**Extended Data Fig. 1 Transcriptomic differences in two inbred lines (Balilla and Yuetai B) and their reciprocal F$_1$ hybrids at three developmental stages. a,** Correlations between seedling length and weight of two inbred lines and their F$_1$ hybrids at three developmental stages. *P* values indicate Pearson correlation. **b,** Analysis of variance of seedling length by genotype and stage. **c,** Numbers of differentially expressed genes between F$_1$ hybrids and their corresponding parents. *P* value is for the independent samples *t* test. **d**, Numbers of differentially expressed and non-differentially expressed genes (DEGs and NDEGs). **e,** Venn diagram of DEGs at 5, 10 and 15 days after sowing (DAS). **f-g**, PCA score plots of Balilla, YB, their F$_1$ hybrids, and middle parent values (MVP) based on NDEGs (**f**, 16,350 genes) and DEGs (**g**, 5,844 genes). **h-i,** Venn diagrams of DEGs between the two inbred lines and between F$_1$ hybrids and corresponding parents. The numbers and percentages marked in red represent the amount of differentially expressed genes from the parental differential genes. **j,** Numbers of up- and downregulated genes. *P* values indicate paired samples *t* tests. **k,** Venn diagrams of differential terms at 15 DAS. The numbers marked in red represent the number of differential terms from the parental differential terms. **l-p,** The significantly enriched terms between two inbred lines and between F$_1$ hybrids and their parents at 15 days after sowing.
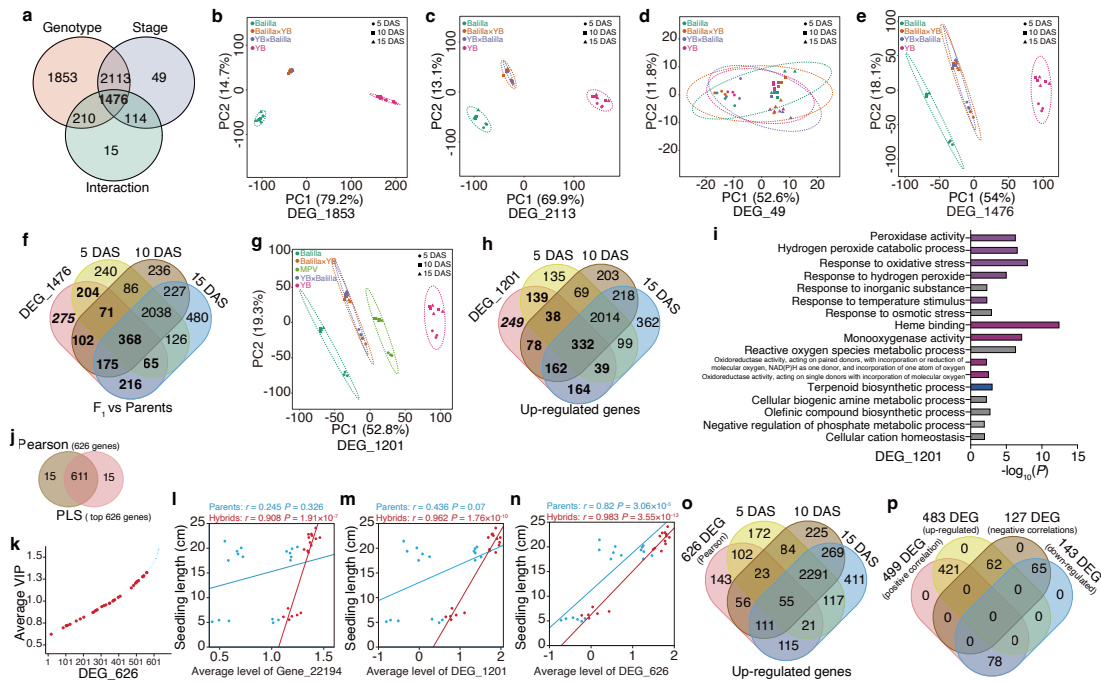
Extended Data Fig. 2 The identification of heterosis-associated genes related to seedling length from the differentially expressed genes between F₁ hybrids and their corresponding parents. a, Venn diagram of differentially expressed genes (DEGs) affected by genotype, stage, and their interaction. **b-e**, PCA score plots of Balilla, Yuetai B (YB), and reciprocal F₁ hybrids based on different types of DEGs. The differentially expressed genes were divided into different types by analysis of variance with genotype and stage as factors. **f**, The 1,476 DEGs included both DEGs between parents and DEGs between F₁ hybrids and parents. **g**, PCA score plot of Balilla, YB, their F₁ hybrids, and the middle parent values (MPVs) based on 1,201 DEGs. **h,** Venn diagram of the 1,201 DEGs with upregulated gene expression levels. **i,** Significantly enriched terms of the 1,201 DEGs. **j,** Venn diagram of 626 significantly correlated genes and 626 top contributing genes according to partial least squares (PLS) analysis of seedling length. The significance level of Pearson correlation was set at 0.05. **k,** Contributions of the 626 top contributing genes. The 37 heterosis-associated genes from 12 terms are marked in red. **l-n,** Correlations between gene expression levels and seedling length in parents and F₁ hybrids. *P* values indicate Pearson correlation. **o,** Venn diagram of the upregulated genes of the 626 DEGs. **p,** Most of the upregulated genes had positive correlations with seedling length.
