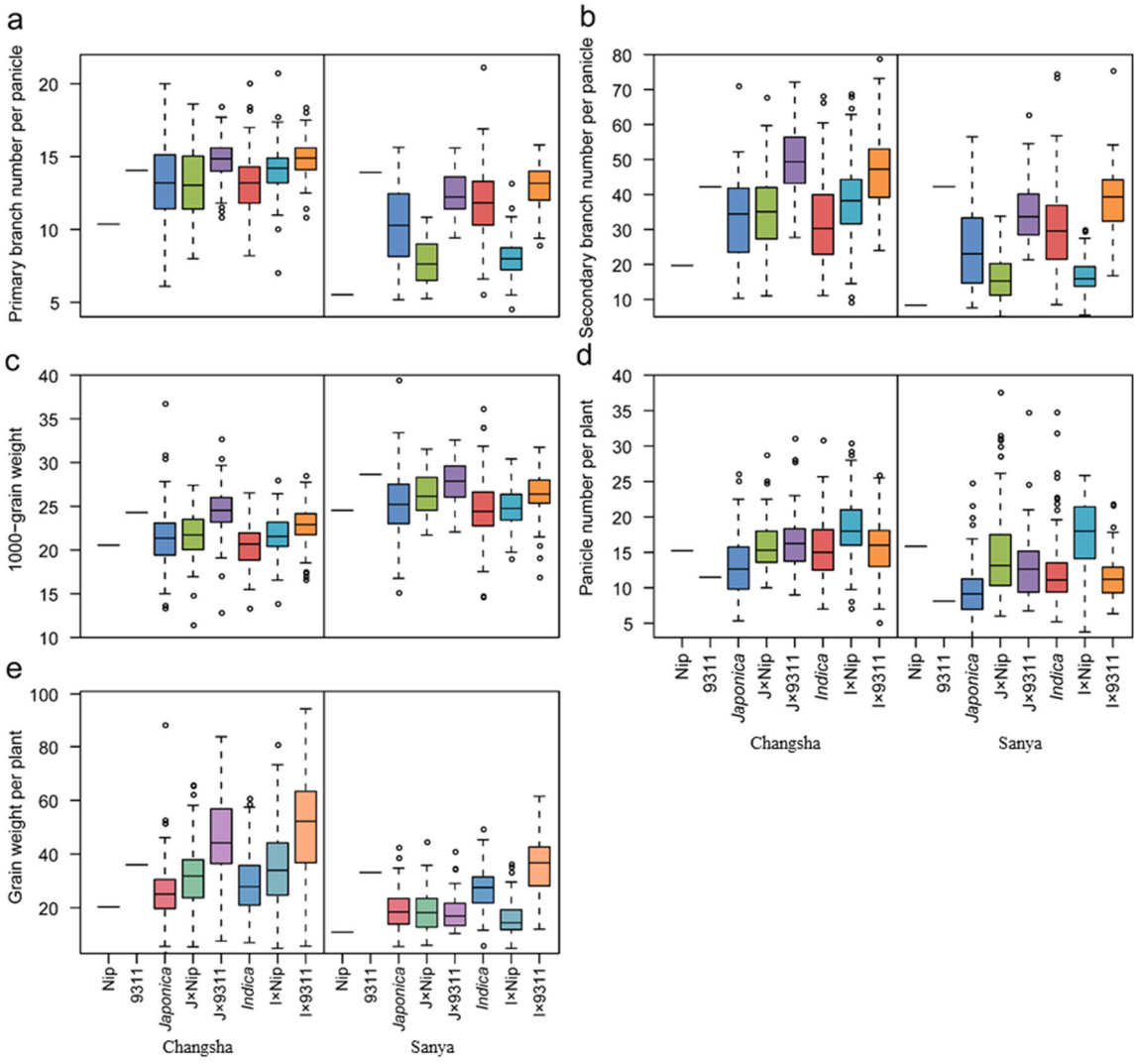


Supplementary Figure 4 Genome-wide heterozygosity for different kinds of combinations. (a) The

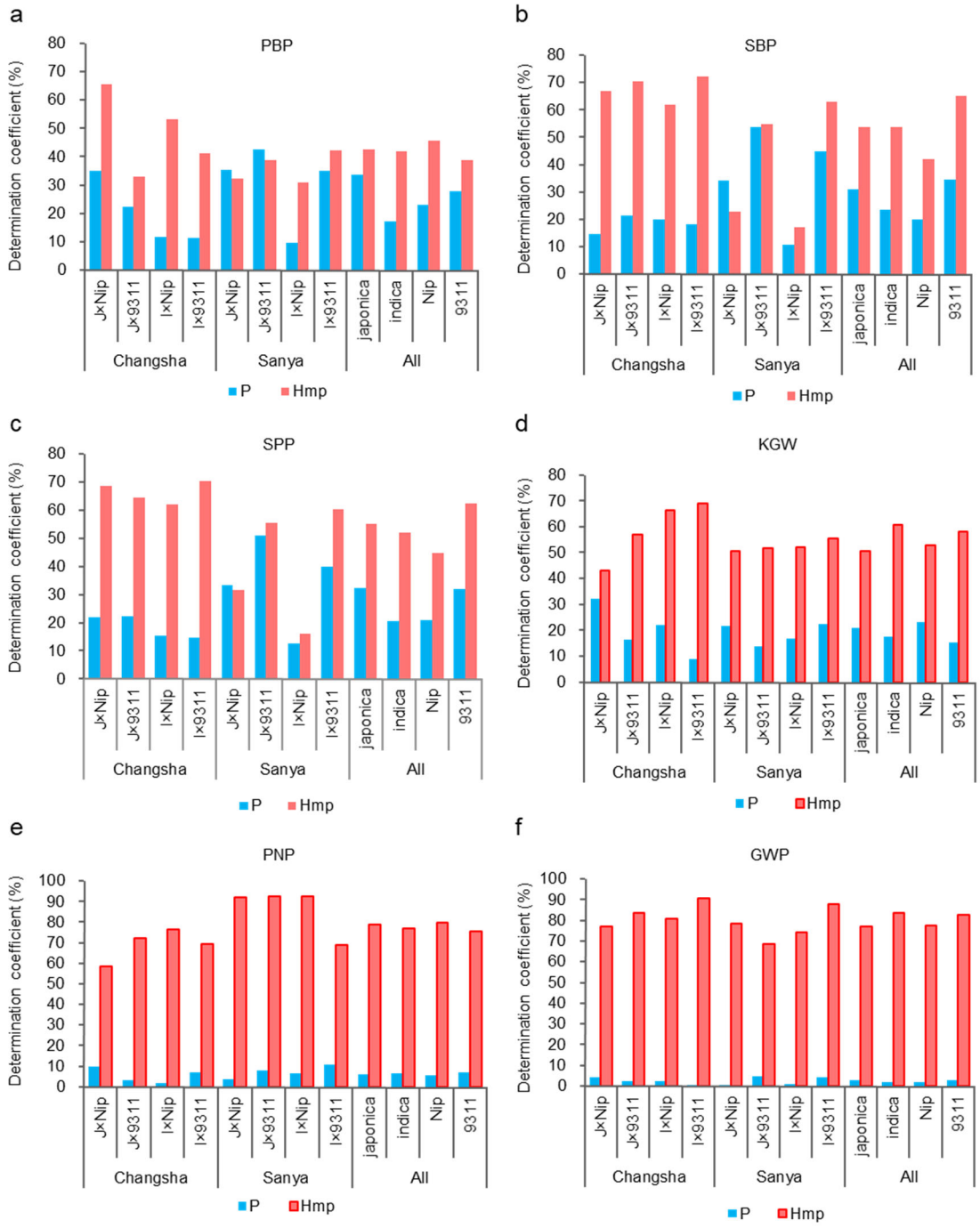
local heterozygosity was calculated in each 200 kb window across the entire rice genome for each combination. The color legend from blue to red denote the level of heterozygosity from low to high. (b) Eeach line represents the average heterozygosity in the whole genome for each combination. Nip, Nipponbare; J×Nip, combination of *Japonica* and Nipponbare; I×Nip, combination of *Indica* and Nipponbare, and the others are similar.



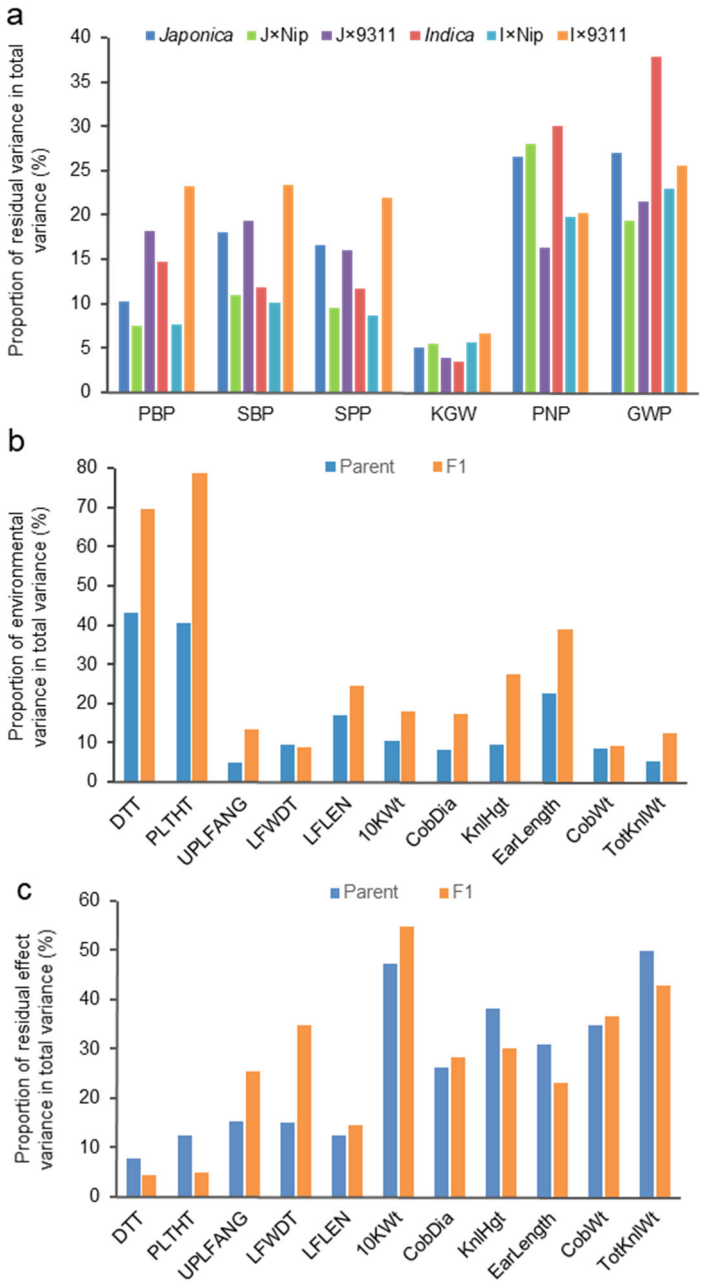
Supplementary Figure 5 Distribution of *Japonica* specific alleles in 9311 and *Indica specific* alleles in Nipponbare. The total number of *Japonica* specific alleles that introgressed into 9311 (blue bar), and that of *Indica* specific alleles that introgressed into Nipponbare (red bar). >0.6, >0.7, >0.8 and >0.9 mean that the allele frequency in one subspecies are higher than 0.6, 0.7, 0.8 and 0.9 but equal and less than 0.4, 0.3, 0.2 and 0.1 in the other subspecies.



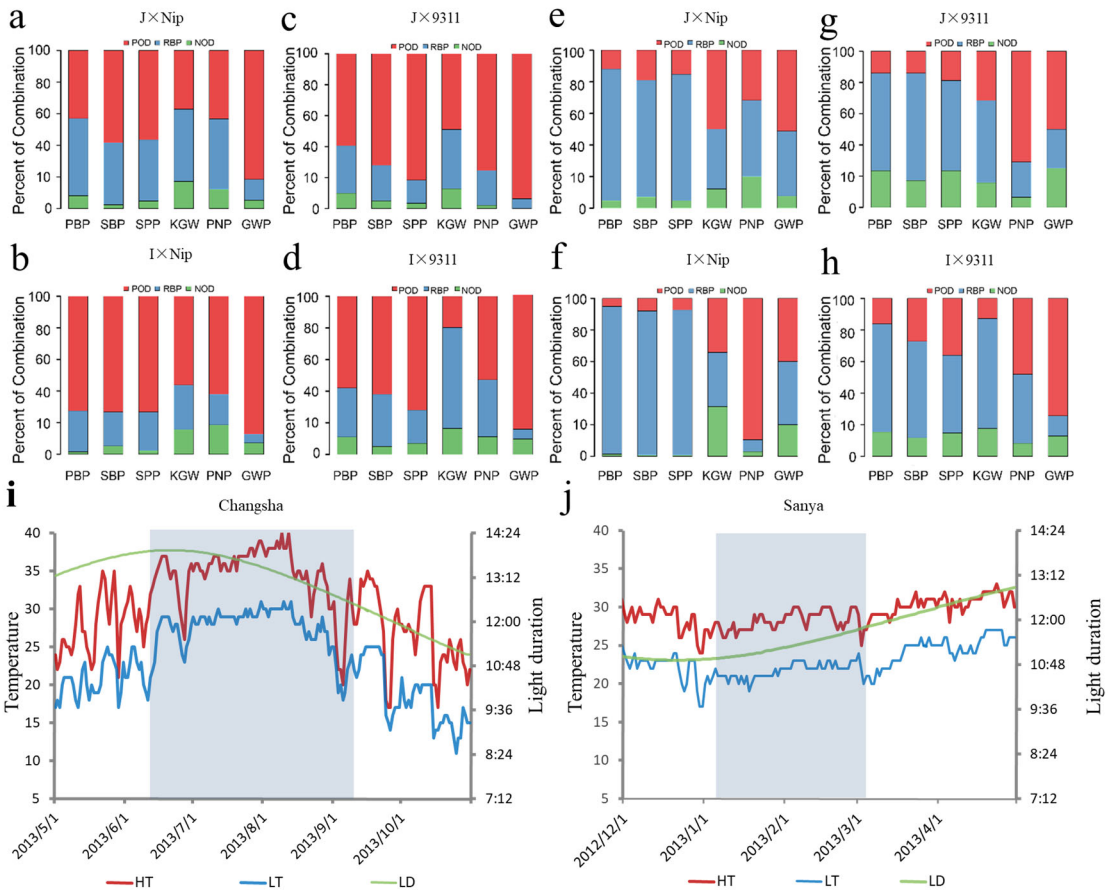
Supplementary Figure 6 The phenotype distribution of parent and F_1 for different yield traits in Changsha and Sanya. Nip, Nipponbare; J×Nip, combination of *Japonica* and Nipponbare; I×Nip, combination of *Indica* and Nipponbare, and the others are similar.



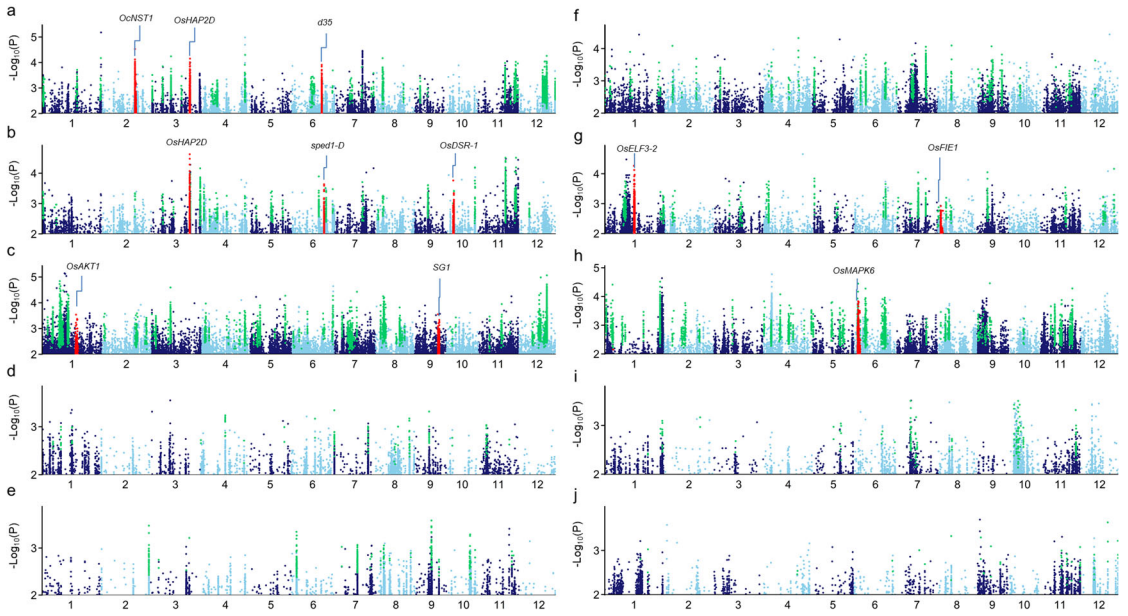
Supplementary Figure 7 Phenotypic contribution of parental inbred genetic basis (P) and middle parent heterosis (Hmp) to hybrids. PBP, primary branch number per panicle; SBP, secondary branch number per panicle; SPP, spikelet number per panicle; KGW, 1000-grain weigth; PNP, panicle number per plant; GWP, grain weight per plant; all means all combinations for *Japonica*, *Indica*, Nipponbare and 9311 in both environments.



Supplementary Figure 8 Supplementary Figure 8 Residual and environmental variance of yield traits in inbred parents and hybrids. (a) The proportion of residual variance of each yield trait estimated in inbred parents and hybrid in rice. (b) Proportion of environmental variance in total variance for yield related traits in panels of inbred parents and hybrids in maize (data were collected from Flint-Garcia., Buckler. et al. 2009). (c) The proportion of residual variance of each yield trait estimated in parents and hybrid in maize. DTT, Days to anthesis; PLTHT, Plant height; UPLFANG, Upper leaf angle; LFWDT, Leaf width; LFLEN, leaf length; 10Kwt, 10 kernel weight; CobDia, cob diameter; KnHgt, kernel height; EarLength, Ear length; CobWt, Cob Weight; TolKnWt, Total Kernel Weight.

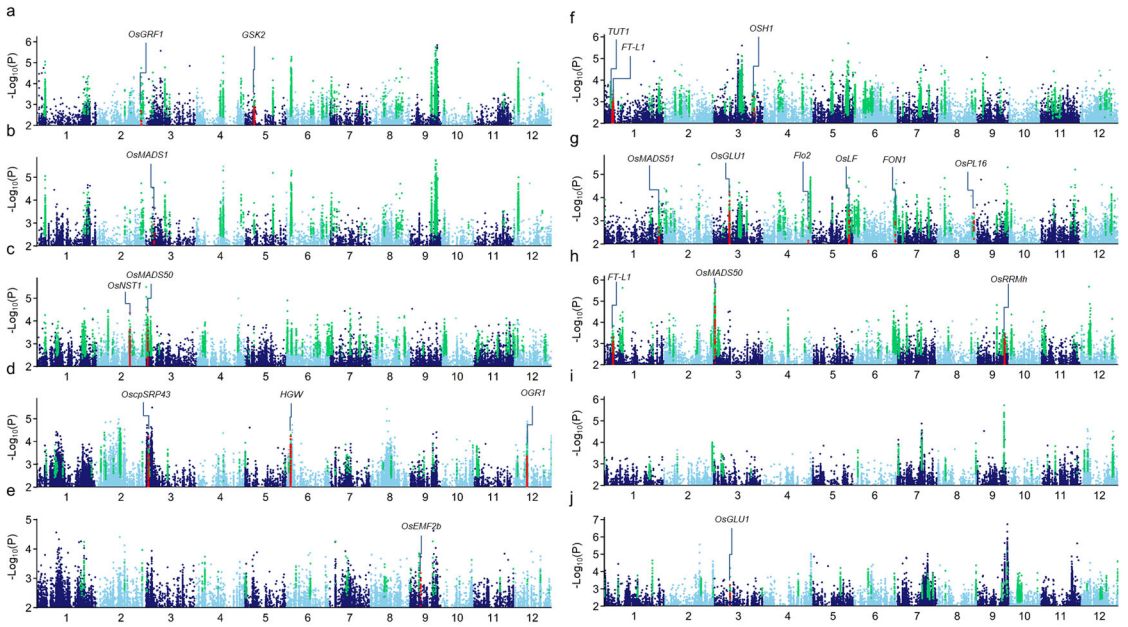


Supplementary Figure 9 The distribution of combinations showing POD, RBP and NOD phenotype for different yield traits and different kinds of combinations under two environments. (a-d) The distribution in Changsha. (e-f) The distribution in Sanya. (i) The temperature and light duration during the growth season in Changsha. (j) The temperature and light duration during the growth season in Sanya. POD, F_1 showing phenotype over the higher parent, referred as positive over-dominant (POD); RBP, F_1 showing phenotype ranging between parents referred as RBP; NOD, F_1 showing phenotype below the lower parent, referred as negative over-dominant (NOD). The grey boxes indicated the stage from reproductive initiation to grain filling. HT, high temperature; LT, low temperature; LD, light duration.



315

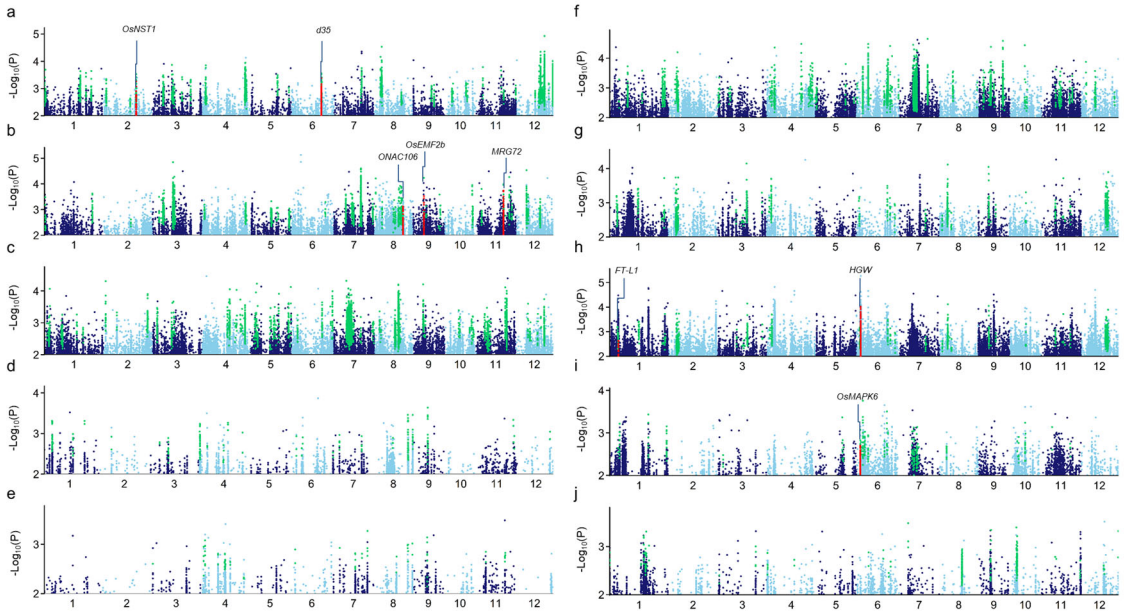
316 **Supplementary Figure 10 Genome-wide association study of primary branch number per panicle**
317 **(PBP) in *Japonica* parents and their combinations using compressed MLM.** (a) Manhattan plots
318 for J×Nip F_1 phenotype in Changsha. (b) Manhattan plots for J×Nip mid-parent heterosis value
319 in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for
320 J×9311 F_1 phenotype in Changsha. (e) Manhattan plots for J×9311 mid-parent heterosis value in
321 Changsha. (f) Manhattan plots for J×Nip F_1 phenotype in Sanya. (g) Manhattan plots for J×Nip
322 mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i)
323 Manhattan plots for J×9311 F_1 in Sanya. (j) Manhattan plots for J×9311 mid-parent heterosis in
324 Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the
325 SNPs in cloned genes with 2 kb promoter.
326



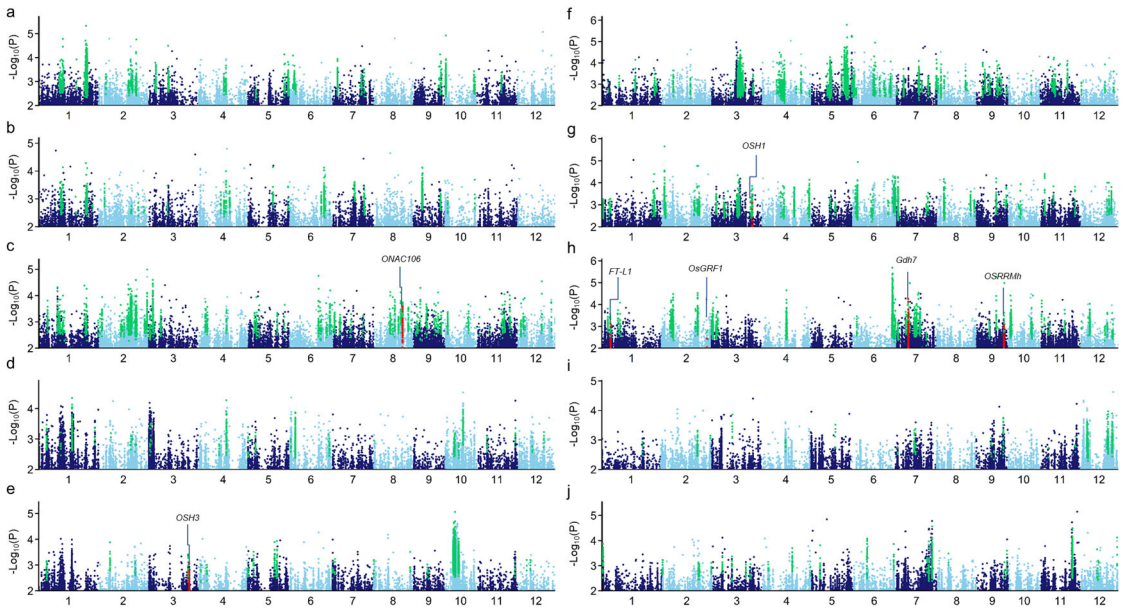
327

328 **Supplementary Figure 11 Genome-wide association study of primary branch number per panicle**

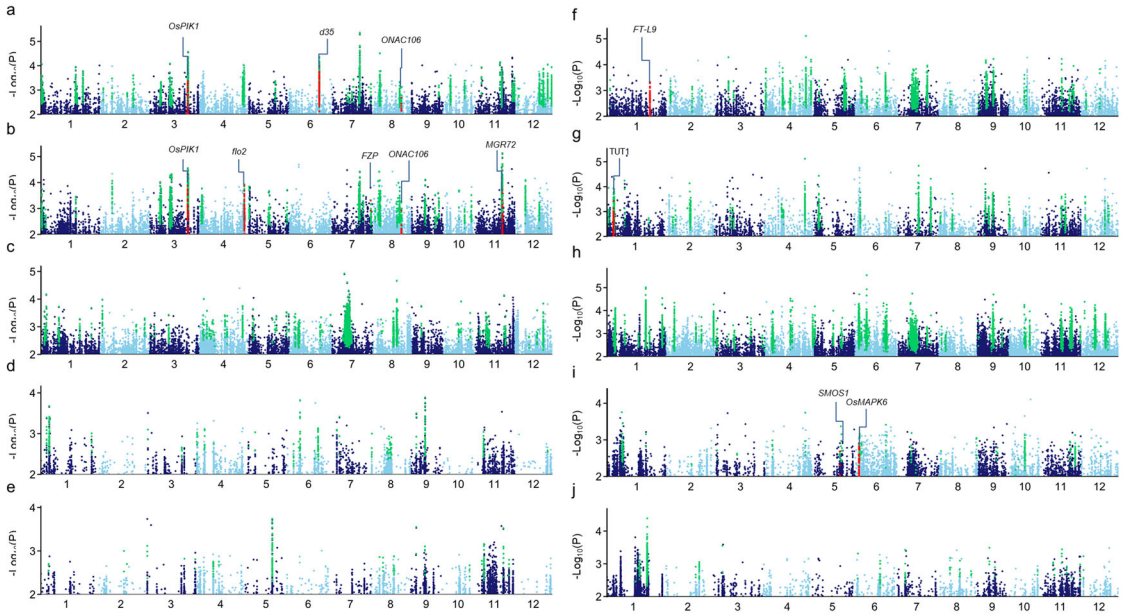
(PBP) in *Indica* parents and their combinations using compressed MLM. (a) Manhattan plots for I×Nip F_1 phenotype in Changsha. (b) Manhattan plots for I×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for I×9311 F_1 phenotype in Changsha. (e) Manhattan plots for I×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for I×Nip F_1 phenotype in Sanya. (g) Manhattan plots for I×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for I×9311 F_1 in Sanya. (j) Manhattan plots for I×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.



Supplementary Figure 12 Genome-wide association study of secondary branch number per panicle (SBP) in *Japonica* parents and their combinations using compressed MLM. (a) Manhattan plots for J×Nip F_1 phenotype in Changsha. (b) Manhattan plots for J×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for J×9311 F_1 phenotype in Changsha. (e) Manhattan plots for J×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for J×Nip F_1 phenotype in Sanya. (g) Manhattan plots for J×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for J×9311 F_1 in Sanya. (j) Manhattan plots for J×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.

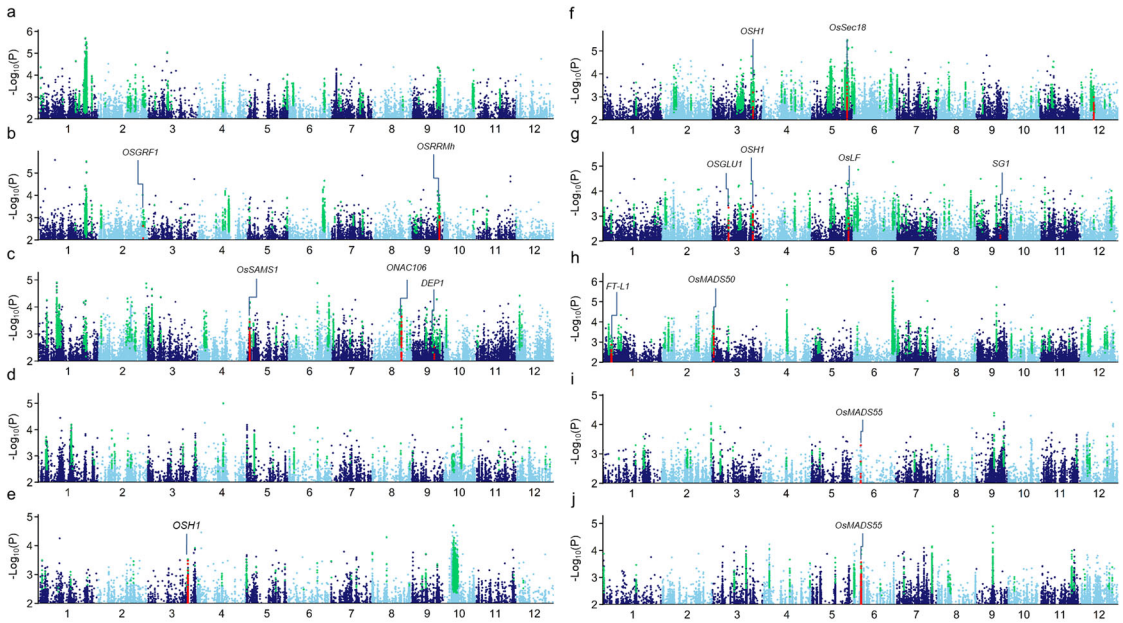


Supplementary Figure 13 Genome-wide association study of secondary branch number per panicle (SBP) in *Indica* parents and their combinations using compressed MLM. (a) Manhattan plots for I×Nip F_1 phenotype in Changsha. (b) Manhattan plots for I×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for I×9311 F_1 phenotype in Changsha. (e) Manhattan plots for I×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for I×Nip F_1 phenotype in Sanya. (g) Manhattan plots for I×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for I×9311 F_1 in Sanya. (j) Manhattan plots for I×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.

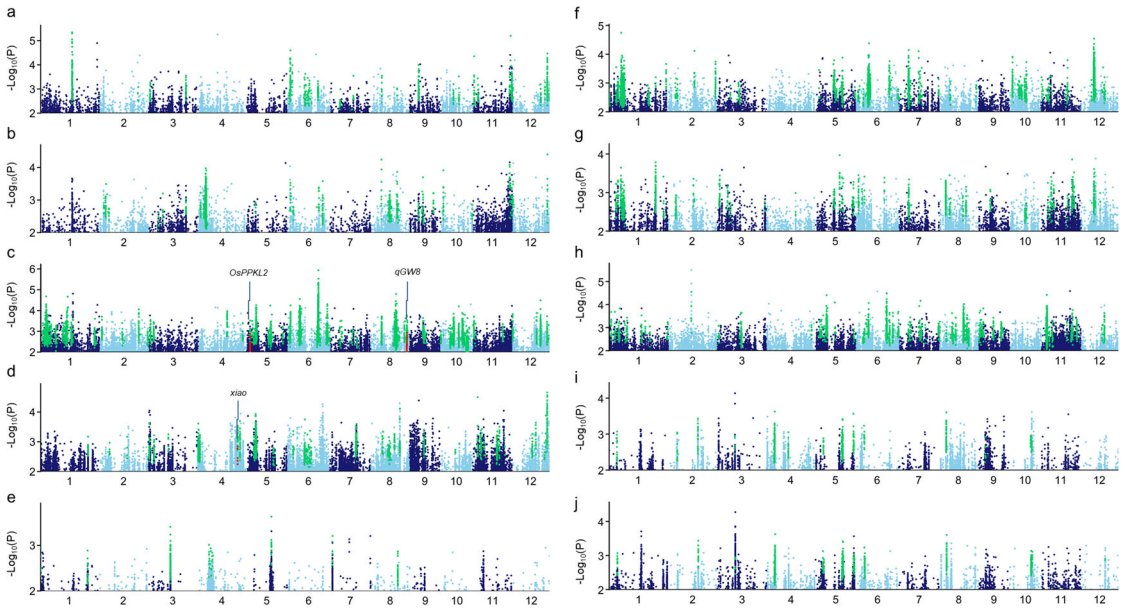


Supplementary Figure 14 Genome-wide association study of spikelet number per panicle (SPP) in

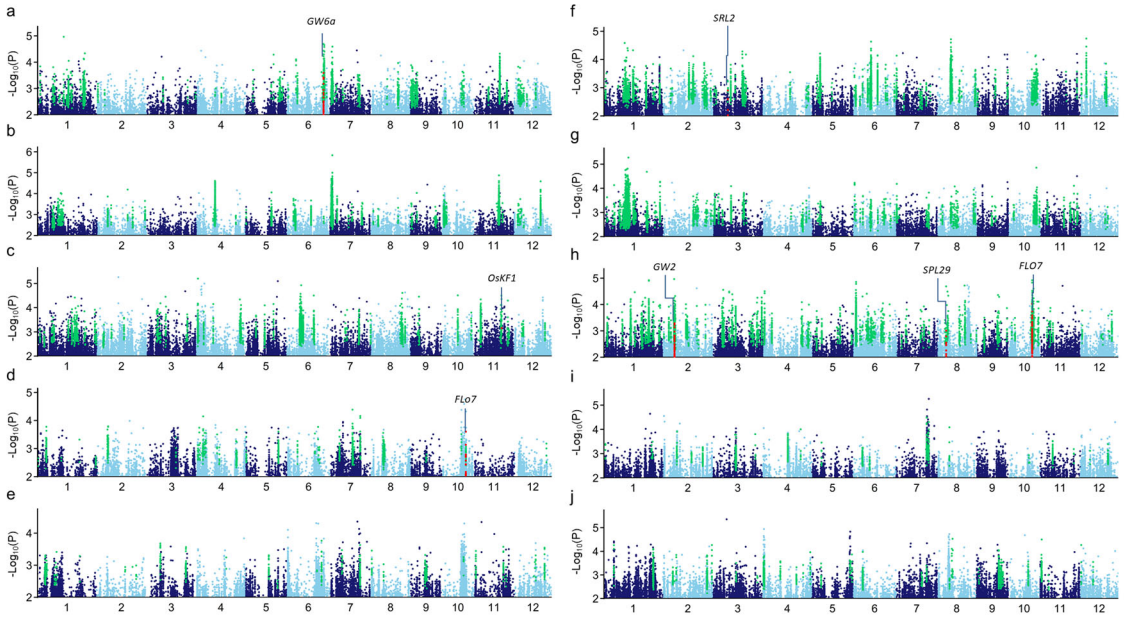
***Japonica* parents and their combinations using compressed MLM.** (a) Manhattan plots for J×Nip F_1 phenotype in Changsha. (b) Manhattan plots for J×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for J×9311 F_1 phenotype in Changsha. (e) Manhattan plots for J×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for J×Nip F_1 phenotype in Sanya. (g) Manhattan plots for J×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for J×9311 F_1 in Sanya. (j) Manhattan plots for J×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.



Supplementary Figure 15 Genome-wide association studies of spikelet number per panicle (SPP) in *Indica* parents and their combinations using compressed MLM. (a) Manhattan plots for I×Nip F_1 phenotype in Changsha. (b) Manhattan plots for I×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for I×9311 F_1 phenotype in Changsha. (e) Manhattan plots for I×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for I×Nip F_1 phenotype in Sanya. (g) Manhattan plots for I×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for I×9311 F_1 in Sanya. (j) Manhattan plots for I×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.

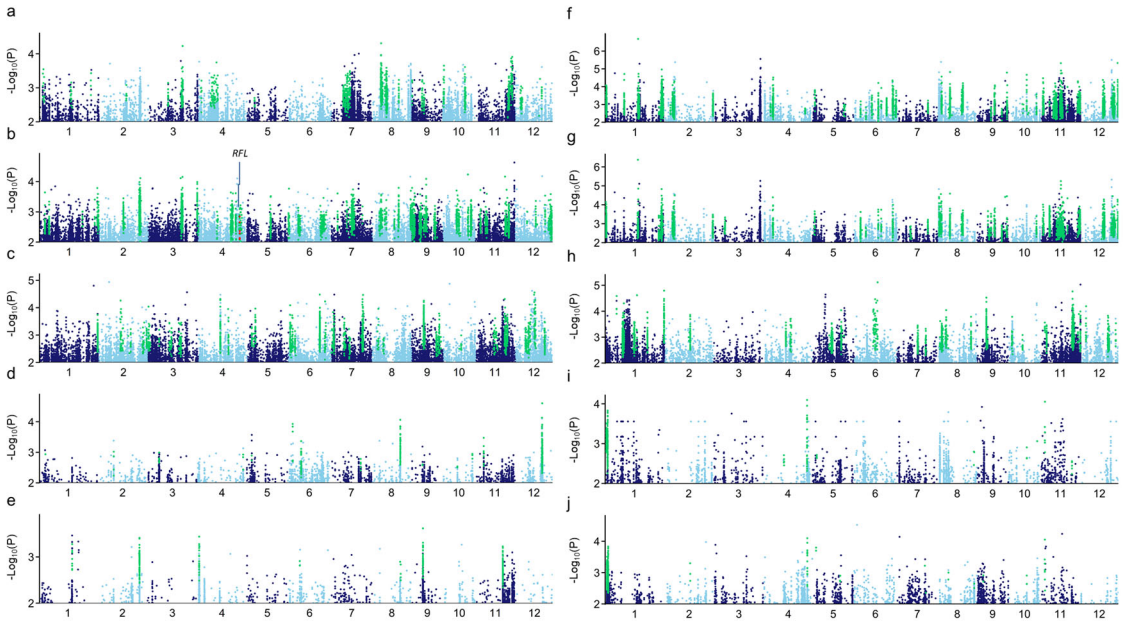


Supplementary Figure 16 Genome-wide association study of 1000-grain weight (KGW) in *Japonica* parents and their combinations using compressed MLM. (a) Manhattan plots for J×Nip F_1 phenotype in Changsha. (b) Manhattan plots for J×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for J×9311 F_1 phenotype in Changsha. (e) Manhattan plots for J×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for J×Nip F_1 phenotype in Sanya. (g) Manhattan plots for J×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for J×9311 F_1 in Sanya. (j) Manhattan plots for J×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.

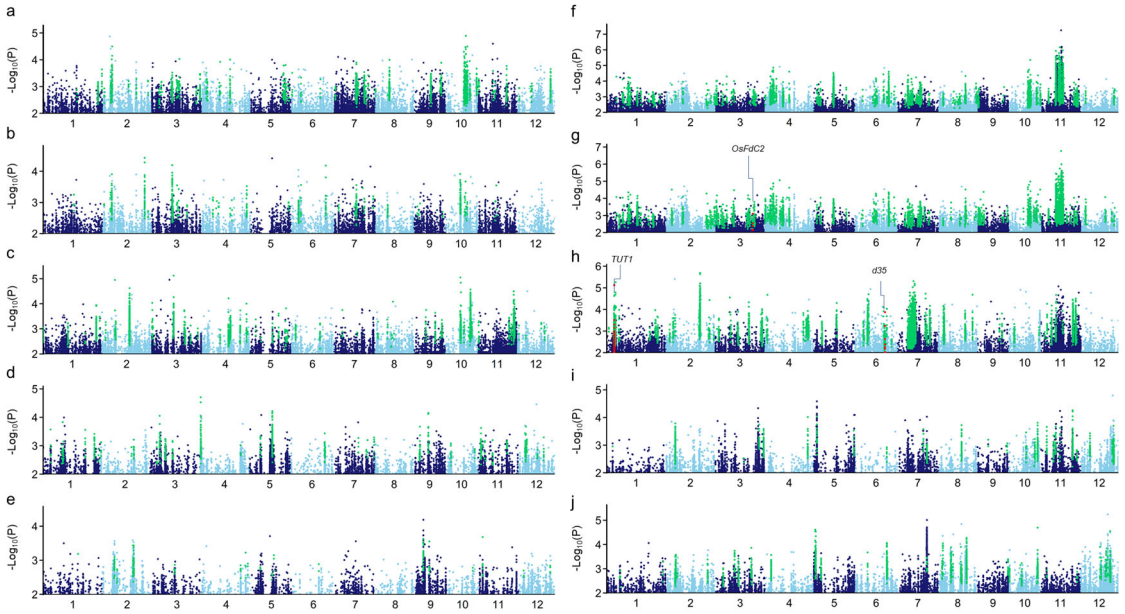


Supplementary Figure 17 Genome-wide association study of 1000-grain weight (KGW) in *Indica* parents and their combinations using compressed MLM. (a) Manhattan plots for I×Nip F_1 phenotype in Changsha. (b) Manhattan plots for I×Nip mid-parent heterosis value in Changsha.

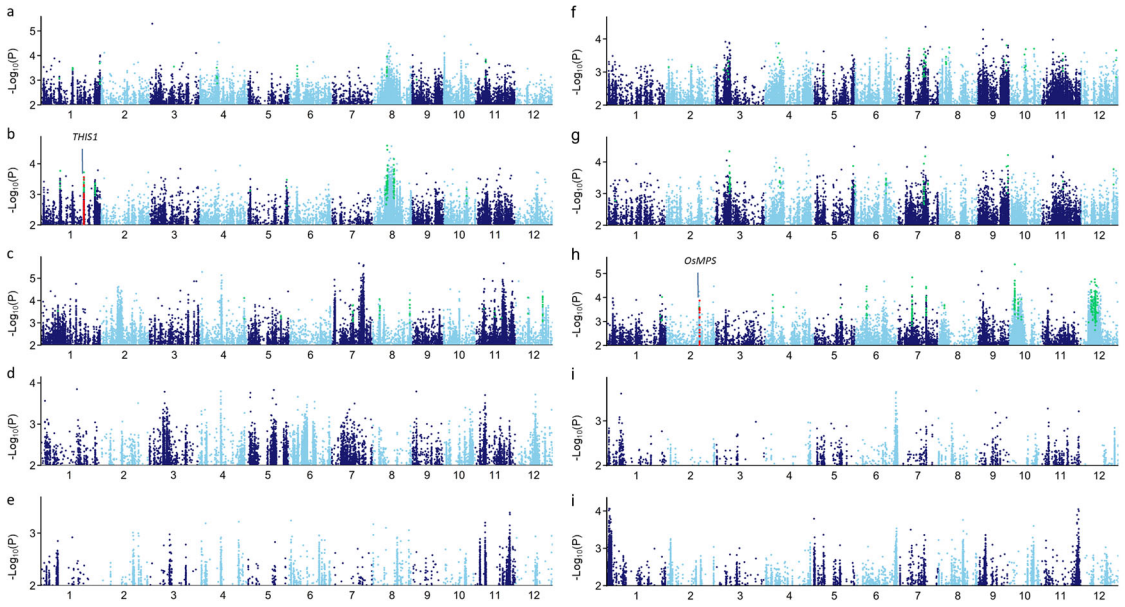
(c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for I×9311 F_1 phenotype in Changsha. (e) Manhattan plots for I×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for I×Nip F_1 phenotype in Sanya. (g) Manhattan plots for I×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for I×9311 F_1 in Sanya. (j) Manhattan plots for I×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.



Supplementary Figure 18 Genome-wide association study of panicle number per plant (PNP) in *Japonica* parents and their combinations using compressed MLM. (a) Manhattan plots for J×Nip F_1 phenotype in Changsha. (b) Manhattan plots for J×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for J×9311 F_1 phenotype in Changsha. (e) Manhattan plots for J×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for J×Nip F_1 phenotype in Sanya. (g) Manhattan plots for J×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for J×9311 F_1 in Sanya. (j) Manhattan plots for J×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.

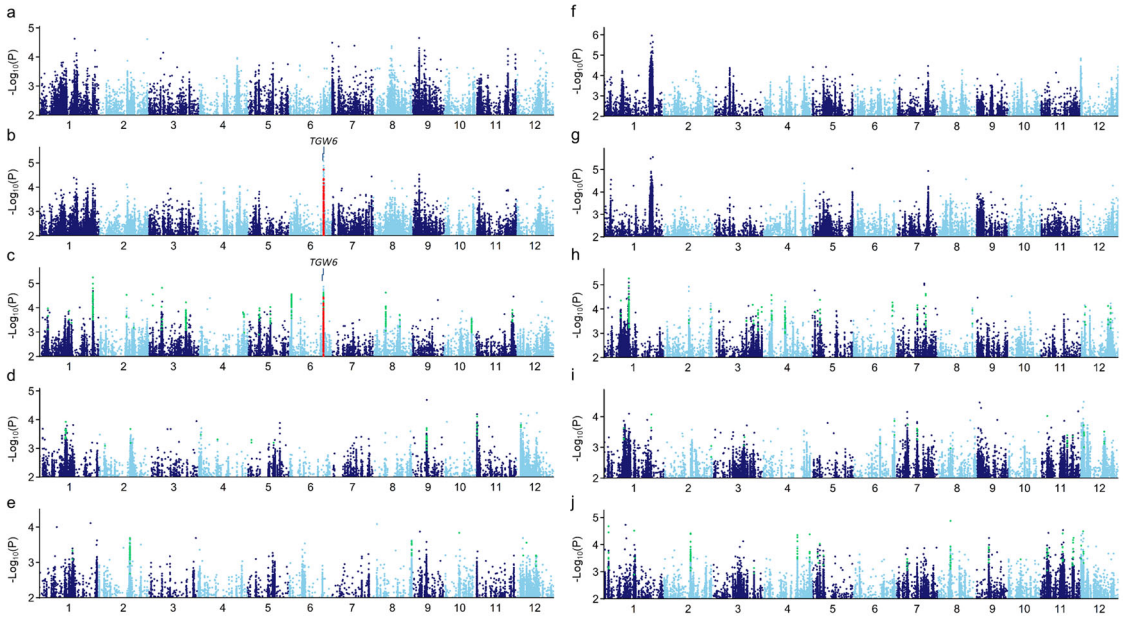


Supplementary Figure 19 Genome-wide association study of panicle number per plant (PNP) in *Indica* parents and their combinations using compressed MLM. (a) Manhattan plots for I×Nip F_1 phenotype in Changsha. (b) Manhattan plots for I×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for I×9311 F_1 phenotype in Changsha. (e) Manhattan plots for I×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for I×Nip F_1 phenotype in Sanya. (g) Manhattan plots for I×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for I×9311 F_1 in Sanya. (j) Manhattan plots for I×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.

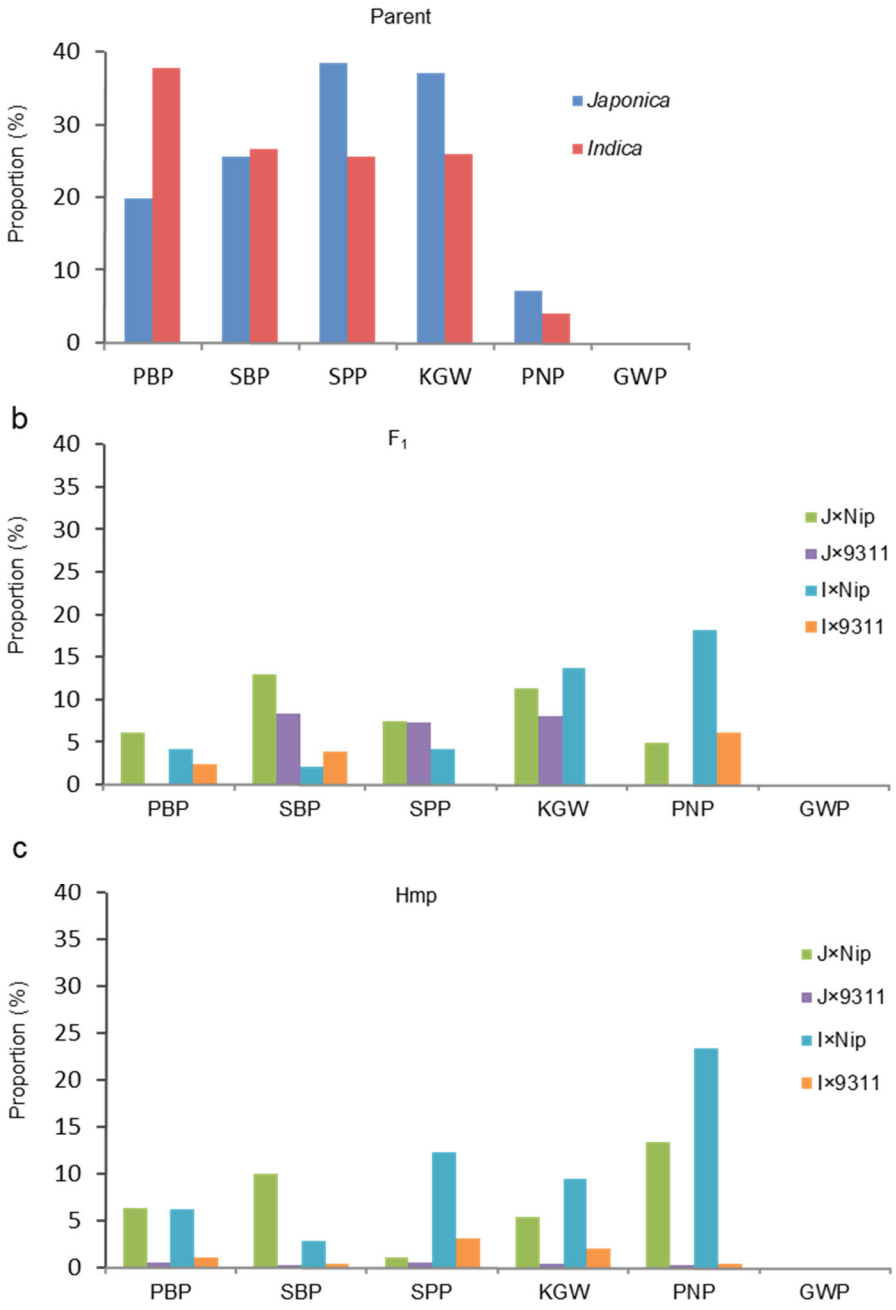


Supplementary Figure 20 Genome-wide association study of grain weight per plant (GWP) in *Japonica* parents and their combinations using compressed MLM. (a) Manhattan plots for J×Nip F_1 phenotype in Changsha. (b) Manhattan plots for J×Nip mid-parent heterosis value in Changsha. (c)

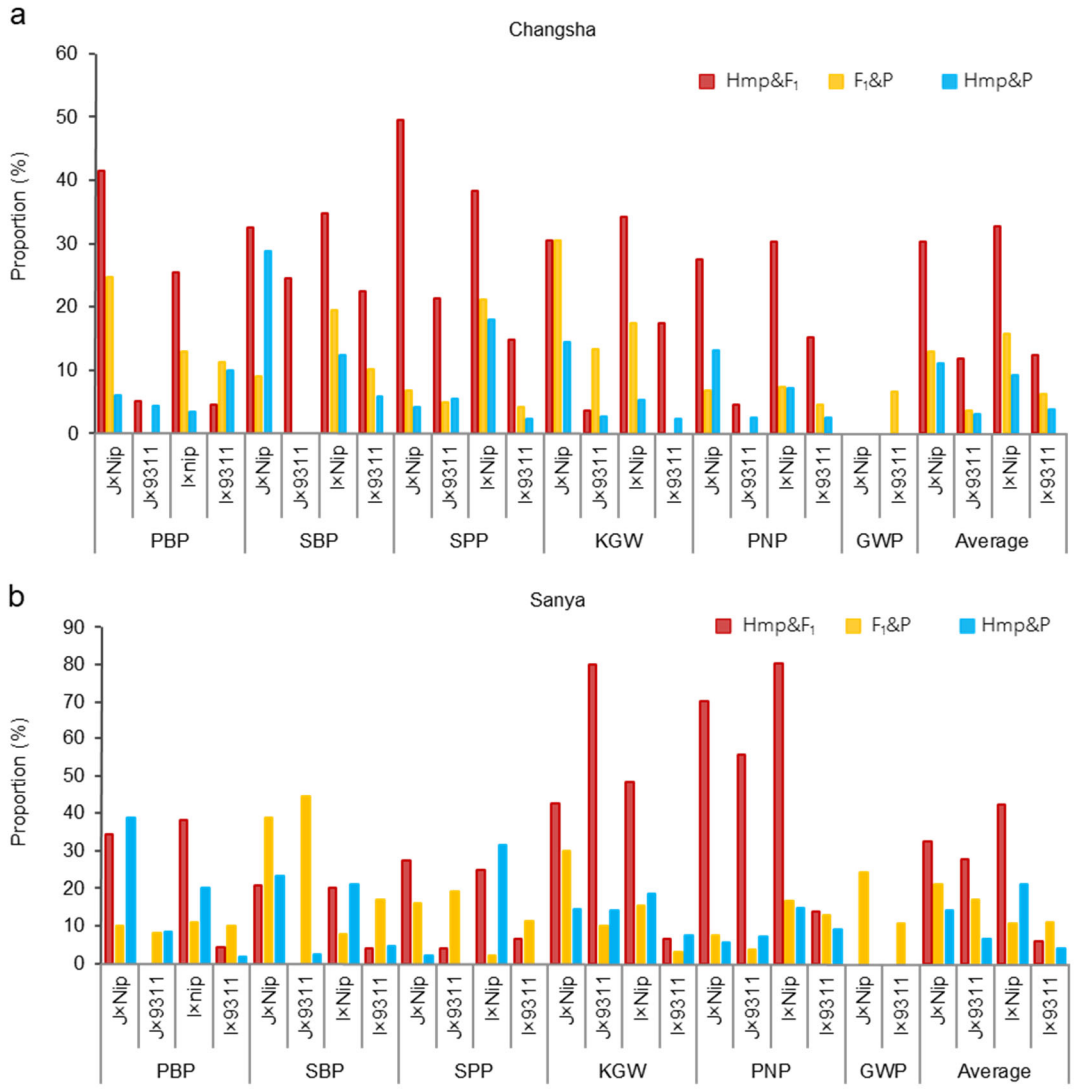
Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for J×9311 F_1 phenotype in Changsha. (e) Manhattan plots for J×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for J×Nip F_1 phenotype in Sanya. (g) Manhattan plots for J×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for J×9311 F_1 in Sanya. (j) Manhattan plots for J×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.