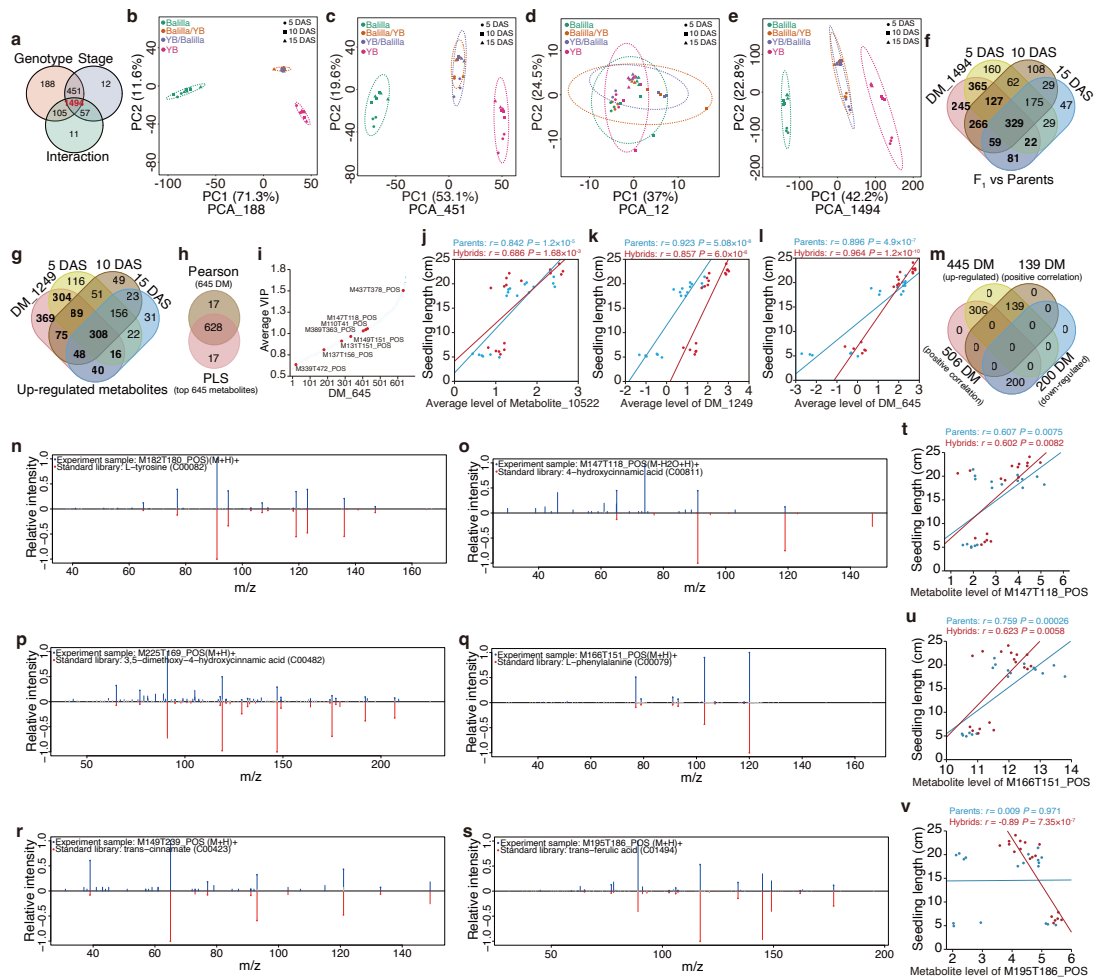
Extended Data Fig. 3 The identification of heterosis-associated metabolites related to seedling length from differential metabolites between F₁ hybrids and their corresponding parents. a, Venn diagram of differential metabolites (DMs) affected by genotype, stage, and their interaction. **b-e,** PCA score plots of Balilla, YB, and their reciprocal F₁ hybrids based on different types of DMs. **f,** Both DMs between parents and between F₁ hybrids and their parents were included in 1,494 DMs. **g,** Most of the 1,249 DMs had upregulated metabolite levels in F₁ hybrids at single or multiple stages. **h,** Venn diagram of 645 significantly correlated genes and 645 top contributing genes according to partial least squares (PLS) analysis of seedling length. The significance level of Pearson correlation was set at 0.05. **i,** Contributions of the 645 top contributing metabolites. The eight metabolites from three enriched pathways are marked in red. **j-l,** Correlations between average metabolite levels and seedling length in parents and F₁ hybrids. **m,** Most of the upregulated metabolites had positive correlations with seedling length. **n-s,** MS/MS spectra of six metabolites. MS/MS spectra of the experimental samples and chemical standards are shown above and below the horizontal line, respectively. **t-v,** Correlations between metabolite levels of three chemically identified metabolites and seedling length. *P* values indicate Pearson correlation in **j-l** and **t-v**.
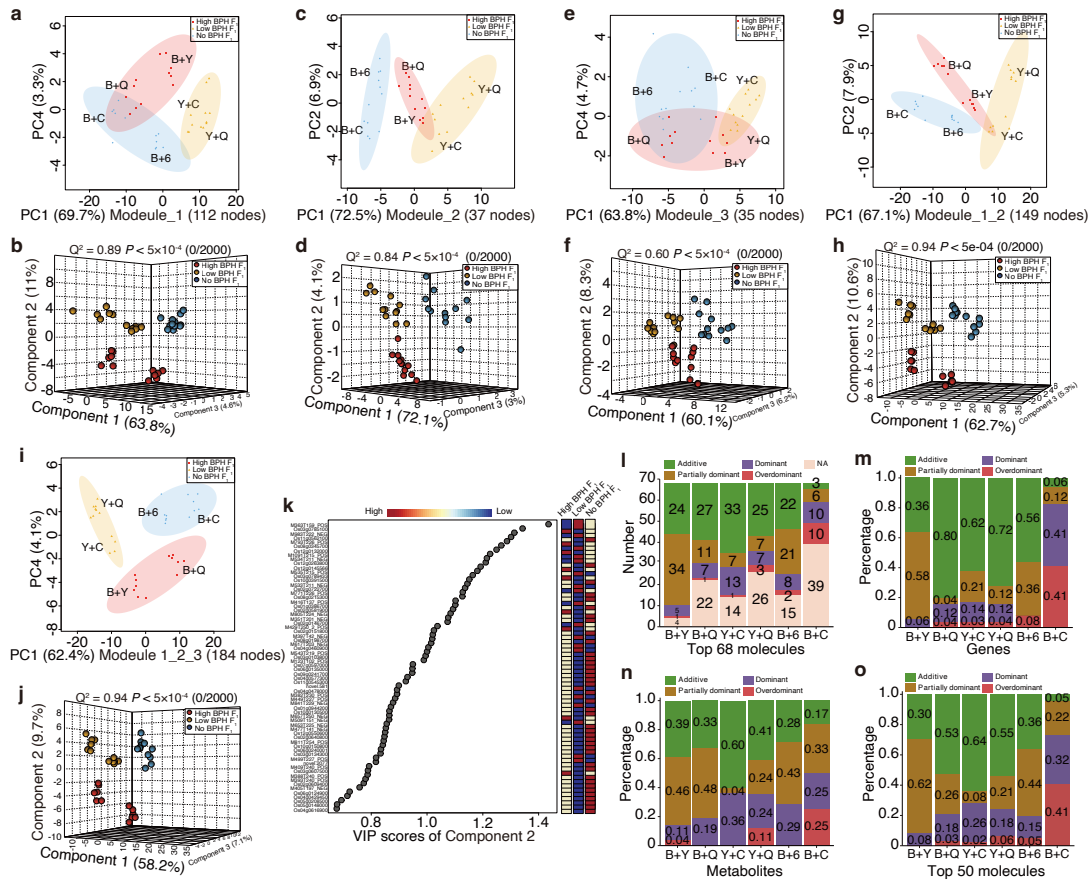
Extended Data Fig. 4 Inheritance patterns of heterosis-associated molecules for seedling length. a-j, PCA and PLS-DA score plots of six pairs of reciprocal F_1 hybrids based on the molecular levels of Module_1, Module_2, Module_3, Module_1 plus Module_2, and the integration of Module_1, Module_2, and Module_3. **k,** Variable importance in projection (VIP) scores for Component 2. The PLS-DA model was built in **h** and the molecular levels of 68 heterosis-associated molecules in the high-, low-, and no better-parent heterosis groups are shown on the right side. **l,** Numbers of different types of inheritance patterns of the 68 heterosis-associated molecules. **m,** Percentages of different inheritance patterns of 39 heterosis-associated genes. **n,** Percentages of different inheritance patterns of 29 heterosis-associated metabolites. **o,** Percentages of different inheritance patterns of the top 50 heterosis-associated molecules.
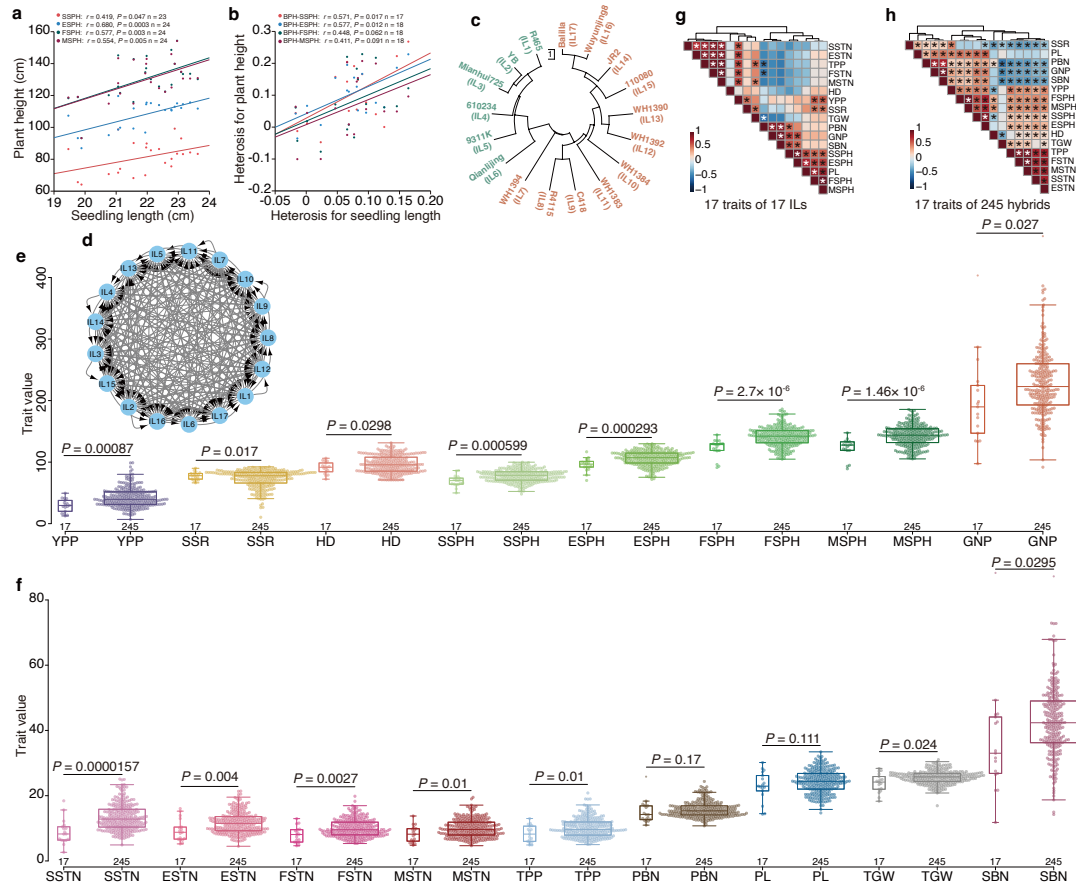
Extended Data Fig. 5 Phenotypes of seventeen agronomic traits of the diallel cross population. a, Correlations between seedling length at 15 days after sowing and plant height at four developmental stages. **b,** Correlations between heterosis for seedling length at 15 days after sowing and heterosis for plant height at four developmental stages. *P* values are for Pearson correlations in **a** and **b**. One of the F_1 hybrid, namely, C418×Qianlijing, was excluded from the analysis because it showed a large difference from its reciprocal hybrid. **c,** Neighbour-joining tree of 17 ILs based on InDel markers. **d,** Diallel cross design of 17 inbred lines (ILs). **e-f,** Comparisons of 17 traits between 17 ILs and 245 corresponding F_1 hybrids. *P* values indicate independent samples *t* tests. **g-h,** Heatmaps of correlations among 17 traits of 17 ILs and among those of 245 F_1 hybrids. Better-parent heterosis = BPH, yield per plant = YPP, secondary branch number = SBN, grain number per panicle = GNP, tiller number per plant = TPP, seed setting rate = SSR, heading date = HD, thousand grain weight = TGW, seedling stage tiller number = SSTN, elongation stage tiller number = ESTN, flowering stage tiller number = FSTN, maturation stage tiller number = MSTN, seedling stage plant height = SSPH, elongation stage plant height = ESPH, flowering stage plant height = FSPH, maturation stage plant height = MSPH, panicle length = PL, and primary branch number = PBN.
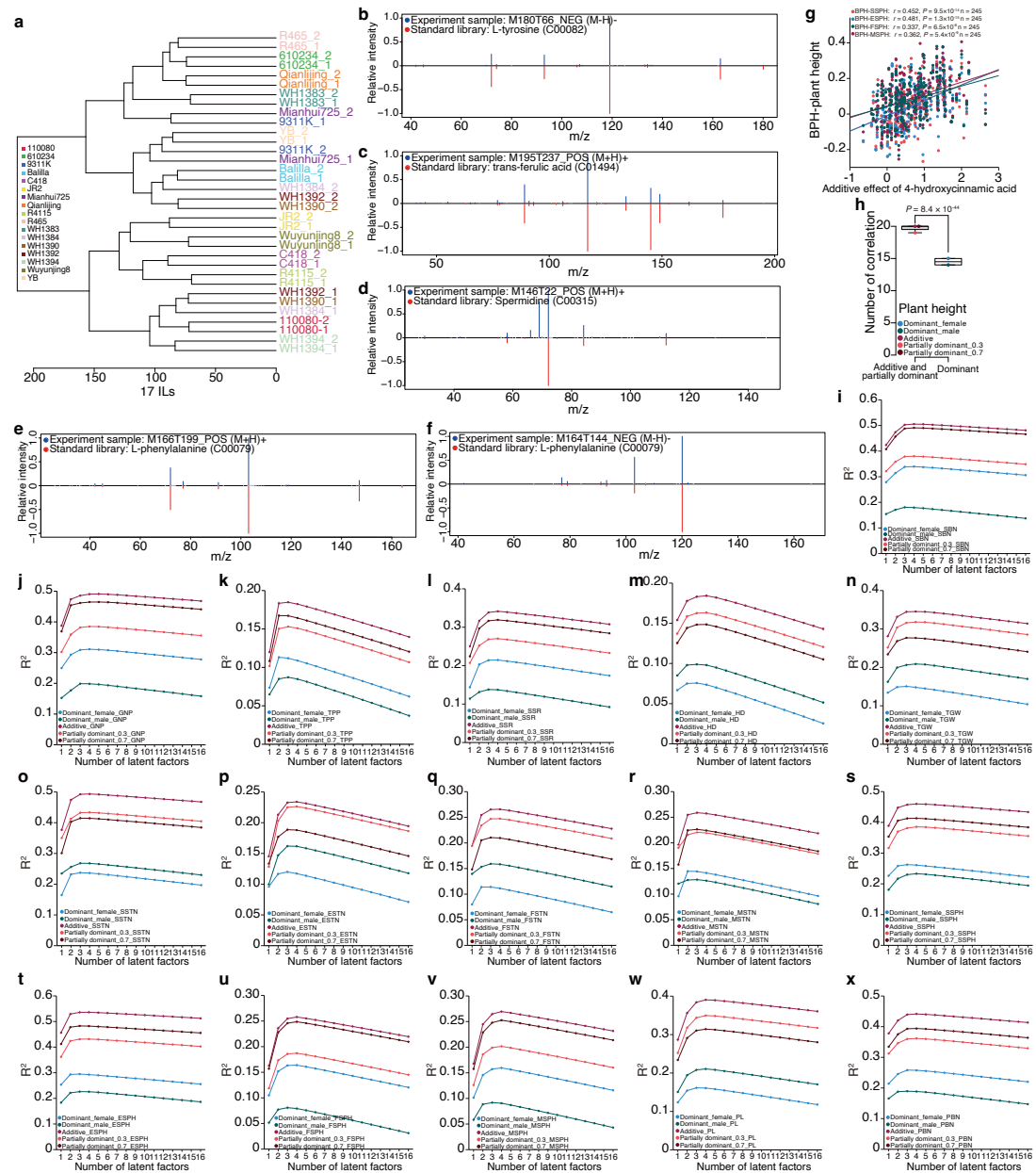
Extended Data Fig. 6 Predictive models for heterosis of 16 traits based on three types of inheritance patterns of 1,306 differential metabolites. a, Dendrogram of 17 ILs based on untargeted metabolite profiles. **b-f,** MS/MS spectra of four metabolites. **g,** Correlations between with additive effect of 