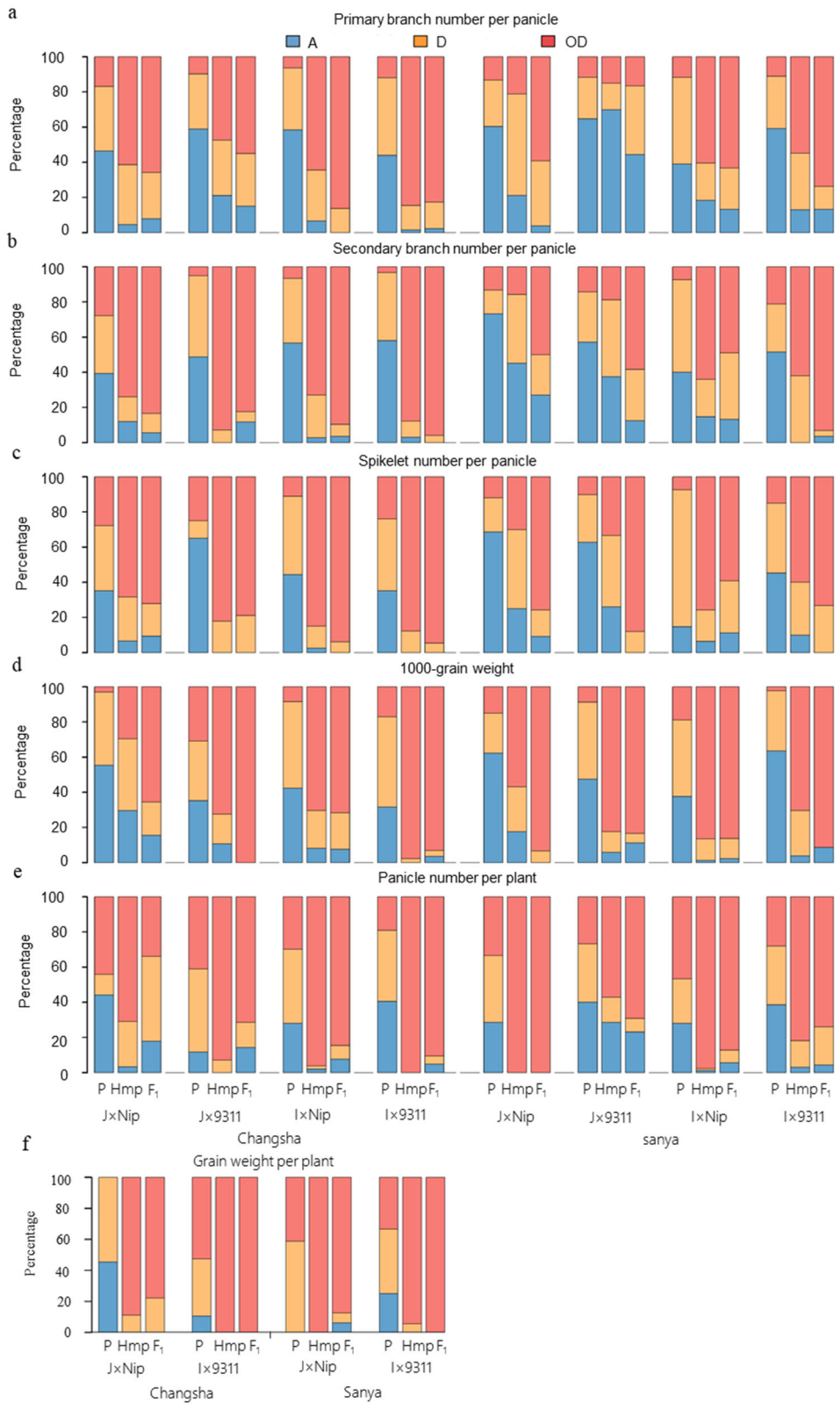
Supplementary Figure 21 Genome-wide association study of grain weight per plant (GWP) in *Indica* parents and their combinations using compressed MLM. (a) Manhattan plots for I×Nip F_1 phenotype in Changsha. (b) Manhattan plots for I×Nip mid-parent heterosis value in Changsha. (c) Manhattan plots for phenotype of parents in Changsha. (d) Manhattan plots for I×9311 F_1 phenotype in Changsha. (e) Manhattan plots for I×9311 mid-parent heterosis value in Changsha. (f) Manhattan plots for I×Nip F_1 phenotype in Sanya. (g) Manhattan plots for I×Nip mid-parent heterosis value in Sanya. (h) Manhattan plots for phenotype of parents in Sanya. (i) Manhattan plots for I×9311 F_1 in Sanya. (j) Manhattan plots for I×9311 mid-parent heterosis in Sanya. Green dots represent the significant SNPs in identified QTLs, red dots represent the SNPs in cloned genes with 2 kb promoter.



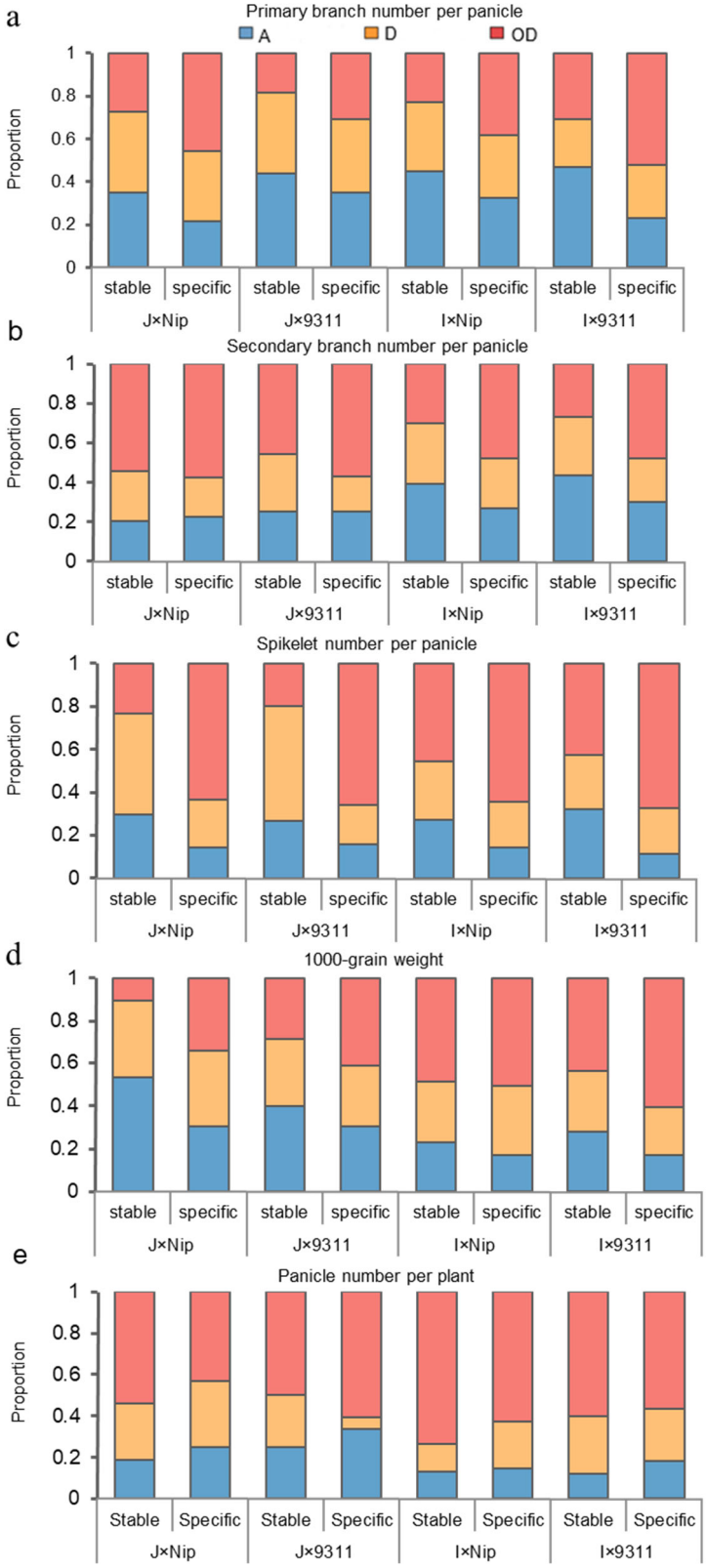
Supplementary Figure 22 The colocalized QTL between two environments for yield and its sub-component traits. (a) The proportion of colocalized QTL between Changsha and Sanya for P_QTL. (b) The proportion of colocalized QTL between Changsha and Sanya for F_1 _QTL. (c) The proportion of colocalized QTL between Changsha and Sanya for Hmp_QTL.



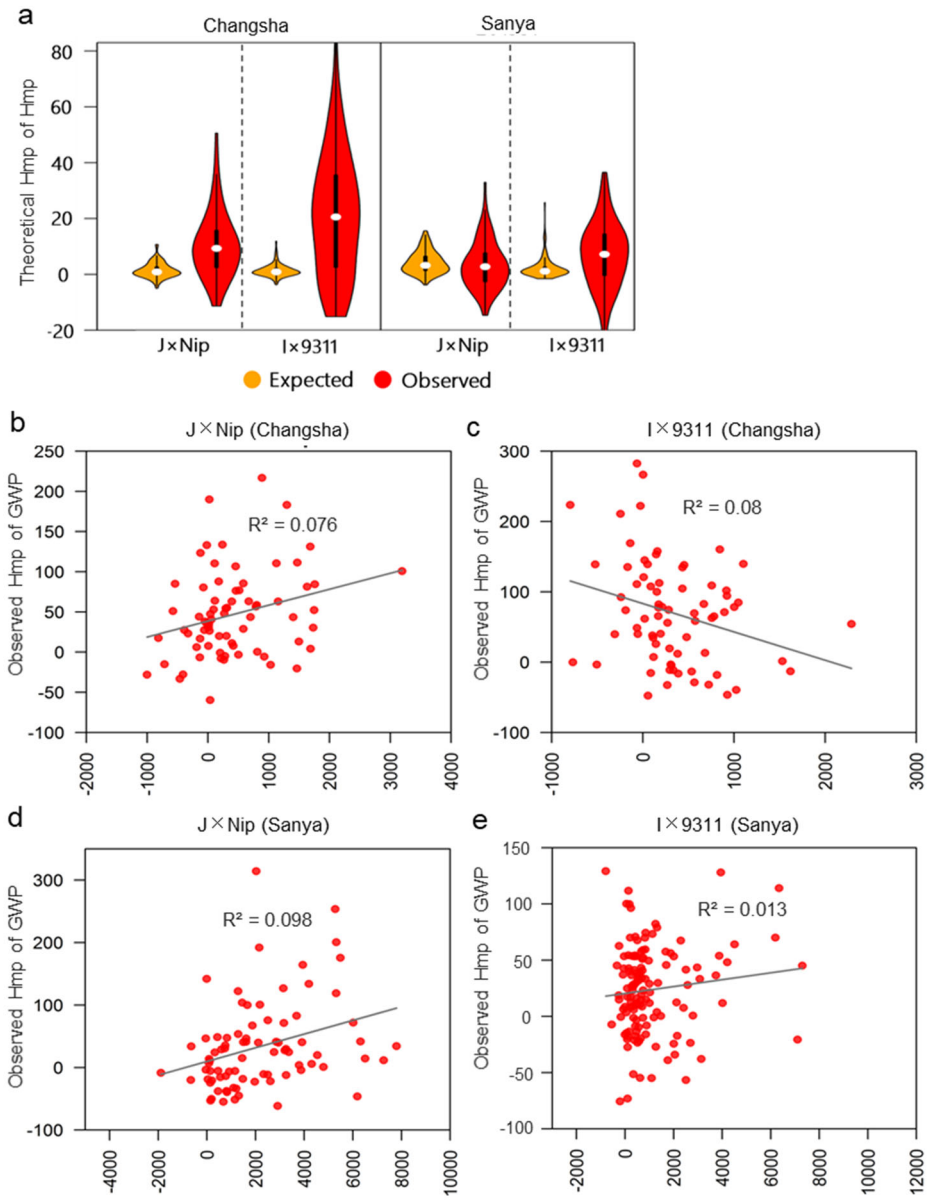
Supplementary Figure 23 The colocalized QTL between F_1_QTL, Hmp_QTL and _QTL for grain yield and its sub-component traits. (a) The proportion of colocalized QTL between F_1_QTL, Hmp_QTL and P_QTL for the trait of PBP, SBP, SPP, KGW, PNP and GWP in Changsha. (b) The proportion of colocalized QTLs between F_1_QTL, Hmp_QTL and P_QTL for the trait of PBP, SBP, SPP, KGW, PNP and GWP in Sanya.



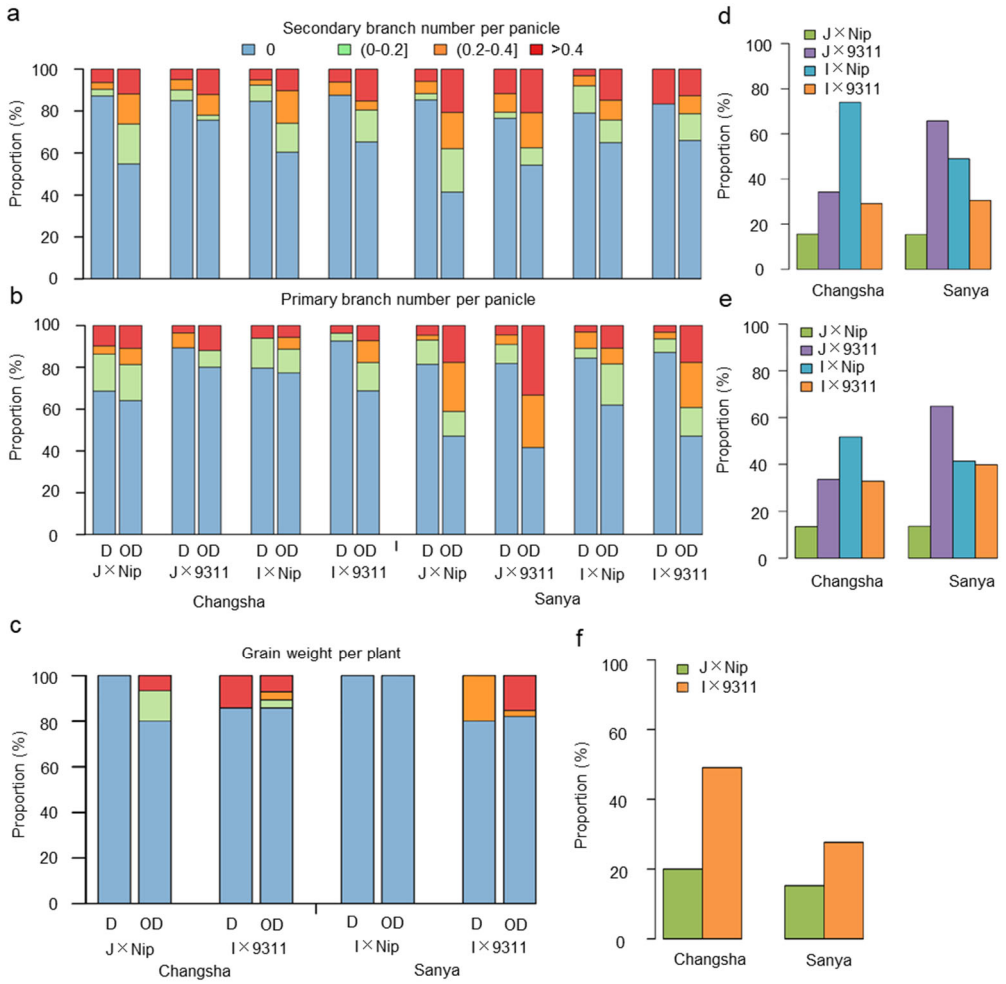
Supplementary Figure 24 The percentage of detected QTLs with additive, dominant and over-dominant effects for each yield trait using the phenotype of parents (P), the phenotype of F_1 (F_1) and the middle parent heterosis value (Hmp) in different combinations under two environments respectively. A, additive (in blue); D, dominant (in orange); OD, over-dominant (in red).



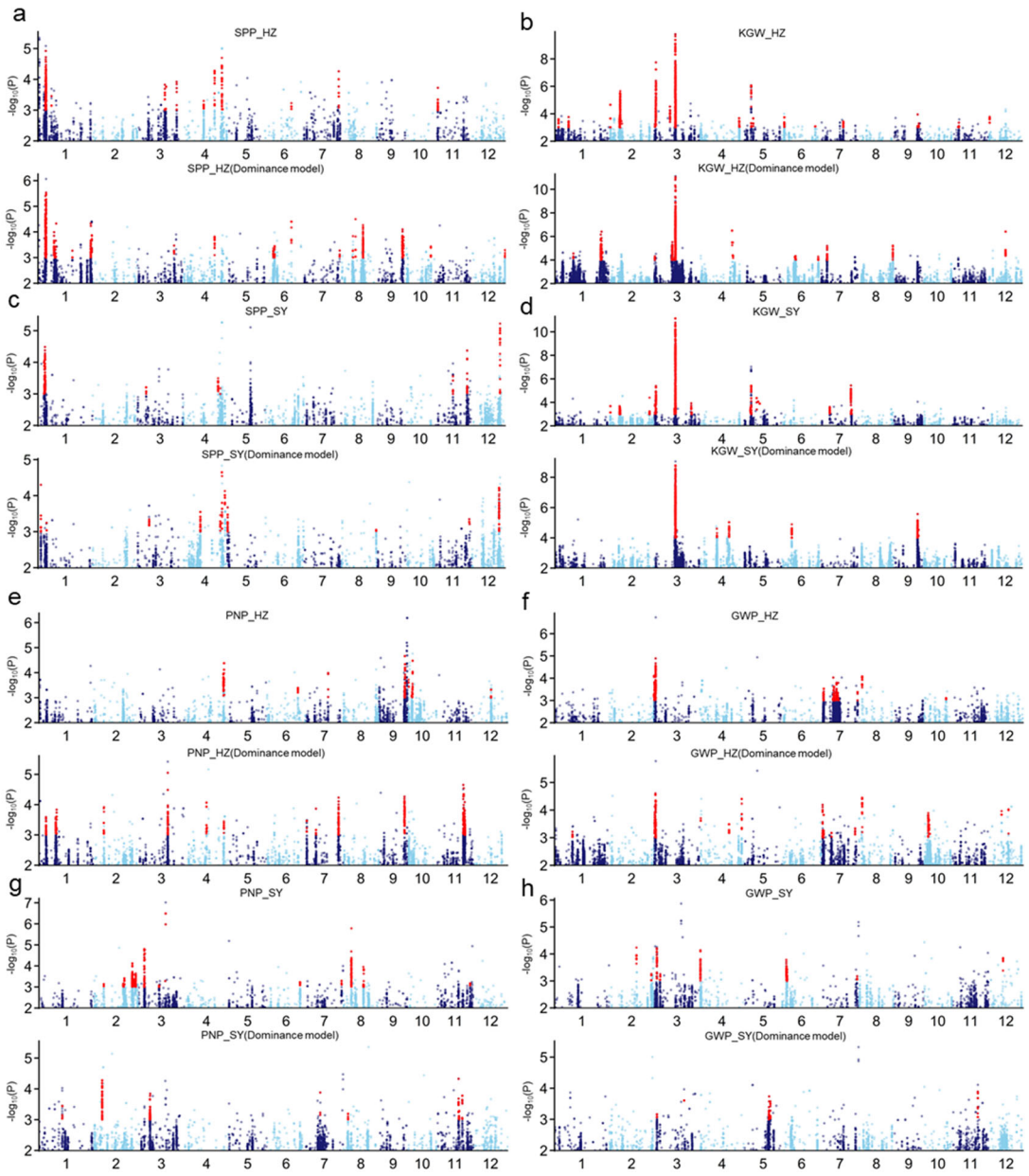
Supplementary Figure 25 The proportion of colocalized QTLs between two environments with additive, dominant and over-dominant effects. Stable, means the colocalized QTLs between two environments; Specific, means the QTLs can only be detected in one environment.



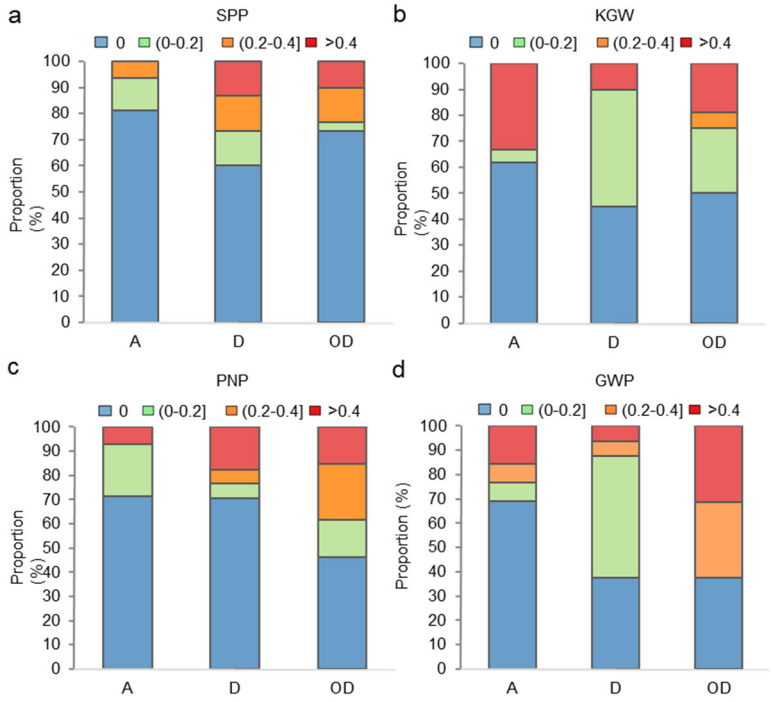
Supplementary Figure 26 Comparison between the observed Hmp of GWP and the theoretical Hmp of GWP estimated according to the multiplicative from additive effect of three main yield components (SPP, KGW and PNP). (a) The violin plot for the observed Hmp of GWP and the theoretical Hmp of GWP. (b-e) The scatter plot for the observed Hmp of GWP and the theoretical Hmp of GWP for different combinations in two environments.



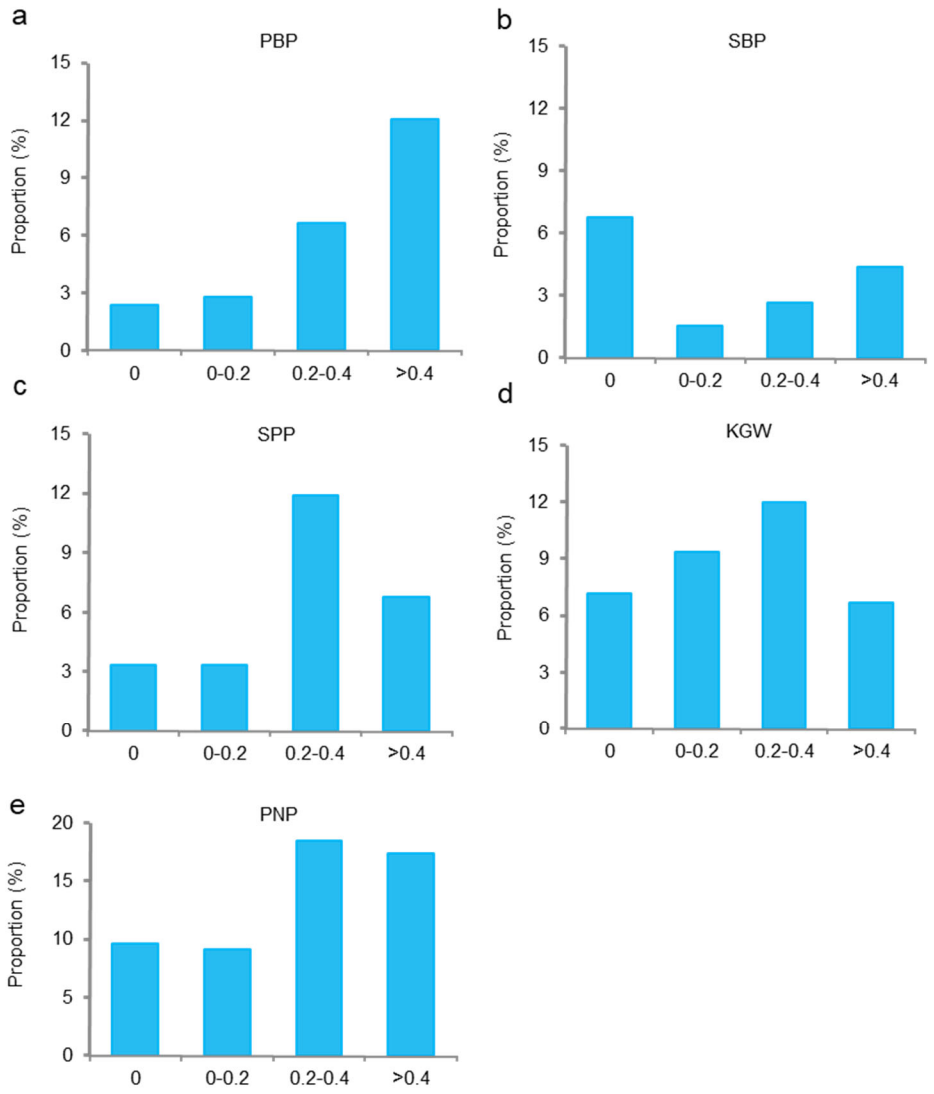
Supplementary Figure 27 The repulsive degree in dominant and over-dominant QTLs. (a-c) The proportion of QTLs with different repulsive degrees; here, the sky blue represents that there is no significant SNP with repulsive additive effects within the QTL, and the light green, orange and red represent the repulsive degree within the range of (0-0.2], (0.2-0.4] and >0.4 within each QTL. (e-f) The average proportion of combinations with repulsive effect alleles per QTL that contains repulsive effect alleles.



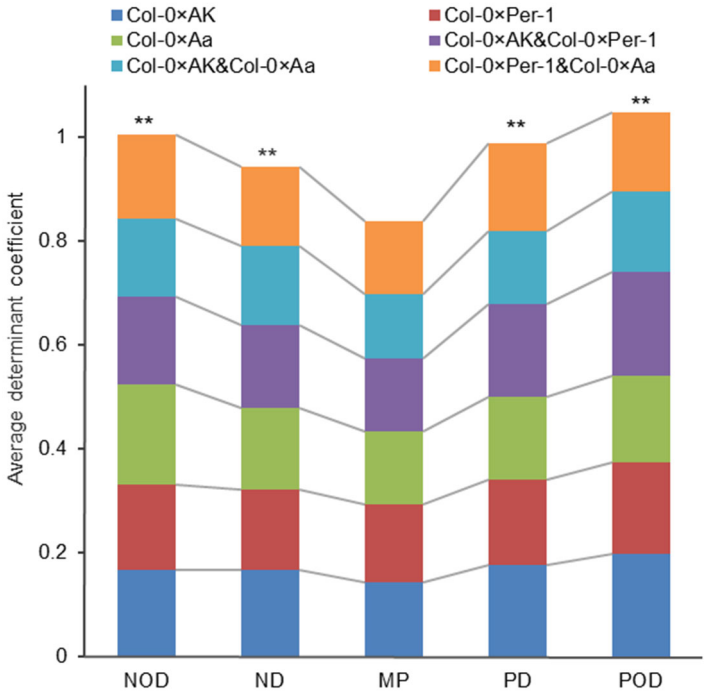
Supplementary Figure 28 Genome-wide association study of yield traits in 1086 three-line system hybrids using compressed MLM. (a) Manhattan plot for SPP in Hangzhou (HZ). (b) Manhattan plot for SPP in HZ by dominance coded genotype. (c) Manhattan plot for KGW in Sanya (SY). (d) Manhattan plot for KGW in SY by dominance coded genotype. (e) Manhattan plot for PNP in HZ. (f) Manhattan plot for PNP in HZ by dominance coded genotype. (g) Manhattan plot for GWP in SY. (h) Manhattan plot for GWP in SY by dominance coded genotype. The red dots are the significant SNPs located within the defined QTL. The data were collected from the published paper (Huang, X.H. et al. Genomic analysis of hybrid rice varieties reveals numerous superior alleles that contribute to heterosis. Nature Communications 6(2015).).



Supplementary Figure 29 The repulsive degree in additive (A), dominant (D) and over-dominant (OD) QTLs identified in 1086 three-line hybrids. The sky blue represents that there is no significant SNP with repulsive additive effects within the QTL, and the light green, orange and red represent the repulsive degree within the range of (0-0.2], (0.2-0.4] and >0.4 within each QTL.