4-hydroxycinnamic acid and heterosis for plant height. *P* values are for Pearson correlation. **h,** Comparisons of additive and partially dominant effects to dominant effect on plant height. *P* values indicate independent samples *t* tests. **i-x,** PLS-based predictive models for heterosis of 16 traits with different inheritance patterns of 1,306 differential metabolites from 17 ILs. The R^2 indicates the adjusted value.
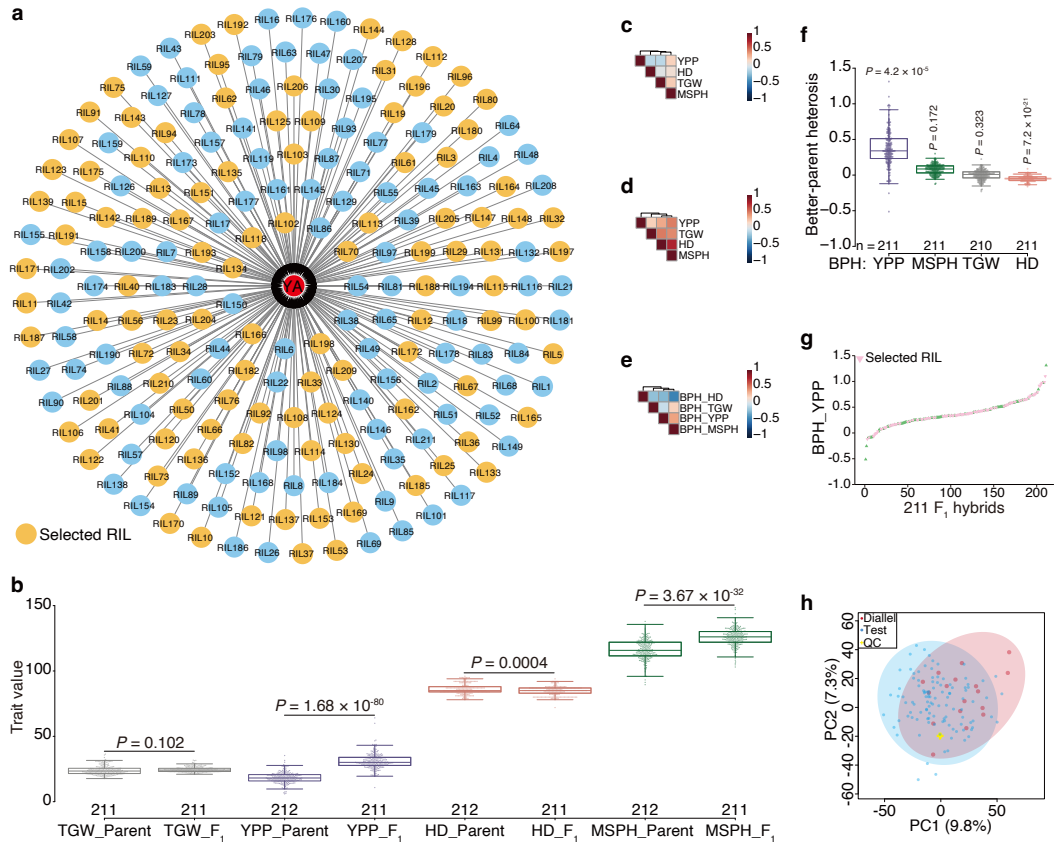
Extended Data Fig. 7 Cross design and four traits of the test cross population. a, Cross design of the test cross population. Recombinant inbred line = RIL. Yuetai A = YA. **b,** Comparisons of four traits between F_1 hybrids and their corresponding parents. *P* values are for independent samples *t* tests. **c-e,** Heatmaps of correlations among four traits of the parents, among four traits of 211 F_1 hybrids, and among better-parent heterosis for four traits of 211 F_1 hybrids. **f,** BPH for four traits of the test cross population. *P* values indicate independent samples *t* tests between the diallel and test cross populations. **g,** Better-parent heterosis for grain yield per plant of F_1 hybrids from 211 RILs. The RILs selected for metabolite profiling analysis are marked in pink. **h,** PCA score plot of parents for the diallel (17 ILs) and test cross (107 RILs and Yuetai A) populations based on untargeted metabolomics data. Quality control = QC.
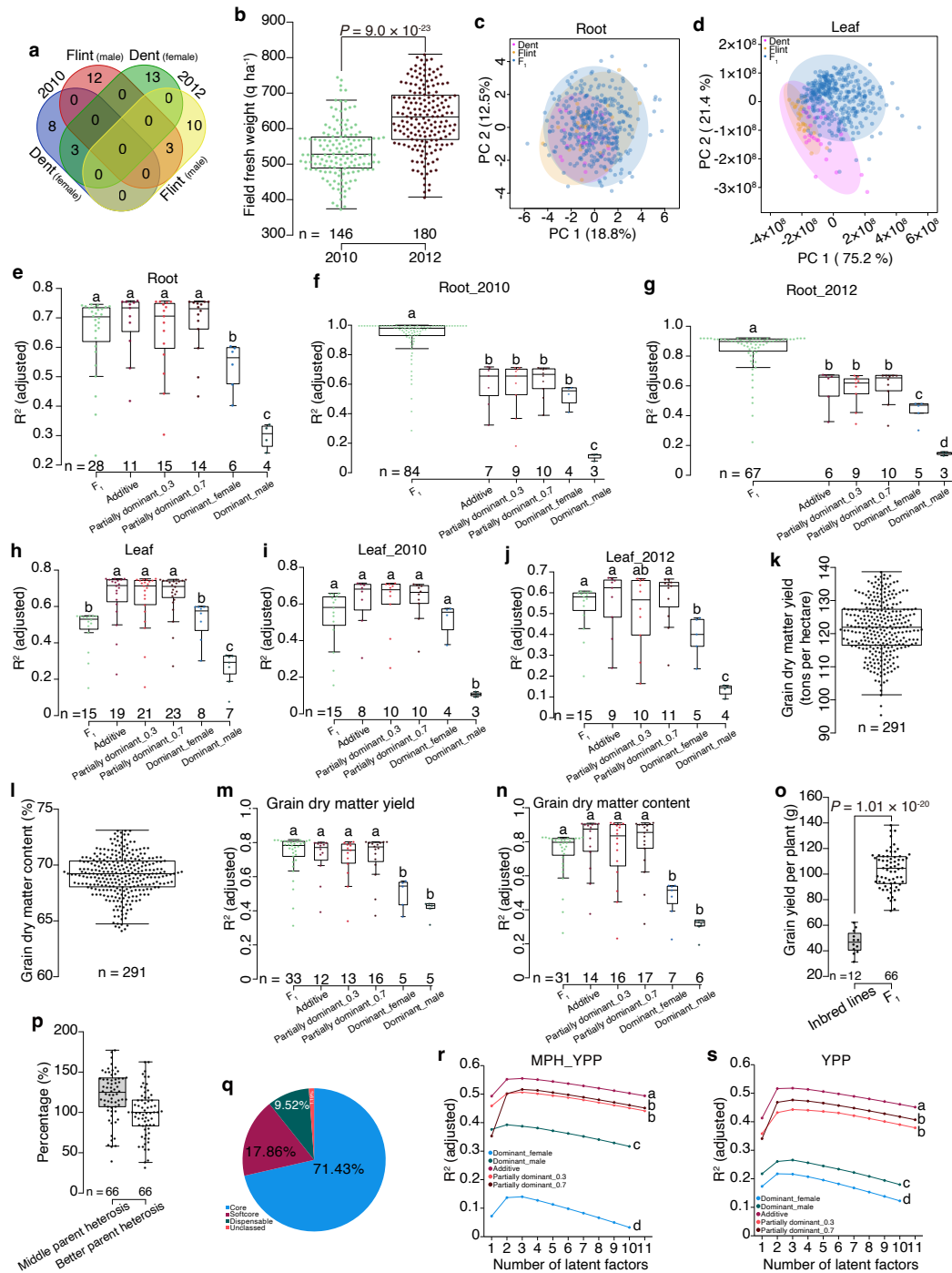
Extended Data Fig. 8 Additive and partially dominant effects contribute to maize heterosis. a, Venn diagram of 24 Dent lines (female parents) and 25 Flint lines (male parents) for two maize hybrid populations in 2010 (146 F_1 hybrids) and 2012 (180 F_1 hybrids). **b,** Field fresh weight of the maize hybrid populations. *P* value indicates independent samples *t* test. **c-d,** PCA score plots of 24 Dent lines, 25 Flint lines, and their 326 F_1 hybrids (Dent × Flint) based on metabolic data obtained from young roots and leaf samples. **e-j,** R^2 of PLS-based predictive models for field fresh weight based on hybrid profiles and different inheritance patterns of metabolic data. Metabolite profiles of the root (165 metabolites) and leaf samples (81 metabolites) were used to analyse the field fresh weight. The two hybrid populations were combined for analysis

and were also analysed independently. **k-l,** Grain dry matter yield and grain day matter content of 291 maize F_1 hybrids. **m-n,** R^2 for grain dry matter yield and grain day matter content with hybrid profiles (F_1) and different inheritance patterns of root metabolic data. **o,** Grain yield per plant of 12 maize inbred lines and their 66 F_1 hybrids. *P* value indicates independent samples *t* test. **p,** Middle-parent and better-parent heterosis of grain yield per plant of 66 F_1 hybrids. **q,** Distribution of 84 differentially expressed genes in 12 inbred lines. **r,** R^2 for middle-parent heterosis of grain yield per plant (MPH_YPP) with different inheritance patterns of 84 differentially expressed genes. **s,** R^2 for grain yield per plant.