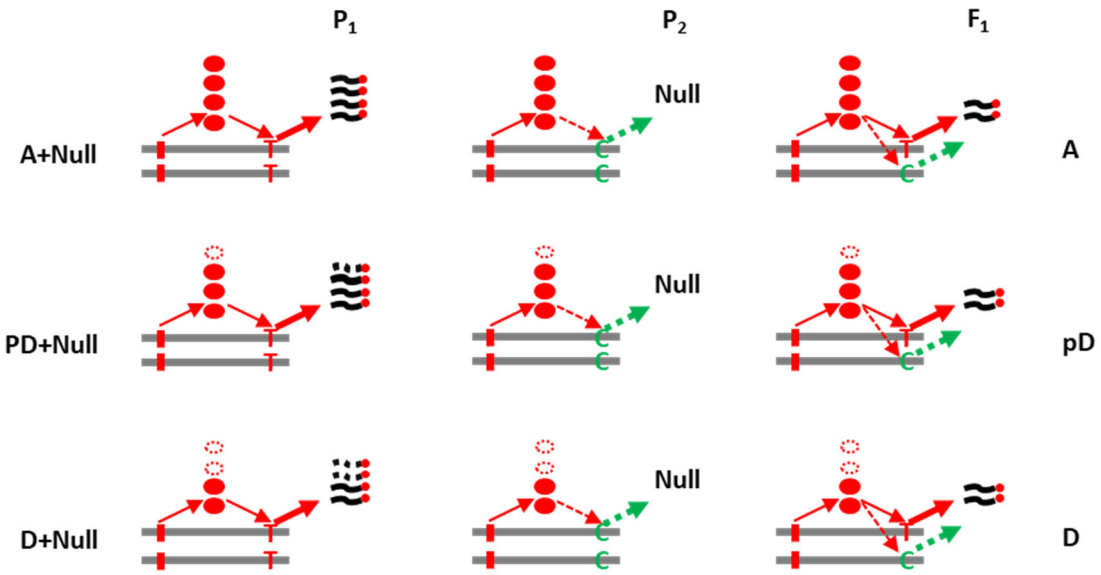


Supplementary Figure 30 The colocalization proportion for over-dominant QTLs with different repulsive degrees (0, 0-0.2, 0.2-0.4 and >0.4) for the trait of PBP (a), SBP (b), SPP (c), KGW (d) and PNP (e).

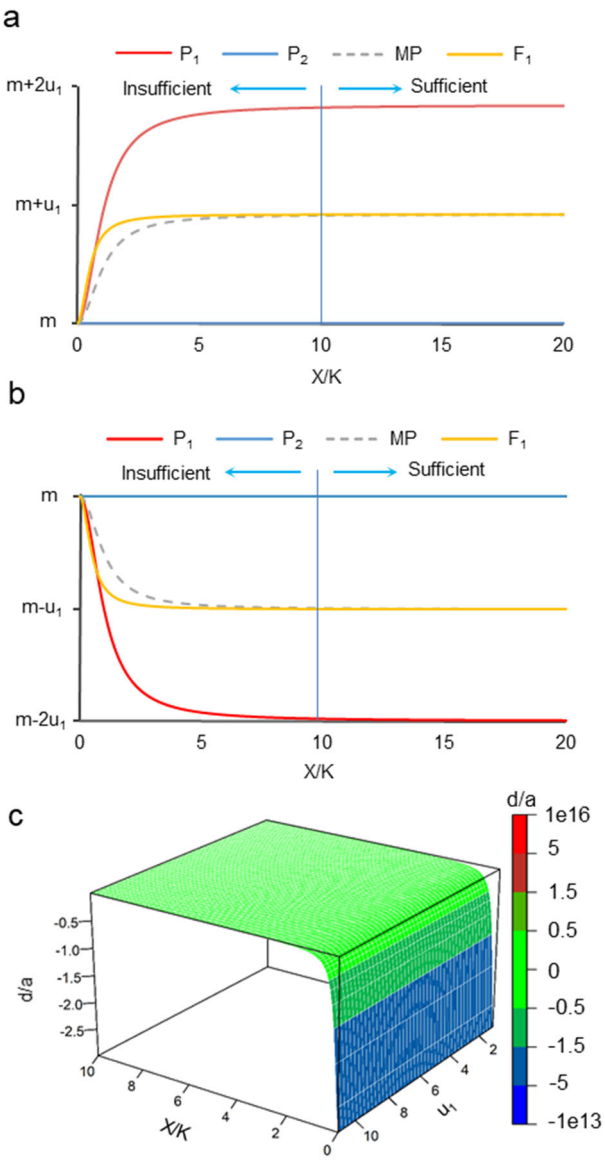


504

505 **Supplementary Figure 31 The average determinant coefficient between genes with different**
506 **expression patterns and their transcription factors in three *A. thaliana* combinations.** The
507 determination coefficient was estimated by six pairs of transcription levels in the first leaf between the
508 gene and its transcription factor across three *A. thaliana* combinations (including Col-0×Per-1, Col-0×Aa,
509 Col-0×Ak) and their parents. Here, Col-0×AK& Col-0×Per-1 refers to those genes that show the same
510 expression pattern in both combinations Col-0×AK and Col-0×Per-1. And the others are similar. We
511 estimated the significant difference of NOD, ND, PD and POD with MP by paired t-test; and "**"
512 marked the significant level at 0.01. The raw data were collected from the published paper (Yang, M. et al.
513 Genomic architecture of biomass heterosis. Proc Natl Acad Sci USA6 (2017)).

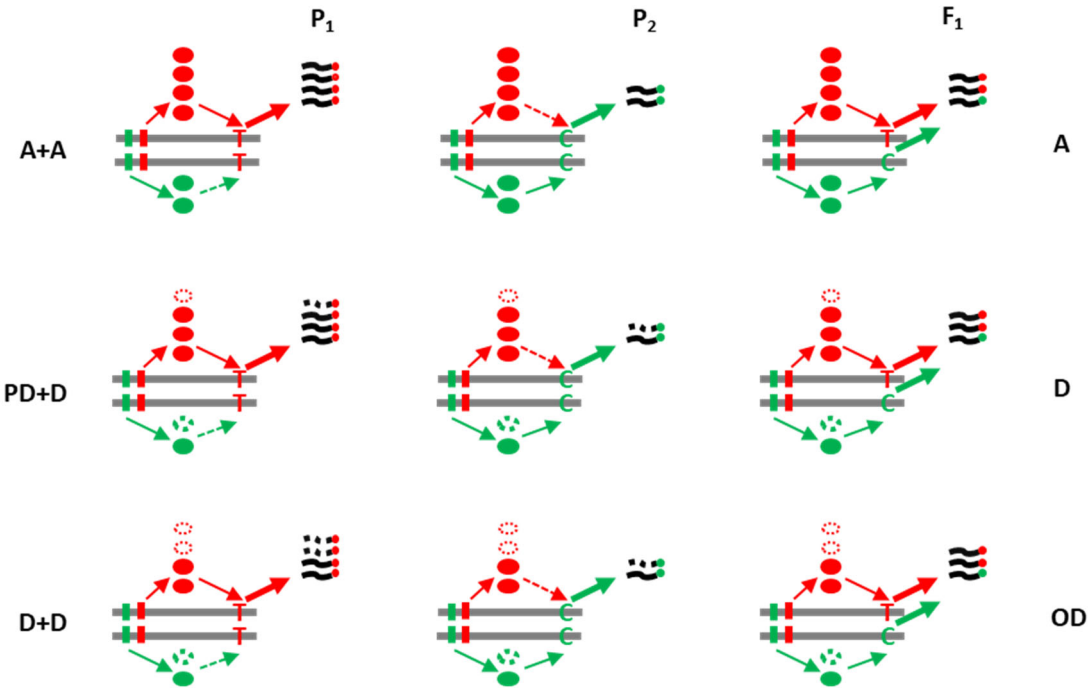


Supplementary Figure 32 The schematic diagram of regulation model for molecular mechanism of additive and dominant effect produced by single site with null allele (C) and one functional allele (T) under one positive regulator background. The grey thick lines show two chromosomes, the red bars on which are two homologous alleles of the regulator that are uniform among P_1, P_2 and F_1. The break and solid pies together represent the required regulator function that can maximize the function of T/T homozygote of the target site, and the solid pies represent different regulator functions and thus provide different backgrounds to the target site. The arrow represents the function process, and the break arrow represents the break function process. The break and solid curves together represent the maximum function of the T/T homozygote in parents or one T allele in F_1, and the solid curves represent the real function. A + Null shows the action mode for each allele and between two alleles (T vs C) that they are independent with T being additive and C being null; and PD+Null and D+Null are similar. A, pD and D mean additive, partial dominant, and dominant effect, respectively.

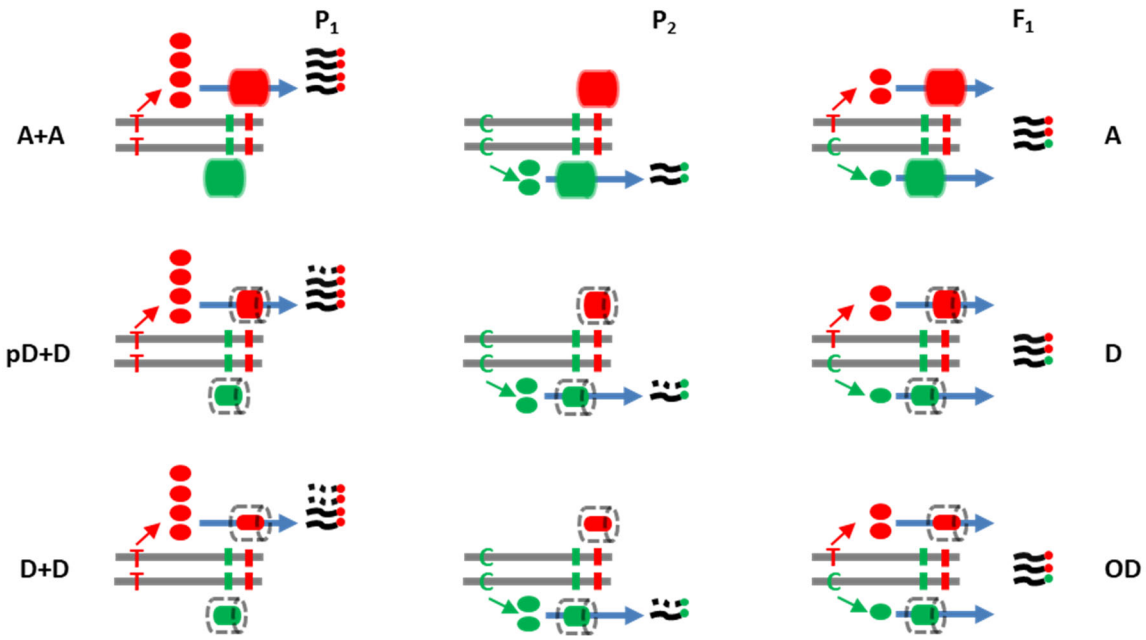


Supplementary Figure 33 The simulated diagram of regulation model for molecular mechanism of additive and dominant effect produced by single site with null allele and one functional allele under one regulator background. (a) The performance of the target site in two parents, F_1 and the middle parent (MP) under the activator background with different sufficiencies (X/K). It was simulated according to Hill function with $\mu_1 = 3$ and $n = 2$. μ_1 means the maximum function at steady state for one functional allele. n is the Hill coefficient. Left arrow represents a relatively insufficient activator background, and the right arrow represents the relatively sufficient activator background. (b) The performance of the target site in two parents, F_1 and the middle parent (MP) under the repressor background with different sufficiencies (X/K). It was simulated according to Hill function with $\mu_1 = 3$ and $n = 2$. μ_1 means the maximum function at steady state for one functional allele. n is the Hill coefficient. (c) The dominant

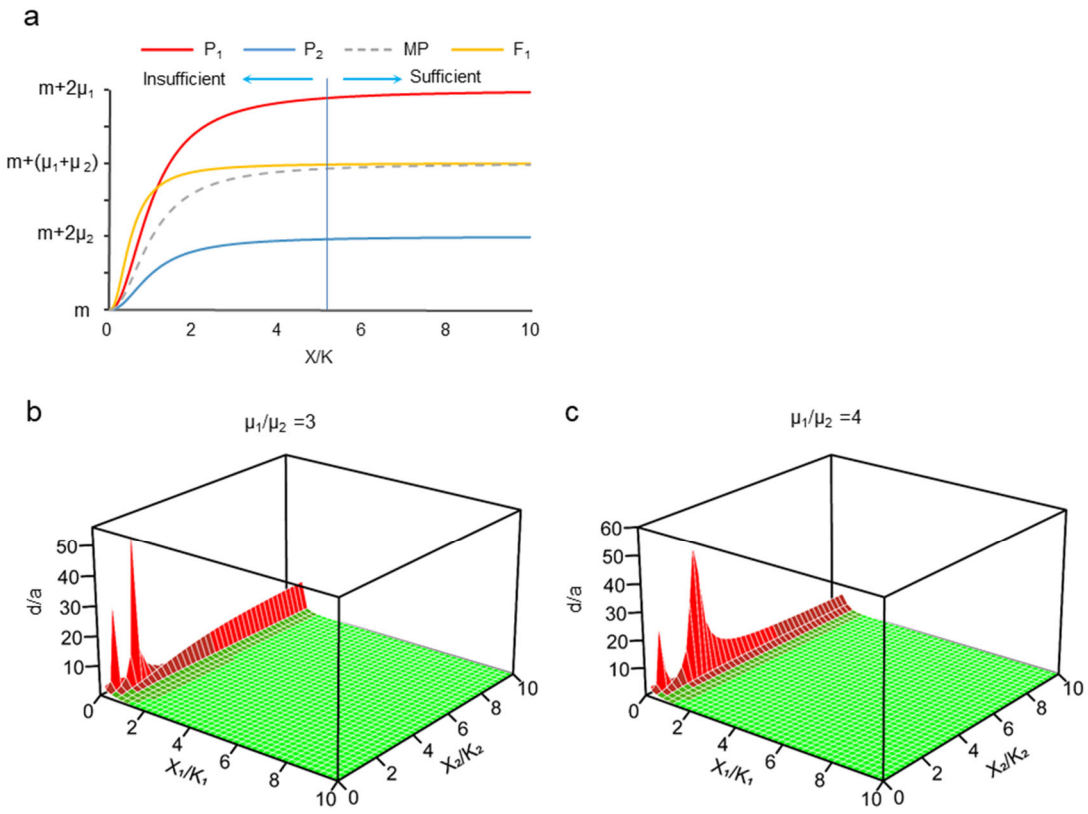
degree of the target site under the repressor background with different sufficiencies (X/K) and different μ_1.



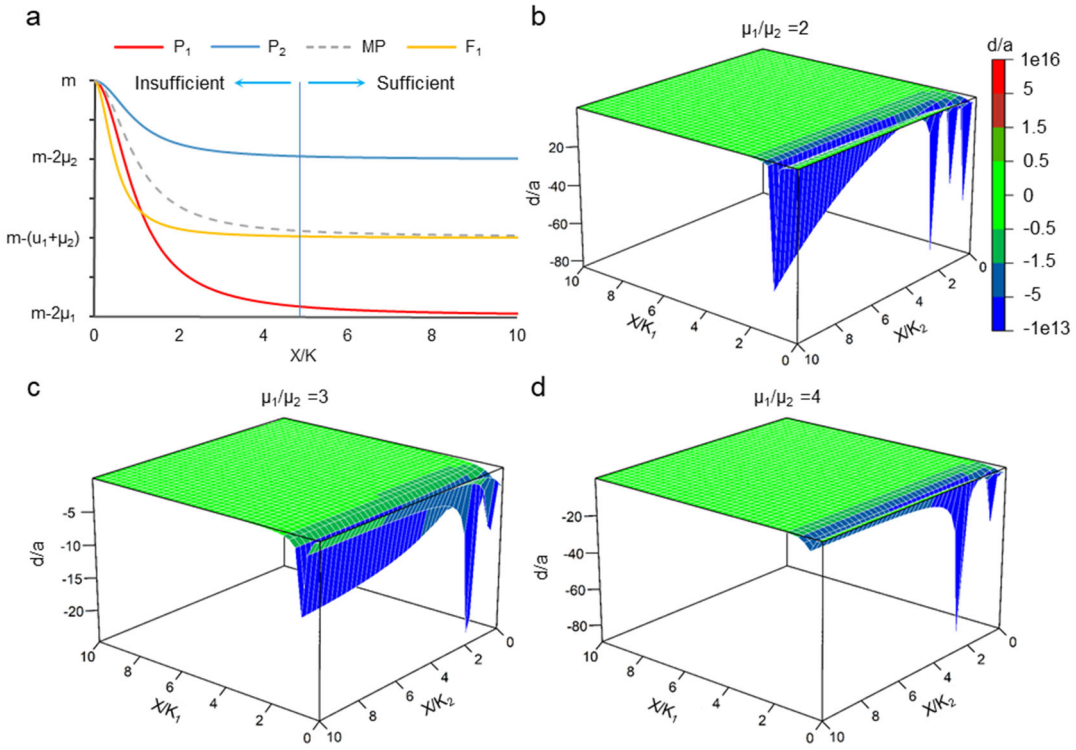
Supplementary Figure 34 The schematic diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the cumulated functions of two alleles of one polymorphic site under two independent positive regulators as the upstream backgrounds. The grey thick lines show two chromosomes, the red and green bars on which represent the homologous alleles of the regulators of T and C alleles at target site, respectively; and the function of these regulators keep constant and among P_1, P_2 and F_1. The break and solid pies together represent the required regulator function that can maximize the function of the homozygote of the corresponding target allele, and the solid pies represent different regulator functions and thus provide different backgrounds to the target allele. The arrow represents the function process. The break and solid curves together represent the maximum function of the homozygote in parents or one allele in F_1, and the solid curves represent the real function; those curves with red or green dots represent the function of allele T or C respectively. A+A shows the action mode for each allele and between two alleles (T vs C) that they are independent and cumulative with T and C both being additive; and PD+D and D+D are similar. A, D and OD mean additive, dominant, and over-dominant effect, respectively.



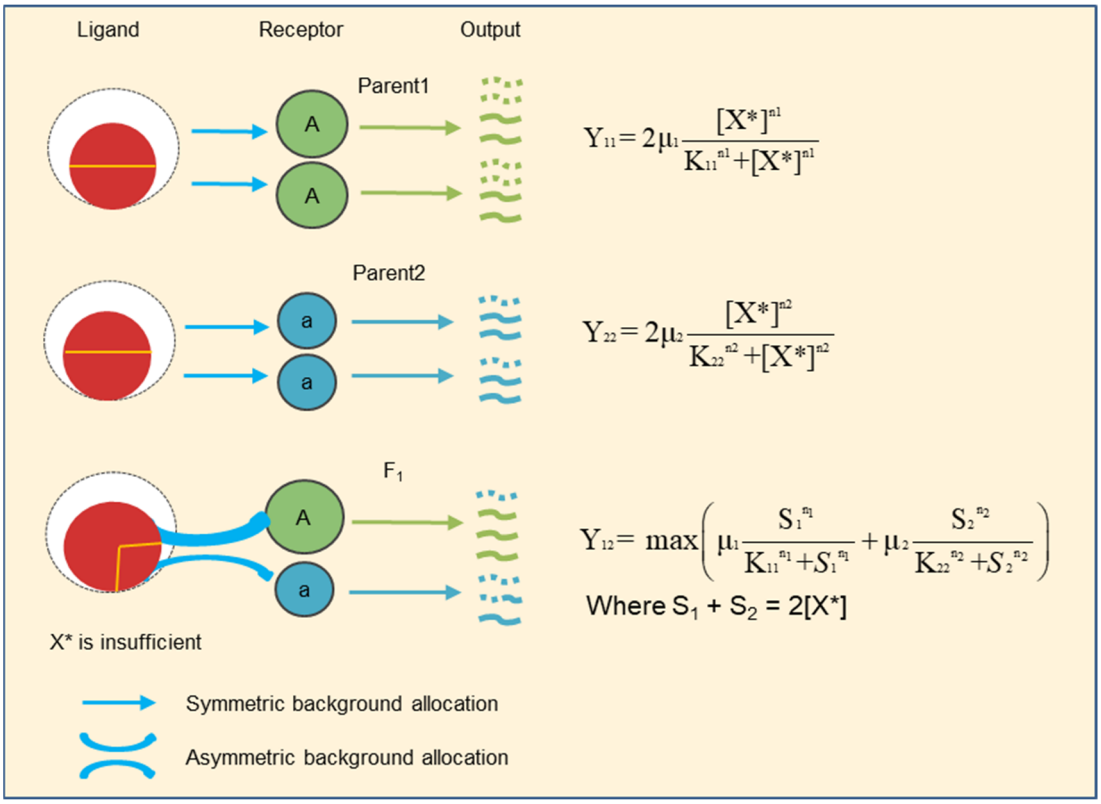
Supplementary Figure 35 The schematic diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the cumulated functions of two alleles of one polymorphic site under two independent positive regulators or responsors as the downstream backgrounds. The grey thick lines show two chromosomes, the red and green bars on which represent the homologous alleles of the regulators or responsors of T and C alleles at target site, respectively; and the function of these regulators or resposors keep constant among P_1, P_2 and F_1. Different numbers of red and green pies represent the maximum products of the homozygotes of allele T and C of target site, respectively. The doted cylinder or the same size of solid cylinder represent the required regulator or responsor function that can transform the full maximum function of the products of the homozygote of the corresponding target allele, with red corresponding to allele T and green to allele C; and the solid cylinders in doted cylinder represent different regulator or response functions and thus provide different backgrounds to the target allele. The arrow represents the function process. The break and solid curves together represent the transformed maximum function of the homozygote in parents or one allele in F_1, and the solid curves represent the real transformed function; those curves with red or green dots represent the transformed function of allele T or C respectively. A+A shows the action mode for each allele and between two alleles (T vs C) that they are independent and cumulative with T and C both being additive; and PD+D and D+D are similar. A, D and OD mean additive, dominant, and over-dominant effect, respectively.



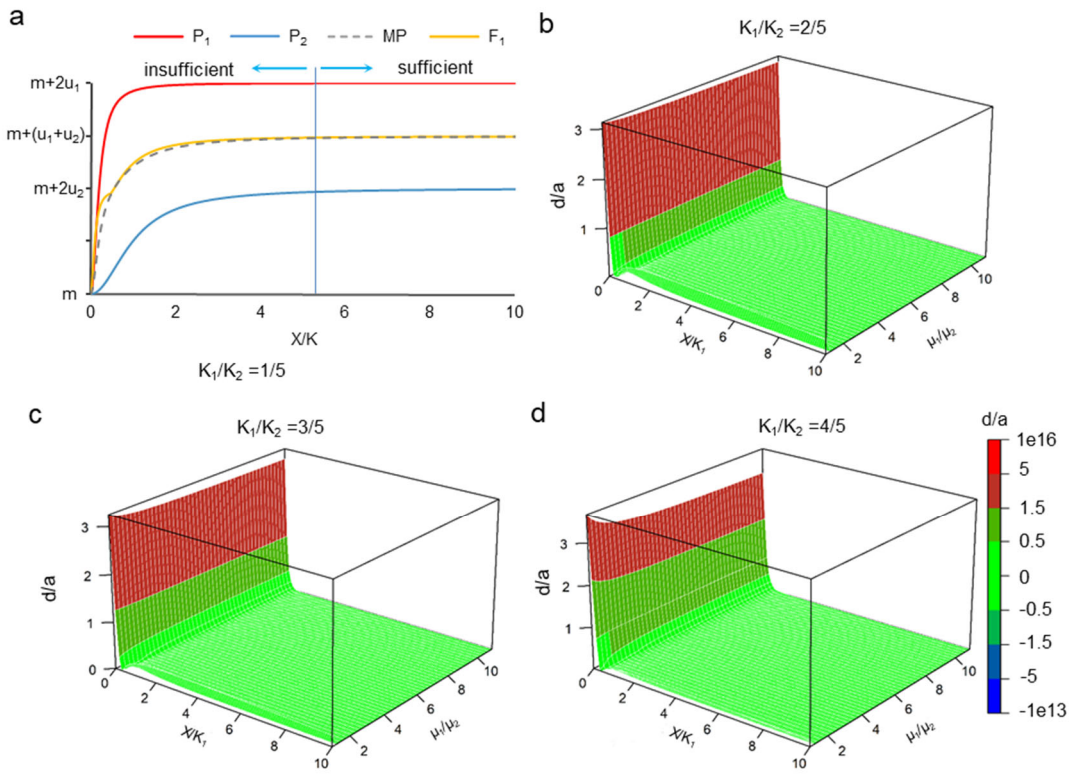
Supplementary Figure 36 The simulated distribution of dominance to additive effect ratio (d/a) with same backgrounds, but the two alleles in F_1 are regulated by different factors in the background for positive regulation. (a) The diagram of the performance of two parent, F_1 and middle parent (MP) under the condition of μ_1 =3, μ_2 =1 and K_1 = K_2 for positive regulation. left arrow means background is relative insufficient and right arrow means background is relative sufficient. (b-c) The simulated distribution of d/a with same homologous backgrounds, but the two alleles in F_1 are regulated by different factors in the background. d/a means degree of dominance effect/additive effect, it can be in both positive and negative direction. μ_1/μ_2 means the ratio of maximum function of P_1 genotype to maximum function of P_2 genotype when their respective background afford to the complete expression of the corresponding homologous genotype; d/a means the degree of dominance to additive effect ratio.



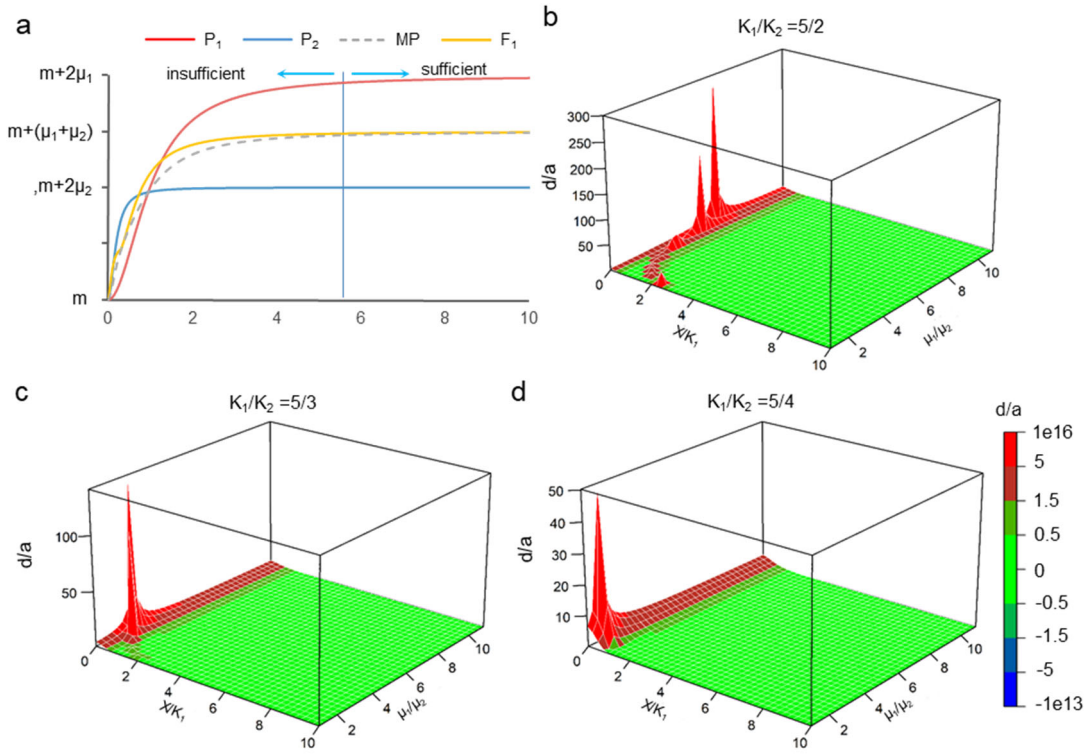
Supplementary Figure 37 The simulated diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the cumulated functions of two alleles of one polymorphic site under two independent negative regulators or responsors as the backgrounds. (a) The performance of the target site in two parents, F_1 and the middle parent (MP) under the repressor background with different sufficiencies (X/K). It was simulated according to Hill function with $\mu_1 = 3$, $\mu_2 = 1$, $K_1 = K_2$ and $n = 2$. μ_1 and μ_2 means the maximum function at steady state for allele with higher function and lower function, respectively. n is the Hill coefficient. Left arrow represents a relatively insufficient repressor background, and the right arrow represents the relatively sufficient repressor background. (b) The dominant degree of the target site under the repressor background with different sufficiencies (X/K) for two alleles of the target site with $\mu_1/\mu_2 = 2$. (c) The dominant degree of the target site under the repressor background with different sufficiencies (X/K) for two alleles of the target site with $\mu_1/\mu_2 = 3$. (d) The dominant degree of the target site under the repressor background with different sufficiencies (X/K) for two alleles of the target site with $\mu_1/\mu_2 = 4$.



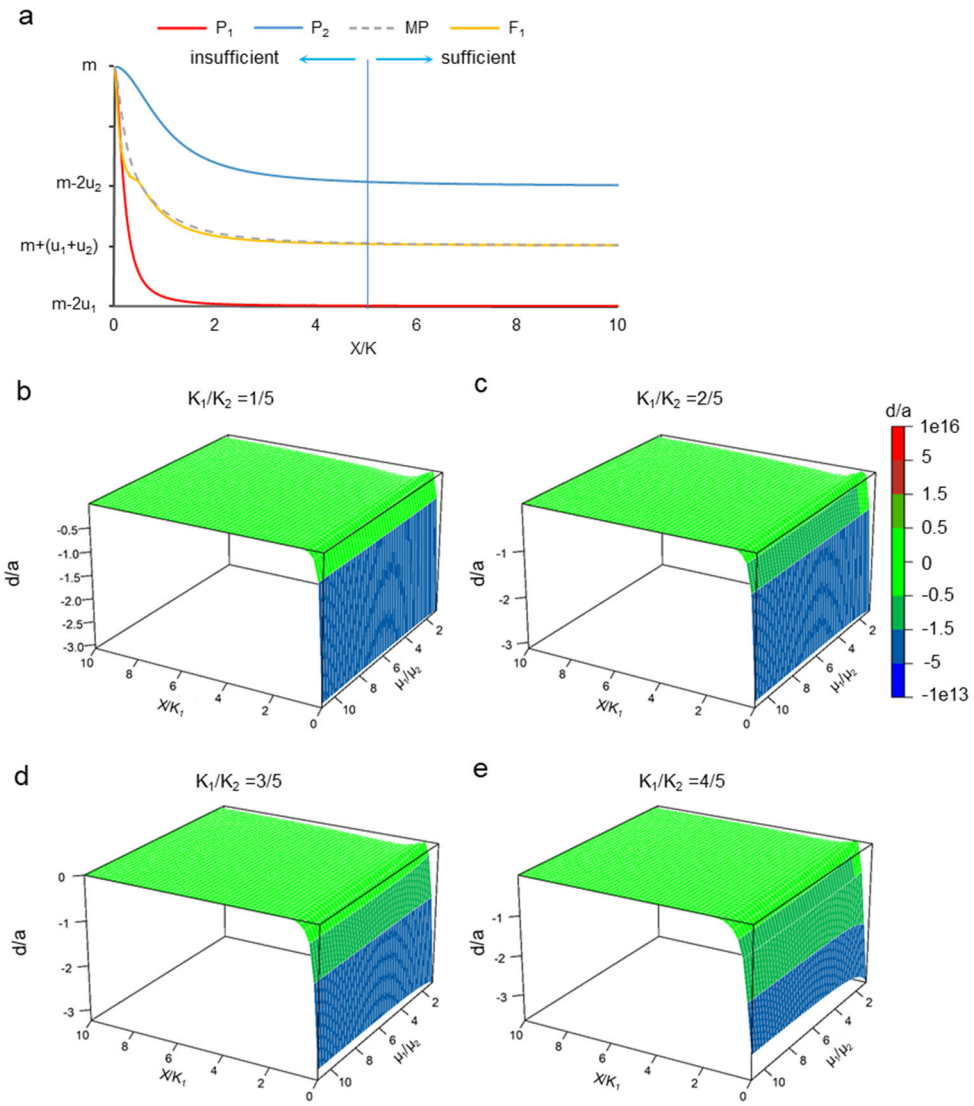
Supplementary Figure 38 The regulation or response model for molecular mechanism of dominant and over-dominant effect produced by single locus with the same background. Here, the hollow circle means the background level that satisfy the full potential of the target. The filled red circle means the background level that homozygous parent actually supplied. The orange line indicates the background allocation. When the background is insufficient and the affinity of the two genes is different, the two allele of A and a will be in a competitive use of limited background, the more competitive gene is likely to get relative more backgrounds in hybrid than that in the original homozygous parents. The parameter of μ, K and n are described in the **Supplementary Note**.



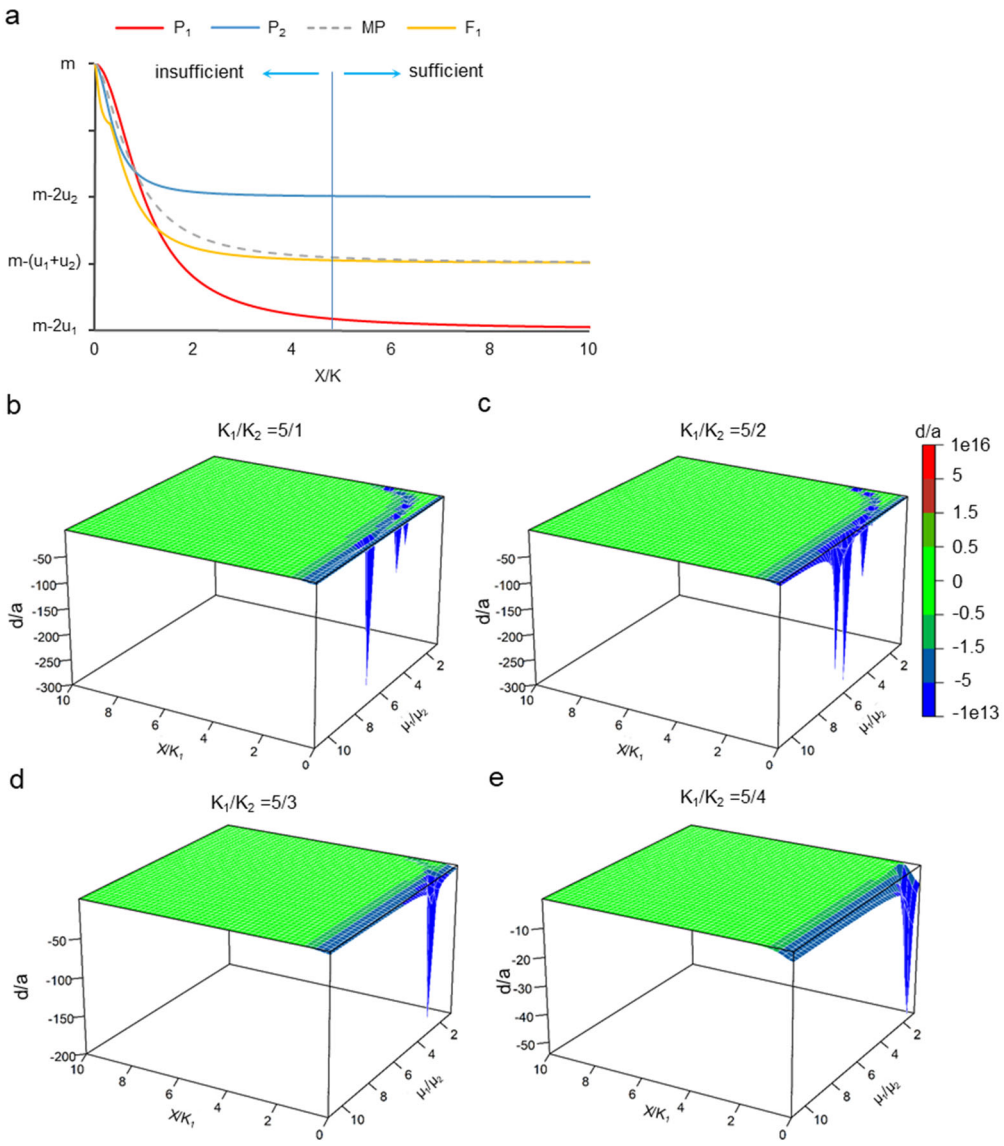
Supplementary Figure 39 The simulated diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the cumulated functions of two alleles of one polymorphic site under the same positive regulators or responsors as the background when allele 1 showing higher maximum function and higher affinity ($\mu_1 > \mu_1$ and $K_1 < K_2$). (a) The performance of the target site in two parents, F_1 and the middle parent (MP) under the activator background with different sufficiencies (X/K). It was simulated according to Hill function with $\mu_1 = 3$, $\mu_2 = 1$, $K_1 = 1$, $K_2 = 5$ and $n = 2$. μ_1 and μ_2 means the maximum function at steady state for allele with higher function and lower function, respectively. n is the Hill coefficient. Left arrow represents a relatively insufficient activator background, and the right arrow represents the relatively sufficient activator background. (b) The dominant degree of the target site under the activator background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 2/5$. (c) The dominant degree of the target site under the activator background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 3/5$. (d) The dominant degree of the target site under the activator background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 4/5$.



Supplementary Figure 40 The simulated diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the cumulated functions of two alleles of one polymorphic site under the same positive regulators or responsors as the background when allele 1 showing higher maximum function but lower affinity ($\mu_1 > \mu_1$ and $K_1 > K_2$). (a) The performance of the target site in two parents, F_1 and the middle parent (MP) under the activator background with different sufficiencies (X/K). It was simulated according to Hill function with $\mu_1 = 3$, $\mu_2 = 1$, $K_1 = 5$, $K_2 = 1$ and $n = 2$. μ_1 and μ_2 means the maximum function at steady state for allele with higher function and lower function, respectively. n is the Hill coefficient. Left arrow represents a relatively insufficient activator background, and the right arrow represents the relatively sufficient activator background. (b) The dominant degree of the target site under the activator background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 5/2$. (c) The dominant degree of the target site under the activator background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 5/3$. (d) The dominant degree of the target site under the activator background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 5/4$.

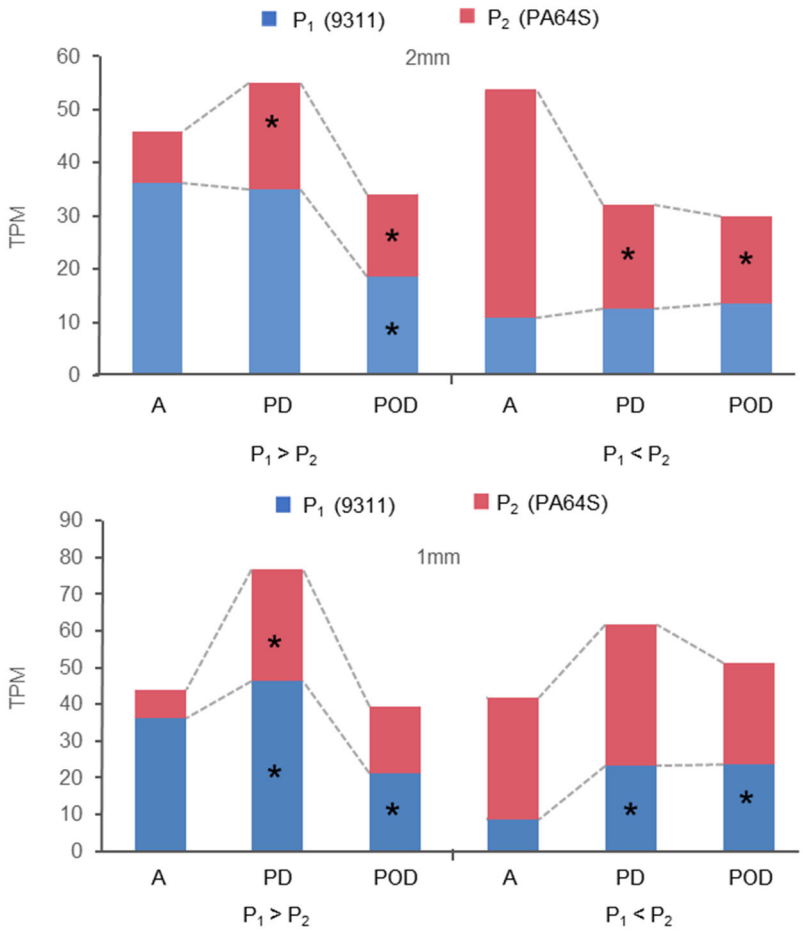


Supplementary Figure 41 The simulated diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the cumulated functions of two alleles of one polymorphic site under the same negative regulators or responsors as the background when allele 1 showing higher maximum function and higher affinity ($\mu_1 > \mu_1$ and $K_1 < K_2$). (a) The performance of the target site in two parents, F_1 and the middle parent (MP) under the repressor background with different sufficiencies (X/K). It was simulated according to Hill function with $\mu_1 = 2$, $\mu_2 = 1$, $K_1 = 1$, $K_2 = 5$ and $n = 2$. μ_1 and μ_2 means the maximum function at steady state for allele with higher function and lower function, respectively. n is the Hill coefficient. Left arrow represents a relatively insufficient repressor background, and the right arrow represents the relatively sufficient repressor background. (b) The dominant degree of the target site under the repressor background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 1/5$. (c) The dominant degree of the target site under the repressor background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 2/5$. (d) The dominant degree of the target site under the repressor background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 3/5$. (e) The dominant degree of the target site under the repressor background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 4/5$.

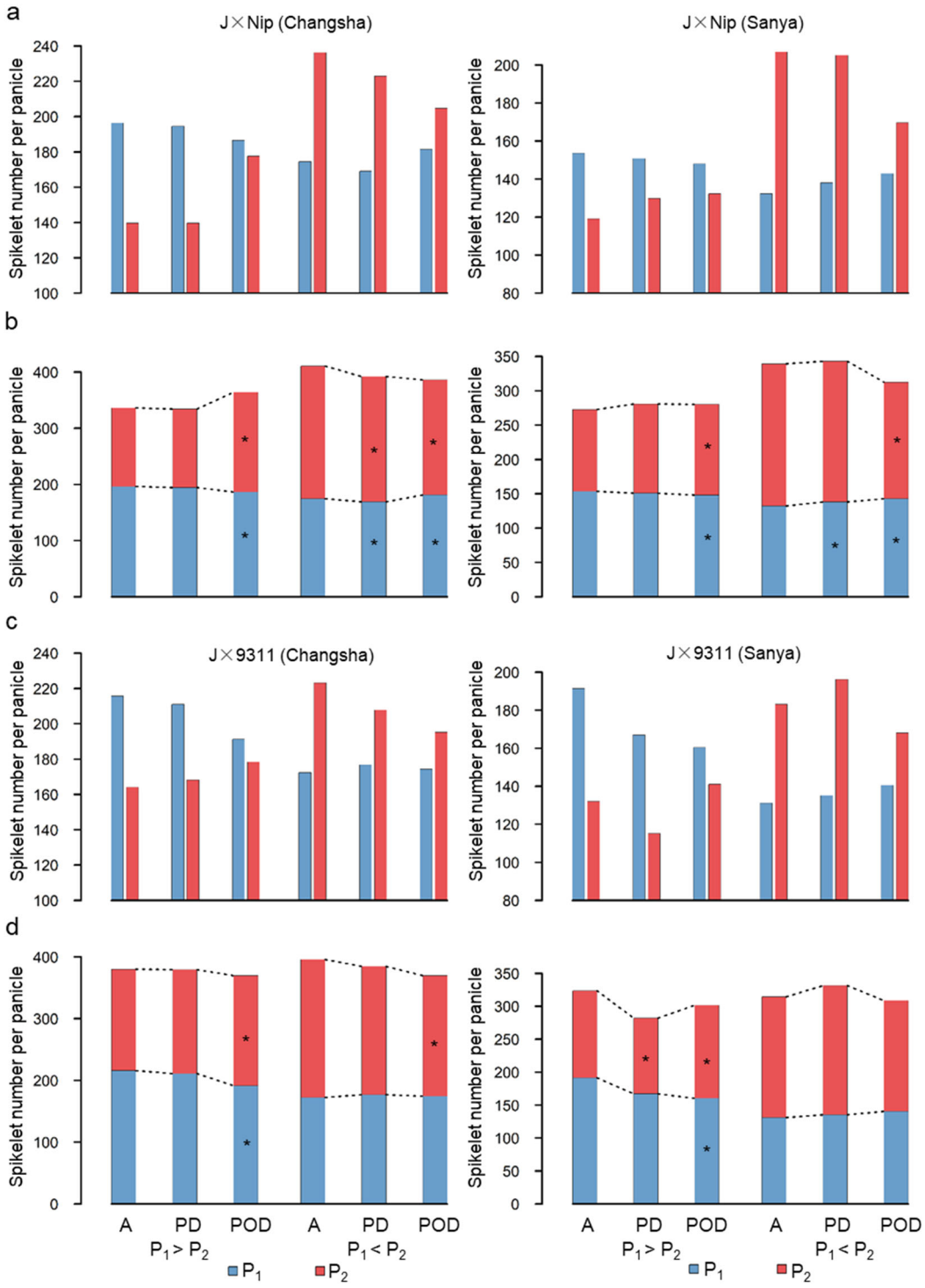


Supplementary Figure 42 The simulated diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the cumulated functions of two alleles of one polymorphic site under the same negative regulators or responsors as the background when allele 1 showing higher maximum function but lower affinity ($\mu_1 > \mu_1$ and $K_1 > K_2$). (a) The performance of the target site in two parents, F_1 and the middle parent (MP) under the repressor background with different sufficiencies (X/K). It was simulated according to Hill function with $\mu_1 = 2$, $\mu_2 = 1$, $K_1 = 5$, $K_2 = 1$ and $n = 2$. μ_1 and μ_2 means the maximum function at steady state for allele with higher function and lower function, respectively. n is the Hill coefficient. Left arrow represents a relatively insufficient repressor background, and the right arrow represents the relatively sufficient repressor background. (b) The dominant degree of the target site under the repressor background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 5/1$. (c) The dominant degree of the target site under the repressor background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 5/2$. (d) The dominant degree of the target site under the repressor background with different sufficiencies for allele with higher function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 5/3$. (e) The dominant degree of the target site under the repressor background with different sufficiencies for allele with higher

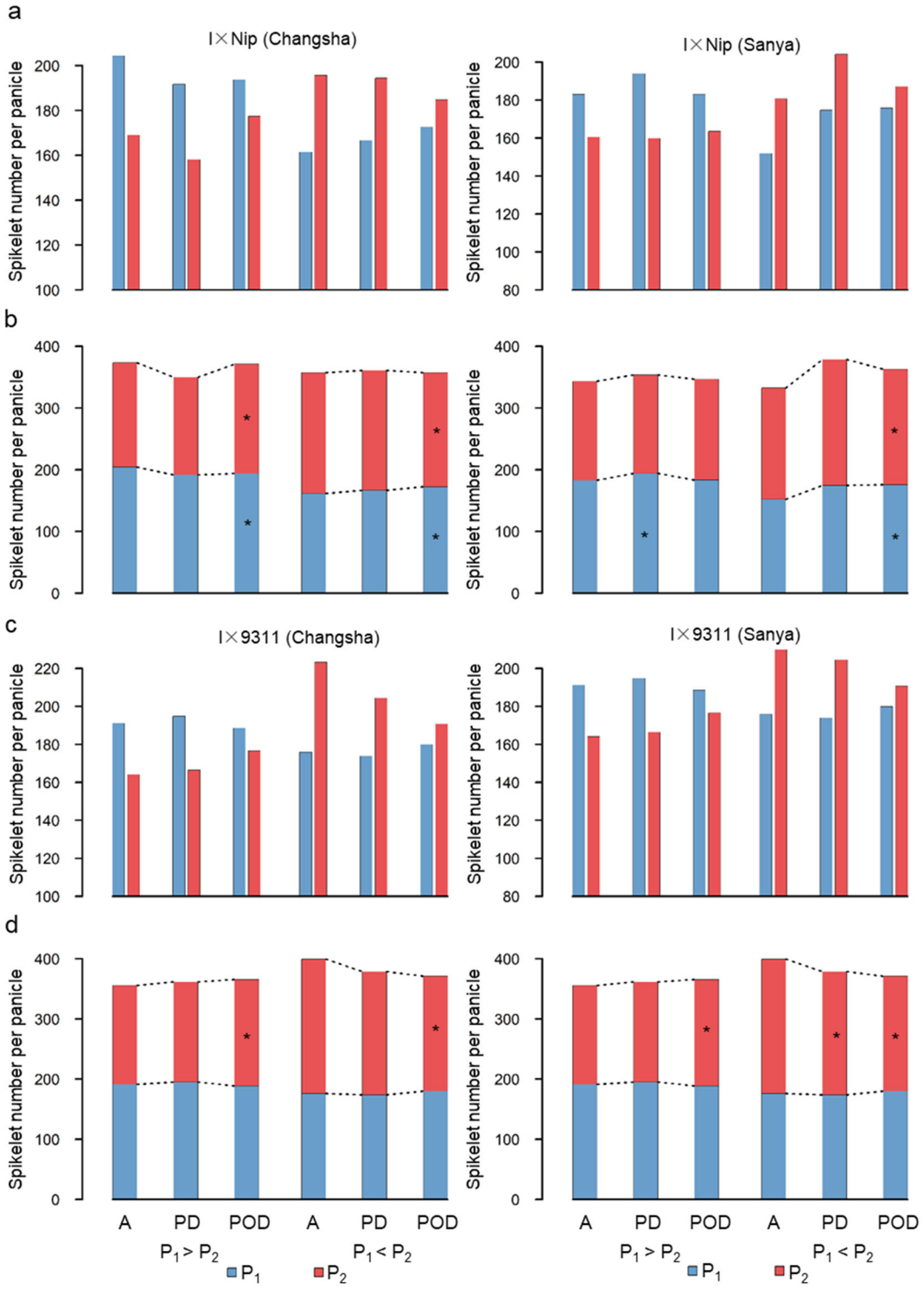
function (X/K_1) of the target site and with different μ_1/μ_2 when $K_1/K_2 = 5/4$.



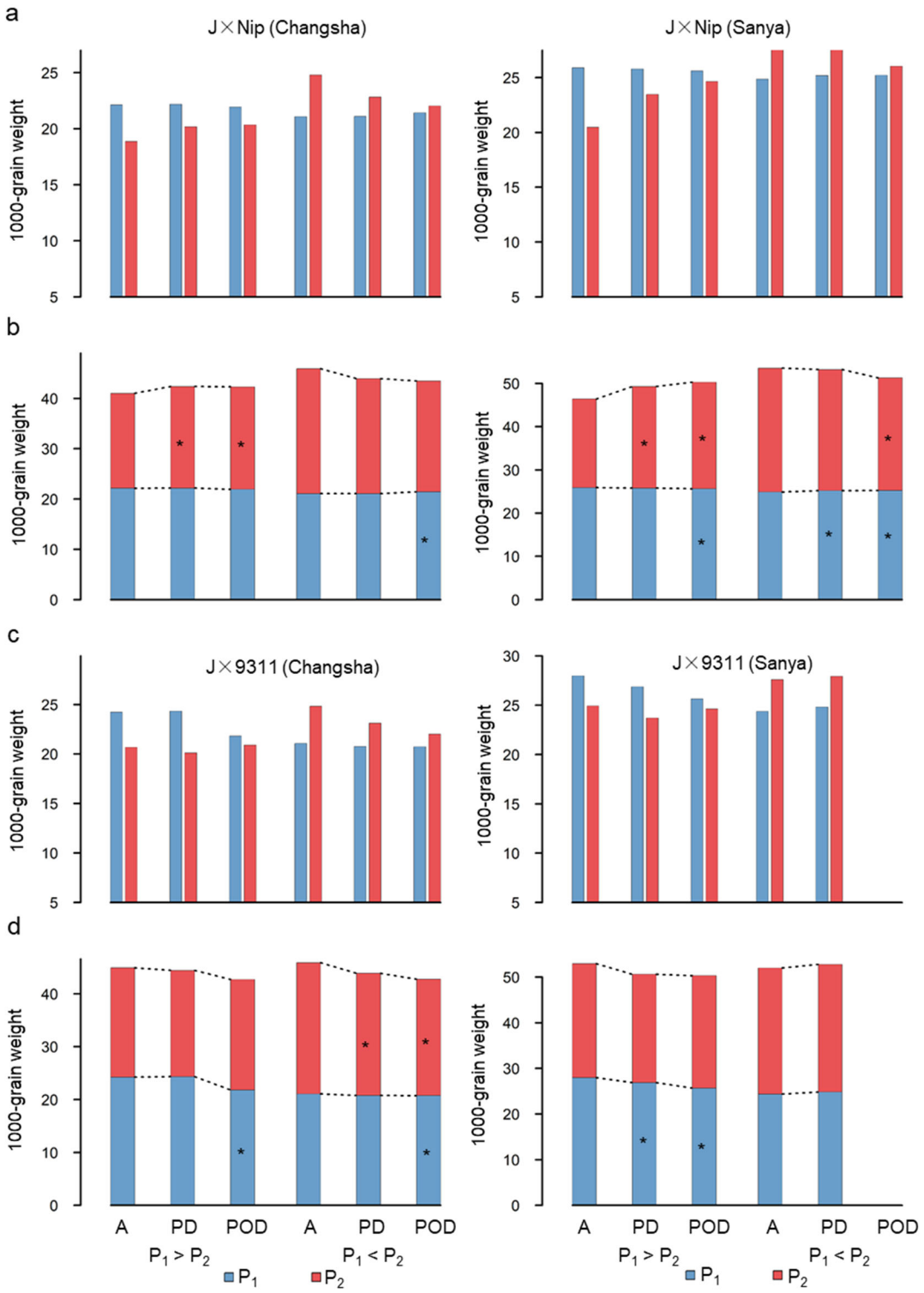
Supplementary Figure 43 The expression level of genes with different expression patterns in 1 mm and 2 mm young panicles of two parents. Here, A, PD and POD mean the expression patterns appearing additive, positive dominant and over-dominant, respectively; the star means significant difference from the genes with additive expression pattern.



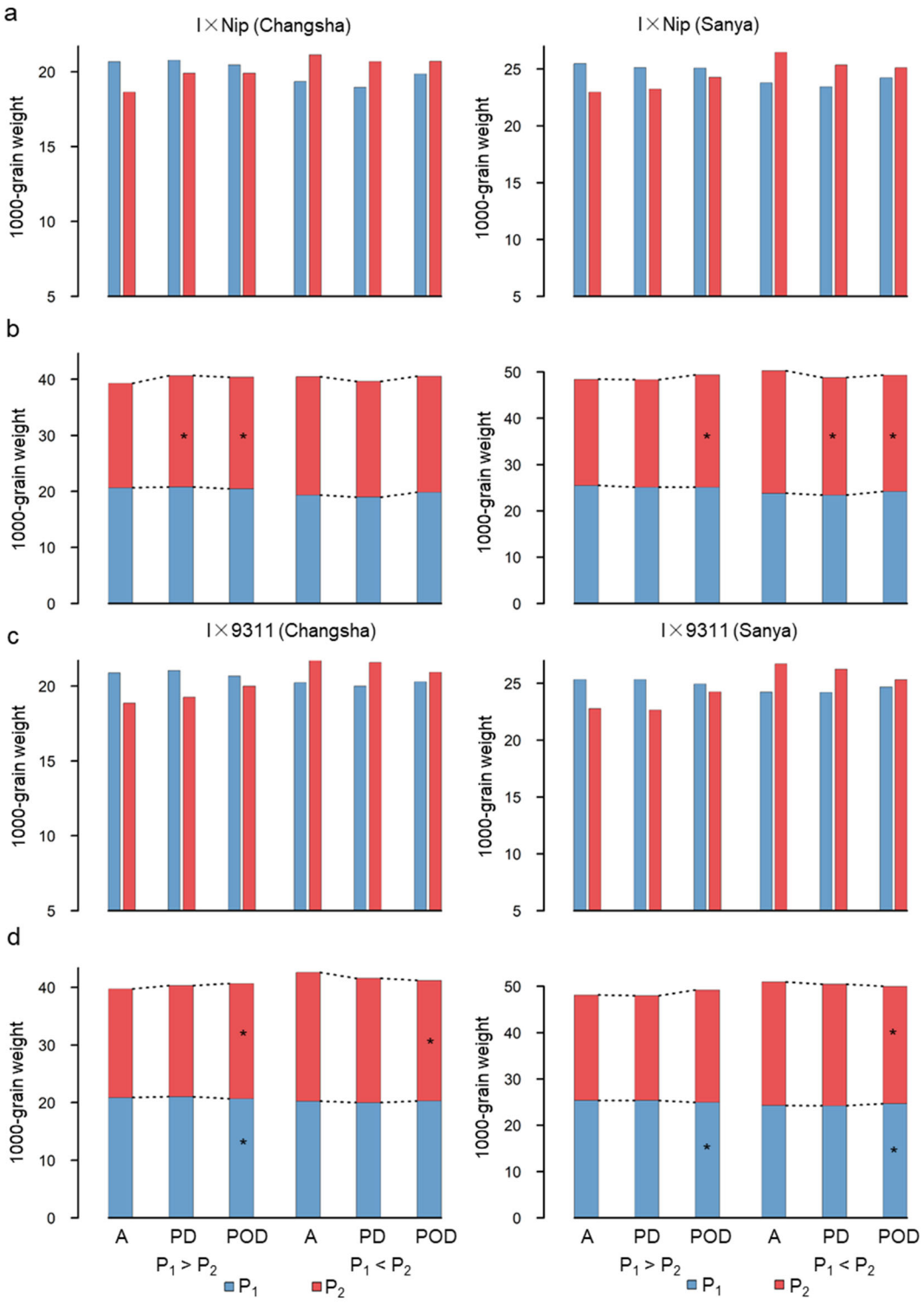
Supplementary Figure 44 The spikelet number per plant (SPP) of parents with non-tester genotype (P_1) and parents with tester genotype (P_2) of the SPP QTLs showing different genetic effect types. Here, A, PD and POD mean the additive, positive dominant and positive over-dominant QTLs, respectively. $P_1 > P_2$ means that P_1 contains the genotype with higher effect of the QTL, and on the contrary $P_1 < P_2$ means that P_2 contains the genotype with higher effect of the QTL. The star means significant difference from the addtive QTLs.