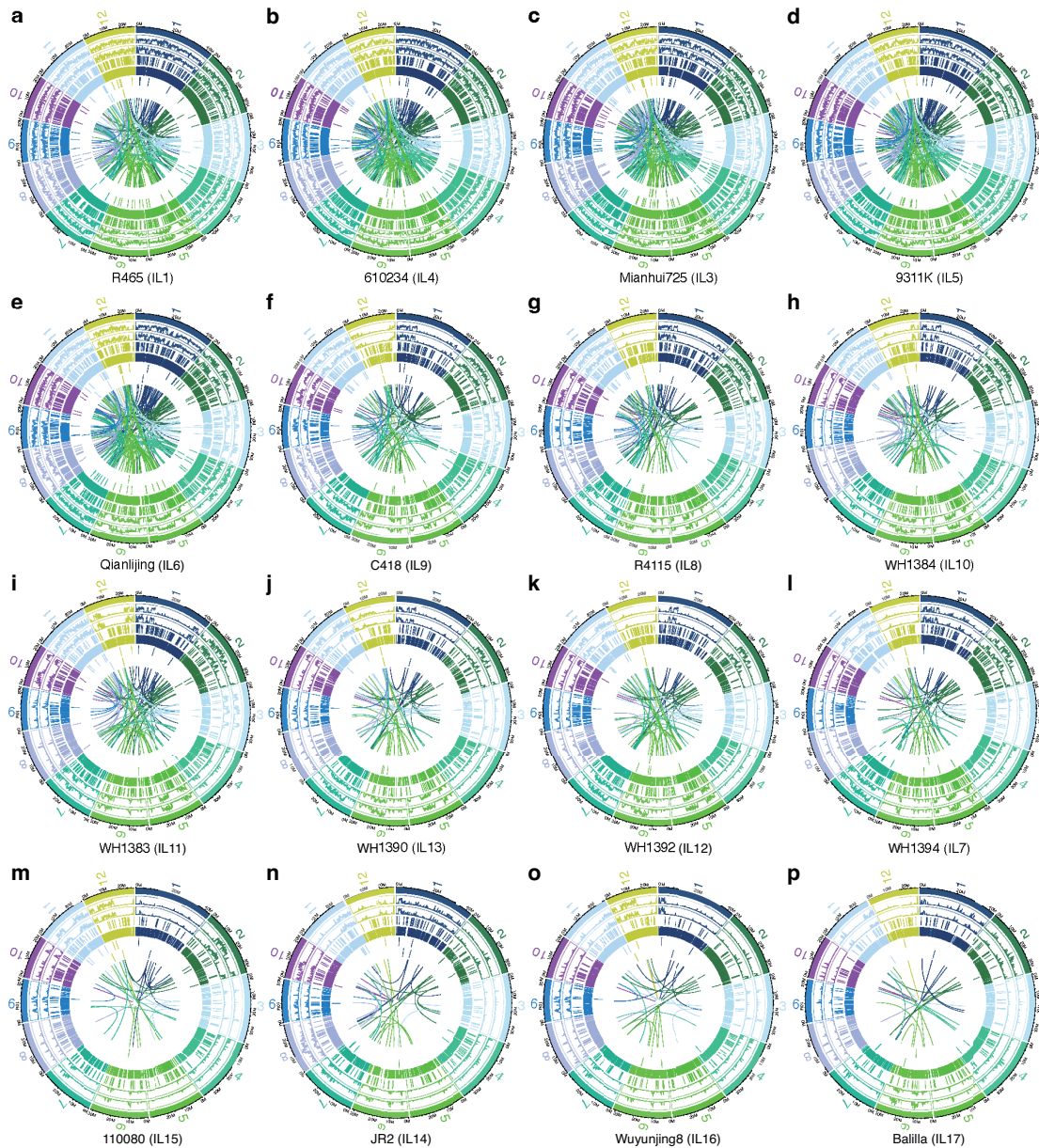
Extended Data Fig. 9 Four types of genomic variants of 16 ILs. Genomic variants, including SNPs, InDels, CNVs, and SVs, were obtained for 17 ILs via genomic analysis. The genomic variants are shown for R465 (IL1, **a**), 610234 (IL4, **b**), Mianhui725 (IL3, **c**), 9311K (IL5, **d**), Qianlijing (IL6, **e**), C418 (IL9, **f**), R4115 (IL8, **g**), WH1384 (IL10, **h**), WH1383 (IL11, **i**), WH1390 (IL13, **j**), WH1392 (IL12, **k**), WH1394 (IL7, **l**), 110080 (IL15, **m**), JR2 (IL14, **n**), Wuyunjing8 (IL16, **o**), and Balilla (IL17, **p**). The chromosome position, SNP density, InDel density, CNV duplication, CNV deletion position, SV insertion position, SV deletion position, and SV inversion position are listed from the outer to the inner circles.