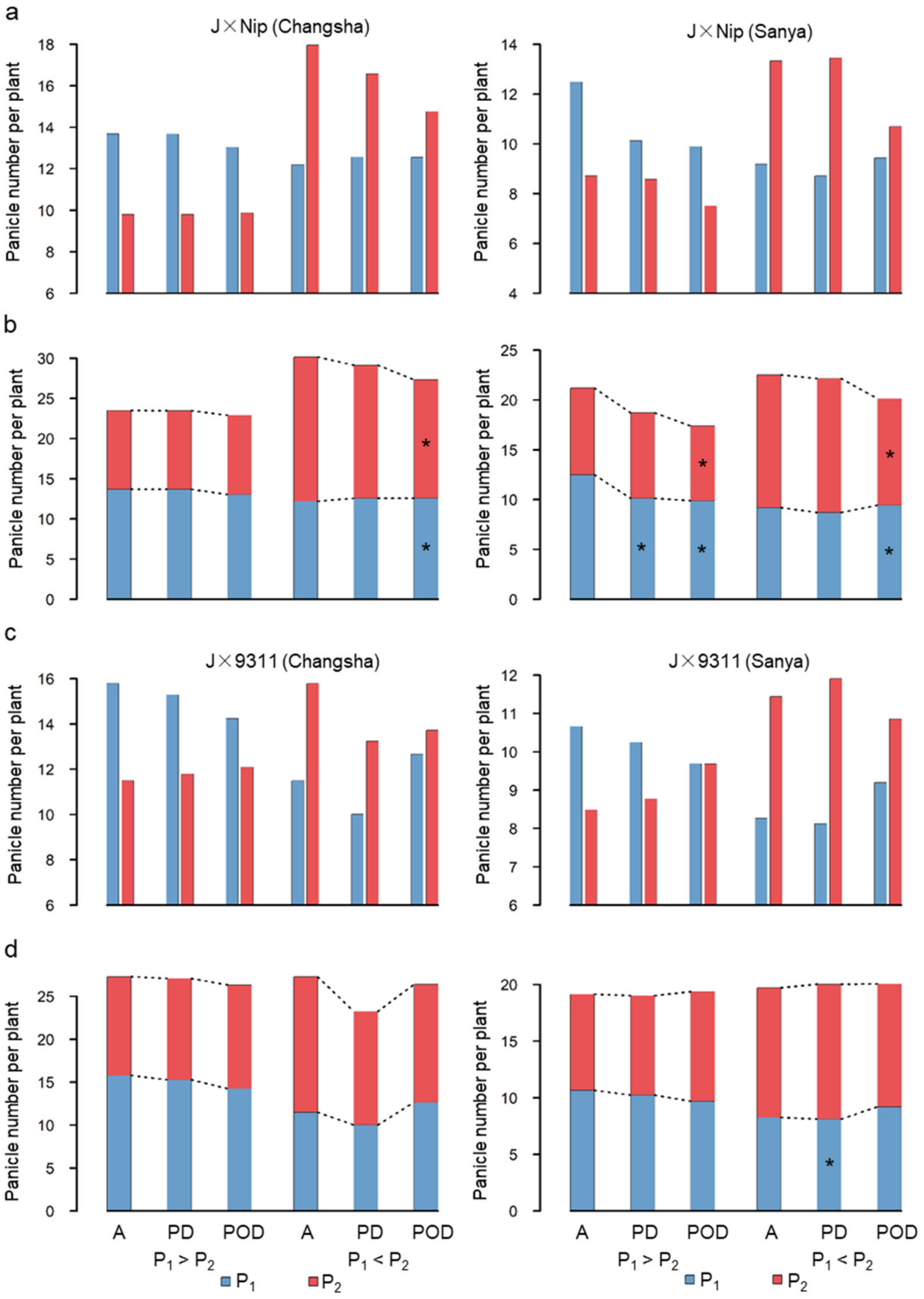
Supplementary Figure 44 (continued)



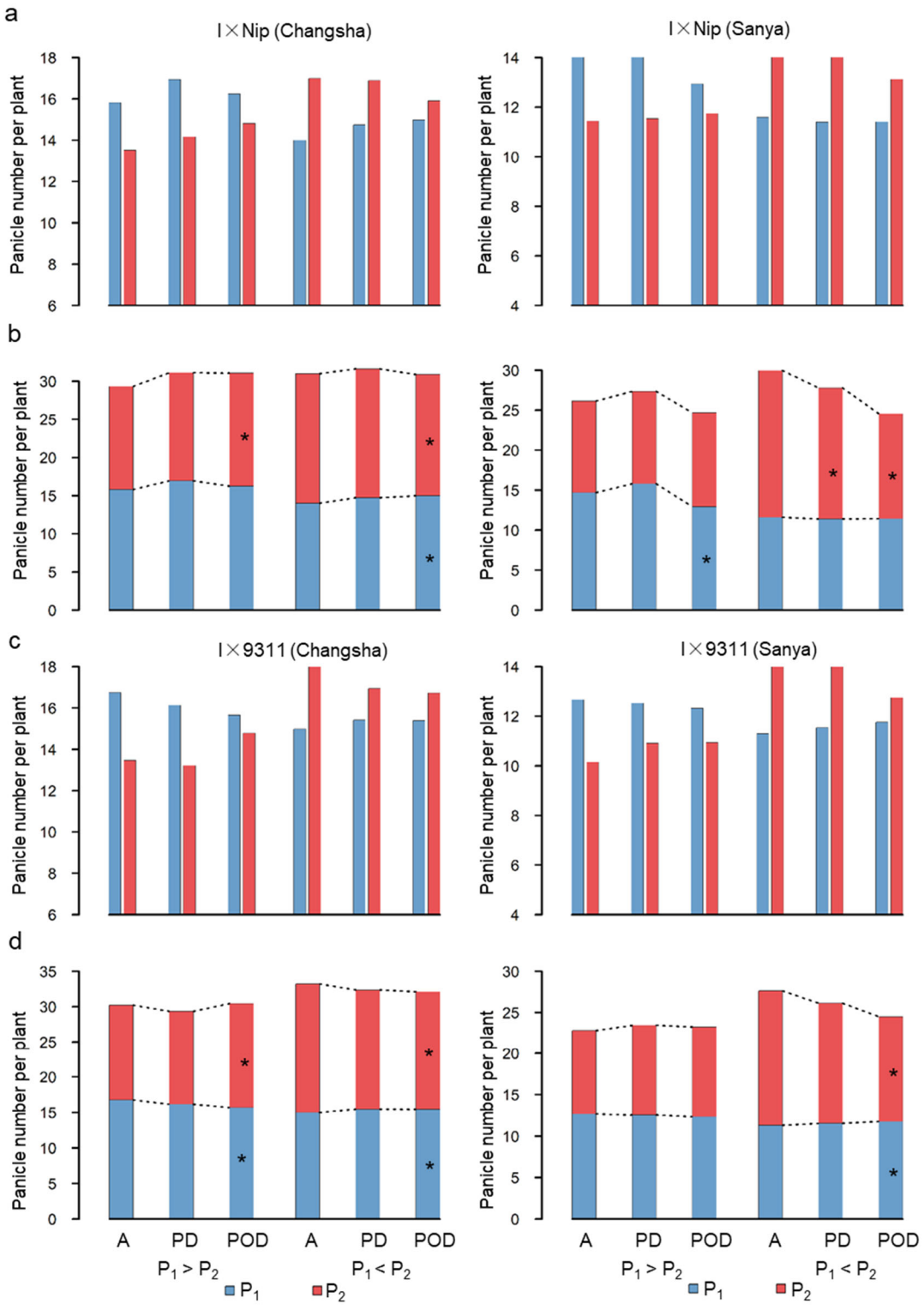
Supplementary Figure 45 The 1000-grain weight (KGW) of parents with non-tester genotype (P_1) and parents with tester genotype (P_2) of the KGW QTLs showing different genetic effect types. Here, A, PD and POD mean the additive, positive dominant and positive over-dominant QTLs, respectively. $P_1 > P_2$ means that P_1 contains the genotype with higher effect of the QTL, and on the contrary $P_1 < P_2$ means that P_2 contains the genotype with higher effect of the QTL. The star means significant difference from the addtive QTLs.



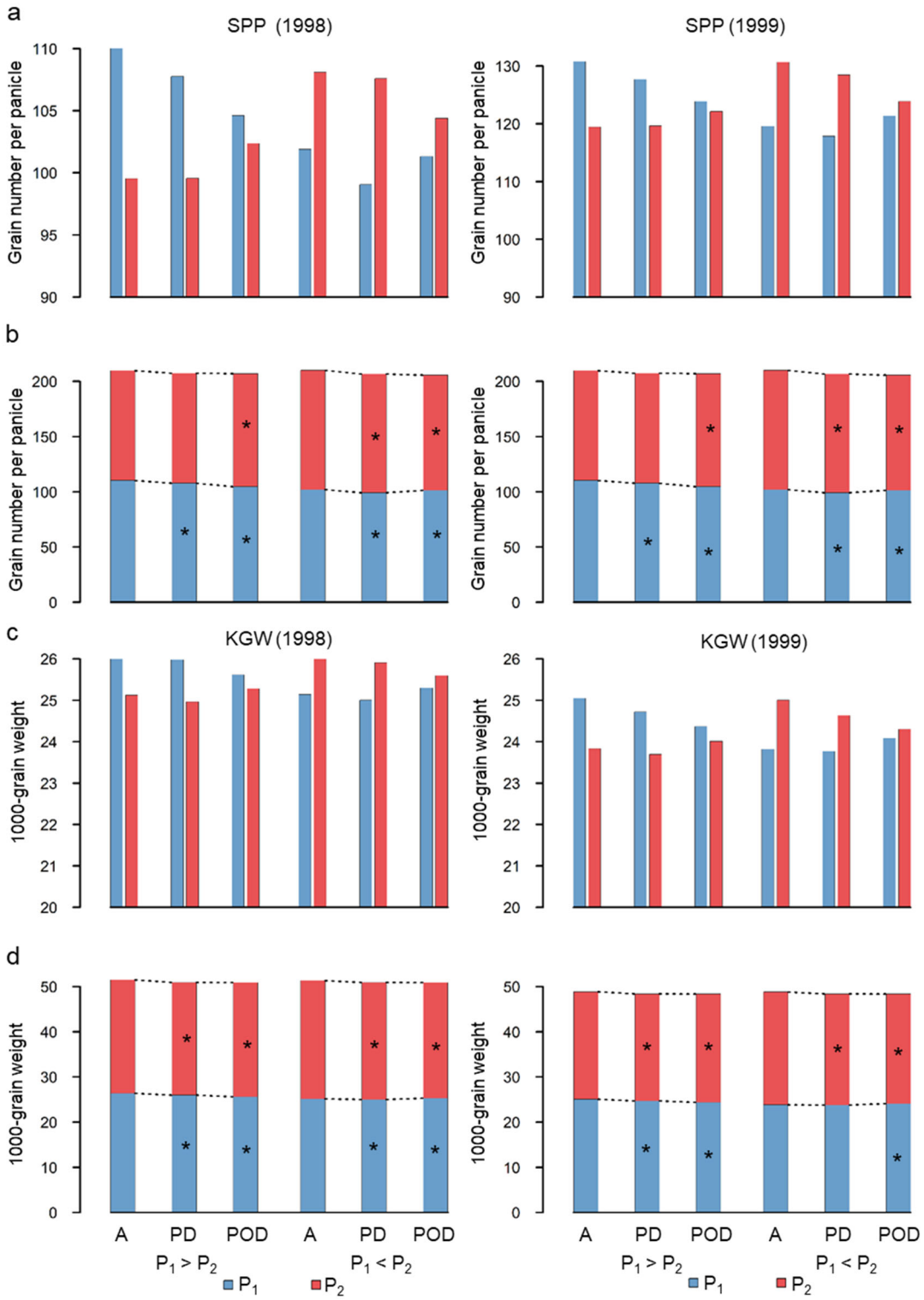
Supplementary Figure 45 (continued)



Supplementary Figure 46 The panicle number per plant (PNP) of parents with non-tester genotype (P_1) and parents with tester genotype (P_2) of the PNP QTLs showing different genetic effect types. Here, A, PD and POD mean the additive, positive dominant and positive over-dominant QTLs, respectively. $P_1 > P_2$ means that P_1 contains the genotype with higher effect of the QTL, and on the contrary $P_1 < P_2$ means that P_2 contains the genotype with higher effect of the QTL. The star means significant difference from the addtive QTLs.

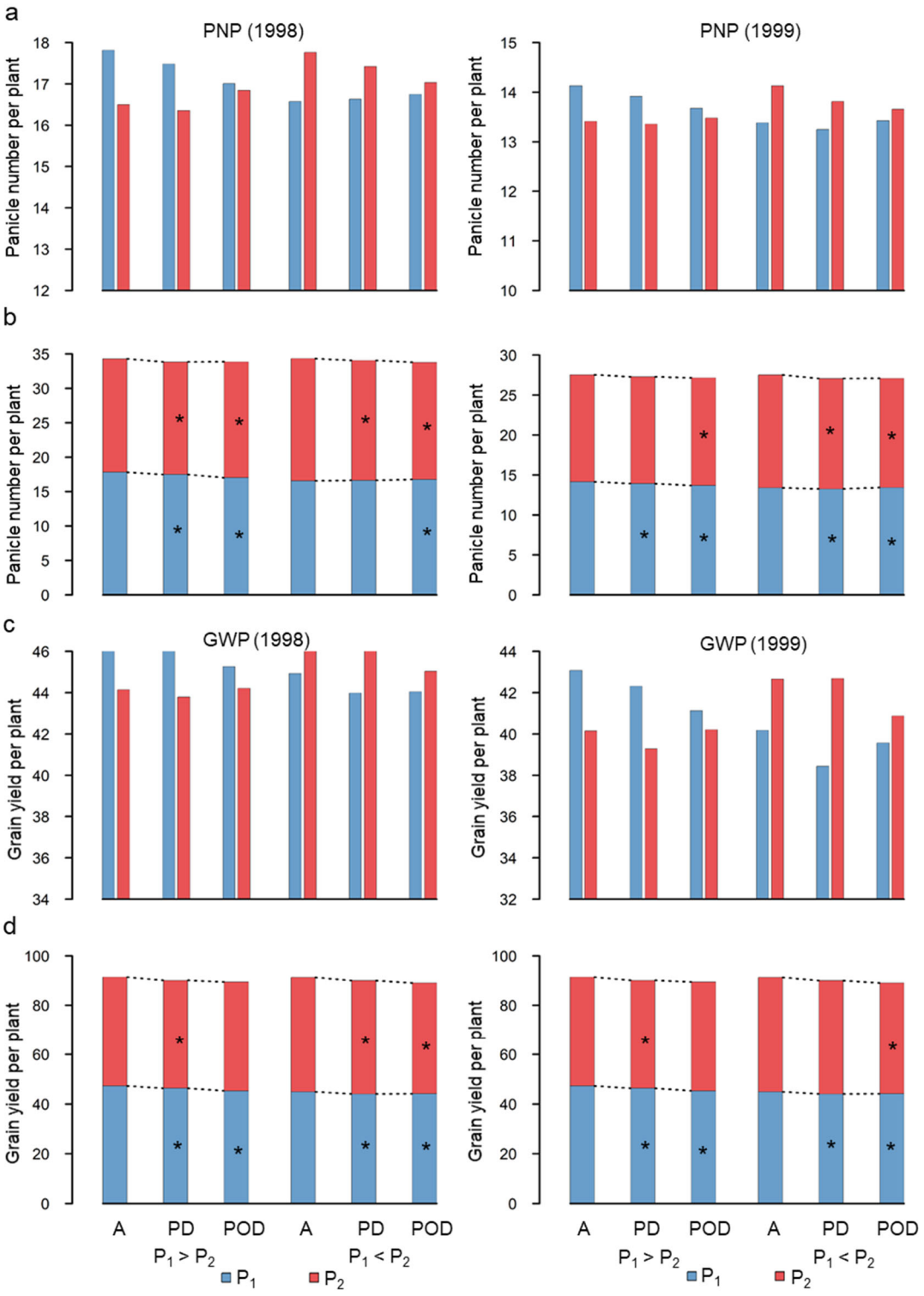


Supplementary Figure 46 (continued)



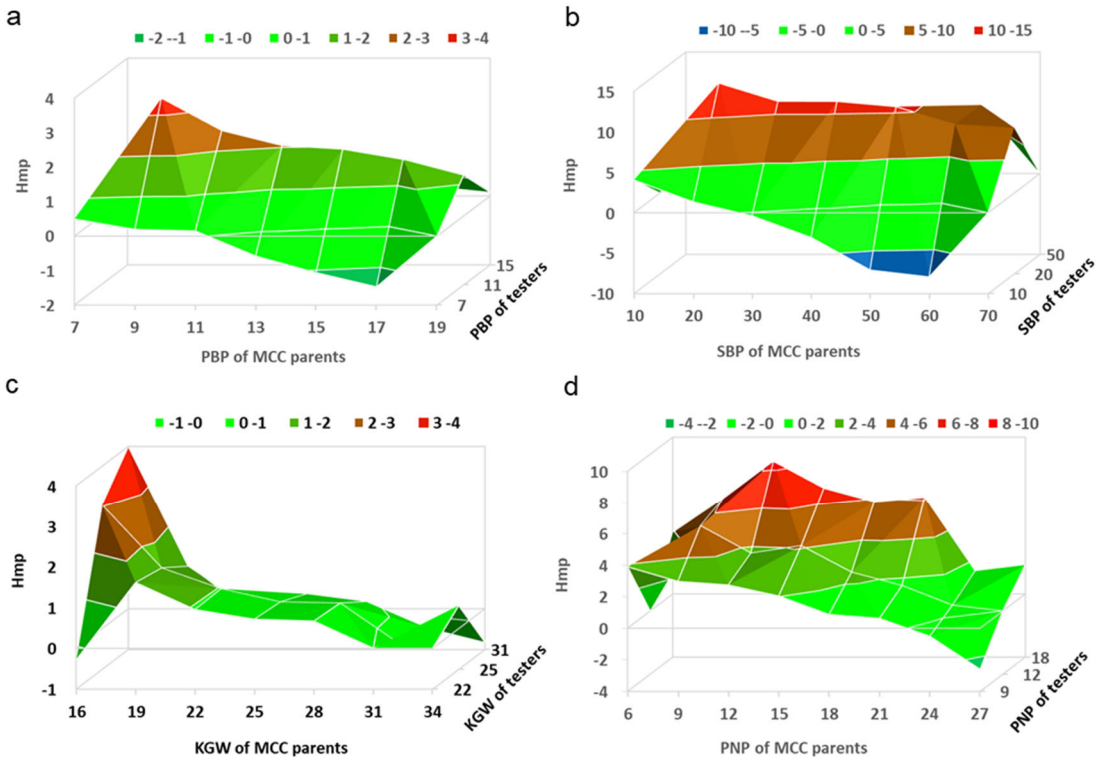
Supplementary Figure 47 The yield traits of lines with Zhenshan97 (P_1) genotype and Minghui63 (P_2) genotype of those QTLs with different genetic effect types in IMF_2 population in 1998 and 1999. Here, A, PD and POD mean the additive, positive dominant and positive over-dominant QTLs, respectively. $P_1 > P_2$ means that P_1 contains the genotype with higher effect of the QTL, and on the contrary $P_1 < P_2$ means that P_2 contains the genotype with higher effect of the QTL. The star means significant difference from the addtive QTLs. The QTLs were identified according to the published data (Zhou, G. et al. Genetic composition of yield heterosis in an elite rice hybrid. Proceedings of the National

719 Academy of Sciences of the United States of America 109, 15847-15852 (2012)).

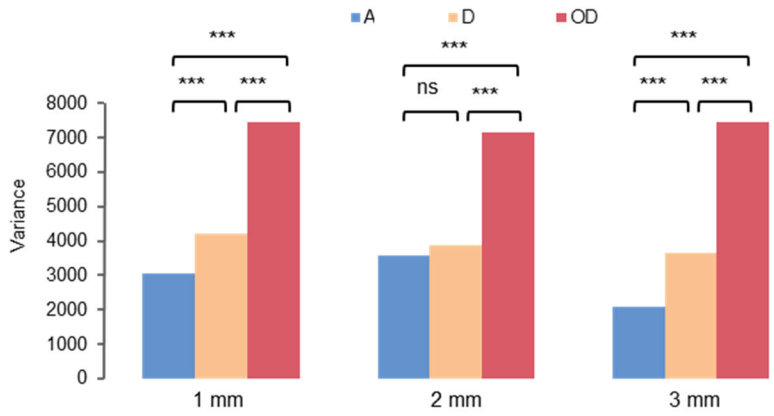


720

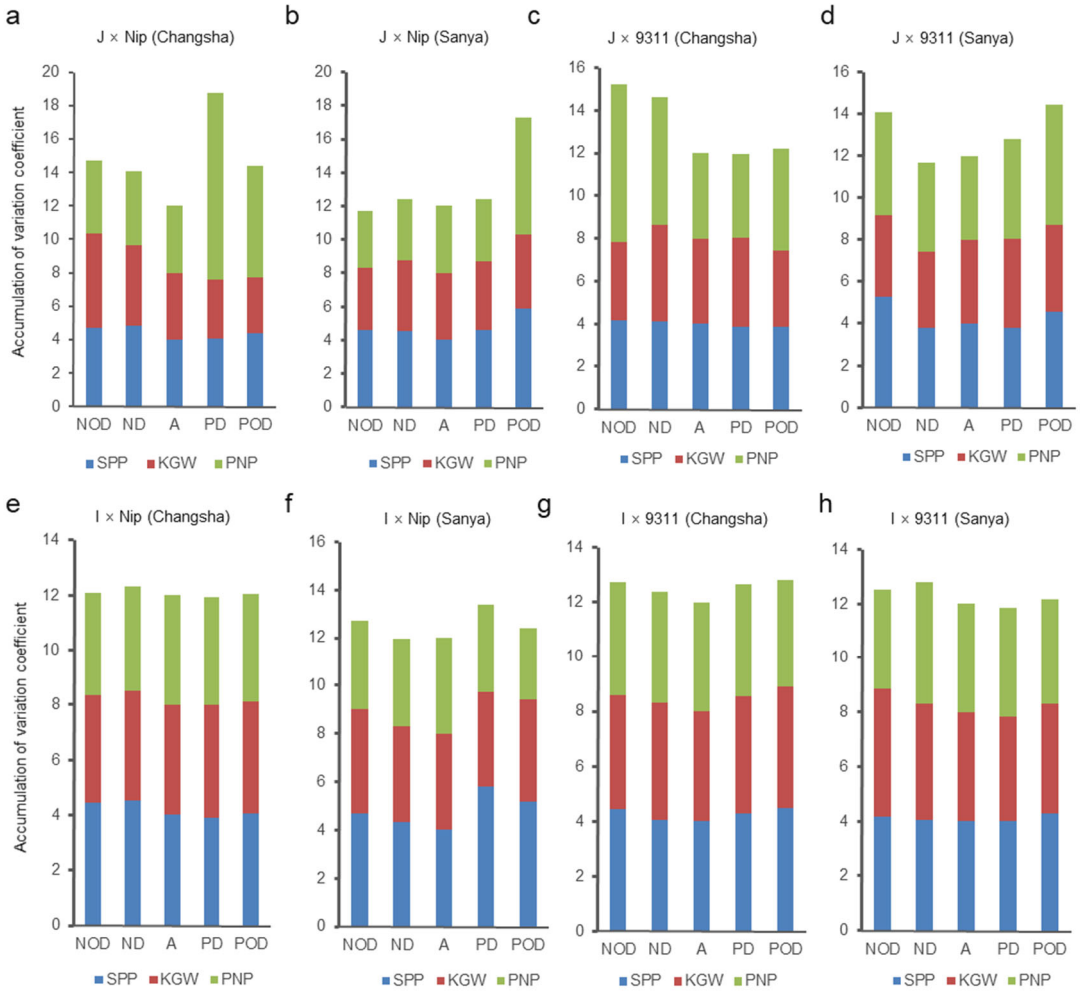
721 **Supplementary figure 47 (continued)**



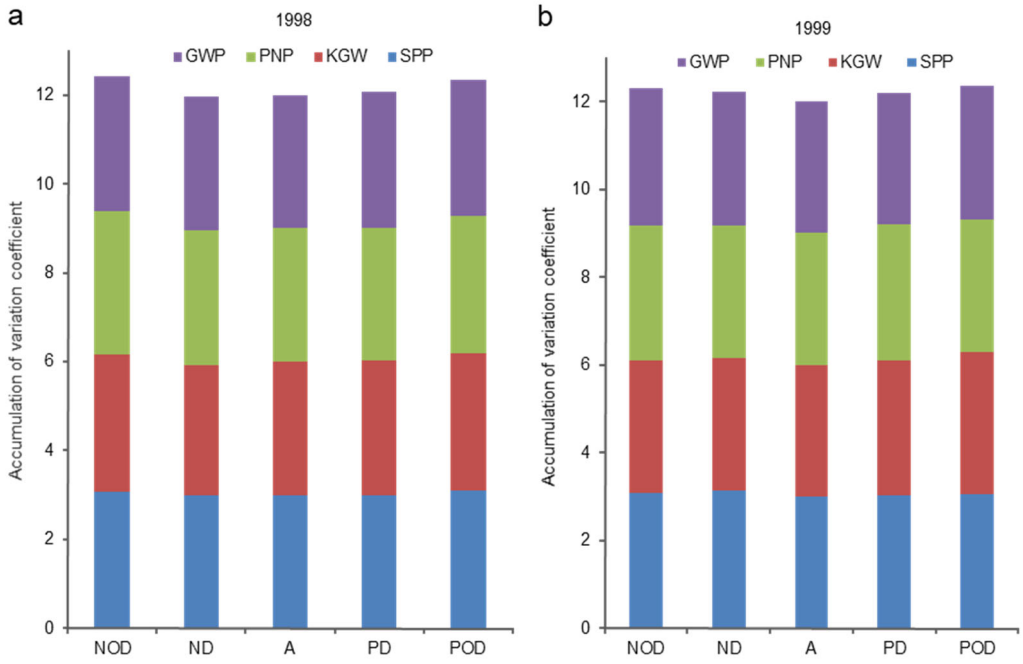
Supplementary Figure 48 The relationship between the middle-parent heterosis (Hmp) and the phenotypes of their parents for primary branch number per panicle (a), secondary branch number per panicle (b), 1000-grain weight (c) and panicle number per plant (d). The results were calculated according to the phenotype of 418 combinations of MCC in Changsha and Sanya.