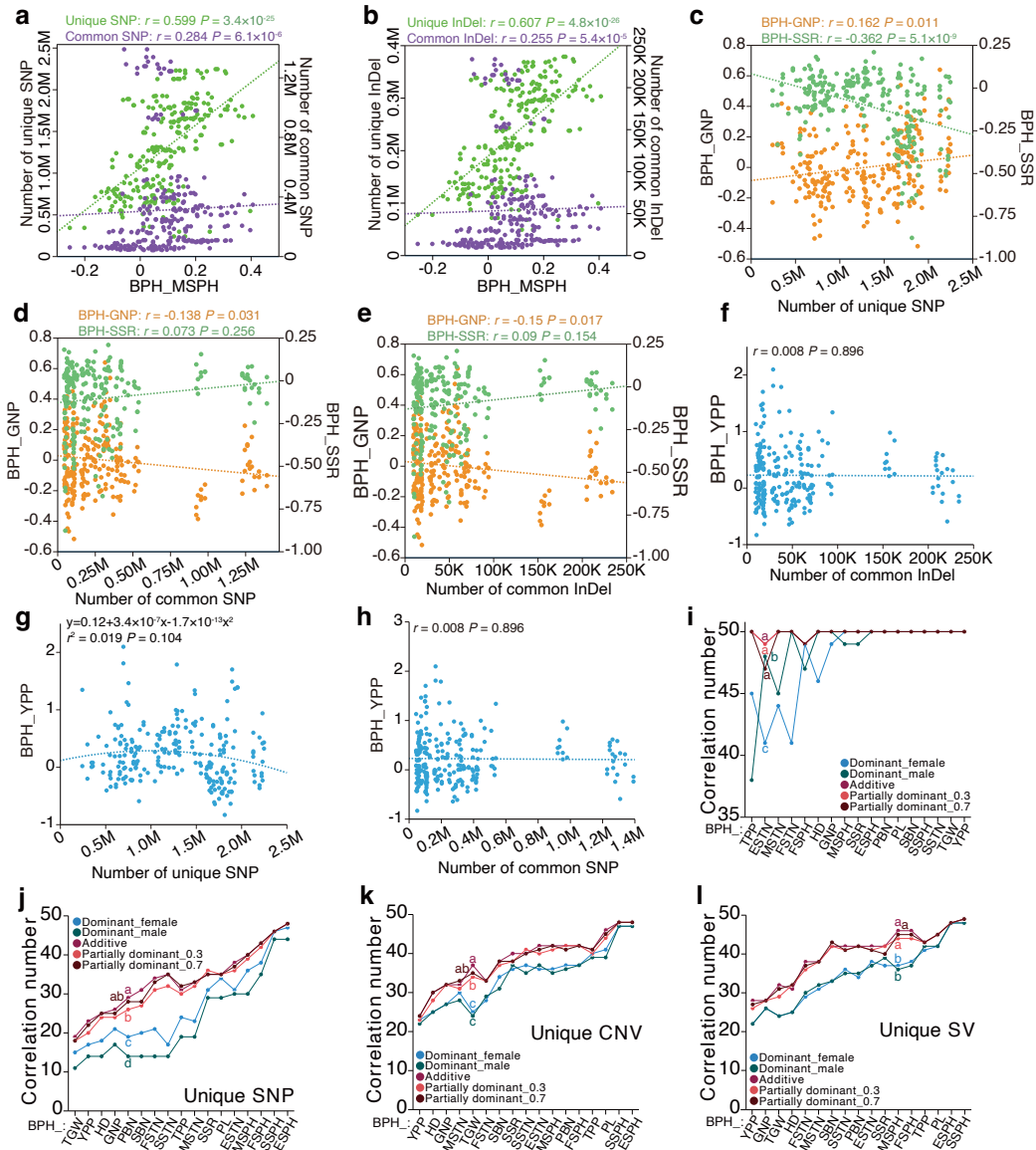
Extended Data Fig. 10 Correlations of genomic variants, heterosis, and different inheritance patterns. a-b, Scatter plots between better-parent heterosis for plant height at the maturation stage and the numbers of parental common and unique SNPs (**a**) and InDels (**b**). **c-d,** Scatter plots between the numbers of SNPs and better-parent heterosis for grain number per panicle (GNP) and the seed setting rate (SSR). *P* values are for Spearman rank correlation. **e,** Scatter plots between the numbers of parental common InDels and heterosis for grain number per plant and seed setting rate. **f,** Scatter plot between the numbers of parental common InDels and heterosis for grain yield per plant. **g-h**, Scatter plots between the number of SNPs and better-parent heterosis for grain yield per plant. *P* value is for quadratic regression in **g** and Spearman rank correlation in **h**. **i,** Correlation numbers for the three inheritance patterns and heterosis. **j-l,** Correlation numbers for the three inheritance patterns and different genomic variants. *P* value indicates Spearman rank correlation in **a-f**, and **h**.