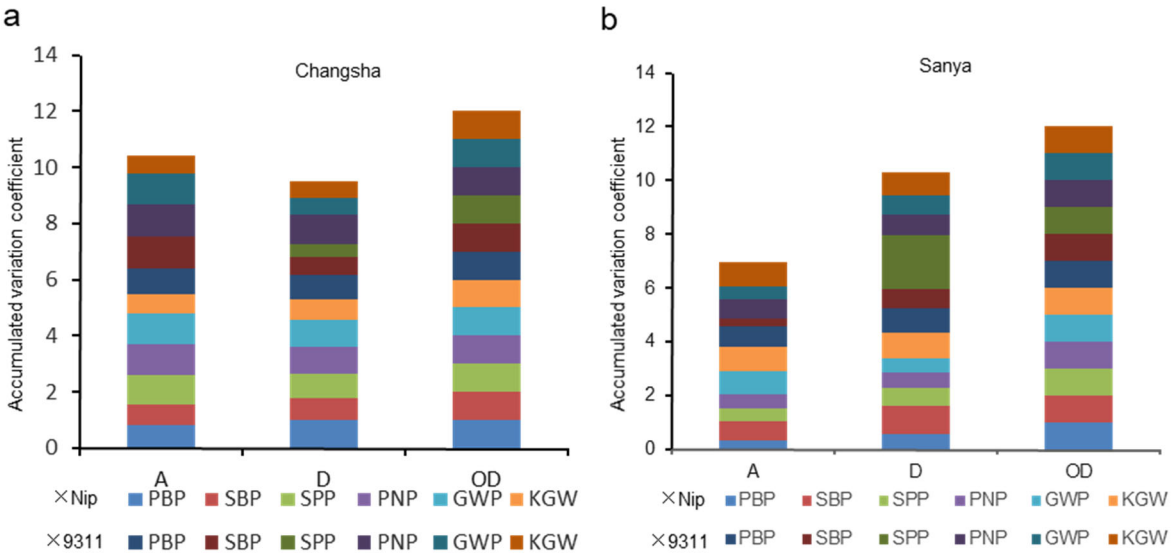
Supplementary Figure 49 The variance of expression levels among different tissues (1 mm, 2 mm and 3mm young panicles) for genes with additive, dominant and over-dominant expression patterns in 1mm, 2mm or 3mm young panicles of hybrids. Triple-star means significant difference with $p < 0.001$; ns means no significance. The variance was estimated from 48 dataset collected from ricexpro (http://ricexpro.dna.affrc.go.jp/)



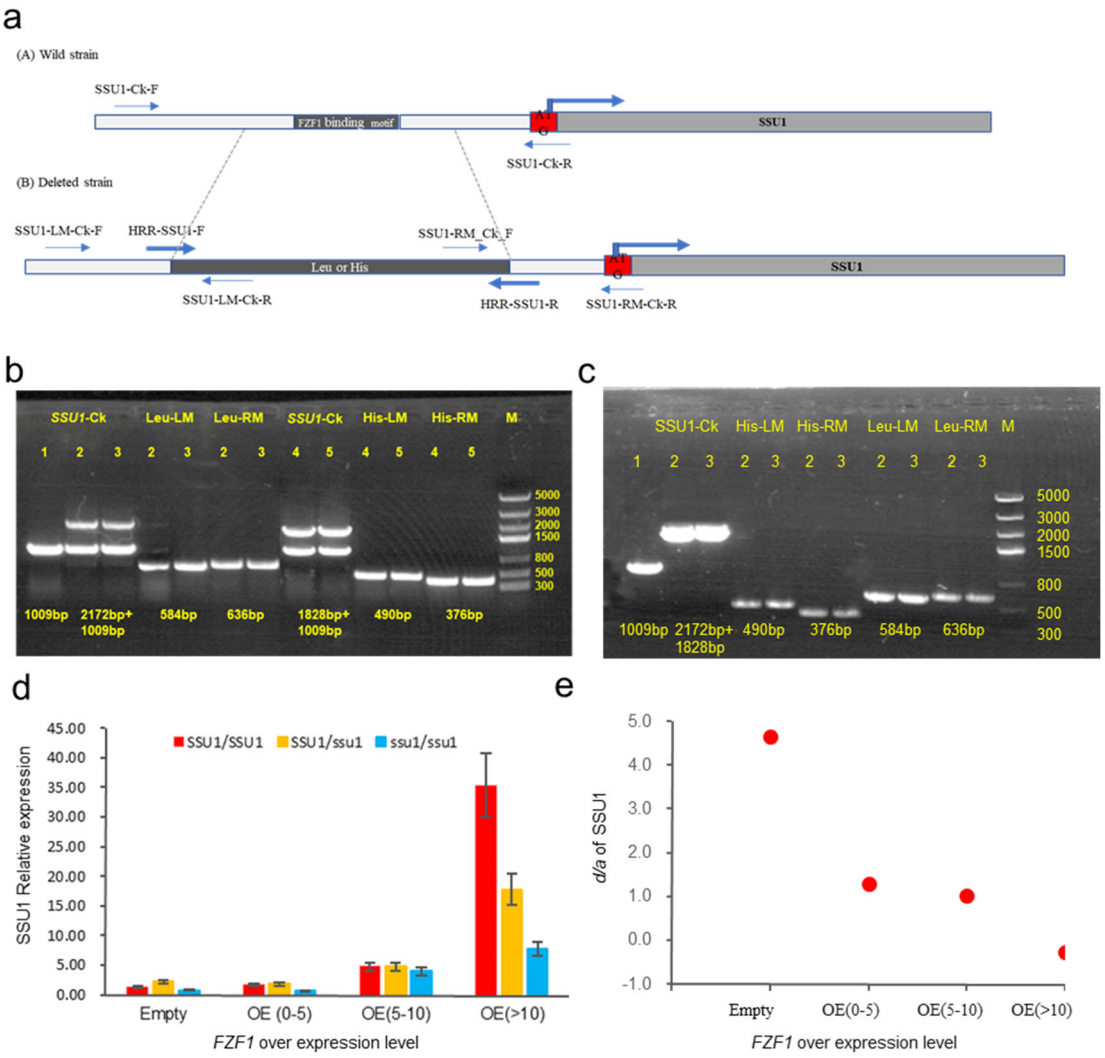
Supplementary Figure 50 The accumulation of average variation coefficient estimated in each identified QTL across four kinds of combination for different type of genetic component. Here NOD, ND, A, PD and POD represent the type of negative over-dominant, negative dominant, additive, positive dominant and positive over-dominant QTLs.



Supplementary Figure 51 The accumulation of average variation coefficient estimated in each identified QTL for different type genetic component in IMF$_2$ population. (a) The comparisons of additive, dominance and overdominance for the accumulation of average variation coefficient estimated in each identified QTL at 1998. (b) The comparisons of additive, dominance and overdominance for the accumulation of average variation coefficient estimated in each identified QTL at 1999.

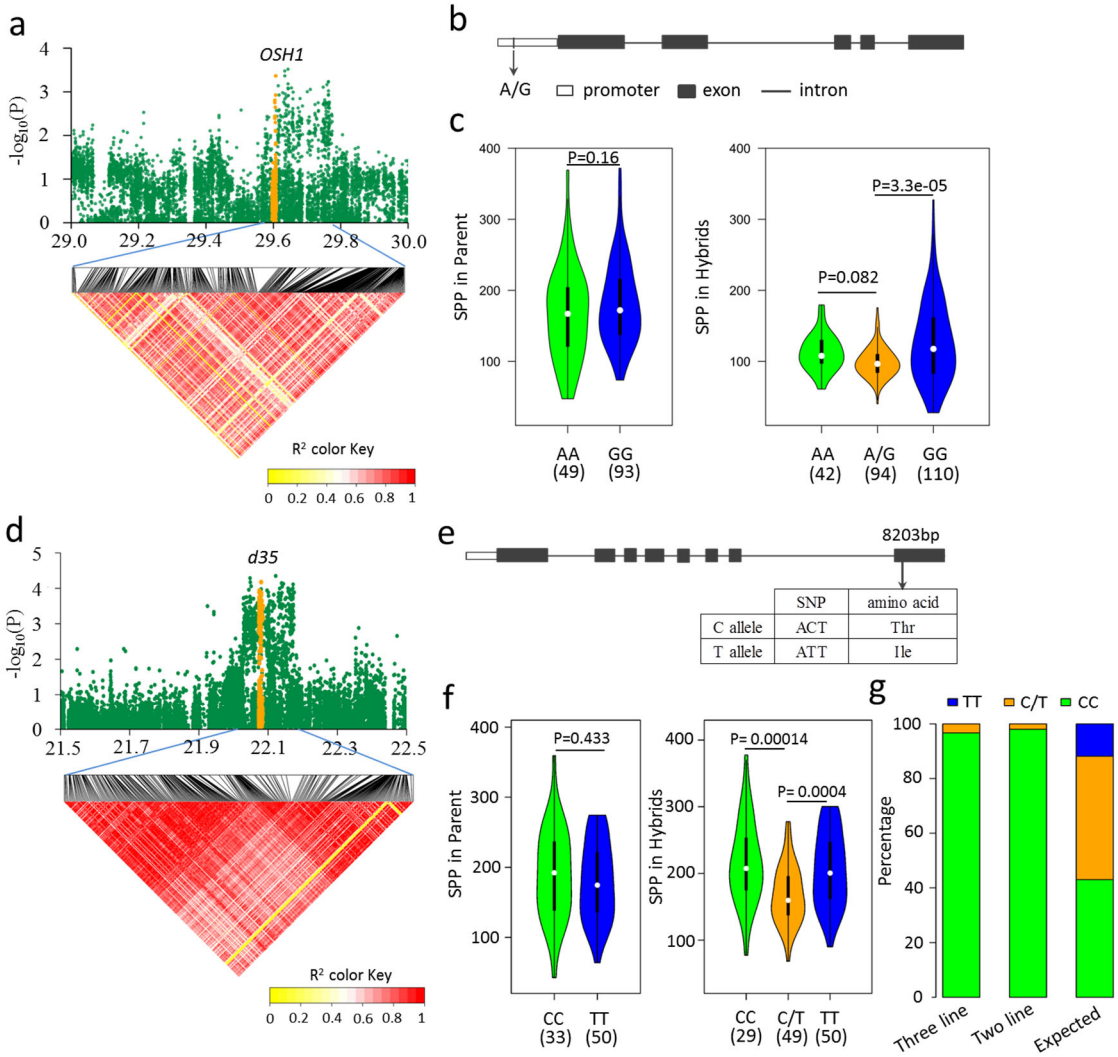


Supplementary Figure 52 The variance of different traits among environments for different degrees of dominant effects. (a) The accumulated variation coefficient of 6 yield related trait in Changsha. (b) The accumulated variation coefficient of 6 yield related trait in Sanya.



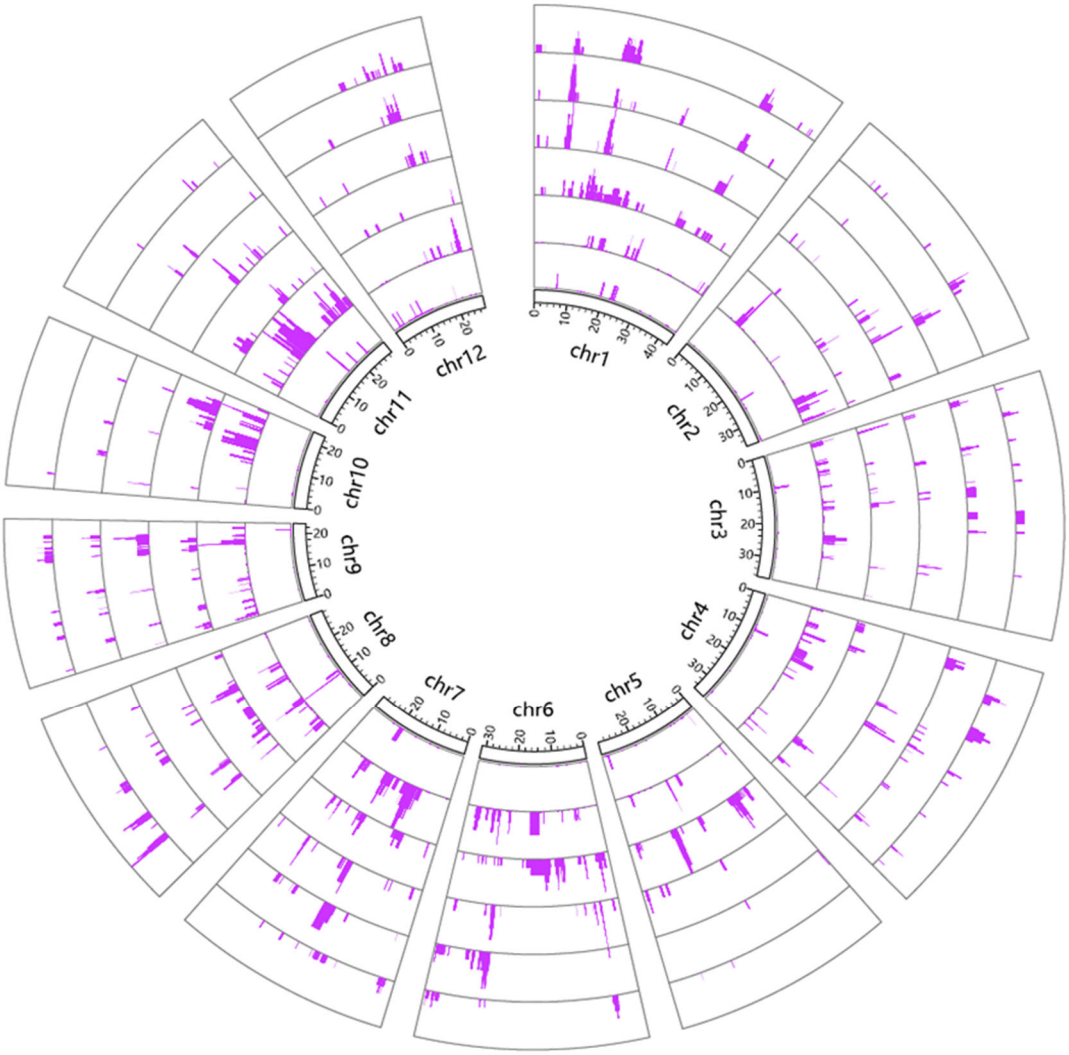
Supplementary Figure 53 Experimental validation of HoIIB model *Saccharomyces cerevisiae.* (a) Primer design principles and requirements. SSU1_Ck-F and SSU1_Ck-R are located on both side of the transcription factor *FZF1* recognition motif in the *SSU1* gene promoter. SSU1-LM-Ck-R and SSU1-RM-Ck-F are located in inserted marker (marker can be Leu or His); HRR-SSU1-F and HRR-SSU1-R had 38bp homologous sequences on both sides of the *FZF1* recognition motif in *SSU1* promoter, and the outer 21bp sequence was the upstream and downstream primers for screening markers genes on the amplified plasmid (pfa6a-leu1mx or pFA6a-His3MX6). (b) 1 was genomic fragment containing the *FZF1* recognition motif in the wild type (BY4743), 2 and 3 were the genomic fragment (2.1kb and 1kb respectively) of heterozygous mutant that one copy of *FZF1* recognition motif was substituted by Leu and the other was remain unchanged; 4 and 5 were the genomic fragment (1.8kb and 1kb respectively) of heterozygous mutant that one copy of *FZF*1 binding motif was substituted by His and the other was remain unchanged; Leu-LM was the primer used to amplify the left DNA fragment of Leu substitution genotype, Leu-RM was the primer used to amplify the right DNA fragment of Leu substitution genotype, His-LM was the primer used to amplify the left DNA fragment of His substitution genotype, Leu-RM was the primer used to amplify the right DNA fragment of Leu substitution genotype. (c) 1 was genomic fragment containing the *FZF1* recognition motif in the wild type (BY4743), 2 and 3 were the genomic fragment (2.1kb and 1kb respectively) of diploid mutant that one copy of *FZF1* recognition motif was substituted by Leu and the other was substituted by His; (d) The relative expression of gene *SSU1* in

different *SSU1* genotypes under different expression levels of its transcription factor (*FZF1*) in *Saccharomyces cerevisiae* BY4743; here, AA, aa and Aa represent the homologous genotype of wild type, the homologous genotype of mutant, and their heterozygous genotype, respectively; OE(0-5) means the strain with upregulated *FZF1* by 0-5 folds, and similar for OE(5-10) and OE(>10), and Empty means the strain with empty vector free of *FZF1*. (e) The dramatically decreased dominance degree of *SSU1* along with the increase of upregulation levels of its transcription factor *FZF1* in *Saccharomyces cerevisiae* BY4743.

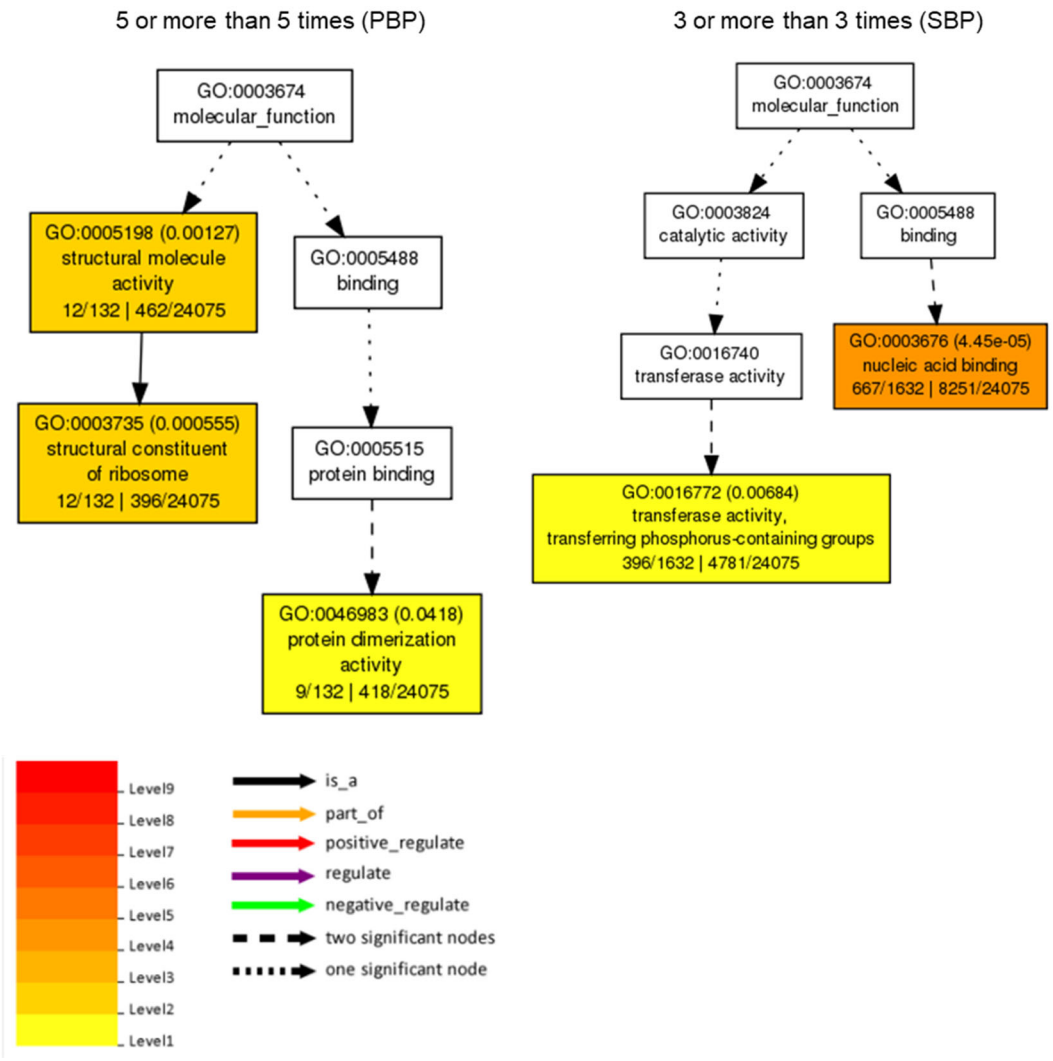


Supplementary Figure 54 The association of previously identified genes related to rice spikelet number per panicle (SPP). (a) The associated QTL and LD heat map within the QTL for SPP gene *OSH1*. The orange dots are the SNPs within gene *OSH1*. (b) The structure and the peak association signal of gene *OSH1*. The SNP A/G on promoter significantly associates with the spikelet number per panicle in F_1 of J×Nip. (c) Violin plots of *OSH1* genotypes for SPP in parents and hybrids. (d) The associated QTL and LD heat map within the QTL for SPP gene *d35*. The orange dots are the SNPs within gene *d35*. (e) The structure and the peak association signal of gene *d35*. The SNP C/T on the last exon significantly associates with the spikelet number per panicle in F_1 of J×Nip. (f) Violion plots of *d35* genotypes for SPP in parents and hybrids. (g) The genotype frequency of the SNP C/T in *d35* in the three-line and two-line

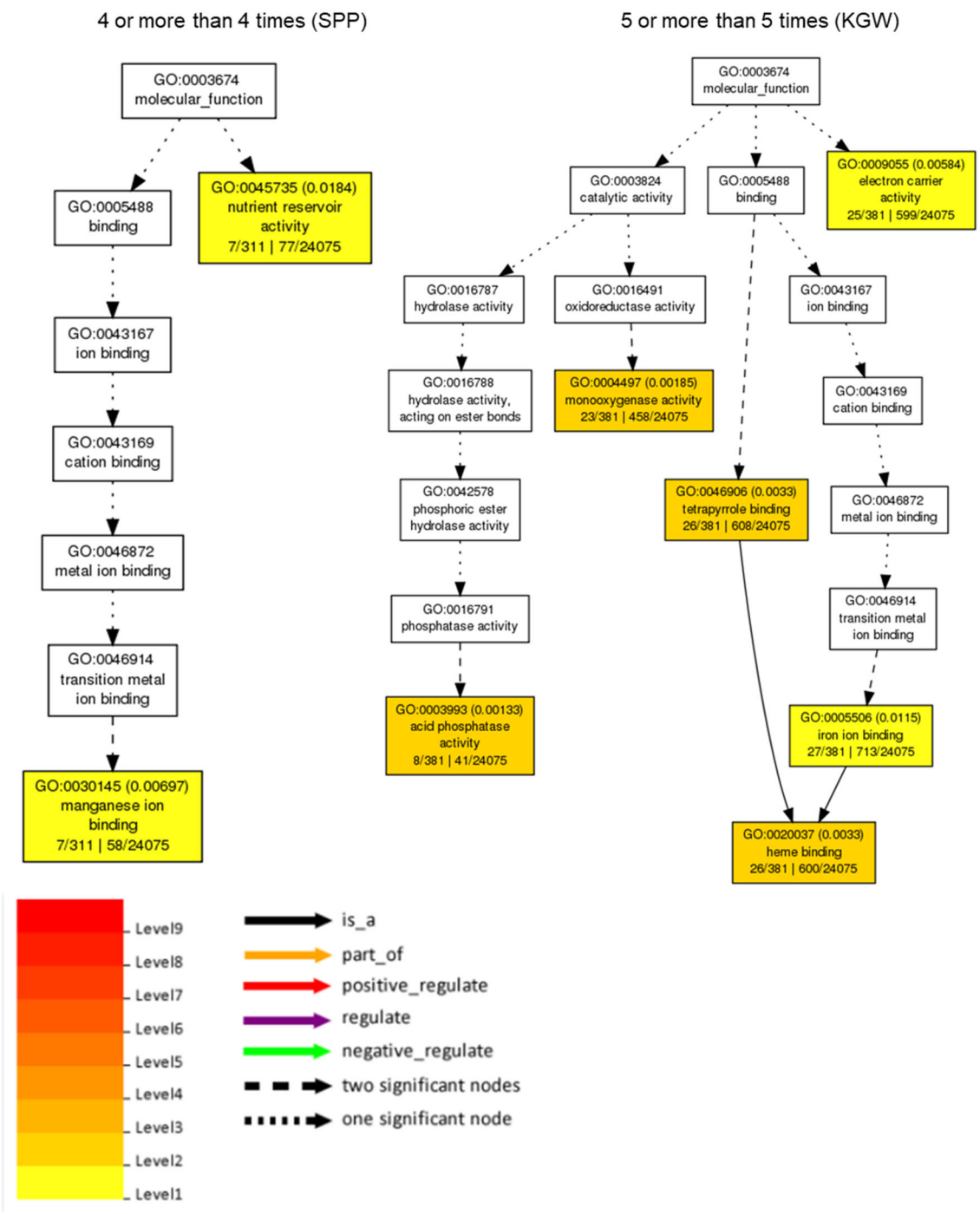
hybrids of *indica* subspecies, and the expected genotype frequency in combinations of *indica* varieties.



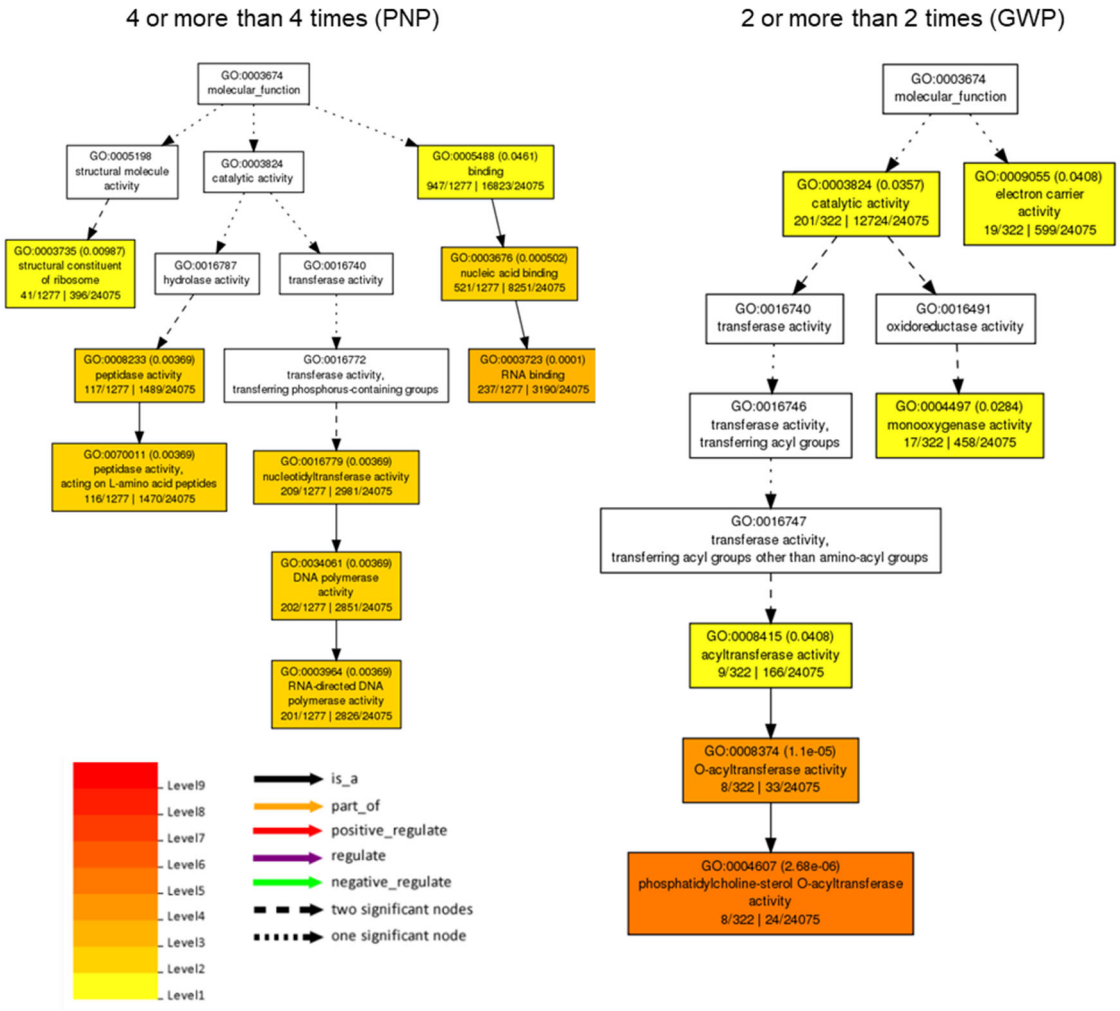
Supplementary Figure 55 The times of associated genes that can be repeatedly identified in the dominance and over-dominance QTLs across four kinds of combinations and two environments. Each line represents one associated gene. From the inner to outer layer, the height of lines represents 2-4 (for GWP), 3-6 (for PNP), 3-8 (for KGW), 3-6 (for SPP), 3-7 (for SBP) and 3-8 (for PBP) times, respectively.



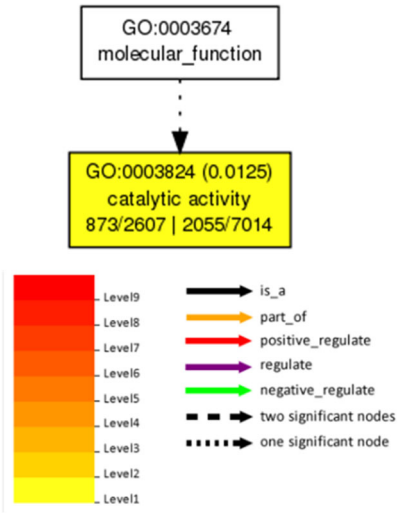
Supplementary Figure 56 The GO enrichment of repeated identified non-additive genes in PBP, SBP, SPP, KGW, PNP and GWP across four kinds of combination and two environments for rice. Only the term of molecular function was showed in figure, the other results were prepared in **Supplementary table 14**.



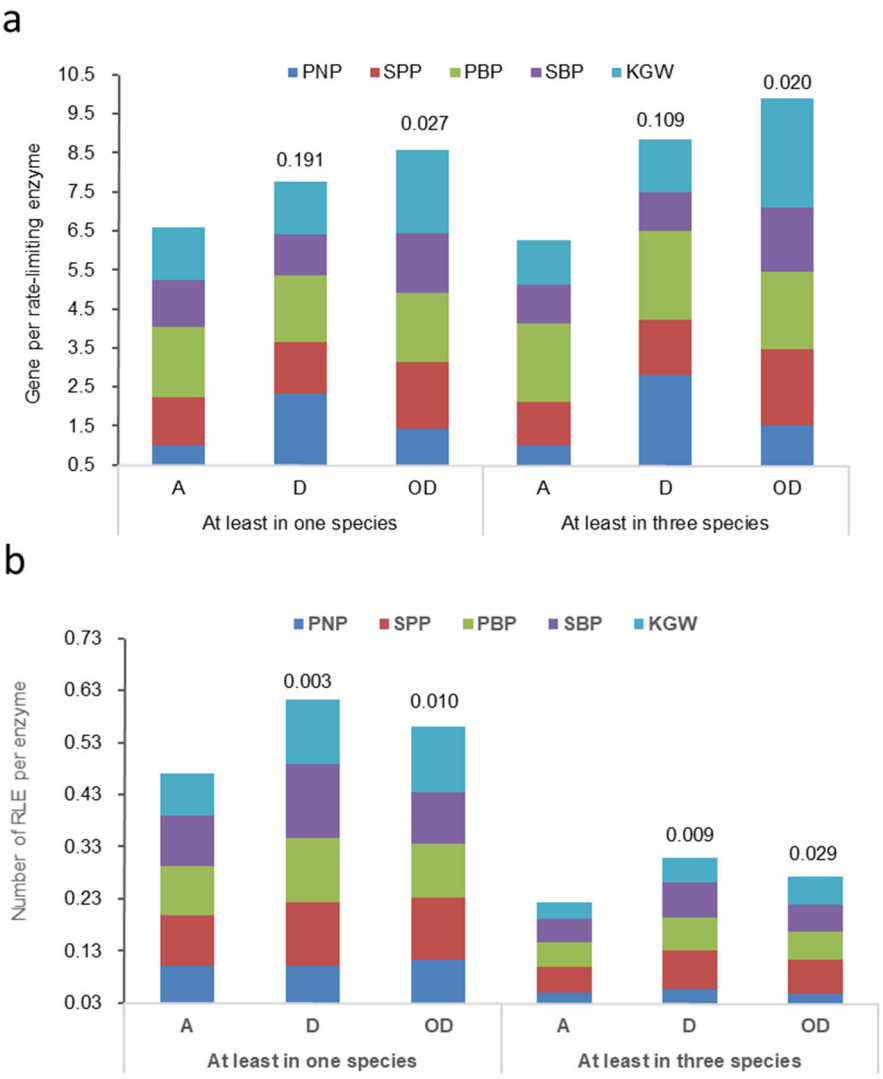
Supplementary Figure 56 continued



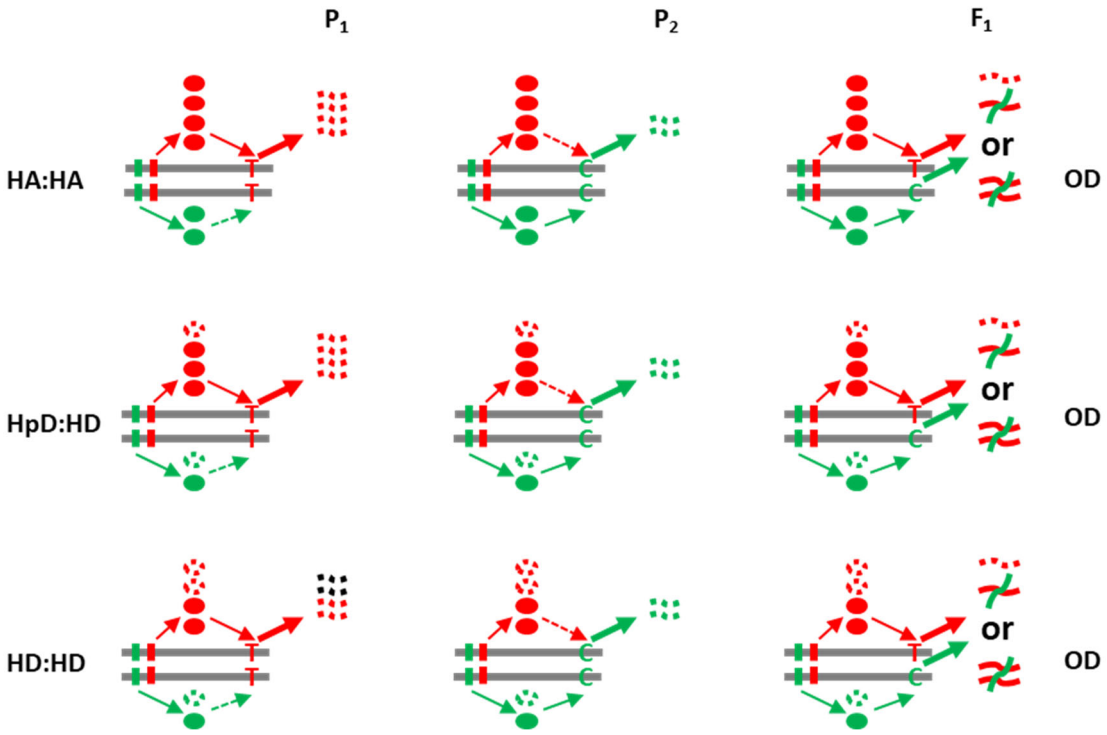
Supplementary Figure 56 continued



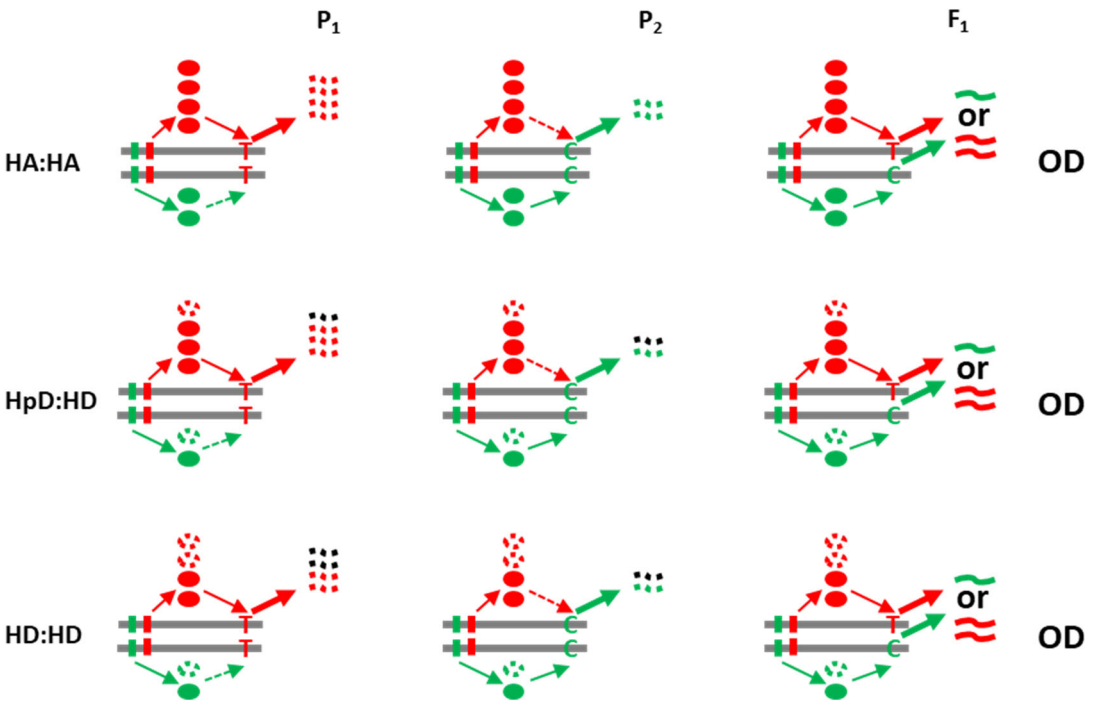
Supplementary figure 57 The GO enrichment of repeated identified genes with non-additive performance in non-lethal deletion yeast strains grown in five media. Only the term of molecular function was showed in the figure.



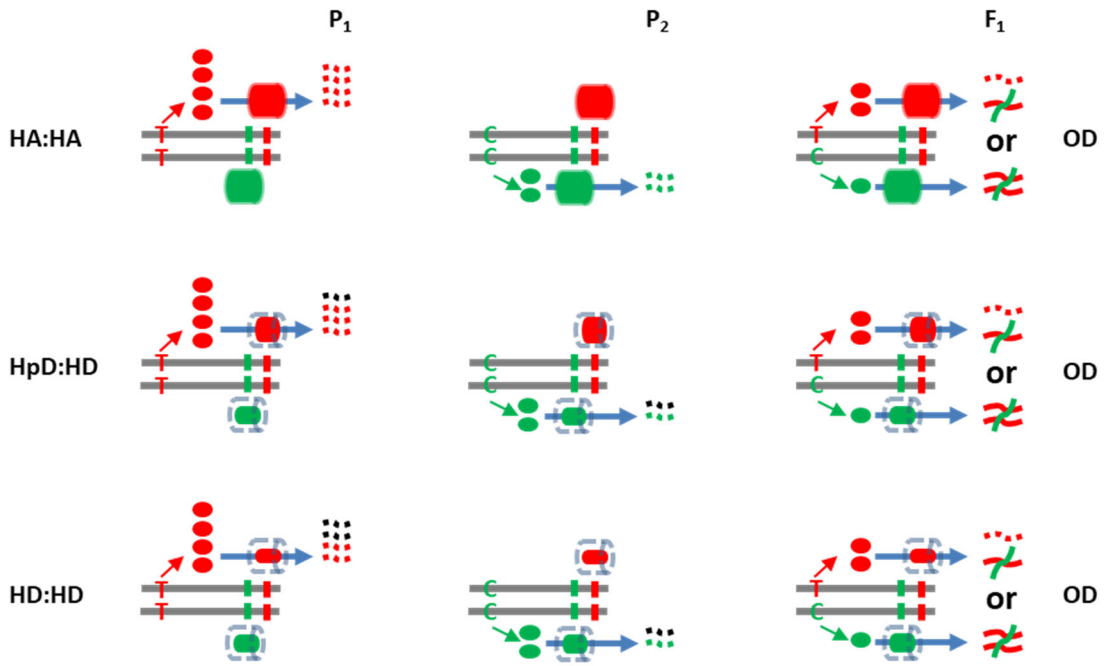
Supplementary figure 58 The distribution of rate-limiting enzymes coded by candidate genes within additive, dominant and over-dominant QTLs related to five yield component trait including PBP, SBP, SPP, KGW and PNP. The rate-limiting enzymes were identified in species yeast, mouse, and Human. (a) the number of genes per enzyme for additive, dominance and over-dominance candidate genes. (b) the number of genes per rate limiting enzyme for additive, dominance and over-dominance candidate genes. The P value at the top of bar means significant difference from the additive.



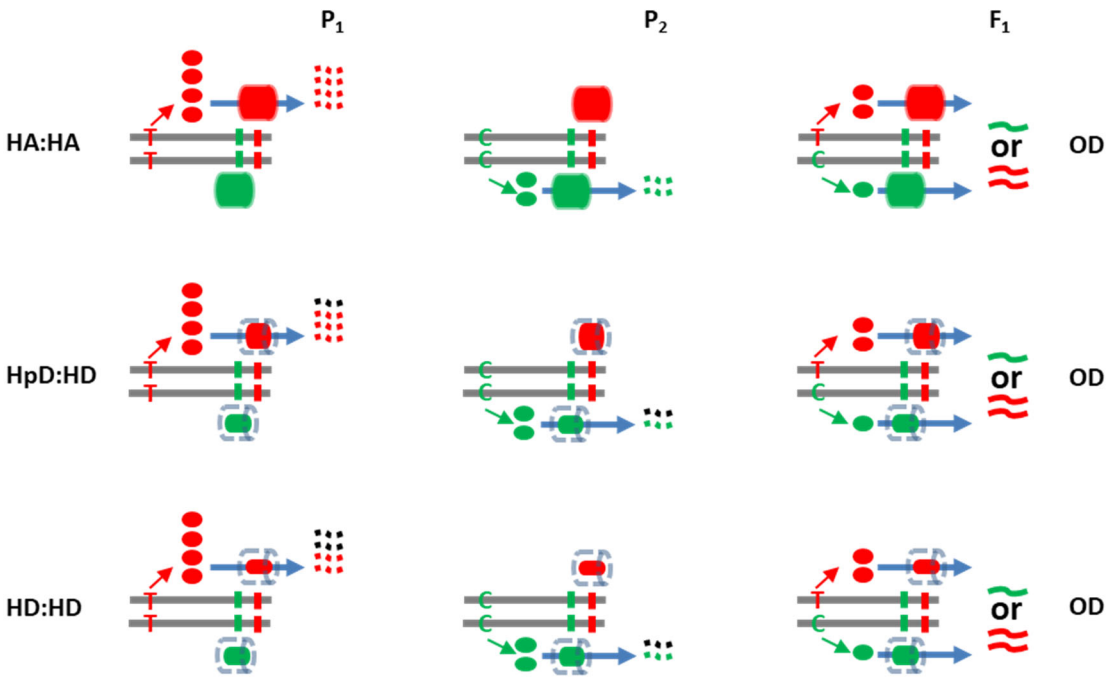
Supplementary Figure 59 The schematic diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the parallel complementation of two alleles of one polymorphic site under two independent regulators as the upstream backgrounds. The grey thick lines show two chromosomes, the red and green bars on which represent the homologous alleles of the regulators of T and C alleles at target site, respectively; and the function of these regulators keep constant among P_1, P_2 and F_1. The break and solid pies together represent the required regulator function that can maximize the output of the homozygote of the corresponding target allele, and the solid pies represent different regulator functions and thus provide different backgrounds to the target allele. The colored and black break curves together represent the maximum output of the homozygote in parents or one allele in F_1. The output of T without C or output of C without T does not take effect, and thus represented by the red or green break curves respectively for allele T or C; and the combined outputs of T and C can only take effect, as indicated by the cross solid curves. HA:HA means the complementation of the hidden additive effect of allele T and the hidden additive effect of allele C, in that the output of allele T is additive but non-functional itself and that of allele C does too; and HpD:HD and HD:HD are similar. HpD means hiden partial dominant.



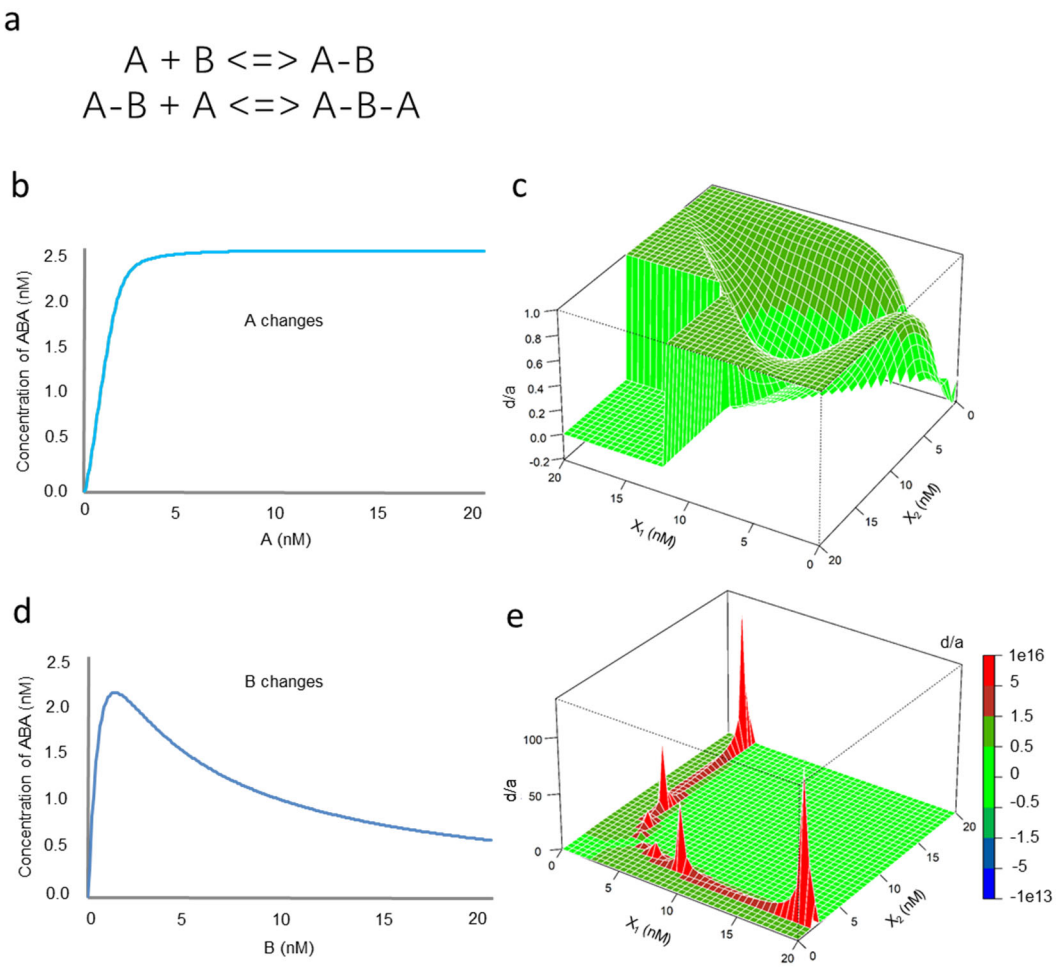
Supplementary Figure 60 The schematic diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the sequential complementation of two alleles of one polymorphic site under two independent regulators as the upstream backgrounds. The grey thick lines show two chromosomes, the red and green bars on which represent the homologous alleles of the regulators of T and C alleles at target site, respectively; and the function of these regulators keep constant among P_1, P_2 and F_1. The break and solid pies together represent the required regulator function that can maximize the output of the homozygote of the corresponding target allele, and the solid pies represent different regulator functions and thus provide different backgrounds to the target allele. The colored and black break curves together represent the maximum output of the homozygote in parents or one allele in F_1. The output of T without C or output of C without T does not take effect, and thus represented by the red or green break curves respectively for allele T or C; and the output of T following that of C or the output of C following that of T can only take effect, as indicated by the solid curves. HA:HA means the complementation of the hidden additive effect of allele T and the hidden additive effect of allele C, in that the output of allele T is additive but non-functional itself and that of allele C does too; and HpD:HD and HD:HD are similar. HpD means hidden partial dominant.



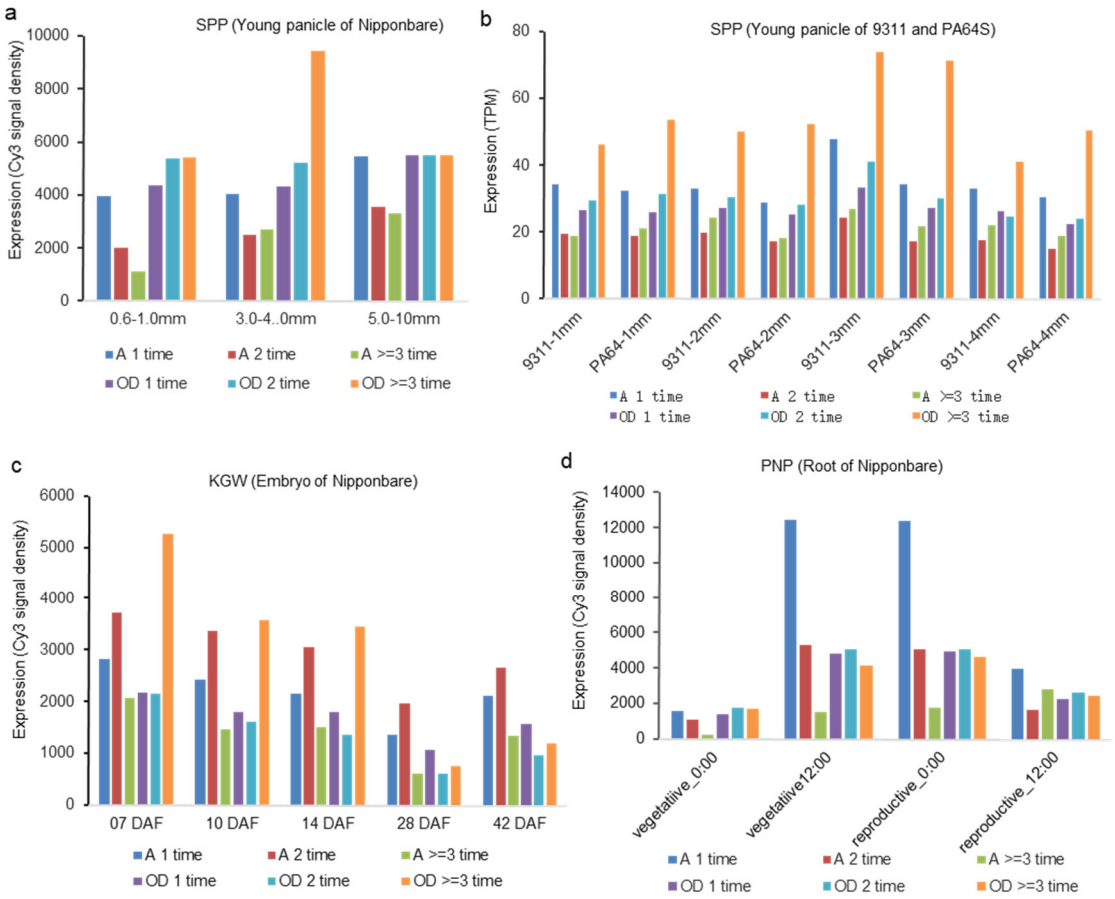
Supplementary Figure 61 The schematic diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the parallel complementation of two alleles of one polymorphic site under two independent regulators or responsors as the downstream backgrounds. The grey thick lines show two chromosomes, the red and green bars on which represent the homologous alleles of the regulators or responsors of T and C alleles at target site, respectively; and the function of these regulators or resposors keep constant among P_1, P_2 and F_1. Different numbers of red and green pies represent the maximum outputs of the homozygotes of allele T and C of target site, respectively. The doted cylinder or the same size of solid cylinder represent the required regulator or responsor function that can transform the full maximum output of the homozygote of the corresponding target allele, with red corresponding to allele T and green to allele C; and the solid cylinders in doted cylinder represent different regulator or response functions and thus provide different backgrounds to the target allele. The arrow represents the function process. The transformed output of T without that of C or the transformed output of C without that of T does not take effect, and thus represented by the red or green break curves respectively for allele T or C; and the combined transformed outputs of T and C can only take effect, as indicated by the cross solid curves. HA:HA means the complementation of the hidden additive effect of allele T and the hidden additive effect of allele C, in that the output of allele T is additive but non-functional itself and that of allele C does too; and HpD:HD and HD:HD are similar. HpD means hidden partial dominant.



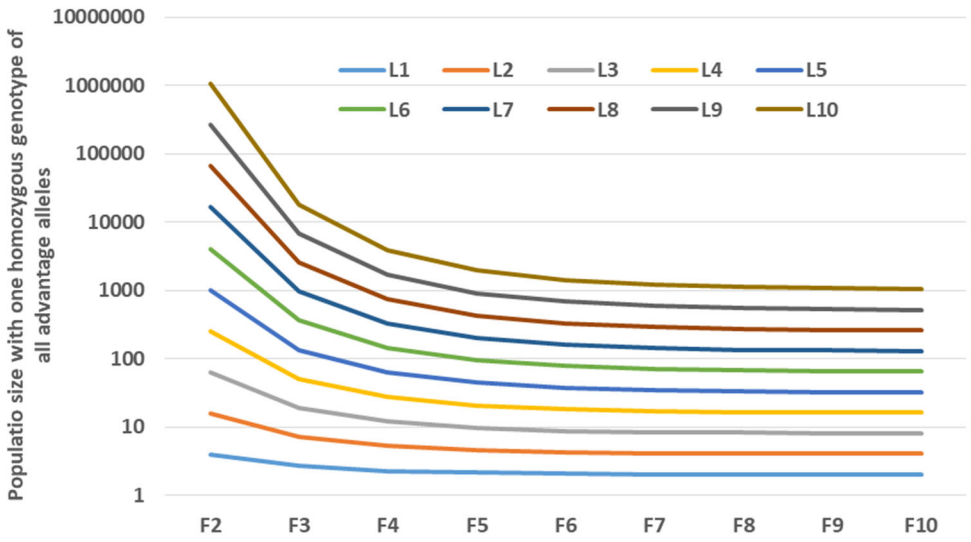
Supplementary Figure 62 The schematic diagram of regulation model for molecular mechanism of additive, dominant and over-dominant effect produced by the sequential complementation of two alleles of one polymorphic site under two independent regulators or responsors as the downstream backgrounds. The grey thick lines show two chromosomes, the red and green bars on which represent the homologous alleles of the regulators or responsors of T and C alleles at target site, respectively; and the function of these regulators or resposors keep constant among P_1, P_2 and F_1. Different numbers of red and green pies represent the maximum outputs of the homozygotes of allele T and C of target site, respectively. The doted cylinder or the same size of solid cylinder represent the required regulator or responsor function that can transform the full maximum output of the homozygote of the corresponding target allele, with red corresponding to allele T and green to allele C; and the solid cylinders in doted cylinder represent different regulator or response functions and thus provide different backgrounds to the target allele. The arrow represents the function process. The transformed output of T without that of C or the transformed output of C without that of T does not take effect, and thus represented by the red or green break curves respectively for allele T or C; and the transformed outputs of T following the transformed output of C or the transformed outputs of C following the transformed out of T can only take effect, as indicated by the solid curves. HA:HA means the complementation of the hidden additive effect of allele T and the hidden additive effect of allele C, in that the transformed output of allele T is additive but non-functional itself and that of allele C does too; and HpD:HD and HD:HD are similar. HpD means hidden partial dominant.



Supplementary Figure 63 The simulated diagram of regulation model for molecular mechanism of additive and dominant effect produced by single site in relation to the assembly of functional trimer ABA. (a) The schematic diagram depicting the assembly of trimer ABA. In this model, the nonfunctional dimer AB can compete component B against the functional trimer ABA. (b) The simulated curve of ABA concentration in the equilibrium state with different concentrations of subcomponent A and keeping subcomponent B constant. The sigmodial curve indicates that the sensitivity of ABA concentration to the concentration of A will decrease along with the decrease of background sufficiency as denoted by the increased ratio between A and B, and a nearly complete dominant effect will occur when the concentration of A reaches nearly 2X that of B. (c) The dominant degree of the timer ABA produced by different concentrations of subcomponent A in two parents (X_1 and X_2). (d) The simulated curve of ABA concentration in the equilibrium state with different concentrations of subcomponent B and keeping subcomponent A constant. The curve indicates that the sensitivity of ABA concentration to the concentration of B will decrease and nearly lost along with the decrease of background sufficiency as denoted by the increased ratio between B and A. (e) The dominant degree of the timer ABA produced by different concentrations of subcomponent B in two parents (X_1 and X_2). All the simulated results indicated that the non-additive effect usually generates under insufficient background. Detail of the model and the values of parameters can be found in **supplementary note.**



Supplementary Figure 64 The expression level of associated candidate genes within additive (A) and over dominant (OD) QTLs. (a) The expression level of associated candidate genes of SPP QTLs in young panicles of Nipponbare. (b) The expression level of candidate genes of SPP QTLs in the young panicles of 9311 and PA64S. (c) The expression level of associated candidate genes of KWG QTLs in the embryo of Nipponbare. (d) The expression level of associated candidate genes of PNP QTLs in the root of Nipponbare. The raw data of gene expression in (a), (c) and (d) were obtained from the database of RiceXpro. 1 time, 2 time and >=3 times means those candidate genes within the QTLs that can be detected in 1, 2, and 3 or more than 3 situations among four combinations under two environments.



Supplementary Figure 65 The theoretical population size with at least one homozygous genotype of all advantage alleles in different generations of two parents with different numbers of polymorphic loci (from 1 to 10). Here, we calculate the population under the hypothesis that there is no linkage between loci and all loci are randomly combined; L1 - L10 mean the locus number from 1 to 10; F2 - F10 mean the self-crossing generations from 2 to 10.